

Pleiotropy-robust Mendelian Randomization

Hans van Kippersluis^{1,2,3} and Cornelius A. Rietveld^{1,3,4,*}

¹Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands, ²Department of Economics, Chinese University of Hong Kong, Hong Kong SAR, ³Tinbergen Institute, The Netherlands, ⁴Erasmus University Rotterdam Institute for Behavior and Biology, Erasmus University Rotterdam, The Netherlands

*Corresponding author. Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA, Rotterdam, The Netherlands. E-mail: nrietveld@ese.eur.nl.

Abstract

Background: The potential of Mendelian Randomization studies is rapidly expanding due to (i) the growing power of GWAS meta-analyses to detect genetic variants associated with several exposures, and (ii) the increasing availability of these genetic variants in large-scale surveys. However, without a proper biological understanding of the pleiotropic working of genetic variants, a fundamental assumption of Mendelian Randomization (the exclusion restriction) can always be contested.

Methods: We build upon and synthesize recent advances in the econometric literature on instrumental variables (IV) estimation that test and relax the exclusion restriction. Our Pleiotropy-robust Mendelian Randomization (PRMR) method first estimates the degree of pleiotropy, and in turn corrects for it. If a sample exists for which the genetic variants do not affect the exposure, and pleiotropic effects are homogenous, PRMR obtains unbiased estimates of causal effects in case of pleiotropy.

Results: Simulations show that existing MR methods produce biased estimators for realistic forms of pleiotropy. Under the aforementioned assumptions, PRMR produces unbiased estimators. We illustrate the practical use of PRMR by estimating the causal effect of (i) cigarettes smoked per day on Body Mass Index (BMI); (ii) prostate cancer on self-reported health, and (iii) educational attainment on BMI in the UK Biobank data.

Conclusions: PRMR allows for instrumental variables that violate the exclusion restriction due to pleiotropy, and corrects for pleiotropy in the estimation of the causal effect. If the degree of pleiotropy is unknown, PRMR can still be used as a sensitivity analysis.

Key words: Mendelian Randomization, pleiotropy, plausibly exogenous

Key messages:

- If genetic variants have pleiotropic effects, causal estimates of Mendelian Randomization studies will be biased.
- Pleiotropy-robust Mendelian Randomization (PRMR) produces unbiased causal estimates in case (i) a subsample can be identified for which the genetic variants do not affect the exposure, and (ii) pleiotropic effects are homogenous.
- If such a subsample does not exist, PRMR can still routinely be reported as a sensitivity analysis in any MR analysis.
- If pleiotropic effects are not homogenous, PRMR can be used as an informal test to gauge the exclusion restriction.

Introduction

Establishing causal effects of treatments or exposures on behavioral and disease outcomes is of great public health importance.¹ The practice of “Mendelian Randomization” (MR) uses genetic variants as instrumental variables (IV) for a certain modifiable exposure in order to estimate the causal effect of that exposure on a certain disease or other outcome.^{2,3} This method has the potential to overcome the traditional biases due to confounding and reverse causality that plague observational studies.⁴ The past decade has seen increasing interest in MR,^{4,5,6,7} and its potential is rapidly developing through the increasing number of Genome-Wide Association Studies (GWAS) that find robust associations between genetic variants and exposures of interest.^{8,9}

The assumptions of MR have been discussed extensively^{2,10,11,12,13}, and the advantages and disadvantages of MR are heavily debated.^{14,15} Recently, there has been much progress in dealing with the disadvantages of MR, yet the critical assumption remains the exclusion restriction: the proposed IV (genetic variant) should not directly affect the outcome.^{14,15} While this assumption can be contested in any IV approach, the assumption is even more critical in the context of MR, since the biological working of genes is usually poorly understood.^{13,16}

In particular, more and more studies show how the same genetic variant affects multiple outcomes through different biological pathways, a phenomenon known as biological (or horizontal) pleiotropy¹⁷, and this violates the exclusion restriction. In contrast, mediated (or vertical) pleiotropy, in which a genetic variant is associated with a certain phenotype and this phenotype is causal for a second phenotype, is not

problematic.⁴ Therefore, we will only focus on biological pleiotropy, for brevity “pleiotropy” from here.

In response to the possible problem of pleiotropy, a couple of MR approaches have been proposed. Davey Smith and Hemani suggest using multiple genetic variants consecutively as IV, and argue that it is increasingly unlikely that 2, 3 or 4 different genetic variants produce the same estimate of the causal effect.⁴ While an appealing argument, this approach still hinges critically on the assumption that at least some of the genetic variants do not violate the exclusion restriction. Moreover, this informal test cannot discriminate between violations of the exclusion restriction and heterogeneous causal effects.¹⁸ Kang et al. propose a method that produces valid estimates in case at least 50% of the combined instrument strength across all variants originates from variants that satisfy the exclusion restriction.¹⁹ While helpful, it is generally not possible to distinguish the valid IVs from the invalid IVs in case the estimates based on different sets of IVs diverge.²⁰ Moreover, this approach, like the weighted median estimator²¹, still requires some of the genetic variants to satisfy the exclusion restriction.

Bowden et al. propose to apply an Egger regression to Mendelian Randomization estimates.²⁰ The idea is that a violation of the exclusion restriction leads to a bias of the MR estimate that is inversely proportional to the first-stage coefficient of the IV on the exposure. Under the assumption that, across all genetic variants, the effect of the IV on the outcome and the effect of the IV on the exposure are uncorrelated (“InSIDE assumption”), this implies that IVs with a stronger effect on the exposure should give less-biased MR estimates. A regression of the MR estimates on the first stage coefficients including an intercept then provides a consistent estimate of the causal effect. As acknowledged by the authors, this InSIDE assumption cannot be tested and

may not hold if the genetic variants used as IVs are correlated with confounders of the association between exposure and outcome.

In this paper we build upon, and combine, two recent advances in the econometric literature on IV estimation that test and relax the exclusion restriction. We introduce the “plausibly exogenous” method²² in MR research to account for pleiotropy. Two studies have previously noticed the possibility of applying this method as a sensitivity analysis in the context of MR,^{12,23} but both papers do not provide any guidance on how to choose the essential input parameters. Our innovation is that we combine this method with another stream of econometric research that designs auxiliary regressions to test for violations of the exclusion restriction.^{24,25,26} The intuition is that in a subsample for which the first stage (that is, the effect of the IV on the exposure) is zero, the reduced form (that is, the effect of the IV on the outcome) should be zero too in case the exclusion restriction is satisfied. While traditionally used merely as a test of the exclusion restriction, there is no earlier notion that the reduced-form estimate obtained in this subsample is exactly the input required for the “plausibly exogenous” method. We term the synthesis of these techniques Pleiotropy-robust Mendelian Randomization (PRMR).

Simulation results show that if a sample exists for which the first stage is zero, and the pleiotropic effects are homogenous, it is possible to obtain unbiased estimates of causal effects using PRMR, even when all genetic instruments violate the exclusion restriction. We empirically illustrate our method by estimating the effect of (i) the number of cigarettes smoked per day on body mass index (BMI), (ii) prostate cancer on subjective health evaluations, and (iii) educational attainment on BMI.

Methods

Mendelian Randomization

In the general case, one is interested in the causal effect β of a certain exposure X on an outcome Y . The idea of MR is that there is a vector of genetic variants G (usually single nucleotide polymorphisms (SNPs), an allele score, or polygenic score) that is known to be correlated with the exposure X , but is assumed to be uncorrelated with other (unobserved) determinants of the outcome Y . Consider the equations (we follow the notation of Bowden et al.²⁰ in matrix notation here)

$$Y = X\beta + G\alpha + \varepsilon^Y \quad (1)$$

$$X = G\gamma + \varepsilon^X \quad (2)$$

where ε^Y and ε^X are composite error terms including unobserved confounders U .

The assumptions of MR are (see Figure 1 for a graphical exposition):^{12,13,16}

1. Relevance: The genetic variants G have an effect on the exposure X : $\gamma \neq 0$.
2. Independence: The genetic variants G are uncorrelated with any confounders of the exposure-outcome relationship.
3. Exclusion restriction: The genetic variants G affect the outcome Y only through the exposure X : $\alpha = 0$.

The use of genetic variants as IV has at least two very attractive properties. First, publicly available GWAS results make it relatively straightforward to select genetic variants G for which $\gamma \neq 0$, i.e. genetic variants that are robustly associated with the exposure of interest. Second, given that genetic variants are randomly distributed at conception, conditional on population stratification variables or family-specific effects, it is usually plausible that the independence assumption holds.²⁷

The exclusion restriction is widely acknowledged as the most problematic assumption of MR.^{14,15} In particular, the existence of pleiotropy would lead to a violation of the exclusion restriction, and $\alpha \neq 0$ in equation (1). In traditional MR it is assumed however that α is equal to 0, which leads to biased estimates of β , the causal effect of interest, in case pleiotropy is present. Moreover, this bias gets amplified by the typically low explanatory power of the genetic variants for the exposure.²⁸

(Figure 1 here)

Pleiotropy-robust Mendelian Randomization (PRMR)

In the “plausibly exogenous” method²² the assumption that $\alpha = 0$ is relaxed, and replaced by a user specified assumption on the plausible value, range or distribution of α . When the prior on α follows a Normal distribution with mean μ_α and variance Ω_α , and the uncertainty about α reduces with the sample size (i.e., “local-to-zero”), then the estimate of the causal effect β in equation (1) is given by:

$$\hat{\beta} \sim N(\beta + A\mu_\alpha, V_{MR} + A\Omega_\alpha A') \quad (3)$$

Where $N()$ indicates the Normal distribution, $A = (X'G(G'G)^{-1}G'X)^{-1}(X'G)$, and β and V_{MR} are the traditional MR point estimate and variance-covariance matrix, respectively. The plausibly exogenous method on itself, however, provides no guidance on the value or distribution of α .

We can, however, estimate the value of α if there is a population subgroup for which the first stage is known to be zero. A recent stream of econometric research emphasizes the identification of these subgroups to test the exclusion restriction.^{24,25,26} An early example is Altonji et al. (2005)²⁴, who investigate the validity of the instrument ‘being Catholic’ to study the effect of attending a Catholic high school on a wide variety of outcomes. They identify a subsample of public eighth graders among

which practically nobody subsequently attends a Catholic high school. Hence, among this subsample the first stage is zero, and any association between the IV (being Catholic) and the outcome reflects a direct effect, indicating a violation of the exclusion restriction.

While this strategy is so far exclusively used as a test of the exclusion restriction, our observation is that the result of this auxiliary regression is exactly the required input for the “plausibly exogenous” method. Consider the reduced form equation that is obtained by substituting (2) into (1):

$$Y = G(\alpha + \beta\gamma) + (\varepsilon^Y + \beta\varepsilon^X) \quad (4)$$

In a subsample for which the first stage is zero ($\gamma = 0$), the reduced form coefficient of the genetic variant is an estimator for α .

Practically, we suggest first estimating the reduced-form (4) in a sample for which $\gamma = 0$, to obtain estimates $\mu_\alpha = \hat{\alpha}$ and Ω_α equal to the squared standard error of $\hat{\alpha}$. Thereafter, plug in μ_α and Ω_α into the plausibly exogenous equation (3), to obtain estimates of the causal effect of interest β . The estimator is easy to obtain in standard software. For example, the user-written command “plausexog” is readily available in STATA.²⁹ While the plausibly exogenous method has been originally developed for continuous outcomes, it is also valid for binary outcomes in as far one is willing to estimate linear probability models for binary outcomes.

Simulation

We present a simulation study to illustrate the performance of regular MR through Two-Stage Least Squares (2SLS), the Inverse Variance Weighting (IVW) method³⁰, MR-Egger regression, and PRMR. Following Bowden et al. (2015)²⁰, we consider the

following four scenarios with varying violations of the MR assumptions (see the appendix for more details):

1. No pleiotropy
2. Balanced pleiotropy, InSIDE satisfied (α parameters take positive and negative values, but are independent of the first stage parameters γ)
3. Directional pleiotropy, InSIDE satisfied (α parameters take only positive values, but are independent of the first stage parameters γ)
4. Directional pleiotropy, InSIDE not satisfied (α parameters take only positive values and are correlated with the first stage parameters γ)

As in Bowden et al. (2015) we assume that all of the SNPs violate the exclusion restriction in scenarios 2-4, and so we do not consider methods that require at least 50% of the SNPs to be valid. In all scenarios, we use a sample size of 1,000; 10,000 simulation runs; 25 genetic variants with minor allele frequency 0.30; and a causal effect β of 0.00 and 0.05.

(Table 1 here)

The simulation results show that in a one-sample setting 2SLS and IVW produce very similar results, and that in case of no pleiotropy and balanced pleiotropy all methods give average estimates of β (almost) equal to the true value. With directional pleiotropy (both when InSIDE is satisfied and violated), the average estimates are biased away from the true value for 2SLS and IVW. MR-Egger performs slightly better, but still produces biased estimates. Not surprisingly, when the nature of pleiotropy is known, PRMR gives unbiased estimates. We acknowledge however that both deviations from homogenous pleiotropic effects, and non-zero first-stage coefficients

among the subsample for which the first stage should be zero, would produce biased estimates for α , and in turn the causal effect β .

Examples

To illustrate our approach, we exploit genetic data from the interim release of the UK Biobank³¹ to study (i) the effect of the number of cigarettes smoked per day (CPD) on BMI; (ii) the effect of prostate cancer on subjective health evaluations; and (iii) the effect of educational attainment on BMI. Following recommendations from the genotyping center, we restrict the analyses in the UK Biobank to 112,338 (52.53% female) conventionally unrelated individuals with “White British” ancestry.³² We use the first 15 principal components (PCs) of the genetic relatedness matrix as provided by UK Biobank to further control for population stratification.^{32,33} All individuals are born between 1934 and 1970. For individuals with missing information on a specific SNP, we impute the mean genotype from the sample. For all measures, we use reported values from the first interview round.

The effect of Cigarettes per Day on BMI

The relationship between smoking and BMI has received considerable attention in the literature^{34,35,36,37,38}, and from these studies Wehby et al. (2012)³⁷ used MR to assess the causal effect of CPD on BMI. In existing GWAS, the SNPs that are robustly associated with smoking measures can mostly be traced back to nicotine dependence.³⁹ This provides a context to apply PRMR: after all, the development of nicotine dependence requires initiating smoking in the first place. Hence, the group of never smokers provides a subsample among which the SNPs do not have an effect on CPD, and we can use this subsample to estimate the direct effect of the SNPs on the outcome measure BMI.

As instrument we use standardized values of rs12914385, the SNP with the strongest statistical association with CPD from the only locus found to be associated with CPD in the GWAS from the Tobacco and Genetic Consortium.⁴⁰ CPD is measured as number of cigarettes currently smoked per day, and BMI is measured in kg/m². OLS results suggest a positive association between CPD and BMI (0.05, $p = 1.90 \times 10^{-17}$). The first stage regression of CPD on the SNP shows a strong positive association ($p = 3.64 \times 10^{-16}$, $F = 66.47$). Among never-smokers, the SNP has a very modest positive association with BMI (0.01, $p = 0.44$). In contrast, among current-smokers, there is a strong negative association with BMI (-0.16 , $p = 8.49 \times 10^{-5}$). This provides evidence that the exclusion restriction is satisfied, and regular MR can be applied. The 2SLS results indicate a negative causal effect of CPD on BMI (-0.24 , $p = 3.50 \times 10^{-3}$), suggesting an effect of CPD on BMI in the opposite direction of the OLS estimate.

The effect of Prostate Cancer on Self-reported Health

Earlier studies have reported on the effect of (prostate) cancer on health outcomes^{41,42,43}, yet none of these studies used Mendelian Randomization. Al Olama et al. (2014)⁴⁴ find 12 autosomal SNPs to be related to prostate cancer at genome-wide significance level among individuals from European descent. Since prostate cancer naturally is only a risk factor among *males*, the first stage (that is, the effect of genetic variants on prostate cancer) among *females* is zero. As a result, the effect of a genetic variant associated with prostate cancer on a certain outcome among females may be used as a reasonable estimate for α .

UK Biobank contains a self-report about whether a person was told by a doctor to have prostate cancer (1) or not (0). Subjective health is measured on a four-point scale ranging from Excellent (4), via Good (3) and Fair (2), to Poor (1). We build an allele score for prostate cancer using the results from the GWAS on prostate cancer,⁴⁴

since the allele score has more power than using the individual SNPs.^{45,46,47} 10 out of the 12 SNPs are available in the genetic data of UK Biobank. The allele score has been standardized to have mean 0 and standard deviation 1.

An OLS regression of subjective health on prostate cancer among males reveals that prostate cancer is negatively associated with subjective health (-0.17 , $p = 9.17 \times 10^{-11}$), and a regression of prostate cancer on the allele score shows that the allele score is positively associated with prostate cancer ($p = 2.45 \times 10^{-3}$, $F = 9.21$). A standard 2SLS regression with the allele score as instrument for prostate cancer produces a large negative effect of prostate cancer on subjective health (-1.26), but $p = 0.54$.

Using PRMR, we find that the allele score is negatively associated with subjective health among females (-0.009 , $p = 1.81 \times 10^{-3}$), and this reduced form estimate is even larger in absolute terms than among males (-0.002 , $p = 0.53$). This suggests that the exclusion restriction is violated and that the MR results are biased. When we plug in the reduced form estimate among females in the plausibly exogenous method we find that the effect of prostate cancer on subjective health is estimated to be positive (4.53) among males, but this estimate is implausibly large and surrounded by a large 95% confidence interval (-0.88 to 9.94 , $p = 0.10$). Hence, we cannot reject a zero effect of prostate cancer on self-reported health.

The effect of Educational Attainment on BMI

The education-health gradient is well-documented and one of the most robust findings in social science.^{45,49} Several studies reported before on the causal effect of educational attainment on BMI, with mixed findings, and apart from one study investigating the reverse effect of BMI on educational attainment⁵⁰, none of them used Mendelian randomization.^{51,52,53,54,55,56}

Educational Attainment (EA) is constructed as years of education, similar as in the recent Educational Attainment GWAS (the UK Biobank data is used as out-of-sample replication sample in that study).⁵⁷ For educational attainment, we build an allele score from the 74 SNPs identified found to be associated with EA.⁵⁷ Of these 74 SNPs, 72 are available in UK Biobank.

An OLS regression of BMI on EA reveals that EA is negatively associated with subjective health (-0.11 , $p = 9.20 \times 10^{-308}$), and a regression of EA on the allele score shows that the allele score is positively associated with EA ($p = 4.10 \times 10^{-122}$, $F = 552.12$). A standard 2SLS regression with the allele score as instrument for EA provides a negative estimate (-0.39) for the causal effect of EA on BMI, with $p = 9.63 \times 10^{-21}$. However, it is likely that the exclusion restriction is violated, since the allele score for EA is, conditional on population stratification controls, gender, and birth year, positively associated with birth weight (0.01 , $p = 0.04$), a known confounder of the relationship between EA and BMI.⁵⁸

In contrast to the previous examples, here it is difficult to find a subsample for which the first stage is zero. Therefore we illustrate how PRMR can be used as a sensitivity analysis to determine how strong the violation of the exclusion restriction should be to render the causal effect β to be 0. Since one usually selects SNPs as IV for their association with the exposure rather than the outcome, it seems plausible that, in absolute value, the standardized first stage effect of the IV on the exposure, γ , is larger than the standardized direct effect of the IV on the outcome, α . Therefore, we define the proportion $0 \leq \lambda \leq 1$, we set μ_α equal to $\lambda \hat{\gamma}$, and the variance Ω_α equal to the squared standard error of $\hat{\gamma}$. Subsequently, we apply equation (3) for different values of λ to find for which λ the corresponding β equals 0. When the first stage estimate has the opposite

sign of the direct effect estimate, one needs to change the sign of $\hat{\gamma}$ in this sensitivity analysis.

In our case, the first stage effect (allele score on EA) has plausibly the opposite sign of the direct effect (allele score on BMI) and hence we set μ_{α} equal to $-\lambda\hat{\gamma}$. The resulting estimates for β are plotted in Figure 2 as a function of λ . The estimate in case of $\lambda = 0$ corresponds to the MR point estimate, but the 95% confidence interval is a little wider because $\Omega_{\alpha} > 0$. Moving along the x -axis, the causal effect of EA on BMI is estimated to be 0 when $\lambda = 0.41$, and the 95% confidence interval includes 0 already when $\lambda = 0.29$. Hence, a relatively mild violation of the exclusion restriction – 29% of the first-stage effect – implies that we cannot reject a zero effect of EA on BMI, producing at best weak evidence that EA causally reduces BMI.

(Figure 2 here)

Conclusion

The fact that the pleiotropic effects of genes are poorly understood makes it difficult to use genetic variants as credible instrumental variables in Mendelian Randomization. In this paper we argued that if (i) one can identify a subsample for which genetic variants do not have an effect on the exposure, and (ii) pleiotropic effects are homogenous, PRMR provides a way to deal with violations of the exclusion restriction due to pleiotropy, and opens the door for several applications in epidemiology and social science research.

A simulation study showed that PRMR clearly outperforms existing methods in case of violations of the exclusion restriction. We illustrated our PRMR approach by estimating the causal effect of cigarettes smoked per day on BMI, and the effect of

prostate cancer on subjective health evaluations. In those two cases it is possible to identify subsamples where the effect of SNPs on the exposure are zero (never smokers and females, respectively), and this allows estimating, and if necessary correcting for, the pleiotropic effect.

We acknowledge that the two requirements for PRMR are not always satisfied. Yet, even if one is not willing to make the assumption of homogenous pleiotropic effects, one can still use the subsample without an effect of the genetic variants on the exposure as a useful test of the exclusion restriction. Moreover, if one cannot identify a subsample without a first stage, as we illustrated for the estimation of the effect of education on BMI, PRMR still allows for an informative sensitivity analysis that could routinely be applied in all MR analyses.

Funding

This work was supported by the Netherlands Organization for Scientific Research [016-145-082 to H.v.K. and 016-165-004 to C.A.R.], and the National Institute on Aging [R01AG037398 to H.v.K.].

Acknowledgments

This research has been conducted using the UK Biobank Resource. We thank Daniel Benjamin, Tom DiPrete, Patrick Turley, Peter Visscher, and Dinand Webbink for helpful comments.

Conflict of interest: None declared.

References

1. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Met Med Res* 2007;**16**:309–330.
2. Davey Smith G, Ebrahim S. ‘Mendelian randomization’: Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;**32**:1–22.
3. Davey Smith G, Ebrahim S. Mendelian randomization: Prospects, potentials, and limitations. *Int J Epidemiol* 2004;**33**:30–42.
4. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;**23**:89–98.
5. Ebrahim S, Davey Smith G. Mendelian randomization: Can genetic epidemiology help redress the failures of observational epidemiology? *Hum Genet* 2007;**123**:15–33.
6. Lawlor DA, Harbord RM, Sterne JAC, Timpson NJ, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;**27**:1133–1163.
7. Burgess S, Timpson NJ, Ebrahim S, Davey Smith G. Mendelian randomization: where are we now and where are we going? *Int J Epidemiol*, 2015;**44**:379–388.
8. Boef AGC, Dekkers OM, Le Cessie S. Mendelian randomization studies: A review of the approaches used and the quality of reporting. *Int J Epidemiol* 2015;**44**:496–511.
9. Swerdlow, DI, Kuchenbaecker, KB, Shah, S, Sofat, R, Holmes, MV, White J, Mindell JS, Kivimaki M, Brunner EJ, Whittaker JC, Casas JP, Hingorani AD. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies *Int. J. Epidemiol.* 2016; doi:10.1093/ije/dyw088.
10. Nitsch D, Molokhia M, Smeeth L, DeStavola BL, Whittaker JC, Leon DA. Limits to causal inference based on Mendelian randomization: A comparison with randomized controlled trials. *Am J Epidemiol* 2006;**163**:397–403.
11. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian randomization and causal inference in observational epidemiology. *PLoS Med* 2008;**5**:e177.
12. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. Methodological challenges in Mendelian randomization. *Epidemiology* 2014;**25**:427–435.
13. Von Hinke S, Davey Smith G, Lawlor DA, Propper C, Windmeijer F. Genetic markers as instrumental variables. *J Health Econ* 2015. Forthcoming. <http://www.sciencedirect.com/science/article/pii/S0167629615001198>.
14. Pickrell J. Fulfilling the promise of Mendelian randomization. *bioRxiv* 2015. DOI: <http://dx.doi.org/10.1101/018150>.
15. Davey Smith G. Mendelian randomization: a premature burial? *bioRxiv* 2015. DOI: <http://dx.doi.org/10.1101/021386>.
16. Glymour MM, Tchetgen EJT, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* 2012;**175**:332–339.
17. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genetic* 2013;**14**:483–495.
18. Imbens G, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica* 1994;**62**:467–75.
19. Kang H, Zhang A, Cai TT, Small DS. Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. *J Am Stat Assoc* 2016. Forthcoming. DOI: 10.1080/01621459.2014.994705.
20. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;**44**:512–525.
21. Bowden J, Davey Smith G, Haycock, PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genetic Epidemiology* 2016;**40**(4):304–314.

22. Conley TG, Hansen CB, Rossi PE. Plausibly exogenous. *Rev Econ Stat* 2012;**94**:260–272.
23. Fletcher JM. The promise and pitfalls of combining genetic and economic research. *Health Econ* 2011;**20**:889–892.
24. Altonji JG, Elder TE, Taber CR. An evaluation of instrumental variable strategies for estimating the effects of Catholic schooling. *J Hum Res* 2005;**40**:791–821.
25. Imbens GW. Instrumental variables: an econometrician's perspective. *Stat Sci* 2014;**29**:323–358.
26. Angrist J, Lavy V, Schlosser A. Multiple experiments for the causal link between the quantity and quality of children. *J Lab Econ* 2010;**28**:773–824.
27. Davey Smith, GD, Lawlor DA, Harbord R, Timpson NJ, Day I, Ebrahim S. Clustered environments and randomized genes: A fundamental distinction between conventional and genetic epidemiology. *PLoS Med* 2007;**4**:e352.
28. Bound J, Jaeger DA, Baker RG. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J Am Stat Assoc* 1995;**90**:443–450.
29. Clarke D. PLAUSEXOG: Stata module to implement Conley et al's plausibly exogenous bounds. *Statistical Software Components S457832*. Boston College Department of Economics, 2014. DOI: <https://ideas.repec.org/c/boc/bocode/s457832.html>.
30. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. *Gen Epidemiol* 2013;**37**(7):658–665.
31. Sudlow C, Gallacher J, Allen A, 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**:e1001779.
32. Marchini J, O'Connell J, Delaneau O., et al. Genotype imputation and genetic association studies of UK Biobank: Interim data release. Technical report, 2015. http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/imputation_documentation_May2015.pdf.
33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;**38**:904–909.
34. Gruber J, Frakes M. Does falling smoking lead to rising obesity? *J Health Econ* 2006;**25**:183–197.
35. Chiolerio A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. *Am J Clin Nutr* 2008;**87**:801–809.
36. Courtemanche C. Rising cigarette prices and rising obesity: Coincidence or unintended consequence? *J Health Econ* 2009;**28**:781–798.
37. Wehby GL, Murray JC, Wilcox A, Lie RT. Smoking and body weight: Evidence using genetic instruments. *Econ Hum Biol* 2012;**10**:113–126.
38. Cawley J, Dragone D, Von Hinke Kessler Scholder S. The demand for cigarettes as derived from the demand for weight loss: A theoretical and empirical investigation. *Health Econ* 2016;**25**:8–23.
39. Bierut LJ. Genetic vulnerability and susceptibility to substance dependence. *Neuron* 2011;**69**:618–627.
40. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010;**42**:441–447.
41. Hewitt M, Rowland JH, Yancik R. Cancer survivors in the United States: Age, health, and disability. *J Gerontol A-Biol* 2003;**58**:M82–M91.
42. Demark-Wahnefried W, Aziz NM, Rowland JH, Pinto BM. Riding the crest of the teachable moment: Promoting long-term health after the diagnosis of cancer. *J Clin Oncol* 2005;**23**:5814–5830.
43. Yabroff KR, Lawrence WF, Clauser S, Davis WW, Brown ML. Burden of illness in cancer survivors: Findings from a population-based national sample. *J Natl Cancer I* 2004;**96**:1322–1330.
44. Al Olama AA, Kote-Jarai Z, Berndt SI, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet*, 2014;**46**:1103–1109.

45. Burgess S, Thompson SG, CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol*, 2011;**40**(3):755-764.
46. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization *Int. J. Epidemiol*, 2013;**42**(4):1134-1144.
47. Davies NM, Von Hinke S, Farbmacher H, Burgess S, Windmeijer F, Davey Smith G. The many weak instruments problem and Mendelian randomization. *Stat Med* 2015;**34**:454–468.
48. Ross CE, Wu C The links between education and health. *Am Soc Rev* 1995;**60**:719–745.
49. Mackenbach JP, Stirbu I, Roskam AJR, et al. Socioeconomic inequalities in health in 22 European countries. *New Engl J Med* 2008;**358**:2468–2481.
50. Tyrrell J, Jones, SE, Beaumont R, Astley CM, Lovell R, Yaghootkar H, Tuke M, Ruth KS, Freathy RM, Hirschhorn JN, Wood AR, Height, body mass index, and socioeconomic status: Mendelian Randomisation study in UK Biobank. *BMJ* 2016;**352**, p.i582.
51. Clark D, Royer H. The effect of education on adult mortality and health: Evidence from Britain. *Am Econ Rev* 2013;**103**:2087–2120.
52. Arendt JN. Does education cause better health? A panel data analysis using school reforms for identification. *Econ Educ Rev* 2005;**24**:149–160.
53. Spasojevic J. Effects of education on adult health in Sweden: Results from a natural experiment. In: *Current Issues in Health Economics (Contributions to Economic Analysis, Volume 290)*. Emerald Group Publishing Limited, 2010.
54. Jürges H, Reinhold S, Salm M. Does schooling affect health behavior? Evidence from the educational expansion in western Germany. *Econ Educ Rev* 2011;**30**:862–872.
55. Kemptner D, Jürges H, Reinhold S. Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany. *J Health Econ* 2011;**30**:340–354.
56. Braakmann N. The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England. *J Health Econ* 2011;**30**:753–763.
57. Okbay A, Beauchamp JP, Fontana M, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;**533**:539-542.
58. Devereux PJ, Black SE, Salvanes KG. From the cradle to the labor market? The effect of birth weight. *The Quarterly Journal of Economics* 2007;**122**(1),409-439.

Tables

Table 1. Performance of Two-stage-least-squares (2SLS), Inverse-variance weighting (IVW), MR-Egger regression and PRMR in simulation study for one-sample Mendelian randomization with a null ($\beta = 0.00$) and a positive ($\beta = 0.05$) causal effect.

Scenario	Mean estimate			
	2SLS	IVW	MR-Egger	PRMR
$\beta = 0.00$				
1: No pleiotropy, InSIDE satisfied	0.00	0.00	0.01	0.00
2: Balanced pleiotropy, InSIDE satisfied	0.00	0.00	0.01	0.00
3: Directional pleiotropy, InSIDE satisfied	0.04	0.04	0.01	0.00
4: Directional pleiotropy, InSIDE violated	0.13	0.13	0.04	0.00
$\beta = 0.05$				
1: No pleiotropy, InSIDE satisfied	0.05	0.05	0.06	0.05
2: Balanced pleiotropy, InSIDE satisfied	0.05	0.05	0.06	0.05
3: Directional pleiotropy, InSIDE satisfied	0.09	0.09	0.06	0.05
4: Directional pleiotropy, InSIDE violated	0.18	0.18	0.09	0.05

Figure 1. Illustrative diagram showing the standard MR model with its assumptions. The genetic effect of instrumental variable G_j on exposure X is γ_j ; the genetic effect on unobserved confounder U is ϕ_j ; the direct genetic effect on the outcome Y is α_j , and the causal effect of the exposure X on the outcome Y is β .

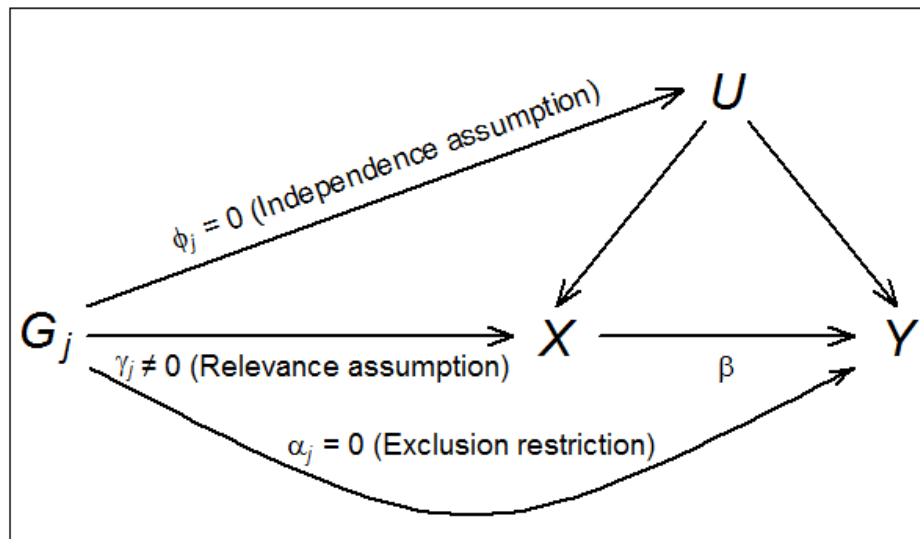


Figure 2. The causal effect of Educational Attainment on BMI, for varying values of λ (the percentage of the standardized effect of the IV on EA which is considered to be the direct effect of the IV on standardized BMI). The grey area represents the 95% Confidence Interval.

