bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# Bias in two-sample Mendelian randomization by using covariable-adjusted summary associations

Fernando Pires Hartwig[1,2]*, Kate Tilling[2,3], George Davey Smith[2,3], Deborah A Lawlor[2,3], Maria Carolina Borges[2,3]

[1]Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil.

[2]Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, United Kingdom.

[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom.

*Corresponding author. Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas (Brazil) 96020-220. Phone: +55 53 981068670. E-mail: fernandophartwig@gmail.com; fh15144@bristol.ac.uk.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# Abstract

Two-sample Mendelian randomization (MR) is increasingly used to strengthen causal inference using observational data. This method allows the use of freely accessible summary association results from genome-wide association studies (GWAS) for a range of traits. Some GWAS studies adjust for heritable covariables in an attempt to estimate direct effects of genetic variants on the trait of interest. One, both or neither of the genetic instrumental variables (IVs)-exposure association or genetic IVs-outcome association may have been adjusted for heritable covariables (referred to as GWAS covariables). However, it is unclear how this may affect two-sample MR analysis. We evaluated this in a simulation study comprising different scenarios that could motivate covariable adjustment in a GWAS. Our results indicate that the impact of covariable adjustment is highly dependent on the underlying causal structure. In the absence of residual confounding between exposure and covariable, between exposure and outcome, and between covariable and outcome, using covariable-adjusted summary associations for two-sample MR may eliminate bias due to horizontal pleiotropy. However, the presence of residual confounding (especially between the covariable and the outcome) leads to bias upon covariable adjustment, even in the absence of horizontal pleiotropy. Bias was lower when the true causal effect of the exposure on the outcome was zero compared to a non-zero causal effect. In an analysis using real data from the Genetic Investigation of ANthropometric Traits (GIANT) consortium and UK Biobank, the direction of the causal effect estimate of waist circumference on blood pressure changed upon adjustment of waist circumference for body mass index. Our findings indicate that using covariable-adjusted summary associations in MR should generally be avoided. When that is not possible, careful consideration of the causal relationships underlying the data (including potentially unmeasured confounders) is required to direct sensitivity analyses and interpret results with appropriate caution.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



**Keywords:** Two-sample Mendelian randomization; Summary results; Genome-wide association study; Bias; Genetic pleiotropy.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# Introduction

Mendelian randomization (MR) uses genetic variants to assess the influence of modifiable exposures on health outcomes (1, 2). As germline genetic variants are generally independent of confounding factors and are determined at conception, MR offers a more robust approach to confounding and reverse causation than other methods used in observational studies (3).

Two-sample MR is an extension to the one-sample MR design, where estimates for the association of genetic variants with exposure and with outcome are derived from different (non-overlapping) samples from the same underlying population (4). These estimates are combined to obtain the causal effect estimate of exposure on outcome (5). Given that genetic variants typically explain a small proportion of the variation in the exposure of interest, large sample sizes are required for adequately-powered MR studies. Therefore, in recent years, two-sample MR has substantially grown in popularity (6) since it capitalizes on the use of publicly-available summary association results from large genome-wide association studies (GWAS).

In GWAS, estimates for the association of genetic variants with the trait of interest are often conditioned on covariables. As an example, GWAS of waist-to-hip ratio have adjusted estimates for body mass index (BMI) (7), GWAS of lung function have adjusted estimates for height and stratified analysis by smoking status (8), and GWAS of birth weight has excluded pre-term births (9). Typically, the aim of conditioning on covariables is estimating the direct effect of genetic variants on the trait (i.e. effects independent of the covariable) or to improve statistical power by reducing residual variance. However, this strategy could introduce bias in GWAS association estimates if the covariable is a collider (or a descendant of a collider) on a pathway linking the genetic variant to the trait of interest. It has previously been shown that conditioning on heritable covariables can bias GWAS association estimates and that the

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



magnitude of this bias is a function of the effect of the genetic variant on the covariable and the correlation structure between the covariable and the trait of interest (10, 11).

Several two-sample MR studies have used summary association data from GWAS that have estimated the effect of genetic variants on the trait of interest conditioned on heritable covariables (e.g. (12-16)). The use of such GWAS data might have biased the findings of these MR studies for the reasons outlined above. In addition, it is challenging to predict the impact of such bias since data from two independent GWAS (one for the exposure and other for the outcome) are used in two-sample MR studies, meaning that the conditional estimates could be restricted to the association of genetic instruments with the exposure, with the outcome, or with both exposure and outcome. Despite the widespread use of covariable-adjusted summary associations (e.g. (12-16)), few considerations have been made about how this could affect the validity of results (e.g. (12-14, 17, 18)), particularly in the context of two-sample MR. In this paper we explore how covariable adjustment in GWAS affects two-sample MR findings using simulated and real data in scenarios that could motivate conditioning on a heritable covariable in a GWAS.

# Methods

## Simulation study

We performed a series of simulations to evaluate the consequences of covariable adjustment in MR. We were interested in evaluating situations where the genetic instrument is marginally associated with both the exposure ($X$) and a covariable ($W$). In these situations, the GWAS analyst might decide to adjust for $W$ as an attempt to estimate the effect of genetic variants on $X$ independent of $W$, thus generating covariable-adjusted summary association results that

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



will be used by two-sample MR analysts, who do not have access to covariable-unadjusted results.

We performed simulations under two main contexts: i) all genetic variants have directionally-consistent direct effects on the same traits (hereafter referred to as "homogeneous genetic variants"); ii) some genetic variants have direct effects only on some traits (hereafter referred to as "heterogeneous genetic variants"). By direct effect, we mean that there is an arrow from the genetic variant directly into the trait in the causal diagram.

Scenarios simulated under situation i) are illustrated in Figure 1. The measured variables are the genetic instrument $Z$, the exposure variable $X$, the outcome variable $Y$ and the covariable $W$ that has been adjusted for in a GWAS. In all situations, $X$ and $W$ are genetically correlated (ie, both are marginally associated with $Z$). The aim is to estimate the causal effect of $X$ on $Y$ using summary GWAS results in a two-sample MR framework. Therefore, there are four possible situations: no adjustment for $W$; adjusted $Z$-$X$ association but unadjusted $Z$-$Y$ association; unadjusted $Z$-$X$ association but adjusted $Z$-$Y$ association; or both $Z$-$X$ and $Z$-$Y$ associations adjusted for $W$.

Figure 1 depicts the assumed causal structures that we evaluated in the simulations. In scenarios A1-A5, $W$ fully mediates the effect of $Z$ on $X$. In scenarios B1-B5, $Z$ independently affects both $X$ and $W$ (in other words, $Z$ is a confounder of the $X$-$W$ association). In scenarios C1-C5, $X$ fully mediates the effect of $Z$ on $W$. In scenarios D1-D5, the effect of $Z$ on $X$ and $W$ is mediated by a common cause $R$, so that the effect of $Z$ on $W$ is correlated with the effect of $Z$ on $X$, even though there is no causal effect from $X$ to $W$ or vice-versa. In all scenarios A1-5 to D1-5, the instrumental variable assumptions hold, so that $Z$ is a valid instrument to estimate the causal effect of $X$ on $Y$. However, this is not the case in scenarios E1-E5 and F1-F5, which are identical to B1-B5 and D1-D5 (respectively), except that $W$ has a direct effect on $Y$ (ie, horizontal pleiotropy). More specifically, the $Z$-$W$ and $Z$-$X$ associations are independent in

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



scenarios E1-E5, meaning that in these scenarios there is horizontal pleiotropy, but the InSIDE (Instrument Strength Independent of Direct Effects) assumption holds. However, in scenarios F1-F5, the InSIDE assumption is violated because Z has an effect on a common cause of $W$ and $X$.

In scenarios A1-F1, there are no unmeasured confounders (other than R). To isolate the implications of unmeasured confounders when controlling for $W$, different confounders were included in different scenarios: $X$-$W$ confounder $U_{X,W}$ in scenarios A2-F2; $W$-$Y$ confounder $U_{X,W}$ in scenarios A3-F3; $X$-$Y$ confounder $U_{X,Y}$ in scenarios A4-F4; and finally all three confounders simultaneously in scenarios A5-F5.

To simulate data under situation ii), the same underlying causal structure used for situation i) was used, with the exception that, in all simulations, there were four non-overlapping subgroups of genetic variants: some with direct effects on $X$ only, some on $W$ only, some on both $X$ and $W$ (but not $R$), and some on $R$ only.

Even though Figure 1 depicts $X$ as a having a non-null causal effect on $Y$, all of these scenarios were also simulated for both a null and non-null causal effect from X to Y. A detailed description of the data generating model is provided in the Supplementary Material.

For analyses using unadjusted variant-exposure summary association results (regardless of whether the variant-outcome associations were adjusted for $W$), we selected variants with an unadjusted association with the exposure achieving a $F$ statistic of 10 or more. The same applies for analysis using adjusted variant-exposure summary association results.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



## Real data example: assessing the causal effect of waist circumference on blood pressure

We conducted an illustrative analysis to explore the impact of covariable adjustment in MR in a real data setting. The exposure of interest was waist circumference (WC), the outcome variables were systolic (SBP) and diastolic blood pressure (DBP), and the covariable was BMI. We selected genetic instruments of unadjusted WC and BMI-adjusted WC from the Genetic Investigation of ANthropometric Traits (GIANT) consortium (7) and calculated the corresponding instrument-BP summary association results using an interim release of UK Biobank data (19). Details on the data sources are provided in the Supplementary Material.

BMI was used as a covariable due to its strong correlation with WC, meaning that variants that affect WC might also affect BMI due to their effect on overall adiposity. Here, we assume that two distinct causal structures (Figure 2) are plausible. In panel A, the genetic variant has a direct effect on a latent variable (which we refer to as adiposity), which manifests itself in measurable constructs such as WC and BMI. Panel B depicts a scenario where the genetic variant has a direct effect on WC rather than on adiposity. Those mechanisms are not mutually exclusive, since different genetic variants can present those different direct effects, or even a single genetic variant can have direct effects on both. Of note, WC, BP, BMI and adiposity are analogous to $X$, $Y$, $W$ and $R$ (respectively) in Figure 1.

We aimed at replicating scenarios in which the summary association results (either unadjusted or adjusted for the covariable) are already available. Therefore, we selected two sets of WC genetic instruments: one using the unadjusted GWAS results; and another using the BMI-adjusted GWAS. In each case, independent genetic variants were selected as WC instruments if P-value $<5\times10^{-8}$. After quality control (described in detail in the Supplementary Materials), 37 single nucleotide polymorphisms (SNPs) and 60 SNPs were retained as genetic instruments for

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



BMI-unadjusted and BMI-adjusted WC, respectively. Prior to analysis, data was harmonised following the steps described in Hartwig et al (6) (as detailed in the Supplementary Material). As in the simulation study, four situations were considered: i) unadjusted instrument-WC and unadjusted instrument-BP associations; ii) adjusted instrument-WC and unadjusted instrument-BP associations; iii) unadjusted instrument-WC and adjusted instrument-BP associations; and iv) adjusted instrument-WC and adjusted instrument-BP associations.

### Statistical analyses

Causal effect estimates were obtained using multiplicative random effects inverse-variance weighting (IVW) (5, 20). In the simulation study, coverage and average causal effect estimates were obtained across 5 000 simulated datasets. Coverage was defined the proportion of times that 95% confidence intervals included the true causal effect.

## Results

### Simulation study

Supplementary Table 1 displays the number of select genetic instruments and mean $F$ statistic of the instrument-$X$ association (among selected instruments), in analyses adjusted and not adjusted for $W$. Since instruments are selected from the exposure GWAS the set of genetic instruments vary according to the analysis. Taking row 1 of Supplementary Table 1 as an example, on average 27 variants were used for analyses using unadjusted instrument-$X$ associations, and 19 variants for analyses using adjusted instrument-$X$ association. The mean $F$ statistic of 293 corresponds to the strength of association between the selected variants (here, 26 on average) and $X$ estimated without adjusting for $W$, while the value of 345 corresponds to the strength of association between the selected variants (here, 19 on average) and $X$ estimated adjusting for $W$.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



**Homogeneous genetic instruments**

For homogeneous genetic instruments, we only evaluated scenarios A5-F5 because in other scenarios adjustment for $W$ yielded in no variants achieving the instrument strength criterion. As shown in Supplementary Table 2, using covariable-adjusted associations did not eliminate bias due to horizontal pleiotropy in presence of unmeasured confounders (scenarios E5 and F-5), and introduced bias even in the absence of horizontal pleiotropy. The bias was mainly driven by the use of covariable-adjusted instrument-$Y$ associations, and was generally higher when the true causal effect was non-zero. Similar trends were observed in the coverage of the 95% confidence intervals.

**Heterogeneous genetic instruments**

Figure 3 displays the bias (see Supplementary Figure 1 for coverage) of the IVW estimate when instrument selection is performed among heterogeneous genetic variants. In the absence of unmeasured confounders (scenario 1), using adjusted instrument-$Y$ associations (especially in combination with adjusted instrument-$X$ associations) eliminated bias due to horizontal pleiotropy. However, the presence of a $W$-$Y$ confounder (scenario 3) resulted in bias in analyses using covariable-adjusted instrument-$Y$ associations even in the absence of horizontal pleiotropy and under the causal null. Moreover, in the presence of a non-null causal effect, the presence of a $X$-$W$ confounder (scenario 2) led to bias in analyses using unadjusted instrument-$X$ and adjusted instrument-$Y$ associations. Similar trends were observed in the coverage of the 95% confidence intervals (Supplementary Figure 1). In general, in cases where $W$ does not have a direct effect on $Y$ (A-D, and thus none of the variants have horizontal pleiotropic effects), the bias was considerably lower in analysis using unadjusted than adjusted instrument-$Y$ association estimates. On the other hand, in the presence of horizontal pleiotropy mediated by $W$, analysis using unadjusted instrument-$Y$ association estimates

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



presented more bias. Bias was larger and coverage lower when the true causal effect was non-zero.

## Real data example

Supplementary Figure 2 displays the correlation between coefficients for SNP-unadjusted and SNP-adjusted WC associations. Coefficients were uncorrelated (r=-0.04) among genetic instruments for unadjusted WC (n = 37) and moderately correlated (r=0.47) among genetic instruments for adjusted WC (n = 60). Distributions of test statistics for SNP-traits associations are shown in Supplementary Figure 3.

In the unadjusted SNP-exposure and SNP-outcome analysis, each standard unit increase in WC was related to an increase of 0.06 mmHg in SBP (95% CI: -0.01, 0.13) and of 0.12 mmHg in DBP (95% CI: 0.05, 0.19). These effects changed in direction when only adjusting the SNP-outcome association for BMI (SBP: -0.09, 95%CI: -0.17, -0.02; DBP: -0.11, 95%CI: -0.19, -0.04). Effect estimates also changed direction, but were consistent with the null when adjusting the SNP-exposure association without (SBP: -0.06, 95%CI: -0.14, 0.01; DBP: -0.02, 95%CI: -0.09, 0.05) or with concomitant adjustments for SNP-outcome associations (SBP: -0.04, 95%CI: -0.11, 0.03; DBP: 0.01, 95%CI: -0.06, 0.08) (Figure 4).

After excluding SNPs associated with BMI ($p < 0.05$), results were similar for the adjusted SNP-exposure analyses. The exclusion of BMI-related SNPs dramatically restricted the number of SNPs included in the unadjusted SNP-exposure analyses (two out of 37 SNPs), and, as a result, estimates are highly imprecise (Figure 5).

# Discussion

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Our results indicate that the impact of covariable adjustment in two-sample MR depends on the causal relation and confounding structure between genetic instruments, exposure, covariable and outcome. In addition, the magnitude and direction of bias will vary depending on whether associations between instrument-exposure, instrument-outcome or both are adjusted for the same covariables. In an analysis using real data from GIANT consortium and UK Biobank, the estimated causal effect of waist circumference on blood pressure changed substantially upon adjustment for BMI.

The strong dependence of the results on the underlying causal structure was expected. In the absence of unobserved common causes (confounders) between exposure-covariable, exposure-outcome and covariable-outcome, covariable adjustment eliminates bias due to horizontal pleiotropy mediated by such covariable (scenarios E and F). However, absence of unobserved confounding is unrealistic in observational studies and is one of the primary motivations for performing MR. In the presence of unobserved confounding, mainly between the covariable and the outcome, covariable adjustment will likely lead to bias even in the absence of horizontal pleiotropy due to collider bias where genetic instruments are marginally associated with the covariable. Therefore, minimising horizontal pleiotropy is generally an invalid justification for using covariable-adjusted summary association results for two-sample MR.

Bias was generally weaker when the true causal effect of the exposure on the outcome is null, and this was consistent among the several scenarios we evaluated. This suggests that covariable adjustment has a lower impact on testing the null hypothesis than on estimating a non-zero causal effect. Indeed, even in the absence of any unmeasured confounding, covariable adjustment may lead to bias when the true causal effect is not null. Therefore, null results from a MR analysis using covariable-adjusted summary association results are generally more reliable (in the sense of being less likely to be a consequence of covariable-adjustment

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



bias) than non-null results, assuming that it is generally unlikely that bias due to using covariable-adjusted summary associations perfectly balances out a given non-null true causal effect.

Using covariable-adjusted instrument-outcome summary associations was also consistently related to more bias compared to using covariable-adjusted instrument-exposure summary associations. The exception was scenarios where there was horizontal pleiotropy. However, since attempting to minimise horizontal pleiotropy via covariable adjustment is generally unjustified due to the likely presence of unmeasured confounders, we are primarily concerned with the scenarios where there is no horizontal pleiotropy. Another result that was generally consistent across the different scenarios was that selecting genetic instruments after covariable adjustment (i.e., using covariable-adjust summary instrument-exposure association results) presented less bias than selection before adjustment. This also minimises other possible problems, such as running into lack of association between the instrument and the exposure upon covariable adjustment, which could compromise precision and lead to weak instrument bias.

To our knowledge, this is the first study to systematically assess the impact of covariable adjustment in (two-sample) MR. Conditioning on a heritable covariable can introduce bias when the covariable is a collider in the pathway between instrument-exposure and/or instrument-outcome. Collider bias can also be introduced in MR studies in other settings, such as in disease progression studies which include a selected (i.e., case-only) group of individuals (17).

It is important to emphasise that using covariable-adjusted GWAS summary association results in two-sample MR studies differs from applying multivariable MR (MVMR) on unadjusted GWAS summary association results to estimate the effect of two or more exposures on an outcome (18). The focus of our paper is to investigate bias in two-sample MR due to using

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



covariable-adjusted summary association results, especially in the case when those are the only available option (e.g., when using summary data from GWAS consortia that only performed covariable-adjusted analyses). We are not primarily concerned with whether using covariable-adjusted summary associations reduces bias due to horizontal pleiotropy, even though our results indicate that this is unlikely to be the case in the presence of unmeasured confounding. When one aims to use covariable data for this aim, MVMR should be preferred, because MVMR uses genetically predicted variations in both the exposure and covariable(s), and is therefore less susceptible to collider bias (21).

One of the strengths of our study was that our simulations covered a wide range of scenarios, thus allowing a detailed evaluation of covariable-adjustment bias in a variety of situations. The simulation study was also complemented with a real data example illustrating the strong influence that covariable-adjustment may have not only on the magnitude, but also on the presence and direction of the causal effect estimate. However, any simulation study is a simplification of a reality that is likely to be much more complex. It is impossible to simulate all possible scenarios that might be of relevance to this topic. Moreover, some results, especially quantitative estimates of bias and coverage, are highly dependent on the data-generating model. Therefore, our results should be interpreted qualitatively, as general indications of some of the main aspects related to covariable-adjustment in MR. It was reassuring that some results (as described above) were consistent across scenarios, which indicates that they may apply generally (although not universally) to covariable-adjusted MR.

In conclusion, our findings indicate that using summary association results adjusted for heritable covariables may lead to bias in two-sample MR due to unmeasured confounding. We recommend avoiding adjustment for such covariables in the context of MR. When only covariable-adjusted data is available, it is important to carefully consider the causal structure underlying the research question to understand the potential impact on the results. In such

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



cases, we recommend that researchers refrain from interpreting the causal estimate too literally (which indeed requires parametric assumptions in addition to the core instrumental variable assumptions even in the absence of covariable adjustment) and focus on testing the causal null hypothesis. To account for horizontal pleiotropy due to measured covariables, MVMR should be preferred over non-genetic covariable adjustment.

# Funding

KT, GDS, DAL and MCB work in the MRC Integrative Epidemiology Unit at the University of Bristol that receives funding from the UK Medical Research Council [MC_UU_00011/1, MC_UU_00011/3 and MC_UU_00011/6]. DAL's contribution is also supported by the European Research Council under the European Union's Seventh Framework Programme [FP/2007-2013] / ERC Grant Agreement [Grant number 669545; DevelopObese] and from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733206 (LifeCycle). MCB is supported by a Medical Research Council (MRC) Skills Development Fellowship [MR/P014054/1]. DAL is a UK National Institute of Health Research Senior Investigator [NF-SI-0611-10196].

# Acknowledgements

This research has been conducted using the UK Biobank Resource under application number 15825. We would like to thank the participants and researchers from the UK Biobank who contributed or collected data.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# Disclosures

**No competing interests:** FPH, KT, GDS, MCB.

**Competing interests:** DAL has received funding from several national and international government and charitable organisations, Roche Diagnostics and Medtronic for research unrelated to that presented in this paper**.**

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# References

1. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133-63.
2. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1-22.
3. Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 2007;4(12):e352.
4. Lawlor DA. Commentary: Two-sample Mendelian randomization: opportunities and challenges. Int J Epidemiol. 2016;45(3):908-15.
5. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658-65.
6. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol. 2016;45(6):1717-26.
7. Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature. 2015;518(7538):187-96.
8. Soler Artigas M, Wain LV, Miller S, Kheirallah AK, Huffman JE, Ntalla I, et al. Sixteen new lung function signals identified through 1000 Genomes Project reference panel imputation. Nat Commun. 2015;6:8658.
9. Horikoshi M, Beaumont RN, Day FR, Warrington NM, Kooijman MN, Fernandez-Tajes J, et al. Genome-wide associations for birth weight and correlations with adult disease. Nature. 2016;538(7624):248-52.
10. Aschard H, Vilhjalmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. Am J Hum Genet. 2015;96(2):329-39.
11. Vansteelandt S, Goetgeluk S, Lutz S, Waldman I, Lyon H, Schadt EE, et al. On the adjustment for covariates in genetic association analysis: a novel, simple principle to infer direct causal effects. Genet Epidemiol. 2009;33(5):394-405.
12. Dale CE, Fatemifar G, Palmer TM, White J, Prieto-Merino D, Zabaneh D, 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):2373-88.
13. Emdin CA, Khera AV, Natarajan P, Klarin D, Zekavat SM, Hsiao AJ, et al. Genetic Association of Waist-to-Hip Ratio With Cardiometabolic Traits, Type 2 Diabetes, and Coronary Heart Disease. JAMA. 2017;317(6):626-34.
14. Borges MC, Oliveira IO, Freitas DF, Horta BL, Ong KK, Gigante DP, et al. Obesity-induced hypoadiponectinaemia: the opposite influences of central and peripheral fat compartments. Int J Epidemiol. 2017.
15. Au Yeung SL, Lin SL, Li AM, Schooling CM. Birth weight and risk of ischemic heart disease: A Mendelian randomization study. Sci Rep. 2016;6:38420.
16. Clarke TK, Adams MJ, Davies G, Howard DM, Hall LS, Padmanabhan S, et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117). Mol Psychiatry. 2017;22(10):1376-84.
17. Davey Smith G, Paternoster L, Relton C. When Will Mendelian Randomization Become Relevant for Clinical Practice and Public Health? JAMA. 2017;317(6):589-91.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



18. Holmes MV, Davey Smith G. Problems in interpreting and using GWAS of conditional phenotypes illustrated by 'alcohol GWAS'. Mol Psychiatry. 2018.
19. UK Biobank. Protocol for a large-scale prospective epidemiological resource 2007 [Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf.
20. Johnson T. Efficient calculation for multi-SNP genetic risk scores Technical Report, The Comprehensive R Archive Network 2013 [Available from: http://cran.rproject.org/web/packages/gtx/vignettes/ashg2012.pd.
21. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2018.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



# Figures

**Figure 1. Causal structures that were assessed in the simulation study. $Z$: genetic instrument; $W$: covariable; $R$: possible direct consequence of $Z$; $X$: exposure; $Y$: outcome: $U$: unmeasured common cause.**

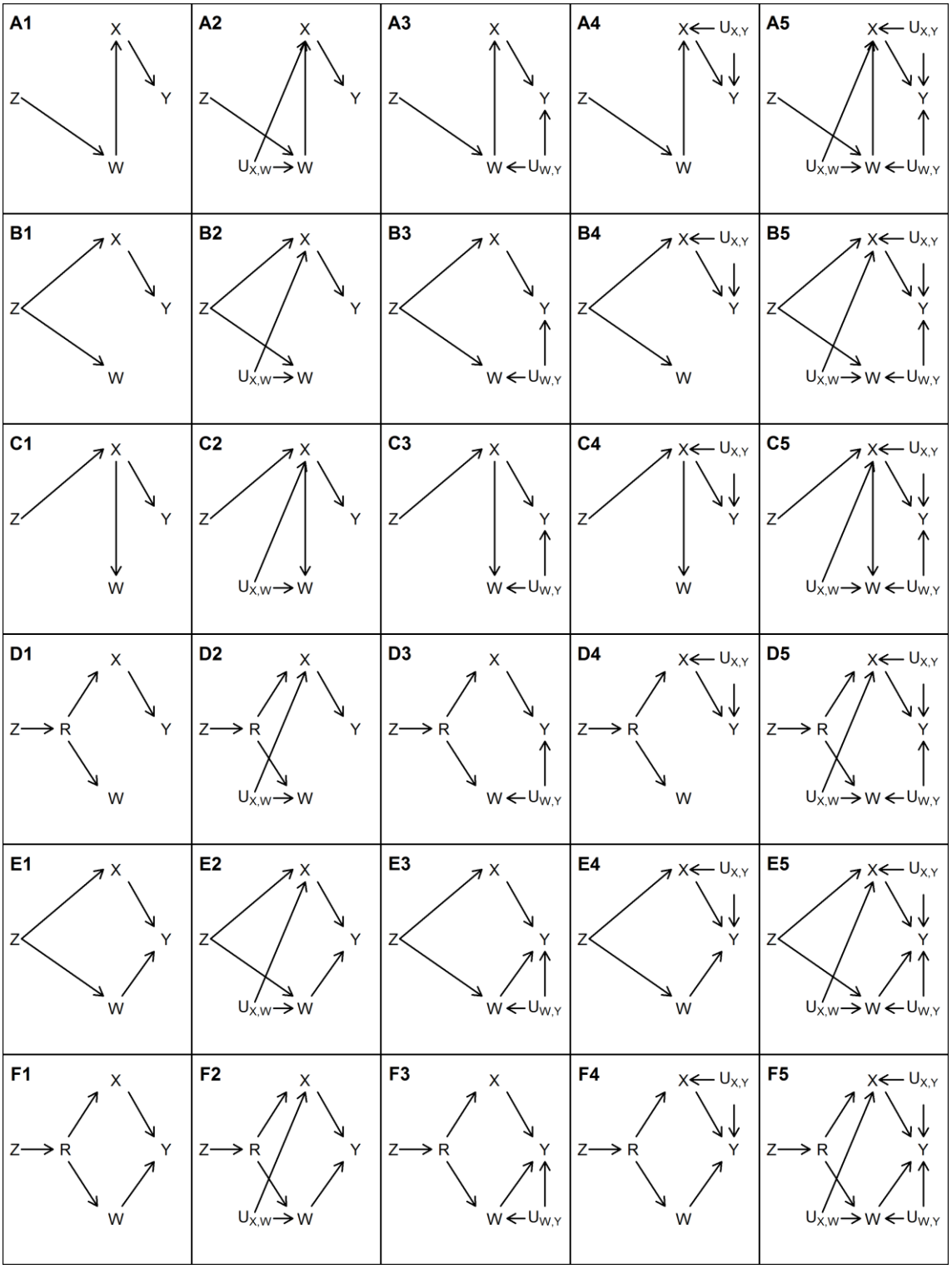
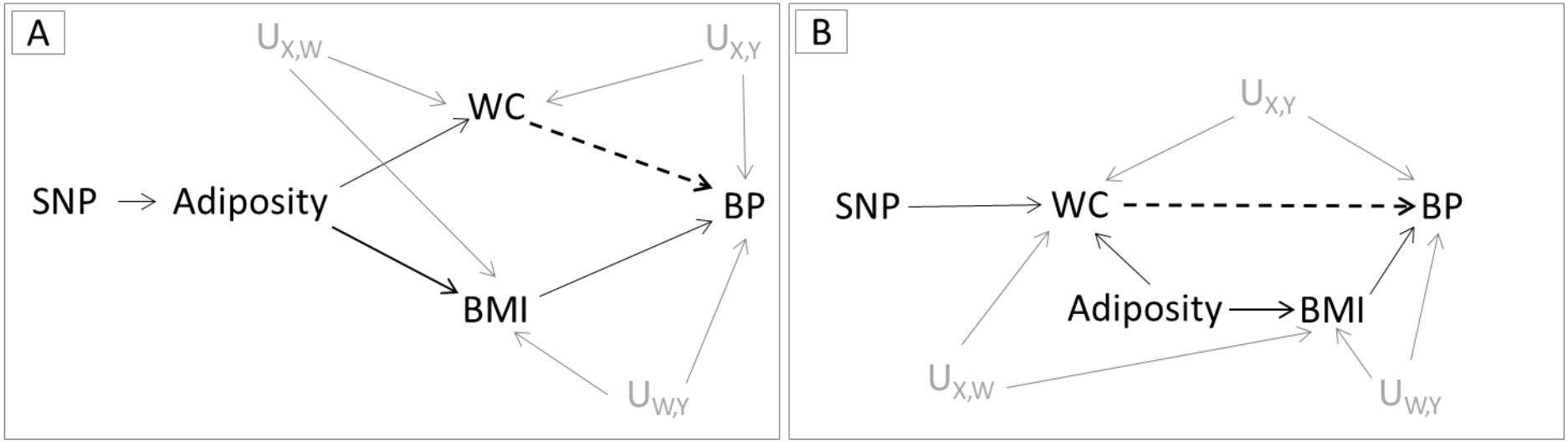


**Figure 2. Causal diagrams representing the assumed causal relationships in the two-sample Mendelian randomization analysis of waist circumference (WC) on blood pressure (BP). The genetic instruments (SNPs) are assumed to influence WC by either affecting overall body adiposity (proxied by BMI) [A] or by specifically changing body fat distribution [B]. The grey solid lines represent the effect of confounders between exposure-outcome ($U_{X,Y}$), exposure-covariable ($U_{X,W}$) and covariable-outcome ($U_{W,Y}$). The dashed lines represent the relationship being tested between WC and BP.**



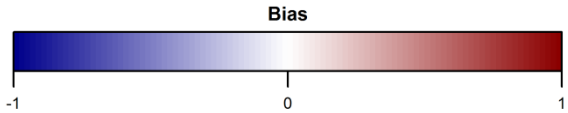

BMI: body mass index; IV: instrumental variable; SNP: single nucleotide polymorphism; U: unmeasured confounders.

bioRxiv preprint doi: https://doi.org/10.1101/816363; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



**Figure 3. Mean bias across 5,000 simulations of the causal effect estimate using heterogeneous genetic instruments.**

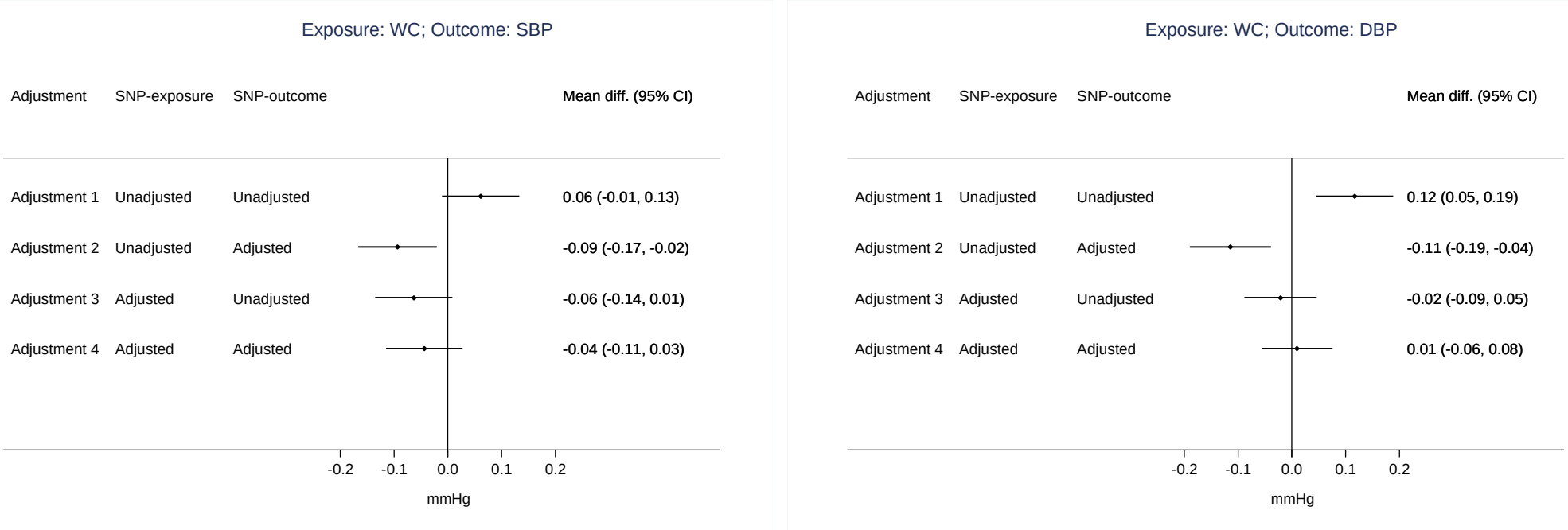
| Scenario | | | Unadjusted SNP-X Unadjusted SNP-Y | Adjusted SNP-X Unadjusted SNP-Y | Unadjusted SNP-X Adjusted SNP-Y | Adjusted SNP-X Adjusted SNP-Y |
|---|---|---|---|---|---|---|
| 1 | β=0 | A | 0.000 | 0.000 | 0.000 | 0.000 |
| | | B | 0.000 | 0.000 | 0.000 | 0.000 |
| | | C | 0.000 | 0.000 | 0.000 | 0.000 |
| | | D | 0.000 | 0.000 | 0.000 | 0.000 |
| | | E | 0.175 | 0.172 | 0.000 | 0.000 |
| | | F | 0.220 | 0.147 | 0.000 | 0.000 |
| | β=0.2 | A | -0.001 | 0.036 | -0.052 | -0.001 |
| | | B | -0.001 | 0.002 | -0.004 | -0.001 |
| | | C | -0.001 | 0.013 | -0.082 | -0.002 |
| | | D | -0.001 | 0.014 | -0.024 | -0.001 |
| | | E | 0.174 | 0.173 | -0.003 | -0.001 |
| | | F | 0.219 | 0.160 | -0.024 | -0.002 |
| 2 | β=0 | A | 0.000 | 0.000 | 0.000 | 0.000 |
| | | B | 0.000 | 0.000 | 0.000 | 0.000 |
| | | C | 0.000 | 0.000 | 0.000 | 0.000 |
| | | D | 0.000 | 0.000 | 0.000 | 0.000 |
| | | E | 0.174 | 0.087 | 0.000 | 0.000 |
| | | F | 0.220 | 0.036 | 0.000 | 0.000 |
| | β=0.2 | A | -0.001 | 0.025 | -0.074 | -0.001 |
| | | B | -0.001 | 0.008 | -0.019 | -0.001 |
| | | C | -0.001 | -0.056 | -0.114 | -0.001 |
| | | D | -0.001 | 0.005 | -0.044 | -0.001 |
| | | E | 0.174 | 0.096 | -0.019 | -0.001 |
| | | F | 0.219 | 0.041 | -0.044 | -0.002 |
| 3 | β=0 | A | 0.000 | 0.000 | -0.108 | -0.077 |
| | | B | 0.000 | 0.000 | -0.078 | -0.077 |
| | | C | 0.000 | 0.000 | -0.169 | -0.033 |
| | | D | 0.000 | 0.000 | -0.099 | -0.066 |
| | | E | 0.175 | 0.171 | -0.078 | -0.077 |
| | | F | 0.219 | 0.146 | -0.098 | -0.065 |
| | β=0.2 | A | -0.001 | 0.037 | -0.160 | -0.078 |
| | | B | -0.001 | 0.002 | -0.082 | -0.077 |
| | | C | -0.001 | 0.013 | -0.250 | -0.035 |
| | | D | -0.001 | 0.014 | -0.122 | -0.067 |
| | | E | 0.173 | 0.172 | -0.081 | -0.077 |
| | | F | 0.218 | 0.160 | -0.122 | -0.067 |
| 4 | β=0 | A | 0.000 | 0.000 | 0.000 | 0.000 |
| | | B | 0.000 | 0.000 | 0.000 | 0.000 |
| | | C | 0.000 | 0.000 | -0.075 | -0.015 |
| | | D | 0.000 | 0.000 | 0.000 | 0.000 |
| | | E | 0.175 | 0.171 | 0.000 | 0.000 |
| | | F | 0.219 | 0.146 | 0.000 | 0.000 |
| | β=0.2 | A | -0.001 | 0.037 | -0.052 | 0.000 |
| | | B | -0.001 | 0.001 | -0.004 | -0.001 |
| | | C | -0.001 | 0.013 | -0.157 | -0.016 |
| | | D | -0.001 | 0.014 | -0.024 | -0.001 |
| | | E | 0.174 | 0.173 | -0.003 | -0.001 |
| | | F | 0.219 | 0.160 | -0.024 | -0.001 |
| 5 | β=0 | A | 0.000 | 0.000 | -0.108 | -0.037 |
| | | B | 0.000 | 0.000 | -0.078 | -0.039 |
| | | C | 0.000 | 0.000 | -0.243 | 0.114 |
| | | D | 0.000 | 0.000 | -0.098 | -0.016 |
| | | E | 0.175 | 0.088 | -0.078 | -0.039 |
| | | F | 0.219 | 0.035 | -0.098 | -0.016 |
| | β=0.2 | A | -0.001 | 0.025 | -0.182 | -0.038 |
| | | B | -0.001 | 0.008 | -0.097 | -0.041 |
| | | C | -0.001 | -0.057 | -0.359 | 0.114 |
| | | D | -0.001 | 0.005 | -0.142 | -0.017 |
| | | E | 0.175 | 0.097 | -0.098 | -0.041 |
| | | F | 0.219 | 0.041 | -0.142 | -0.017 |




$\beta$: True causal effect of the exposure ($X$) on the outcome ($Y$).

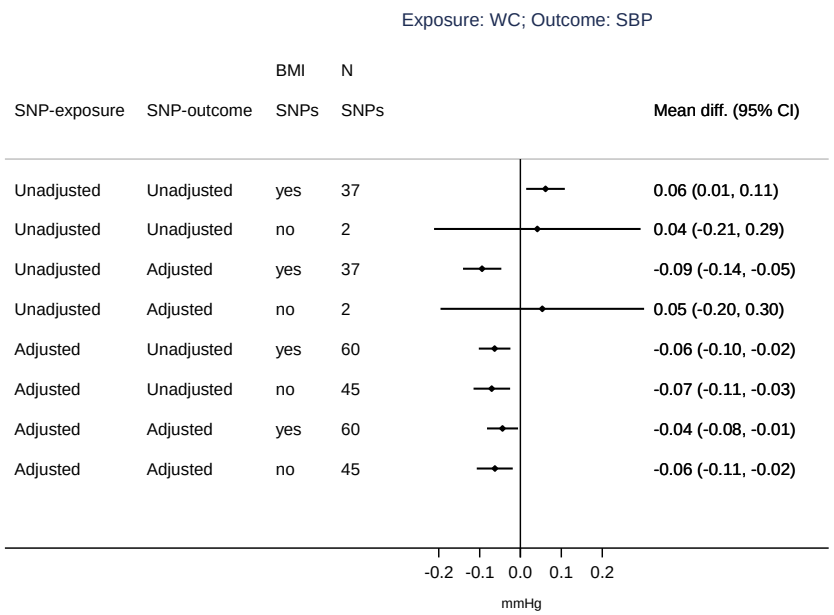
Scenarios A-F assume different causal relationships among $X$, $Y$, the instrument ($Z$), the covariate ($W$) and a common cause of $X$ and $W$ affected by $Z$ ($R$). In scenarios A1-F1, there are no unmeasured confounders. In scenarios A2-F2, there is an unmeasured common cause of $X$ and $W$. In scenarios A3-F3, there is an unmeasured common cause of $W$ and $Y$. In scenarios A4-F4, there is an unmeasured common cause of $X$ and $Y$. In scenarios A5-F5, all these three unmeasured confounders are present. The scenarios are illustrated in Figure 1 and described in detail in the "Simulation study" section.

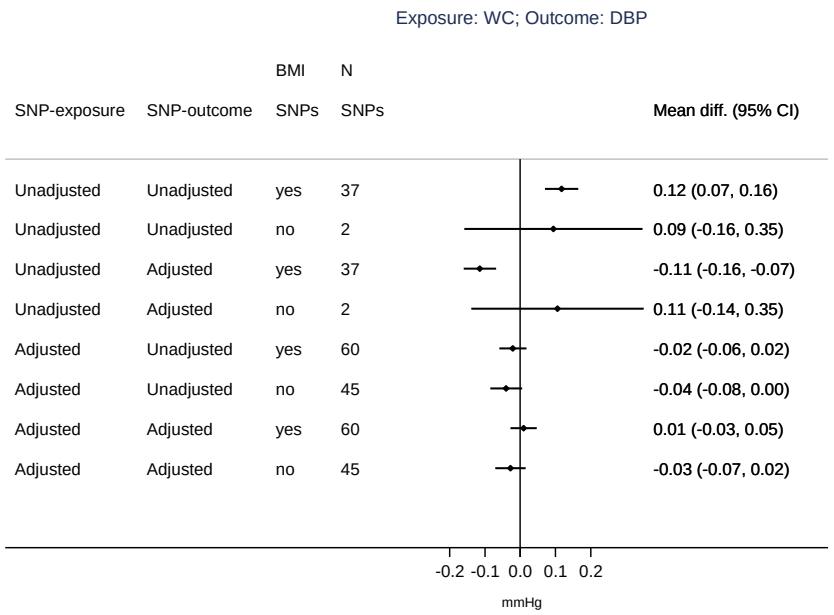
**Figure 4. Two-sample Mendelian randomization estimates of the effect of waist circumference (WC) on systolic blood pressure (SBP) or diastolic blood pressure (DBP) for different combinations of adjustments for BMI in SNP-exposure or SNP-outcome association.**




Effect estimates are expressed as mean difference, and 95% CI, of SBP or DBP (in mmHg) per standard unit increase in WC. 37 SNPs and 60 SNPs were used as instruments for waist circumference unadjusted and adjusted for BMI, respectively.

**Figure 5. Two-sample Mendelian randomization estimates of the effect of waist circumference (WC) on systolic blood pressure (SBP) or diastolic blood pressure (DBP) including (BMI SNPs: yes) or removing (BMI SNPs: no) SNPs associated with BMI.**




Effect estimates are expressed as mean difference, and 95% CI, of SBP or DBP (in mmHg) per standard unit increase in WC. 37 SNPs and 60 SNPs were used as instruments for WC unadjusted and adjusted for BMI, respectively. After removing instruments associated with BMI (p-value < 0.05), 2 SNPs and 45 SNPs were retained as instruments for unadjusted and BMI-adjusted WC, respectively.