medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Bias from heritable confounding in Mendelian randomization studies

Eleanor Sanderson[1,2], Dan Rosoff[1,3,4], Tom Palmer[1,2], Kate Tilling[1,2], George Davey Smith[1,2]* and Gibran Hemani[1,2]*

1. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.
2. Population Health Sciences, University of Bristol, Bristol, UK.
3. Section on Clinical Genomics and Experimental Therapeutics, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, USA
4. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

Corresponding author: Eleanor.sanderson@bristol.ac.uk
*These authors contributed equally to the project and are joint last author

Abstract

Mendelian randomization (MR) is an approach to causal inference that utilises genetic variants to obtain estimates of the causal effect of an exposure on an outcome in the presence of unobserved confounding. MR relies on a set of assumptions to obtain unbiased effect estimates, one of these assumptions is that there is no pathway from the genetic variants to the outcome that does not act through the exposure. Increasing genome-wide association study (GWAS) sample sizes for the exposure enables discovery of instrumental variables with smaller effect sizes. We illustrate through simulations how smaller effect sizes could arise from genetic variants that act through traits that have greater liability to confound an exposure-outcome relationship. When such genetic variants are selected as instruments this can bias the MR effect estimate obtained from that instrument in the same direction as the confounded observational association but with larger magnitude. Through simulation we illustrate how the total bias of the MR estimates increases across a range of standard MR estimation methods increases as the proportion of the genetic instruments that are associated with the confounder increases. However, if such heritable confounders are known and can be instrumented, the confounder free effect estimate can be obtained through applying a pre-estimation filtering to standard MR methods, removing instruments that explain more variation in that confounder than the exposure, or by estimating effects through multivariable MR. We highlight the potential for SNPs identified in GWAS to be associated with potential confounders through examination of a recent GWAS of C-Reactive Protein. Finally, we illustrate our approach through estimation of the causal effect of age at menarche on type 2 diabetes, hypothesising that the MR effect estimate may be biased by confounding due to the inclusion of genetic variants associated with early life adiposity as instruments.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Introduction

Confounding in epidemiological studies can bias observed estimated associations between traits so that they cannot be interpreted as the causal effect of the exposure trait on the outcome. Adjusting for confounders will remove this bias provided that all confounders are known and measured without error, which is rarely plausible.[1] Mendelian Randomisation (MR) is an alternative approach to causal effect estimation which utilises genetic variants to infer causality of the exposure on the outcome that is not biased by that confounding.[2, 3] MR is most often implemented by using genetic variants associated with the exposure within an instrumental variable (IV) estimation framework.[4]

IV estimation relies on three key assumptions to be able to test for a causal effect of the exposure on the outcome. These assumptions are (IV1) the instrument is associated with the exposure, (IV2) there is no confounding of the instrument and the outcome and (IV3) there is no effect of the instrument on the outcome that doesn't act through the exposure. These assumptions are illustrated in **Figure 1A.** In MR genetic variants associated with the exposure of interest are used as IVs. These are usually single nucleotide polymorphisms (SNPs) identified through genome-wide association studies (GWAS).[5] GWAS estimate the association between millions of genetic variants and a trait of interest. Variants associated in a GWAS with the exposure trait at a predetermined level are then selected to be used as IVs in the MR estimation. The level chosen is typically set to select all independent variants where the p-value for the estimated association between the variant and the exposure is smaller than $5x10^{-8}$. This level is chosen as it is the Bonferroni corrected 5% significance level, given the effective number of independent tests being conducted in a GWAS.

The polygenic nature of complex traits, for which there is substantial empirical support[6], posits that the majority of heritable variation is driven by a very large number of genetic factors, each with very small effect sizes. To discover such variants GWAS has been growing in sample size rapidly in recent years, especially through the emergence of large scale biobanks. The most recent GWAS of body mass index (BMI) included approximately 700,000 individuals and identified 941 independent SNPs associated with BMI.[7] The most recent GWAS of educational attainment included approximately 3 million individuals and identified nearly 4000 independent SNPs.[8] Similarly, this growing availability of large sample sizes means that for some traits where previously no SNPs were identified as genome-wide significantly associated with the trait (i.e. no SNPs had a p-value $< 5x10^{-8}$) large studies can now identify enough SNPs to conduct MR estimation.

For MR analyses involving exposures that are themselves complex traits, it is well understood that genetic variants associated with the exposure are mediated through other phenotypic layers.[9] Such mediation opens the possibility of horizontal pleiotropy if those mediating traits also have effects on the outcome through pathways independent of the exposure. If some, or all, of the variants associated with the exposure exhibit horizontal pleiotropy the MR effect estimates can be biased through violation of assumption IV3. There are a number of different forms that horizontal pleiotropy can take and so bias MR.[4] Correlated pleiotropy is used to describe pleiotropy that causes bias in MR which is correlated with the magnitude of association between the SNPs and the exposure. Correlated pleiotropy can occur through heritable confounding, misspecification of the primary phenotype or reverse causation, illustrated in **Figure 1B.** The distinction between heritable

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

confounding and misspecification of the primary phenotype (i.e. the forms of pleiotropy illustrated by **Figure 1B(ii)** and **1B(iii)**) is not always clear, especially as one may lead to the other. If the genetic variant for a heritable confounder is taken to be an instrument for the exposure that the heritable confounder influences then this misspecification of the primary phenotype will clearly lead to erroneous inference, with the extreme case being when the upstream phenotype is the outcome under investigation, as in **Figure 1B(iv).** As each of these mechanisms lead to the same bias in the resulting MR estimates and misspecification of the primary phenotype could be considered to be an extreme form of heritable confounding. Note however that for there to be bias in the MR estimation the path from the confounder to the outcome is required. If there is no effect of the confounder on the outcome that does not act through the exposure then MR estimates will not be biased. Throughout the rest of this paper we consider correlated pleiotropy driven by genetic variants associated with a confounder of the exposure and outcome being included within the set of instruments used in an MR analysis and refer to this as heritable confounding.

There remains a question of whether increasing GWAS sample sizes will lead to increased horizontal pleiotropy through heritable confounding, i.e. are smaller genetic effects for the exposure more likely to also have an independent path to the outcome for any given exposure-outcome relationship? One mechanism by which this could arise is if genetic variants that have effects on the exposure that are mediated through other traits tend to be those that are more phenotypically connected in the overall genotype-phenotype map. Such variants would be likely to have smaller effects on a trait and so only become discoverable as GWAS sample sizes increase. One implication of this is that the genetic variants with the largest effect sizes in the GWAS may be the least likely to be pleiotropic instruments.[10] Although the extent to which this will hold will depend on the genetic architecture of the trait. If all of the pathways from those upstream mediating traits to the outcome act through the exposure then the MR effect estimates will not be biased and the increased GWAS sample size will increase the power of the MR estimate to detect causal effects. However, when one (or more) of the upstream traits is a confounder of the exposure and outcome this will lead to violation of the IV3 assumption and biased MR effect estimates.

There are a number of MR methods available that are robust to violations of assumption IV3, each of which assume particular forms for the structure of the pleiotropy[4]. The most commonly used methods are MR Egger[11], weighted Median[12] and weighted Mode[13]. MR Egger assumes that the size of any pleiotropic effect of each SNP on the outcome is uncorrelated with the SNP-exposure association. Under correlated pleiotropy this assumption will not hold as the size of the pleiotropic effect will be associated with the size of the SNP-exposure association. Weighted Median assumes that more than half of the SNPs (weighted by their inverse variance weights) are not pleiotropic. Weighted Mode assumes that the modal effect estimate obtained by the weighted SNPs is not biased by pleiotropy and so assumes that a plurality of the SNPs are not pleiotropic. The assumptions made by these methods mean that the weighted mode and weighted median approaches will be robust to a small proportion of the SNPs included in the estimation being associated with a confounder of the exposure and outcome but will both be biased as the proportion of SNPs associated with confounders of the exposure and outcome increases. IVW and MR Egger will be biased if any of the SNPs are associated with the confounder, although the level of the bias will depend on the proportion of the SNPs used as instruments that are associated with the

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

confounder, and the presence/absence of another confounder balancing the bias in the opposite direction. This result has been illustrated elsewhere.[14]

Alternative methods of MR estimation are available that adjust for correlated pleiotropy under particular assumptions of the structure the pleiotropy takes.[14-21] These approaches are often computationally intensive, using Bayesian estimation approaches applied to a large number of genetic variants. They also impose an assumption of a structure of the relationship between the pleiotropic variants and other variants, often assuming that the plurality of the instruments are valid.

Here we propose two alternative approaches to estimation in the presence of correlated pleiotropy when the confounder or confounders are hypothesised and measured; including the confounder in a multivariable MR (MVMR) and removing the SNPs associated with the confounder through confounder-based Steiger filtering. Each approach relies on knowledge of the likely confounders, however when this is unknown the other traits associated with the genetic instruments, and so most likely to be confounders, can be identified through PheWAS analysis.[22] The advantage of these approaches are that they are simple to implement and already widely used in MR analyses. They therefore provide a simple approach to explore the potential for bias from heritable confounding in MR studies.

The remainder of this paper is structured as follows; we first compare the bias from confounding in IV estimation to linear regression. We explore the question of whether increasing GWAS sample sizes could increase the rate of correlated pleiotropy through a simulation of a network of traits explained by set of genetic variants. We then present a simulation study to illustrate the bias induced by correlated pleiotropy and how MVMR and confounder-based Steiger filtering can correct the bias. We consider the example of C-reactive protein (CRP) and explore how many potential confounders are likely to be associated with SNPs identified in the most recent large GWAS of CRP. Finally, we consider estimation of the causal effect of age at menarche on risk of type 2 diabetes and show that in this analysis the univariable IVW MR estimate is likely to be biased by correlated pleiotropy from SNPs associated with childhood adiposity.

Results

Bias in IVW MR estimates compared to linear regression

In this section we consider a simple model of a relationship between an exposure (X) and an outcome (Y) which is confounded by a single unobserved confounder (U). The exposure is associated with a single instrument (Z) that also has an effect on the outcome via the unobserved confounder. This single instrument could be considered to be a genetic risk score constructed from all SNPs identified as strongly associated with the exposure in a GWAS. We additionally assume that all relationships are linear and there are no interactions. We focus on the setting with a single heritable confounder for simplicity and clarity however in any application there would also be a number of non-heritable confounders as well as potentially more than one heritable confounder.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

The results given here therefore only compare the results for bias from a single heritable confounder and not the overall bias that would be expected. The model we consider can be written as;

$$Y = \beta_{0y} + \beta_{xy}X + \beta_{uy}U + v_y$$
$$X = \beta_{0x} + \pi_{zx}Z + \beta_{ux}U + v_x$$
$$U = \beta_{0u} + \pi_{zu}Z + v_u$$

Where v_y, v_x and v_u are all uncorrelated random error terms. This model is illustrated in **Figure 2A**.

In this model the estimate of the effect of the exposure on the outcome from linear regression is given by;

$$\hat{\beta}_{LR} = \frac{\sum XY}{\sum X^2}$$

$$\hat{\beta}_{LR} = \frac{\sum X(\beta_{xy}X + \beta_{uy}U + v_y)}{\sum X^2}$$

$$\hat{\beta}_{LR} = \beta_{xy} + \frac{\sum X(\beta_{uy}U + v_y)}{\sum X^2}$$

$$\hat{\beta}_{LR} = \beta_{xy} + \beta_{uy}\frac{\sum XU}{\sum X^2} + \frac{\sum Xv_y}{\sum X^2}$$

As we have assumed no other confounders in this simple model v_y is independent of the exposure X and so $E\left(\sum Xv_y\right) = 0$. We can additionally define δ_{ux} as the estimate that would be obtained from a regression of X on U. Note: this will not be the same as β_{ux} due to the confounding of U and X by Z but will estimate the correlation between X and U. Therefore;

$$E\left(\hat{\beta}_{LR}\right) = \beta_{xy} + \beta_{uy}\delta_{ux} \#(1)$$

This can alternatively be expressed as;

$$E\left(\hat{\beta}_{LR}\right) = \beta_{xy} + \frac{\beta_{uy}}{\beta_{ux}}R^2_{ux}$$

Where R^2_{ux} is the variation in X explained by U. i.e. the true causal effect of the exposure on the outcome plus a bias term that depends on the effect of the confounder on the outcome and the correlation between the confounder and the exposure.

The estimate of the effect of the exposure on the outcome from IV estimation is given by;

$$\beta_{IV} = \frac{\sum ZY}{\sum ZX}$$
$$\hat{\beta}_{IV} = \frac{\sum Z(\beta_{xy}X + \beta_{uy}U + v_y)}{\sum ZX}$$
$$\hat{\beta}_{IV} = \beta_{xy} + \frac{\sum Z(\beta_{uy}U + v_y)}{\sum ZX}$$

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$\hat{\beta}_{IV} = \beta_{xy} + \beta_{uy}\frac{\sum ZU}{\sum ZX} + \frac{\sum Zv_y}{\sum ZX}$$

$$\hat{\beta}_{IV} = \beta_{xy} + \beta_{uy}\frac{\sum ZU}{\sum Z(\pi_{zx}Z + \beta_{xu}U + v_x)} + \frac{\sum Zv_y}{\sum ZX}$$

$$\hat{\beta}_{IV} = \beta_{xy} + \beta_{uy}\frac{\sum Z(\pi_{zu}Z + v_u)}{\sum Z(\pi_{zx}Z + \beta_{xu}(\pi_{zu}Z + v_u) + v_x)} + \frac{\sum Zv_y}{\sum ZX}$$

$$\hat{\beta}_{IV} = \beta_{xy} + \beta_{uy}\frac{\sum(\pi_{zu}Z^2 + Zv_u)}{\sum(\pi_{zx}Z^2 + \beta_{xu}(\pi_{zu}Z^2 + Zv_u) + Zv_x)} + \frac{\sum Zv_y}{\sum ZX}$$

The instrument Z is independent and exogenous so taking expectations of this expression, $E\left(\sum Zv_y\right) = 0$, $E(\sum Zv_x) = 0$ and $E(\sum Zv_u) = 0$. Therefore;

$$E\left(\hat{\beta}_{IV}\right) = \beta_{xy} + \beta_{uy}\frac{\sum \pi_{zu}Z^2}{\sum(\pi_{zx}Z^2 + \beta_{xu}\pi_{zu}Z^2)}$$

$$E\left(\hat{\beta}_{IV}\right) = \beta_{xy} + \beta_{uy}\frac{\pi_{zu}}{\pi_{zx} + \beta_{xu}\pi_{zu}} \quad \#(2)$$

This is the true causal effect of the exposure on the outcome plus a bias term that depends on the ratio of the pleiotropic effect of the instrument on the outcome ($\beta_{uy}\pi_{zu}$) to the effect of the instrument on the exposure ($\pi_{zx} + \beta_{ux}\pi_{zu}$). i.e. the larger the pleiotropic effect as a proportion of the effect of the instrument on the exposure the larger the bias of the IV estimator. The bias in the MR estimate from a single heritable confounder will be less than the bias of the linear regression estimator from the same confounder when;

$$\frac{\beta_{uy}\pi_{zu}}{\pi_{zx} + \beta_{ux}\pi_{zu}} < \beta_{uy}\delta_{ux}$$

The smaller the pleiotropic effect as a proportion of the effect of the instrument on the exposure the more likely this is to hold. Under a simpler model of heritable confounding, where U is a confounder of X and Y and has its own independent genetic profile in which Z influences only one path, i.e. $\pi_{zx} = 0$ the bias term for the IV estimator reduces to $\frac{\beta_{uy}}{\beta_{ux}}$. The implication of this is that if the genetic variants influence the exposure X only through the confounder U the bias in the MR estimate from that confounder will always be greater than the bias in the observational association from the same confounder.

To illustrate the relationship between the bias of linear regression and IV estimation across different levels of pleiotropy we simulated the model given in **Figure 2A** varying the effect of the instrument on the confounder but holding all other parameters constant. We estimated β_{xy} using linear regression and the two-stage least squares IV estimator. The mean bias of each estimator across 1000 repetitions for varying levels of confounding is given in **Figure 2B**. These results illustrate how the bias of the IV estimator increases as the pleiotropic effect via the confounder increases and can

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

be substantially more than the bias from linear regression. Note that the exact position and slope of these lines will depend on the other parameters in the model. **Figures 2C** and **2D** show the same bias plots as **Figure 2B** with an additional non-heritable confounder added to the model. The direction of the confounding from this additional confounder is varied from positive (Figure 2C) to negative (Figure 2D). These plots highlight how the overall bias of each estimator will depend on other factors as well as the level of heritable confounding and the bias of the estimate from MR may not always act in the same direction as the bias of the estimate from linear regression even in the presence of heritable confounding due to other confounding.

Increasing GWAS sample size may increase correlated pleiotropy

GWAS sample sizes of complex traits are continually growing, pursuing a well-worn strategy to uncover hundreds or thousands of smaller effect sizes that in aggregate will explain a large polygenic component of phenotypic variation. A genetic variant inducing correlated pleiotropy must influence a trait that is a parent to both exposure and outcome. We hypothesised that such a genetic variant would generally be associated with more traits upstream of the exposure than valid instruments, and therefore would have relatively smaller effects on the exposure than valid instruments. The consequence of such a relationship would be that increasing GWAS sample sizes to identify more instruments of smaller effect sizes could lead to increasing bias in MR estimates.

To evaluate this hypothesis we performed simulations in which we consider that all hypothesised exposure-outcome relationships exist in a broader genotype-phenotype network. The network represents a directed acyclic graph where all nodes are traits except those that have no parents, which are genetic variants. The simulation proceeds by randomly choosing two traits, one as exposure and one as outcome, and then evaluating a) which variants are instruments for the exposure, and whether or not the instruments act through confounders in the graph, and b) the path length from instrument to exposure. The length of the edge between any pair of nodes has been fixed to be the same length for all edges.

Figure 3 illustrates that in this model instruments that act via confounders tend to have smaller effects due to more traits lying on the path from the genetic variant to the exposure. For traits where our simulation is a realistic model of the underlying biology an implication of this result is that as sample sizes grow for GWAS, the rate of discovery of genetic instruments that are liable to act via confounders will grow. However, it is possible that this will not hold for all traits, for some traits it may be that the strongest effects are via confounders (e.g. because those confounders have very strong effects on the exposure) and the genetic variants with smaller effects will be the unconfounded ones. It is therefore important to consider if this model is likely to hold for each exposure trait considered.

Simulating the impact of genetic confounding on MR estimates

We simulated a model with an exposure and outcome which are confounded by a third (potentially unobserved) heritable trait. Here exposure has no causal effect on the outcome however, linear regression of the exposure on the outcome would give a positive estimate of the association between the exposure and outcome due to confounding. The exposure is influenced by two traits,

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

one simply mediates genetic effects onto the trait (i.e. influences on X), whereas the other is a heritable confounder of the exposure and the outcome. Both mediator and confounder are each associated with a set of genetic variants. This model is illustrated in **Figure 4**, which corresponds to the scenario illustrated in **Figure 1B(ii)**. We have included only a single heritable confounder in these models, no non-genetically influenced confounders and have assumed that all other IV assumptions hold. The results therefore illustrate the bias resulting from that confounder and not the overall bias that would be expected in either the linear regression or the MR estimates.

The genetic variants associated with the mediator are valid instruments for estimating the effect of the exposure on the outcome whereas genetic variants associated with the confounder are not as there is a path from them to the outcome that does not act via the exposure. Therefore, MR using just the set of genetic variants associated with the mediator (G_1) would estimate the true null causal effect of the exposure on the outcome. However, if the researcher mistakenly includes genetic variants in G_2 as instruments for X, bias in the MR will be introduced through the pleiotropic pathway. We use this simulation to illustrate how MR is biased under correlated pleiotropy and explore the ability of commonly used pleiotropy robust methods to identify this bias when different proportions of the total set of genetic variants are in sets G_1 and G_2.

Results from this simulation are given in **Figure 4** and **Supplementary Tables 1** and **2**. When the total proportion of the variation in the exposure is explained by genetic variants that act on that exposure through the confounder is moderate IVW and MR Egger are biased across all p-value selection thresholds, however both weighted median and weighted mode correctly estimate the causal effect of the exposure on the outcome (**Figure 4, Supplementary Table 1**). When almost all of the SNPs are associated with the confounder all univariable methods of estimation are biased and give results that are consistent with each other (**Figure 4, Supplementary Table 1**). With a stricter p-value cut-off, equivalent to a smaller sample size, MR Egger is very imprecise. This is expected as a stricter p-value cut off will mean fewer SNPs are included in the estimation and MR Egger has low power when the number of SNPs included in the estimation is small. In all cases both MVMR and the univariable MR approaches with Steiger filtering between the exposure and confounder applied gave unbiased estimates of the causal effect of the exposure on the outcome. However, each of these methods requires the confounder to be known (/suspected) and have GWAS summary statistics available.

In these simulations there is a single confounder of the exposure and outcome of interest biasing the estimates obtained. These results therefore show how the MR effect estimates are biased by that confounder in the same direction as the confounded observational association and may be substantially more biased than the observational association. In practice it is likely that there may be many potential confounders, some of which will be associated with a subset of the genetic variants selected as instruments. The overall bias in the linear regression will depend on the direction and magnitude of the bias from all of the confounders and so is likely to be substantially more than the bias in the MR estimation from confounding. The results obtained here would apply to each of the heritable confounders, and both Steiger filtering and MVMR could be applied to each of them in turn to identify and reduce that bias.

Table 1 shows the mean number of genetic variants in each simulation that were selected for the univariable MR estimation methods (IVW, MR Egger, weighted median and weighted mode). When

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

the mediator is associated with as many SNPs as the confounder in the data generating process the final proportion of SNPs that are associated with the confounder in the MR estimation is small due to the effects of those SNPs on the exposure being smaller. Weighted median and weighted mode MR estimation methods are therefore unbiased as their assumptions regarding the proportion of SNPs unaffected by pleiotropy are satisfied. When most of the SNPs are associated with the confounder fewer SNPs are selected for estimation. However, most of them are associated with the confounder and so violate the IV assumptions leading to bias in all of the univariable MR methods of estimation.

Example: Heritable confounding in genetic variants associated with C-Reactive Protein

To illustrate how some GWAS may be detecting genetic variants associated with pleiotropic effects, we considered a single potential exposure that has been widely used as an exposure in MR studies; C-Reactive Protein (CRP), a marker of inflammation.[23-25] These studies have considered CRP using either all GWAS significant loci or using a cis MR approach, making use only of the significant loci that lie close to the CRP gene.[26] Cis-MR has the advantage of reducing the probability of high levels of pleiotropy however it has lower power than standard genome-wide MR approaches and limited ability to investigate how sensitive the results are to that pleiotropy. A recent GWAS has identified a large number of genetic variants associated with CRP across the genome, many of which may be associated with CRP due to their influence on traits which are upstream of CRP. We therefore explore the extent of heritable confounding that these additional SNPs may add to MR studies. We made use of the FUMA platform[27] to identify other traits that have been associated with the genomic loci in a recent large GWAS of CRP.[28]

The genetic signature of CRP was comprised of 728 lead SNPS (i.e., SNPs associated with CRP levels at $p< 5x10^{-8}$ and LD $r^2 < 0.1$) in 266 genomic loci and we identified 33,573 SNP -> trait associations among these lead SNPs or related SNPs in strong LD (LD $r^2 > 0.6$) in the GWAS Catalog. The full list of these associated traits is given in the Supplementary material Section 2. Many of the SNPs identified as associated with CRP in the GWAS are likely to be associated with CRP through the traits we have identified. Therefore, these traits are potential heritable confounders which may bias any MR estimation which includes CRP as an exposure or outcome. We therefore selected 5198 traits for follow up analysis. These traits can be broadly categorised into: anthropometry measures related to height and weight, bone mineral density, glycaemic measures, proteins, metabolites and smoking behaviour traits (full list given in Supplementary Material Section 2). For each of these traits we estimated whether MR showed evidence of a causal effect of the trait on CRP. We applied Wald ratio for traits with a single SNP available as an instrument, inverse variance weighting for all traits with two or more SNPs available as instruments. Where possible (>3 SNPs) we also applied MR Egger, Weighted Median and Weighted Mode. **Table 2** summarises the proportion of the analyses conducted that indicated evidence of a causal effect using a heuristic threshold to identify associations that indicate a potential causal effect ($p<0.05$). These results indicate that approximately 20% of the traits identified as associated with CRP-SNPs had an effect on CRP and so are potential confounders in any analysis including CRP. These effects were consistent across different pleiotropy robust MR estimators. These traits have the potential to bias any MR estimate in

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

which CRP is an exposure through heritable confounding and indicate the importance of considering this form of bias.

Example: Estimation of the effect of age at menarche on risk of type 2 diabetes

We illustrate the potential for genetic variants associated with a confounder to bias MR estimation through a example estimating the causal effect of age at menarche on risk of type 2 diabetes. Here, we expect that adiposity confounds age at menarche and type 2 diabetes, and with a well-powered GWAS for age at menarche, some of the instruments will arise via body mass index and therefore MR estimation of this effect is likely to be biased. The hypothesised DAG is given in **Figure 5** and results are given in **Table 3**. Univariable MR analysis shows an estimated causal effect of age at menarche on type 2 diabetes using IVW [odds ratio for a category increase in age at menarche on the risk of type 2 diabetes: 0.76, 95% confidence interval: 0.65 to 0.89]. However, this result is not consistent across the robust MR methods.

We also applied Steiger filtering in which 9 of 181 instruments for age at menarche appeared to primarily influence childhood adiposity were removed, leading to an attenuated effect estimate [OR: 0.90, 95% CI 0.79 to 1.02] (**Figure 6**). Including childhood adiposity in a MVMR more substantially attenuates the effect estimate obtained [OR: 0.99, 95% CI: 0.78 to 1.09]. As childhood adiposity is an established risk factor for both age at menarche and – through adulthood BMI - type 2 diabetes, these results are likely to reflect the model given in **Figure 2b** and suggest that the result seen in the naïve IVW estimation is likely to be biased by confounding from SNPs associated with childhood adiposity. As has been shown elsewhere, there is an overlap in the SNPs associated with childhood and adulthood adiposity.[29] As adiposity is a well-established risk factor for Type 2 Diabetes the SNPs associated with childhood adiposity will also be associated with Type 2 Diabetes through adulthood adiposity.

Discussion

In this paper we have shown how correlated pleiotropy due to heritable confounding can bias MR effect estimates when the genetic variants selected as instruments are associated with a confounder of the exposure and outcome of interest. We show that the bias in MR from including instruments associated with a heritable confounder could be greater than the bias in the linear regression from that confounder. In our network simulations we show, under a plausible model of biological effects, how genetic variants associated with confounders of an exposure and outcome may be associated with more traits upstream of the exposure than genetic variants which are associated with the exposure but not confounders of the exposure and outcome. The implication of this is that, as GWAS sample sizes increase these confounding genetic variants are increasingly likely to be detected as strongly associated with the exposure. This will increase the presence of correlated pleiotropy in MR effect estimates which use genetic variants strongly associated with the exposure in a GWAS as instruments for that exposure. Through simulations where there is a mixture of valid instruments or those acting via confounders, we show that bias caused by a specific confounder in MR estimation could be larger than the bias in the observation association caused by the same confounder and that

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

this problem is likely to increase as sample sizes increase. Although it is likely that there will also be non-heritable confounding biasing the results from linear regression these results highlight how MR results may be increasingly biased as sample sizes for GWAS grow.

If a large proportion of the genetic instruments are associated with a confounder of the exposure and outcome, correlated pleiotropy will bias all of the standard pleiotropy robust methods of estimation. In this case each pleiotropy robust method is biased in the same way as the IVW estimator, potentially leading to false confidence in the results obtained. These results are likely to particularly apply to exposures that are downstream of a phenotype (such as BMI) which is associated with many genetic variants and has an influence on a wide range of traits, or where there are relatively few genetic variants directly associated with the exposure of interest. In these cases a high proportion of the genetic variants used may be associated with one or more heritable confounders, increasing the overall bias in the MR estimates.

We show that this bias from the inclusion of genetic variants associated with a confounder could be corrected through the application of either MVMR including the confounder, or Steiger filtering to remove any genetic variants that explain more variation in the confounder than the exposure. We therefore recommend that researchers applying MR with summary statistics generated from large GWAS studies check the association of their SNPs with the main potential confounders of the exposure and the outcome through a PheWAS analysis. Any genetic variant-trait associations found should then be checked for whether the trait involved is likely to be a confounder.[30] Steiger filtering, removing SNPs that explain more variation in the confounder than the exposure, and/or MVMR also including the confounder can then be applied to the MR analysis as an additional robustness test. A substantial difference between the results obtained is an indication of bias in the original MR analysis. However, in practice it is likely that confounders will not be identifiable/measured and there will be a number of confounders acting in this way and so such an approach does not guarantee the remove of all bias from heritable confounding.

Recently novel MR methods have been proposed that attempt to adjust for correlated pleiotropy.[14, 16, 18, 19, 31, 32] Each of these methods depends on different assumptions about the nature of the correlated pleiotropy and applies a different approach to correct for that pleiotropy. When standard MR methods are suspected to be biased by correlated pleiotropy then applying one or more of those approaches will help to identify whether correlated pleiotropy is present.

We have previously shown that adjustment for a potential confounder in the GWAS for the exposure will not mitigate bias in the MR estimates.[33, 34] This is because the adjustment for the confounder generates an inverse association between those genetic variants and the exposure. This biases the results obtained from the GWAS the opposite direction to the direction of effect of the genetic variant on the confounder. Previous work showed that this could be corrected by inclusion of the confounder the GWAS was adjusted for in a MVMR estimation[33] however all of the results described here would equally apply in that setting.

Where the exposure has a clear biological mechanism by which some genetic variants affect the exposure, comparing the results obtained for the MR estimation with all genetic variants to the

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

results obtained for the genetic variants which have a known biological pathway can help to support the results or highlight the potential for bias in the MR estimate. Here we used the example trait of CRP to explore how many of the genetic variants detected in a recent GWAS were likely to be associated with confounders. These results showed a high level of pleiotropy among the variants identified, with many of the traits identified showing evidence of having an effect on CRP and so being potential confounders. These results highlight how prevalent heritable confounding bias is likely to be for many traits. CRP is an exposure for which alternative 'cis-MR' approaches could be used, making use only of the genetic variants identified within the CRP gene region. Where possible such an approach is likely to be more robust to heritable confounding than using the full set of genome-wide significantly associated variants and so both approaches should be used if possible.

In our illustrative application we estimated the effect of age at menarche on risk of Type 2 diabetes. The standard IVW effect estimate showed a negative causal effect with later age at menarche estimated to be protective for risk of type 2 diabetes. This effect however attenuated in all of the robust methods as well as when Steiger filtering or MVMR was applied to adjust for childhood adiposity. Childhood adiposity is an established risk factor for both age at menarche[22, 35, 36] and type 2 diabetes, through adulthood adiposity, [29, 37-40] and these analyses support the effect observed in the IVW estimation being driven by a set of SNPs that are associated with the confounder childhood adiposity being included as instruments for age at menarche.

The results given here highlight why it is important to consider whether the IV assumptions are likely to hold even when a range of MR estimates give consistent results for the effect of the exposure on the outcome. This is particularly true when there is a strongly genetically predicted potential confounder for the exposure and outcome, such as BMI, that has not been accounted for in the MR estimates or where the total number of genetic variants associated with the exposure of interest is likely to be small. This checking could be done in systematically through PheWAS analysis or through expert subject knowledge of the likely confounders.[30]

While these results show that the MR results can be more biased than linear regression from the exclusion of a single heritable confounder there are many other sources of bias for both types of estimation that should be considered when determining how much weight to give to the results from each type of analysis. It is likely that there will be a number of confounders of the exposure and outcome, if only some of these are heritable in the way described here then the MR estimates may still be less biased from confounding than the observational linear regression estimates. Additionally, MR can overcome bias and effect attenuation from measurement error that occurs in linear regression. However, MR is subject to other forms of bias such as population stratification and pleiotropy of other forms which should also be considered in any analysis.

A limitation of the results here is how we have set up the structure of relationships between genetic variants and phenotypes. In our network simulations we assumed that every edge has the same strength of effect. In our simulations of the bias of the different MR estimators we have assumed the same distribution in size of the association between the relevant genetic variants and the exposure and confounder. These assumptions are relatively strong and are not likely to be realistic for every trait. For some traits the confounders may have much stronger genetic effects than the exposure, which will accentuate the biases described here. For other traits the direct effect of the genetic

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

variants on the exposure will be much larger than for the confounders, meaning that bias of the type described here will be less prevalent. Which of these is likely to hold, along with how strongly the confounders are likely to affect the outcome, will determine how much of a problem the results described here are likely to be in any particular case. Each MR study conducted should therefore consider individually how likely there is to be bias from heritable confounding in that setting and what the main confounders causing that bias may be.

Methods

Network simulations

We use the randomDAG function in the R/dagitty package to simulate a DAG with an arbitrary number of nodes, and a fixed probability that any set of nodes is connected. Nodes with no parents in the simulated DAG are identified as genetic variants, and 500 trait pairs are selected at random, in each pair assigning one to be the hypothesised exposure and one to be the hypothesised outcome. We then identify the path length for every genetic variant for the exposure, and also determine whether they act via a confounder, or otherwise (i.e. directly or through other traits that mediate the association between the genetic variant and the exposure and have no paths to the outcome). The effect size of the genetic variant on the exposure is β^L where L is the path length from genotype to exposure, and for simplicity every edge in the DAG has the same effect $\beta = 0.2$. These simulations were conducted for a set of parameters of graph size N=75, 100, 125, 150 and graph density based on probability of an edge occurring ranging from p = 0.01, 0.02, …, 0.1. Each parameter combination was repeated 10 times.

Simulations evaluating MR methods

We generated the data in this model for two separate samples. In each sample 400 SNPs were generated using a binomial distribution. The proportion of these SNPs associated with the mediator and the confounder varied across the simulations, varying between 50% associated with each trait to 12.5% associated with the mediator and 87.5% associated with the confounder. The SNP-trait (i.e. mediator/confounder) association was drawn from a normal distribution with mean 0 and standard deviation of 0.04 for each SNP, orientated so all SNPs had a positive association with the relevant trait.

The confounder was generated as the sum of all the confounder - SNP effects plus a random error term. The mediator was generated as the sum of all mediator – SNP effects plus a random error term. The exposure was generated as the sum of the mediator, the confounder and a random error term. As shown earlier, SNPs associated with a mediator may be more strongly associated with the exposure than SNPs associated with a confounder. Therefore, we down weighted the confounder – exposure association by 25% relative to the mediator – exposure association (by multiplying them by 0.75) to mimic this more distant association. Results with this down weighting varied (to 50% and 0%) are given in the supplementary material. Finally, the outcome was generated as the effect of the confounder and a random error term.

In one of the samples generated we estimated the association between all of the SNPs and each of the exposure and confounder to generate GWAS summary statistics. This was repeated in the

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

second sample for the SNPs and the outcome. We then select all SNPs (from either set) that were significantly associated with the exposure in the first sample at a defined p-value threshold in these summary statistics to estimate the effect of the exposure on the outcome using summary-data MR methods. Following what is currently standard practice in summary-data MR the same sample was used for the SNP-exposure association in estimation as had been used to identify the SNPs associated with the exposure. We applied four common univariable MR methods of estimation IVW[41], MR Egger[11], weighted median[12] and weighted mode[13]. We additionally estimated the model using multivariable MR (MVMR)[42, 43] including the confounder as an additional exposure, including all SNPs significantly associated with the confounder at the same p-value threshold. All univariable methods of estimation were repeated with Steiger filtering[44] applied between the exposure and confounder to remove SNPs that explain more variation in the confounder than the exposure.

With all other things fixed increasing the sample size will decrease the standard error in the SNP-trait association proportionally across all of the SNPs. This will lead to more SNPs being selected based on a fixed p-value threshold cut-off. Therefore, to model varying the sample size in the GWAS while keeping all other aspects of the data constant we varied the p-value threshold used to select SNPs for use as instruments in the MR across our simulations with larger p-values being equivalent to a larger sample size.

Simulation details

A set of 400 SNPs were generated as;

$$G_{ij} \sim B(2, 0.4)$$

For individual i and SNP j. The SNP exposure effects were generated as;

$$\pi_j \sim N(0, 0.04^2)$$

The confounder was generated as;

$$C_i = \pi_{jc} G_{ij} + v_c$$

Where v_c is a randomly normally distributed error term with mean 0 and standard deviation 1. π_{jc} is an SNP effects vector for the confounder where the elements of π_j associated with the mediator (rather than the confounder) have been set to 0.

The mediator was generated as;

$$M_i = \pi_{jm} G_{ij} + v_c$$

Where v_c is a randomly normally distributed error term with mean 0 and standard deviation 1. π_{jm} is an SNP effects vector for the confounder where the elements of π_j associated with the confounder (rather than the mediator) have been set to 0.

The exposure was generated as;

$$X_i = M_i + 0.75C_i + v_x$$

Where $v_x \sim N(0, 1)$.

The outcome was generated as;

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$$Y_i = 0.5C_i + v_y$$

Where $v_y \sim N(0,1)$.

This data was generated for two different samples and summary statistics were generated by estimating the linear regressions;

$$C_i = \alpha_c + \beta_{c_j} G_{ij} + \epsilon_c$$

and

$$X_i = \alpha_x + \beta_{xj} G_{ij} + \epsilon_x$$

For all j in the first sample and

$$Y_i = \alpha_y + \beta_{y_j} G_{ij} + \epsilon_y$$

For all j in the second sample, to obtain estimates of β_{cj}, β_{xj} and β_{yj}.

All models were simulated with a sample size of 100,000 for each of the exposure and outcome samples and 1000 repetitions.

Heritable confounding in genetic variants associated with C-Reactive Protein

C-reactive protein (CRP) is a widely studied acute phase hepatic protein and inflammation marker. Early observational studies consistently found associations between elevated CRP levels with increased cardiovascular disease (CVD) risk, motivating interest in targeting CRP as a therapeutic for CVDs.[45] In contrast with the observational findings, human genetics studies leveraging SNPs in the *CRP* locus revealed no direct causative links between CRP and cardiovascular diseases (CVDs),[28, 46, 47] indicating that CRP may not be the inflammatory driver of heart disease, which diminished the interest in developing CRP-targeted therapies.[46, 47] It appears that the associations between CRP and CVD observed in observational studies were likely influenced by confounding factors from other biomarkers and risk factors causally related to CVDs.[45, 48] More recent GWASs of circulating CRP levels have been conducted,[28, 49-51] uncovering hundreds of variants across the genome linked to CRP levels. Many of the loci are involved in metabolic and immune response pathways,[28] and the CRP GWAS demonstrates strong genetic correlation with cardiometabolic risk factors and biomarkers such as LDL cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides.[51] Correspondingly, MR studies instrumenting CRP levels with SNPs located throughout the genome have identified consistent, adverse relationships CRP and various CVD outcomes,[52-55] replicating the biased observational estimates and suggesting that heritable confounding with other CVD biomarkers and risk factors may be present.

We therefore sought to explore the pleiotropic relationships of the genomic loci associated with CRP to assess whether there is evidence for heritable confounding. SNPs associated with CRP levels were identified from in a recent GWAS of >500,000 individuals with European genetic ancestry.[28] This GWAS identified 266 loci as associated with CRP, 211 of which were novel. We made use of the FUMA platform[27] to identify other traits which have also been associated the 266 CRP genomic loci. We first identified the lead SNPs the CRP loci (independent SNPs LD $r^2 < 0.1$, and GWAS P-values

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

$< 5\times10^{-8}$) and performed a look up in the GWAS Catalog[56] (look up was performed on April 10th, 2024) on 728 lead SNPs and 3,077 independent SNPs in strong LD with the lead SNPs (GWAS P-values $< 5\times10^{-8}$ and LD $r^2 > 0.1$ with the respective lead SNP).

The resulting list of associated traits was curated to prioritise biomarkers and risk factors over diseases. From this curated list we conducted a summary-data MR analysis of each trait on to CRP levels. Where only one SNP was available as associated with the exposure trait of interest the Wald ratio was calculated. For traits with 2 or more SNPs the IVW MR effect estimate was obtained and for traits with 3 or more SNPs the MR Egger[11], Weighted Mode[13] and Weighted Median[12] estimates were also obtained.

Estimation of the effect of age at menarche on type 2 Diabetes

All data used were downloaded from MRC IEU OpenGWAS[57]. SNP – exposure associations were extracted for age at menarche for data on recalled age at which periods started for women in UK Biobank (n = 243,944) obtained through the IEU GWAS pipeline[58, 59]. SNP – exposure associations were obtained for childhood adiposity for a previously conducted GWAS using data on self-reported childhood adiposity for men and women in UKBiobank.[29] SNP – outcome associations for Type 2 diabetes was taken from a GWAS from FinnGen[60] to avoid sample overlap with UKBiobank.

We applied the univariable summary data estimation approaches; IVW, MR Egger, weighted median and weighted mode and additionally MVMR adjusting for BMI. For each approach we selected all SNPs associated with the relevant exposure at genome-wide significance (p-value for the SNP-exposure association $< 5\times10^{-8}$). For the MVMR we selected all SNPs associated with either age at menarche or childhood adiposity at the same threshold. Each univariable method was applied twice, once including all SNPs and once with Steiger filtering applied to remove SNPs that explained more variation in childhood adiposity than age at menarche. Steiger filtering was not applied to the MVMR estimation due to the inclusion of childhood adiposity as an exposure. Selected SNPs were then clumped to retain only independent SNPs ($r^2 < 0.001$ based on a reference LD panel). All analyses were done in R using the 'TwoSampleMR' package with default parameters for SNP selection and clumping applied.[61]

Acknowledgements

ES, TP, GH, GDS, KT work in a unit funded by the MRC [MC_UU_00032/01, MC_UU_00032/02].

Data availability

Summary statistics used in all analyses are available at OpenGWAS study IDs ukb-b-3768, finn-b-E4_DM2, ukb-b-4650.

Code availability

The code for the simulations and analyses is available at https://github.com/eleanorsanderson/confounding.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

References

1. Davey Smith, G. and A.N. Phillips, *Correlation without a cause: an epidemiological odyssey.* International Journal of Epidemiology, 2020. **49**(1): p. 4-14.
2. Davey Smith, G. and S. Ebrahim, *'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*.* International Journal of Epidemiology, 2003. **32**(1): p. 1-22.
3. Richmond, R. and G. Davey Smith, *Mendelian randomization: concepts and scope.* Cold Spring Harbour Perspectives in Medicine, 2021.
4. Sanderson, E., et al., *Mendelian randomization.* Nature Reviews Methods Primers, 2022. **2**(1): p. 6.
5. Uffelmann, E., et al., *Genome-wide association studies.* Nature Reviews Methods Primers, 2021. **1**(1): p. 59.
6. O'Connor, L.J., et al., *Extreme Polygenicity of Complex Traits Is Explained by Negative Selection.* The American Journal of Human Genetics, 2019. **105**(3): p. 456-476.
7. Yengo, L., et al., *Meta-analysis of genome-wide association studies for height and body mass index in ~ 700000 individuals of European ancestry.* Human molecular genetics, 2018. **27**(20): p. 3641-3649.
8. Okbay, A., et al., *Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals.* Nature genetics, 2022. **54**(4): p. 437-449.
9. Boyle, E.A., Y.I. Li, and J.K. Pritchard, *An Expanded View of Complex Traits: From Polygenic to Omnigenic.* Cell, 2017. **169**(7): p. 1177-1186.
10. Zhang, W., et al., *MR Corge: Sensitivity analysis of Mendelian randomization based on the core gene hypothesis for polygenic exposures.* bioRxiv, 2024: p. 2024.07.18.604191.
11. Bowden, J., G. Davey Smith, and S. Burgess, *Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.* International journal of epidemiology, 2015. **44**(2): p. 512-525.
12. Bowden, J., et al., *Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator.* Genetic epidemiology, 2016. **40**(4): p. 304-314.
13. Hartwig, F.P., G. Davey Smith, and J. Bowden, *Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption.* International journal of epidemiology, 2017. **46**(6): p. 1985-1998.
14. Morrison, J., et al., *Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics.* Nature Genetics, 2020: p. 1-7.
15. Cheng, Q., et al., *MR-LDP: a two-sample Mendelian randomization for GWAS summary statistics accounting for linkage disequilibrium and horizontal pleiotropy.* NAR genomics and bioinformatics, 2020. **2**(2): p. lqaa028-lqaa028.
16. Xue, H., X. Shen, and W. Pan, *Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects.* The American Journal of Human Genetics, 2021. **108**(7): p. 1251-1269.
17. Xu, S., W.K. Fung, and Z. Liu, *MRCIP: a robust Mendelian randomization method accounting for correlated and idiosyncratic pleiotropy.* Briefings in Bioinformatics, 2021.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

18. Brown, B.C. and D.A. Knowles, *Welch-weighted Egger regression reduces false positives due to correlated pleiotropy in Mendelian randomization.* The American Journal of Human Genetics, 2021. **108**(12): p. 2319-2335.
19. Cheng, Q., et al., *Mendelian randomization accounting for complex correlated horizontal pleiotropy while elucidating shared genetic etiology.* Nature Communications, 2022. **13**(1): p. 6490.
20. Cheng, Q., et al., *MR-Corr2: a two-sample Mendelian randomization method that accounts for correlated horizontal pleiotropy using correlated instrumental variants.* Bioinformatics, 2022. **38**(2): p. 303-310.
21. Lin, Z., Y. Deng, and W. Pan, *Combining the strengths of inverse-variance weighting and Egger regression in Mendelian randomization using a mixture of regressions model.* PLoS genetics, 2021. **17**(11): p. e1009922.
22. Millard, L.A.C., et al., *MR-PheWAS: hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization.* Scientific Reports, 2015. **5**(1): p. 16645.
23. Zhu, M., et al., *C-reactive protein and cancer risk: a pan-cancer study of prospective cohort and Mendelian randomization analysis.* BMC Medicine, 2022. **20**(1): p. 301.
24. Zhou, A. and E. Hyppönen, *Vitamin D deficiency and C-reactive protein: a bidirectional Mendelian randomization study.* International Journal of Epidemiology, 2022. **52**(1): p. 260-271.
25. Han, X., et al., *Using Mendelian randomization to evaluate the causal relationship between serum C-reactive protein levels and age-related macular degeneration.* European Journal of Epidemiology, 2020. **35**(2): p. 139-146.
26. Schmidt, A.F., et al., *Genetic drug target validation using Mendelian randomisation.* Nature communications, 2020. **11**(1): p. 3255.
27. Watanabe, K., et al., *Functional mapping and annotation of genetic associations with FUMA.* Nature Communications, 2017. **8**(1): p. 1826.
28. Said, S., et al., *Genetic analysis of over half a million people characterises C-reactive protein loci.* Nature Communications, 2022. **13**(1): p. 2198.
29. Richardson, T.G., et al., *Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study.* bmj, 2020. **369**.
30. Yang, Q., et al., *Exploring and mitigating potential bias when genetic instrumental variables are associated with multiple non-exposure traits in Mendelian randomization.* European Journal of Epidemiology, 2022.
31. Xu, S., W.K. Fung, and Z. Liu, *MRCIP: a robust Mendelian randomization method accounting for correlated and idiosyncratic pleiotropy.* Briefings in Bioinformatics, 2021. **22**(5).
32. Qi, G. and N. Chatterjee, *Mendelian randomization analysis using mixture models for robust and efficient estimation of causal effects.* Nature Communications, 2019. **10**(1): p. 1941.
33. Gilbody, J., et al., *Multivariable MR can mitigate bias in two-sample MR using covariable-adjusted summary associations.* medRxiv, 2022: p. 2022.07. 19.22277803.
34. Hartwig, F.P., et al., *Bias in two-sample Mendelian randomization when using heritable covariable-adjusted summary associations.* International journal of epidemiology, 2021. **50**(5): p. 1639-1650.
35. Vabistsevits, M., et al., *Deciphering how early life adiposity influences breast cancer risk using Mendelian randomization.* Communications Biology, 2022. **5**(1): p. 337.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

36. Mumby, H.S., et al., *Mendelian randomisation study of childhood BMI and early menarche.* Journal of obesity, 2011. **2011**.
37. Wainberg, M., et al., *Homogeneity in the association of body mass index with type 2 diabetes across the UK Biobank: A Mendelian randomization study.* PLOS Medicine, 2019. **16**(12): p. e1002982.
38. Lyall, D.M., et al., *Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study.* JAMA Cardiology, 2017. **2**(8): p. 882-889.
39. Holmes, Michael V., et al., *Causal Effects of Body Mass Index on Cardiometabolic Traits and Events: A Mendelian Randomization Analysis.* The American Journal of Human Genetics, 2014. **94**(2): p. 198-208.
40. Dale, C.E., et al., *Causal associations of adiposity and body fat distribution with coronary heart disease, stroke subtypes, and type 2 diabetes mellitus: a Mendelian randomization analysis.* Circulation, 2017. **135**(24): p. 2373-2388.
41. Burgess, S., A. Butterworth, and S.G. Thompson, *Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data.* Genetic Epidemiology, 2013. **37**(7): p. 658-665.
42. Burgess, S., F. Dudbridge, and S.G. Thompson, *Re:"Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects".* American journal of epidemiology, 2015. **181**(4): p. 290-291.
43. Sanderson, E., et al., *An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings.* International journal of epidemiology, 2019. **48**(3): p. 713-727.
44. Hemani, G., K. Tilling, and G.D. Smith, *Orienting the causal relationship between imprecisely measured traits using GWAS summary data.* PLoS genetics, 2017. **13**(11): p. e1007081.
45. Liu, C. and C. Li, *C-reactive protein and cardiovascular diseases: a synthesis of studies based on different designs.* European Journal of Preventive Cardiology, 2023. **30**(15): p. 1593-1596.
46. Collaboration, C.R.P.C.H.D.G., *Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data.* Bmj, 2011. **342**: p. d548.
47. Zacho, J., et al., *Genetically Elevated C-Reactive Protein and Ischemic Vascular Disease.* New England Journal of Medicine, 2008. **359**(18): p. 1897-1908.
48. Markozannes, G., et al., *Global assessment of C-reactive protein and health-related outcomes: an umbrella review of evidence from observational studies and Mendelian randomization studies.* European Journal of Epidemiology, 2021. **36**(1): p. 11-36.
49. Ligthart, S., et al., *Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders.* Am J Hum Genet, 2018. **103**(5): p. 691-706.
50. Zhang, K., et al., *Genetic association of circulating C-reactive protein levels with idiopathic pulmonary fibrosis: a two-sample Mendelian randomization study.* Respiratory Research, 2023. **24**(1): p. 7.
51. Koskeridis, F., et al., *Pleiotropic genetic architecture and novel loci for C-reactive protein levels.* Nature Communications, 2022. **13**(1): p. 6939.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

52. Qing, X., et al., *Mendelian randomization analysis identifies a genetic casual association between circulating C-reactive protein and intracerebral hemorrhage.* Journal of Stroke and Cerebrovascular Diseases, 2024. **33**(2).
53. Habibi, D., et al., *Effect of C-reactive protein on the risk of Heart failure: a mendelian randomization study.* BMC Cardiovasc Disord, 2023. **23**(1): p. 112.
54. Kuppa, A., et al., *C-Reactive Protein Levels and Risk of Cardiovascular Diseases: A Two-Sample Bidirectional Mendelian Randomization Study.* Int J Mol Sci, 2023. **24**(11).
55. Prins, B.P., et al., *Investigating the Causal Relationship of C-Reactive Protein with 32 Complex Somatic and Psychiatric Outcomes: A Large-Scale Cross-Consortium Mendelian Randomization Study.* PLOS Medicine, 2016. **13**(6): p. e1001976.
56. Sollis, E., et al., *The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource.* Nucleic Acids Res, 2023. **51**(D1): p. D977-d985.
57. Elsworth, B., et al., *The MRC IEU OpenGWAS data infrastructure.* bioRxiv, 2020: p. 2020.08.10.244293.
58. Elsworth, B., et al., *MRC IEU UK Biobank GWAS pipeline version 2.* University of Bristol, 2018. **10**.
59. Sudlow, C., et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**(3): p. e1001779.
60. Kurki, M.I., et al., *FinnGen: Unique genetic insights from combining isolated population and national health register data.* medRxiv, 2022: p. 2022.03.03.22271360.
61. Hemani, G., et al., *The MR-Base platform supports systematic causal inference across the human phenome.* eLife, 2018. **7**: p. e34408.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Figures

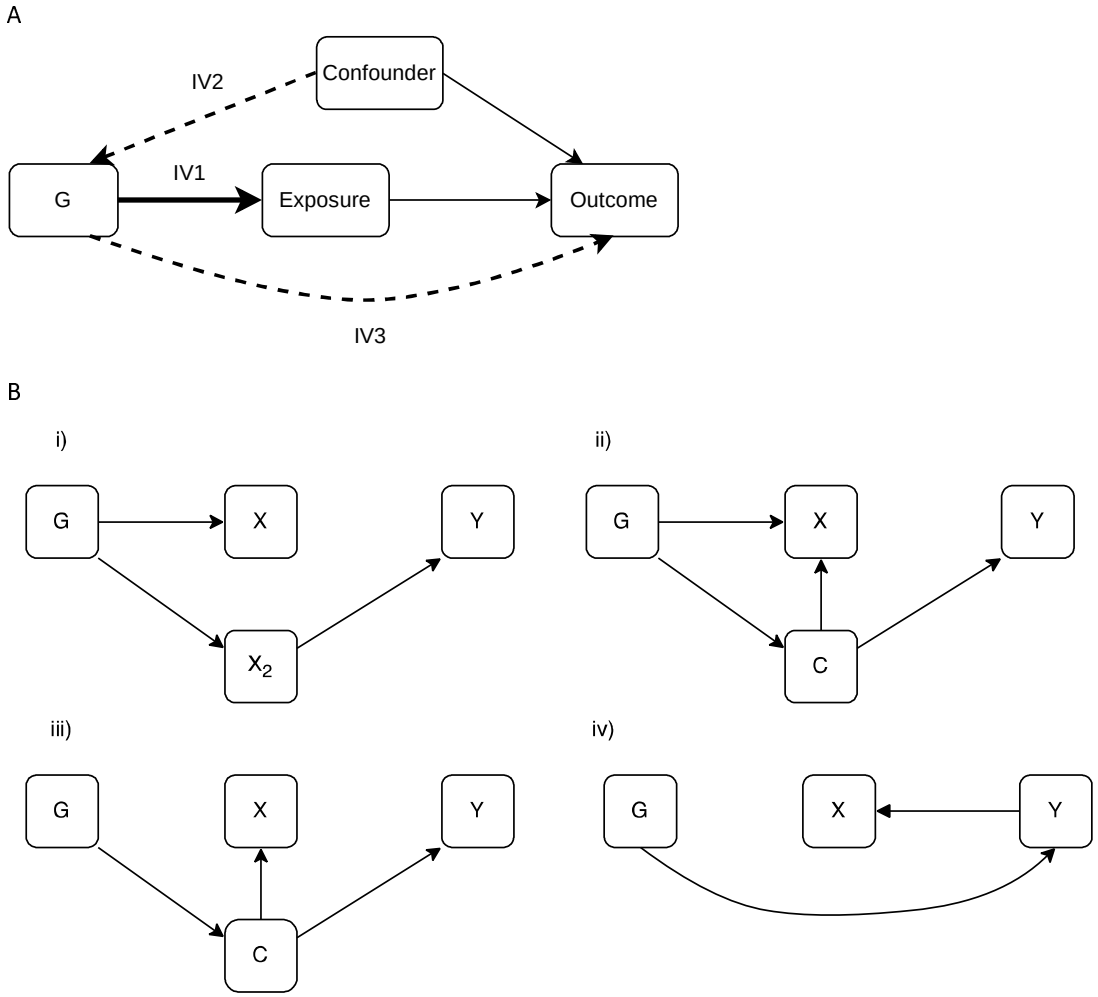

Figure 1: Instrumental Variable assumptions and potential mechanisms that can generate correlated pleiotropy and bias MR.

Figure illustrating the IV assumptions and the different mechanisms that can cause correlated pleiotropy in MR studies in the absence of a causal effect of the exposure on the outcome. X is the exposure of interest, Y is the outcome, C is a confounder of the exposure and outcome. G is a set of SNPs associated with the exposure, some of which may be associated through another phenotype.

- *A. Illustration of the main IV assumptions; (1) the IV should be associated with the exposure, (2) there should be no confounding of the IV and the outcome and (3) there should be no pathway from the IV to the outcome that doesn't act through the exposure. The dashed lines represent associations that should not exist for the IV assumptions to be satisfied.*
- *B. i) Uncorrelated horizontal pleiotropy, some of the SNPs in G are associated with Y through X_2 ii) and iii) show forms of heritable confounding and mis-specification of the primary phenotype. In ii) some of the SNPs in G are associated the exposure through the confounder, in iii) ALL of the SNPs in G are associated with the exposure through the confounder. iv) reverse causation.*

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

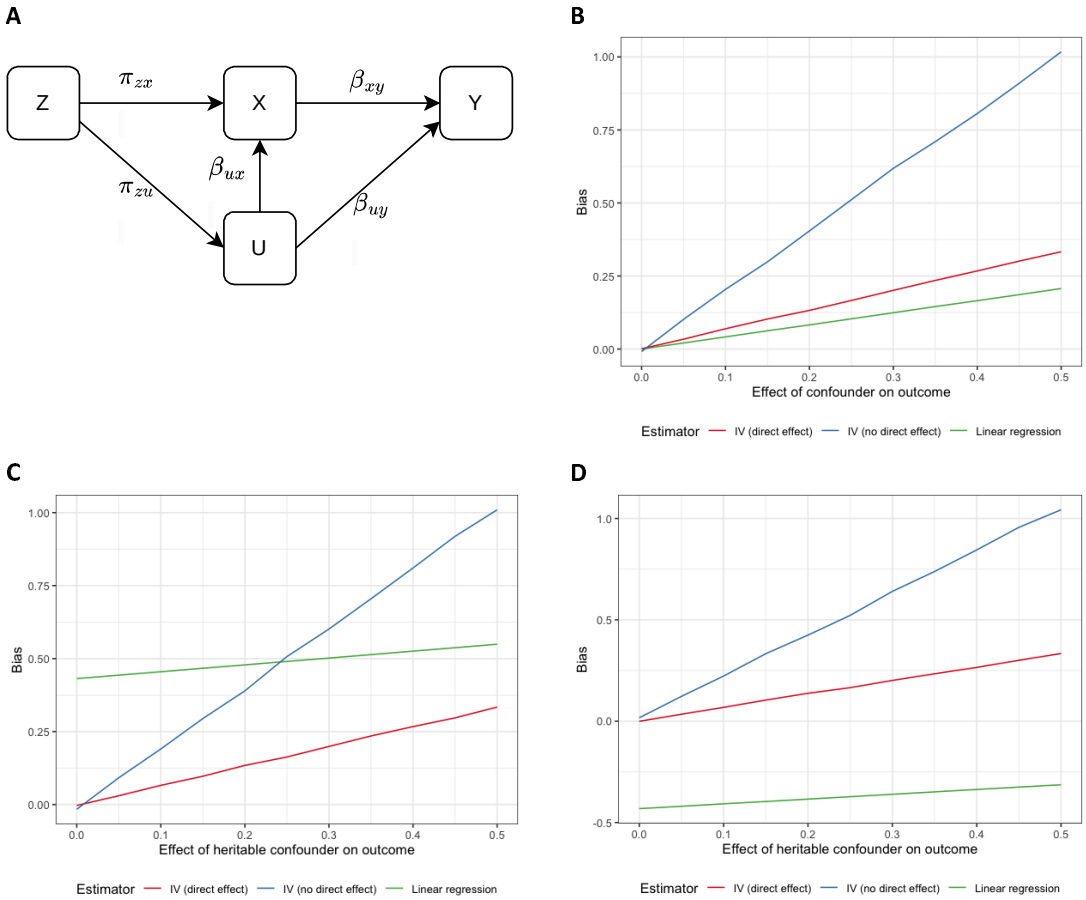

Figure 2 – The Bias of linear regression and IV estimators when instruments have a pleiotropic effect through a single heritable confounder.

A. *A simple model with a single IV associated with an exposure and a confounder. This is used in obtaining the expressions for bias given above and for generating the simulation results in (b).*
B. *Bias from estimation by linear regression and two-stage least squares for the model in (a) estimated assuming U is unobserved and there is no other confounding. Linear Regression gives the bias of the observational association, IV (direct effect) sets and gives the bias of the IV estimator, IV (no direct effect) sets and gives the bias of the IV estimator. The relative positions of these plots depends on the other parameters in the model and the assumption of no other (non-heritable) confounding.*
C. *Bias plots given in B with the addition of a positive non-heritable confounder of the exposure and outcome.*
D. *Bias plots given in B with the addition of a negative non-heritable confounder of the exposure and outcome.*

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

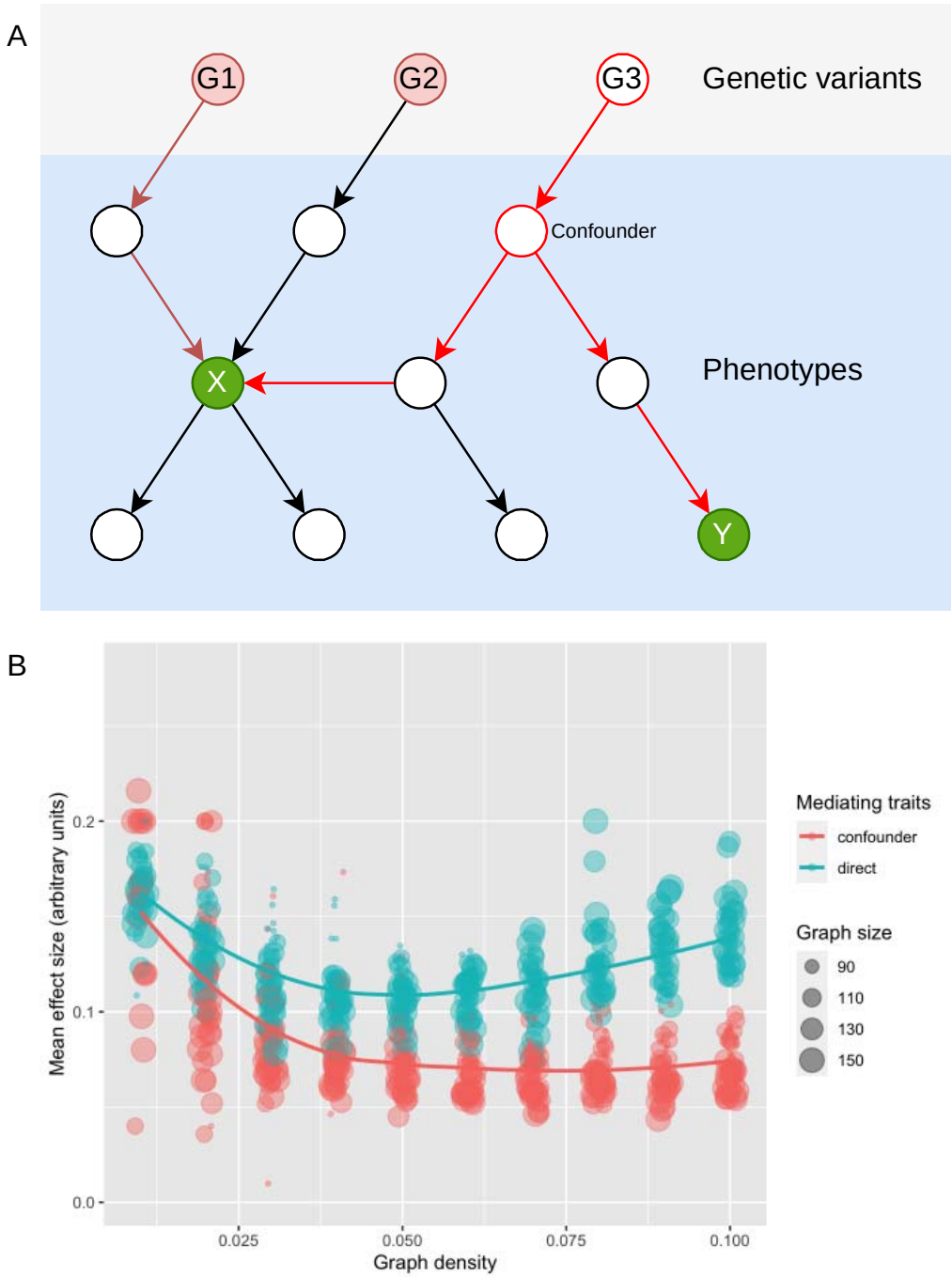

Figure 3. Network-based simulations of the origin of instruments.

A. Illustration of a random network. For each possible trait pair we assign an exposure (X) and an outcome (Y). In this simplified example instrument G1 is considered to act 'directly' because there is no confounding path between X and Y, whereas G2 is considered to act via a confounding path. The effect size of each instrument is determined based on all edges having a uniform effect of 0.1.

B. The mean effect of instrument on the exposure (y-axis) across simulations (x-axis and point size), stratified by whether they act 'directly' or via confounding paths. We observe that in general direct instruments have larger effects than those acting via confounders.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

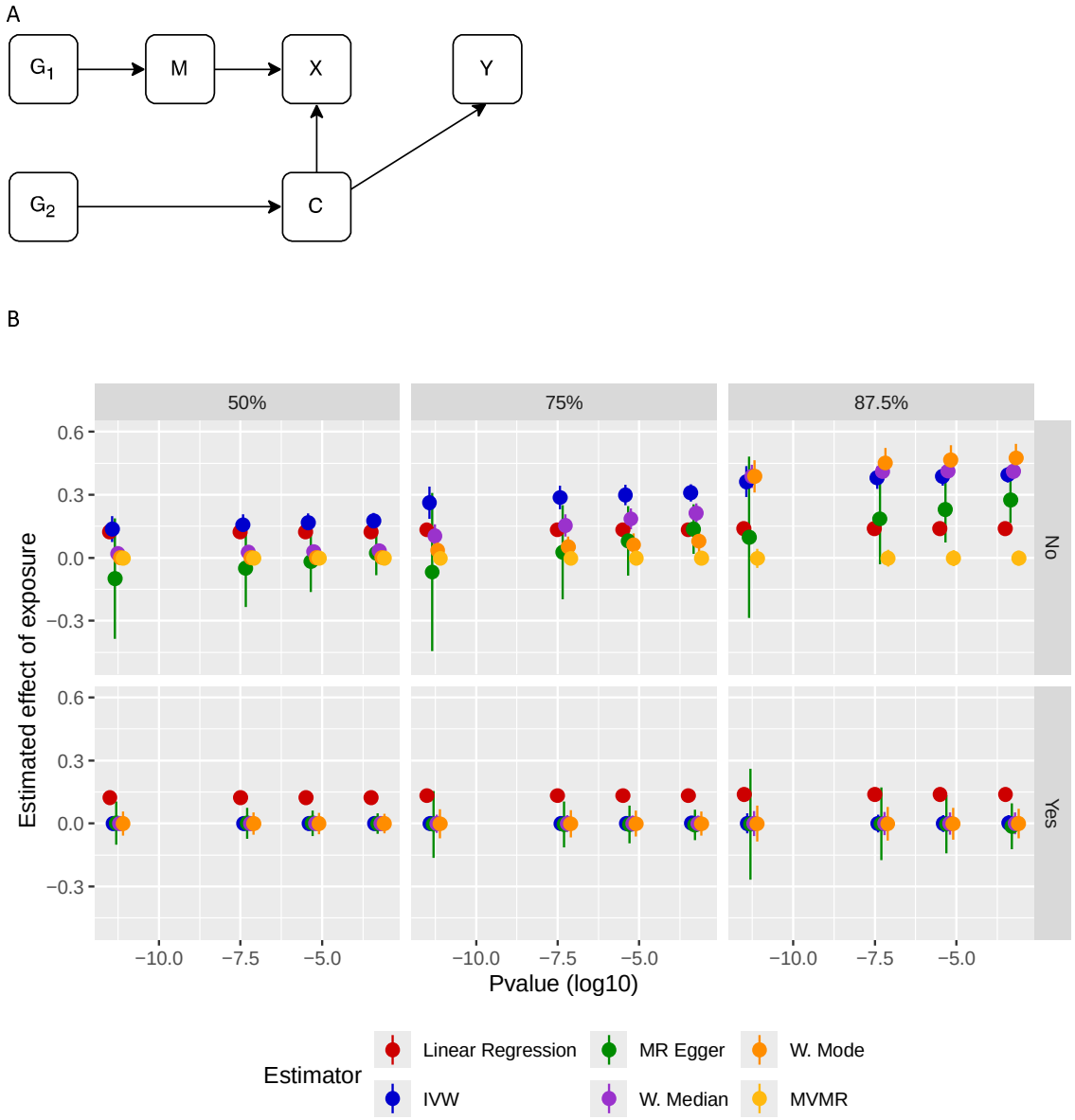

Figure 4 – Simulation of bias from a single heritable confounder in linear regression and MR.

A. The model used to generate the data in the simulation. X is the exposure of interest, Y is the outcome, C is a (potentially observable) confounder of the exposure and outcome, M is a mediator upstream of X. G1 is a set of SNPs associated with the mediator, G2 is a set of SNPs associated with the confounder. SNPs in G1 may be associated with the exposure through the mediator and SNPs in G2 may be associated with the exposure through the confounder. There is no causal effect of the exposure on the outcome.

B. B Results for the simulations varying; the proportion of the SNPs in G2 and whether Steiger filtering is applied. Total number of SNPs – 400, % given is the % in G2, the rest are valid SNPs for X (i.e. G1). All SNPs that are significantly associated with X in the simulated data are included in the univariable MR estimation and all SNPs significantly associated with X or C are included in the MVMR estimation. There is no causal effect of X on Y, the effect of M on X is 1, of C on Y is 0.4 and of C on X is 0.75. n=10000, reps=1000. The effect of X on Y is estimated by; Linear regression, IVW, MR Egger, weighted median, weighted mode and MVMR including C as an additional exposure. Steiger filtered results includes all univariable estimators with Steiger filtering applied to remove SNPs that explain more variation in C than X.

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

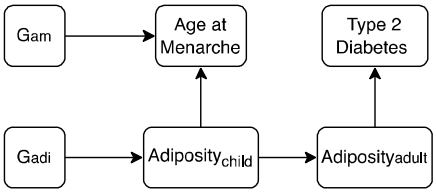

Figure 4 – Illustrative diagram for the hypothesised relationships between adiposity, age at menarche and Type 2 Diabetes. *G_{am} is a set of genetic variants associated with age at menarche, G_{adi} is a set of genetic variants associated with childhood adiposity.*

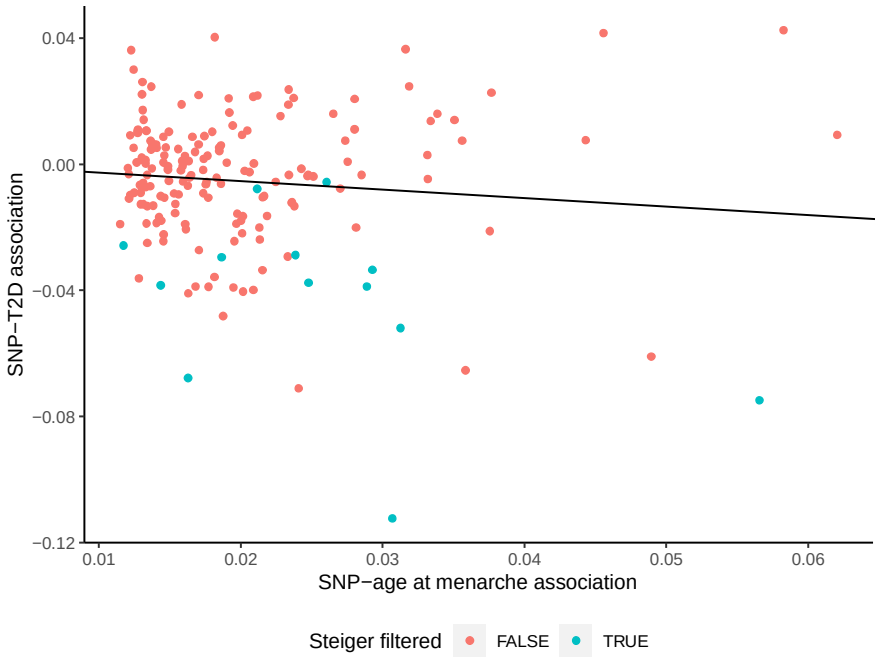

Figure 5 – SNP-exposure and SNP-outcome association for SNPs included in the univariable MR estimation in estimation of the effect of age at menarche on Type 2 Diabetes *SNP-exposure and SNP-outcome associations for the SNPs used in the MR estimation of the effect of age at menarche on type 2 diabetes (T2D). All SNPs are genome-wide significantly associated with age at menarche, those SNPs that showed a stronger association with childhood adiposity than age at menarche are highlighted in blue.*

medRxiv preprint doi: https://doi.org/10.1101/2024.09.05.24312293; this version posted September 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Tables

Table 1 – The number of SNPs selected in each simulation.

P-value	*No. SNPs*	*Exposure*	*Confounder*	*% for confounder*
50% of simulated SNPs associated with confounder				
$5x10^{-12}$	57	39	18	31%
$5x10^{-8}$	97	61	36	37%
$5x10^{-6}$	131	79	52	40%
$5x10^{-4}$	182	104	78	43%
87.5% of simulated SNPs associated with confounder				
$5x10^{-12}$	42	10	32	76%
$5x10^{-8}$	79	15	64	80%
$5x10^{-6}$	112	20	92	82%
$5x10^{-4}$	164	27	137	84%

The number of SNPs selected as associated with the exposure (out of the 400 generated) based on the p-value cutoff used for selection. Exposure gives the number of those that were generated as associated with the exposure and Confounder the number of those that were generated as associated with the confounder but selected as associated with the exposure. Numbers are means across all iterations, rounded to the nearest whole number.

Table 2 – Proportion of traits that show evidence of a causal effect on C-Reactive Protein in MR

	No of tests	**% P < 0.05**	**% P < 0.01**
Wald ratio/IVW	3770	21.7%	13.9%
MR Egger	1480	27.5%	19.2%
Weighted median	1478	29.0%	19.6%
Weighted mode	1478	22.7%	13.7%

Results from MR estimation of the effect of up to 5202 traits on CRP. Traits were selected from those strongly associated with at least one lead SNP in a recent CRP GWAS. Wald ratio was used for traits with 1 SNP available as an instrument and IVW was used for traits with 2 or more SNPs available as instruments. MR Egger, Weighted median and Weighted mode where additionally applied for all traits with 3 or more SNPs available as instruments.

Table 3 – MR effect estimates for age at menarche on Type 2 diabetes.

	All SNPs included			***Steiger filtering applied***		
	OR	**95% CI**	**p-value**	**OR**	**95% CI.**	**p-value**
IVW	0.76	[0.65, 0.89]	0.0005	0.90	[0.79, 1.02]	0.11
MR Egger	0.86	[0.55, 1.32]	0.51	1.21	[0.83, 1.75]	0.32
Weighted median	0.94	[0.81, 1.10]	0.43	1.03	[0.87, 1.21]	0.74
Weighted mode	1.05	[0.83, 1.33]	0.68	1.07	[0.83, 1.37]	0.61
MVMR	0.99	[0.78, 1.09]	0.33	-	-	-

OR: Effect estimates given in odds ratios for an increase in type 2 diabetes risk for a category increase in age at menarche. Univariable MR - 181 SNPs, with Steiger filtering – 168 SNPs, MVMR 273 SNPs.