# Mendelian randomization: genetic anchors for causal inference in epidemiological studies

George Davey Smith

Gibran Hemani

MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, School of Social and Community Medicine, Bristol, UK

Observational epidemiological studies have contributed to identifying important modifiable causes of disease, but also have a poor record of separating causal from non-causal associations. Many examples exist of apparently robust epidemiological associations between behavioural, pharmacological or physiological measures and disease risk which, when subjected to randomised controlled trials, do not deliver the anticipated health benefits. These include many nutritional factors (e.g. several vitamins) pharmacological agents (e.g. hormone replacement therapy) and circulating biomarkers (e.g. HDL cholesterol). In these situations it appears likely that confounding and various biases generate the associations, and that these cannot be adequately accounted for by apparently robust study design and statistical adjustment.[[1]](#endnote-1) [[2]](#endnote-2) The recognition of these problematic aspects of epidemiological investigation have led to the application of a series of methods aimed at improving causal inference in observational epidemiological studies[[3]](#endnote-3) [[4]](#endnote-4), including the use of genetic variants as indicators of exposure that are not subject to the influences that vitiate conventional study designs, an approach known as Mendelian randomization (MR)[[5]](#endnote-5) [[6]](#endnote-6). We will not repeat the many detailed reviews that now exist of MR5 [[7]](#endnote-7) [[8]](#endnote-8) [[9]](#endnote-9) [[10]](#endnote-10) [[11]](#endnote-11) [[12]](#endnote-12) [[13]](#endnote-13) , nor summarise the hundreds of empirical studies applying the technique to a wide range of exposures and disease outcomes, rather, after a brief summary of the foundational principles, we will outline recent developments and potential future directions of the field.

## Mendelian randomization: basic principles

In practice, MR can be used to achieve two slightly different objectives. Its first role is concerned with making unbiased estimates in ordinary least squares regression (OLS). Supposing that the independent variable has an effect on some outcome, but it is also correlated with other (unobserved) variables that too have an effect on the outcome, the estimate of the effect will be biased. MR offers a way to overcome this problem. The second function is perhaps of more fundamental importance in modern biology: making inference about the causal relationship between correlated variables.

Suppose that trait A and trait B are correlated, it follows that if this correlation arises because A is causing B, then any variable that causes trait A should also cause trait B. So the key to inferring a causal relationship between A and B is to identify an ‘instrument’ which is known to cause A. Biologists are in a privileged position in this regard because it is no exaggeration to say that virtually all traits of interest are at least partially influenced by genetic effects, and genetic effects can serve as excellent instruments for a number of reasons. First, and most crucially, it is universally acknowledged that in a genetic association the direction of causation is from the random genetic polymorphism inherited at conception leading to the subsequent trait of interest, and not vice versa. Second, using environmental exposures as instruments for inferring causality (common in conventional observational studies in epidemiology), but this can be problematic because environmental exposures are often associated with a wide range of behavioural, social, and physiological factors that confound associations[[14]](#endnote-14). Functional genetic variants on the other hand can be precise indicators for a particular trait. Third, genetic variants and their effects are subject to relatively little measurement error or bias. Finally, in the era of genome wide association studies (GWAS) and high throughput genomic technologies they are routinely obtainable for huge sample sizes for a vast number of traits.

### Analogy between Mendelian randomization and randomised controlled trials

An intuitive way to understand how MR can be used to infer causality is to compare it to a randomised controlled trial (RCT). In RCTs the study participants are randomly allocated one or another treatment, thus any potential confounders between treatment and outcome are severed, and if the treatment truly has an effect then the causal inference is unambiguous. MR simulates this scenario for us. Supposing a particular allele has an increasing effect on trait A, and trait A causes trait B. From Mendel’s first law of inheritance, we know that the alleles passed from parents to offspring are randomly selected, and in a way subjects who inherit the allele are in effect being randomly allocated a high dosage of trait A, while those who don’t inherit the allele are being randomly allocated a low dosage of trait A. Thus, by stratifying by genotype we are severing any potential associations between confounders that may underlie correlations between trait A and trait B, and if the random allocation to genotype class has an association with trait B then trait A can be deemed causal.

Empirical evidence that there is lack of confounding of genetic variants with factors that confound exposures in conventional observational epidemiological studies comes from several sources. For example, consider the virtually identical allele frequencies in the British 1958 birth cohort and British blood donors[[15]](#endnote-15). Blood donors are clearly a very selected sample of the population, whereas the 1958 birth cohort comprised all births born in one week in Britain with minimal selection bias. Blood donors and the general population sample would differ considerably with respect to the behavioural, socio-economic and physiological risk factors that are the confounding factors in observational epidemiological studies. However, they hardly differ in terms of allele frequencies. Similarly, we have demonstrated the lack of association between a range of SNPs of known phenotypic effects and nearly 100 socio-cultural, behavioural and biological risk factors for disease16.

### Mendelian randomization and instrumental variables analysis

Conventional instrumental variable (IV) analysis requires that the instruments are valid, and in order to be valid they must meet three conditions. Assuming trait A causes B, an instrument must be

1. associated with the causal trait;
2. associated with the consequential trait only through the causal trait; and
3. unrelated to unobserved confounders that influence traits A and B.

In MR condition (1) is inherently straightforward to test, but empirically proving (2) and (3) is often problematic. For example, if the variant is functionally pleiotropic, or if it is in linkage disequilibrium (LD) with another variable that causes the consequential trait directly, this can lead to erroneous causal estimation. We discuss ways in which to overcome this issue, but assuming all conditions are met, then the unbiased estimate of the effect of the causal trait on the consequential trait can be made using two-stage least squares (2SLS) regression. In stage 1 the fitted values are obtained from , where is the instrument for A. In stage 2 the regression coefficient is obtained from . thus represents an asymptotically unbiased estimate of the effect of A on B that occurs due to alone, so in order to translate this into the full effect of A on B we calculate , where the potentially biased is obtained from . There exist several implementations of 2SLS in open source software.

### Examples and limitations of Mendelian randomization studies

In table 1 we provide examples of MR studies that cover the range of applications from biomarkers (CRP, LDL cholesterol), physiological measures (BMI, blood pressure), behaviours (alcohol, smoking), and maternal influences (alcohol).

## Recent extensions to basic Mendelian Randomization

### Use of multiple variants to increase power and test assumptions

Ideally MR is performed using a single variant whose biological effect is understood on the trait for which it is an instrument. But this comes with two potential limitations that can be mitigated by increasing the number variants used as instruments.

First, the genetic effect may not be particularly large, resulting in a weak instrument. By increasing the number of variants the proportion of variance explained by the instrument increases, thus improving precision in 2SLS[[16]](#endnote-33).

Second, the variant could be pleiotropic or in LD with a variant that affects the outcome, thus violating conditions (2) or (3) for being a valid instrument. One can interrogate this potential caveat by using multiple instruments. For example it would be increasingly improbably that two, three, or more independent instruments all result in the same conclusion due to the colocalisation of genetic factors for both traits. For a convincing example demonstrating the causal influence of low-density lipoprotein cholesterol (LDL-C) on coronary heart disease (CHD) see Figure 1 where 9 polymorphisms from 6 genes independently give rise to nearly identical results in MR[[17]](#endnote-34).

Typically genetic variants are only used as instruments if they are significantly detected and replicated in GWAS. But predictive power is often improved when SNPs that do not reach significance thresholds are also included, the rationale being that they may be false negatives due to small effect size[[18]](#endnote-35). Using this approach has potential to dramatically improve the power of MR, but caution should be made due to the increased chance of including pleiotropic effects.[[19]](#endnote-36)

### Two sample Mendelian randomization

It is often the case that an observational association between two variables exists, but high measurement costs or lack of appropriate biospecimens leads to prohibitively small datasets with both measurements and instruments. Methods have been developed to perform MR when the exposure variable and the outcome variable are measured in two independent datasets[[20]](#endnote-37).

Another scenario in which two sample MR can be used is if the data in which MR is being performed is the same as is being used to identify instruments. GWAS is known to lead to overestimation of genetic effect sizes due to the phenomenon of winner’s curse, and this can lead to bias in MR. Subsetting the data into two (or more) samples for estimation and testing can mitigate this problem. This method has been applied to demonstrate that childhood adiposity causes a reduction in physical activity[[21]](#endnote-38).

### Bidirectional Mendelian randomization

A major limitation of the MR is that it is difficult to distinguish between an exposure causing an outcome and an outcome causing a trait with only a single genetic variant because the genetic variant could be biologically causing either variable. For example, atheroma and BMI influence C-reactive protein (CRP) levels, and apparent misleading causal effects can be generated if the genetic variant is mistaken as being a variant for the exposure (e.g. an atheroma or obesity variant taken to be one for CRP)[[22]](#endnote-39).

Aside from focusing on instruments for which there exists some degree of biological understanding, another method is to use bi-directional MR. Here, instruments are required for both variables and MR is performed in both directions. If trait A causes trait B, then the instrument will be significantly associated with both A and B. However, a second instrument specific to trait B, will be associated with trait B, and not with trait A. This method was used in a recent paper that demonstrated that obesity is a causal factor in 25(OH)D levels, concluding that population level interventions to reduce BMI are expected to reduce instances of vitamin D deficiency[[23]](#endnote-40).

### Factorial Mendelian randomization

The manner by which causes of disease act together to increase disease risk can have important public health implications, as above additive effects lead to the clustering of risk factors generating a greater burden of disease in the population. For example it has been suggested that the risk of liver disease associated with the combination of obesity and heavy alcohol consumption is greater than multiplicative [[24]](#endnote-41), and adverse trends for these two risk factors would be predicted to generate considerable increases in liver disease. It is difficult to estimate such effects, however, as confounding can be magnified when examining two already confounded risk factors. By analogy with factorial randomised controlled trials, where separate randomisation of different treatments allows characterization of interactions between them[[25]](#endnote-42), Mendelian randomization studies can investigate disease risk associated with combinations of genetic variants and through this obtain unconfounded estimates of the effect of co-occurrence of the two risk factors that the genetic variants are taken to be indicators of.

### Multiphenotype Mendelian randomization

In some situations genetic variants tend to be associated with more than one intermediate phenotype, and estimating the causal effect of one particular intermediate phenotype is rendered problematic. A case in point are categories of blood lipids that associated with coronary heart disease, in particular HDL cholesterol and triglycerides, which are highly inversely correlated and cannot be reliably separated using purely statistical approaches in observational studies[[26]](#endnote-43). The large number of genetic variants related to HDL-C and triglycerides generally associate with both, but to varying degrees[[27]](#endnote-44) and thus genetic instruments created with multiple variants do not purely relate to one of the two lipids. In this context regression methods can be applied to attempt to separate the effects; two independent studies utilizing this approach suggested that the causal influence of triglycerides was robust, whereas the apparent protective effect of HDL-C was not[[28]](#endnote-45) [[29]](#endnote-46). Unlike in the situation with factorial Mendelian randomization there is a dependence on attempting to statistically separate effects, which reintroduces problems in conventional observational studies. Both the appropriateness of different statistical approaches and whether reliable answers can be obtained in the multiphenotype context remain areas of active investigation.

### Hypothesis free Mendelian randomization

The majority of MR studies have been focused on testing hypotheses that arose from established associations between traits from observational studies. But is this only the tip of the iceberg? In a study of unprecedented size, mining the records of 110 million patients uncovered 2909 associations between Mendelian diseases and complex traits, the majority of which were previously unreported[[30]](#endnote-47). As high-throughput ‘omics technologies continue to reduce in price, datasets with comprehensive phenotyping are destined to grow, and increasingly comprehensive catalogues of genetic variants that can be used as instruments will emerge. Indeed, there already exist reports for methylome-wide searches for causal mediators for various traits, for example increased folate in red blood cells in cord blood leads to a (mostly) positive influence on methylation at seven sites across the genome[[31]](#endnote-48).

## Conclusions

Deconstructing observational correlations into causal relationships is an elusive problem at the heart of biological understanding, pharmaceutical development, and medical practice. MR is a statistically robust method for this endeavour, whose scope for application widens as the cost of data generation continues to reduce.

## Acknowledgements

Thanks to Professor Sheila Bird who (in 2002) suggested the term “factorial Mendelian randomization” and to Dr Tom Palmer who suggested the term “multiphenotype Mendelian randomization”

**Table 1** - Limitations of Mendelian randomization

|  |  |  |
| --- | --- | --- |
| Limitation | Role in Mendelian randomization (MR) studies | Approaches to evaluating or avoiding the limitation |
| Low statistical power | MR studies are often of low power and effect estimates are imprecise because of this | Increase sample size and/or combine genetic variants so they explain more of the variance of the intermediate phenotype |
| Reverse causation | A genetic variant may be causing the disease outcome which in turn causes the biomarker, or the causal direction could be in the opposite direction. 2SLS will not distinguish between these cases | Bi-directional MR can be used to distinguish between the two causal models. |
| Population stratification | Spurious associations used as instruments can lead to faulty causal inference | Restrict analyses to ethnically homogeneous groups, families and/or apply correction methods using ancestrally informative markers or principal components from genome wide data |
| Reintroduced confounding though pleiotropy | A genetic variant may directly influence more than one post-transcriptional process. Known to be the case for some genetic variants | When possible utilise cis variants with respect to the intermediate phenotype under study, as these may be less likely to have pleiotropic effects. Apply multiple instrument approaches with more than one independent genetic variant as unlikely pleiotropy will generate the same associations for different instruments |
| Linkage disequilibrium (LD) induced confounding | LD is crucial in genetic association studies as it allows marker SNPs to proxy for un-genotyped causal SNPs. However this can reintroduce confounding if LD leads to the association of SNPs related to more than one post-transcriptional process. This case will be similar to the pleiotropy situation | Studies can be carried out in populations with different LD structures. Approaches to avoiding distortion by pleiotropy will also counter problems due to LD |
| Canalization / developmental compensation | During development compensatory processes may be generated that counter the phenotypic perturbation consequent on the genetic variant utilized as an instrument | No general approach developed, although context–specific biological knowledge can be appealed to. Period of lifecourse when influence of genetic variation on IPs emerge can indicate whether canalization could, in principle, be an issue |
| Lack of genetic variants to proxy for modifiable exposure of interest | No reliable genetic variant associations for many intermediate phenotypes of interest, although an increasing number of these now identified | Continued genome wide and sequencing based studies |
| Complexity of associations | Without adequate biological knowledge misleading inferences regarding intermediate phenotypes and disease may be drawn | Increased biological understanding of genotype – phenotype links |

Figure 1: Effect of lower LDL-C on risk of CHD (taken from Ference 2012)

## References

1. Davey Smith G, Ebrahim S. Epidemiology - is it time to call it a day? Int J Epidemiol 2001;30:1-11. [↑](#endnote-ref-1)
2. Fewell Z, Davey Smith G, Sterne JAC. The impact of residual and unmeasured confounding in epidemiological studies; a simulation study. Am J Epidemiol 2007;166:646-655. [↑](#endnote-ref-2)
3. Lipsitch M, Tchetgen ET, Cohen T. [Negative Control Exposures in Epidemiologic Studies](http://journals.lww.com/epidem/Citation/2012/03000/Negative_Control_Exposures_in_Epidemiologic.28.aspx" \o "Negative Control Exposures in Epidemiologic Studies). Epidemiology. 2012;23:351-352 [↑](#endnote-ref-3)
4. Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological findings yield robust results? Basic and Clinical Pharmacology and Toxicology 2008;102:245-256. [↑](#endnote-ref-4)
5. Davey Smith G, Ebrahim S. ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiology 2003;32:1-22. [↑](#endnote-ref-5)
6. Podolsky SH, Davey Smith G. Park’s story and Winters’ tale: alternate allocation clinical trials in turn of the Century America. JLL Bulletin: Commentaries on the history of treatment evaluation. Journal of Royal Society of Medicine 2011; 104:262-8 (updated version on http://www.jameslindlibrary.org/illustrating/articles/parks-story-and-winters-tale-alternate-allocation-clinical-tr). [↑](#endnote-ref-6)
7. Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. Int J Epidemiol 2004;33:30-42. [↑](#endnote-ref-7)
8. Sheehan NA, Didelez V, Burton PR, Tobin MD. Mendelian Randomisation and Causal Inference in Observational Epidemiology. PLoS Med 2008;5: e177 [↑](#endnote-ref-8)
9. Lawlor DA, Harbord RM, Sterne JAC, Timpson NJ, Davey Smith G. Mendelian randomization : Using genes as instruments for making causal inferences in epidemiology. Statistics in Medicine 2008; 27:1133-1163 [↑](#endnote-ref-9)
10. Bochud M, Rousson V. Usefulness of Mendelian Randomization in Observational Epidemiology. International Journal of Environmental Research and Public Health. 2010; 7(3):711-728. [↑](#endnote-ref-10)
11. VanderWeele TJ.; Tchetgen Tchetgen, EJ.; Cornelis, M. [Methodological Challenges in Mendelian Randomization](http://journals.lww.com/epidem/Abstract/2014/05000/Methodological_Challenges_in_Mendelian.14.aspx" \o "Methodological Challenges in Mendelian Randomization). Epidemiology. 2014;25:427-435 [↑](#endnote-ref-11)
12. Richmond RC, Davey Smith G, Ness AR, den Hoed M, McMahon G, Timpson NJ. Assessing causality in the association between child adiposity and physical activity levels: A Mendelian randomization analysis. PLoS Medicine 2014; doi:10.1371/journal.pmed.1001618 [↑](#endnote-ref-12)
13. Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. Genes & Nutrition 2011;6:27–43. [↑](#endnote-ref-13)
14. Davey Smith G, Lawlor DA, Harbord R, Timpson NJ, Day I, Ebrahim S. Clustered Environments and Randomized Genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Medicine 2007;4:1985-1992. PMCID: PMC2121108 [↑](#endnote-ref-14)
15. Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 447:661–678. [↑](#endnote-ref-15)
16. Marie-Jo A Brion, Konstantin Shakhbazov, Peter M Visscher. Calculating statistical power in Mendelian randomization studies. Int. J. Epidemiol. (2013) 42 (5): 1497-1501 doi:10.1093/ije/dyt179 [↑](#endnote-ref-33)
17. Ference BA, Yoo W, Alesh I, Mahajan N, Mirowska KK, Mewada A, Kahn J, Afonso L, Williams KA Sr, Flack JM. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. J Am Coll Cardiol. 2012 Dec 25;60(25):2631-9 [↑](#endnote-ref-34)
18. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009 Aug 6;460(7256):748-52. [↑](#endnote-ref-35)
19. Evans DM, Brion MJ, Paternoster L, Kemp JP, McMahon G, Munafò M, Whitfield JB, Medland SE, Montgomery GW; GIANT Consortium; CRP Consortium; TAG Consortium, Timpson NJ, St Pourcain B, Lawlor DA, Martin NG, Dehghan A, Hirschhorn J, Davey Smith G. Mining the human phenome using allelic scores that index biological intermediates. PLoS Genet. 2013 Oct;9(10):e1003919. [↑](#endnote-ref-36)
20. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiol. 2013 Oct 1;178(7):1177-84 [↑](#endnote-ref-37)
21. Richmond RC, Davey Smith G, Ness AR, den Hoed M, McMahon G, Timpson NJ. Assessing causality in the association between child adiposity and physical activity levels: a Mendelian randomization analysis. PLoS Med. 2014 Mar 18;11(3):e1001618 [↑](#endnote-ref-38)
22. Bowden J, Vansteelandt S. [Mendelian randomization analysis of case-control data using structural mean models](http://onlinelibrary.wiley.com/doi/10.1002/sim.4138/abstract). STATISTICS IN MEDICINE 2011;30:678-694 [↑](#endnote-ref-39)
23. Vimaleswaran KS, Berry DJ, Lu C, Tikkanen E, Pilz S, Hiraki LT, Cooper JD, Dastani Z, Li R, Houston DK, Wood AR, Michaëlsson K, Vandenput L, Zgaga L, Yerges-Armstrong LM, McCarthy MI, Dupuis J, Kaakinen M, Kleber ME, Jameson K, Arden N, Raitakari O, Viikari J, Lohman KK, Ferrucci L, Melhus H, Ingelsson E, Byberg L, Lind L, Lorentzon M, Salomaa V, Campbell H, Dunlop M, Mitchell BD, Herzig KH, Pouta A, Hartikainen AL; Genetic Investigation of Anthropometric Traits-GIANT Consortium, Streeten EA, Theodoratou E, Jula A, Wareham NJ, Ohlsson C, Frayling TM, Kritchevsky SB, Spector TD, Richards JB, Lehtimäki T, Ouwehand WH, Kraft P, Cooper C, März W, Power C, Loos RJ, Wang TJ, Järvelin MR, Whittaker JC, Hingorani AD, Hyppönen E. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. PLoS Med. 2013;10(2):e1001383 [↑](#endnote-ref-40)
24. Hart C, Morrison DS, Batty GD, Mitchell RJ, Davey Smith G. Effect of body mass index and alcohol consumption on liver disease: analysis of data from two prospective cohort studies. BMJ 2010;34:c1240. PMCID: PMC2837144 [↑](#endnote-ref-41)
25. Montgomery 2003 [↑](#endnote-ref-42)
26. Phillips A, Davey Smith G. How independent are "independent" effects? Relative risk estimation when correlated exposures are measured imprecisely. J Clin Epidemiol 1991;44:1223-31. [↑](#endnote-ref-43)
27. Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat. Genet.2013;45:1274–1283 [↑](#endnote-ref-44)
28. Do R, Willer CJ, Schmidy EM et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. Nature Genetics 2013;45, 1345-1352 doi:10.1038/ng.2795 [↑](#endnote-ref-45)
29. Holmes MV, Asselbergs FW, Palmer TM, et al. Mendelian randomization of blood lipids for coronary heart disease. Eur Heart J 2014, doi:10.1093/eurheartj/eht571 [↑](#endnote-ref-46)
30. Blair DR, Lyttle CS, Mortensen JM, Bearden CF, Jensen AB, Khiabanian H, Melamed R, Rabadan R, Bernstam EV, Brunak S, Jensen LJ, Nicolae D, Shah NH, Grossman RL, Cox NJ, White KP, Rzhetsky A. A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. Cell. 2013 Sep 26;155(1):70-80 [↑](#endnote-ref-47)
31. Binder AM, Michels KB. The causal effect of red blood cell folate on genome-wide methylation in cord blood: a Mendelian randomization approach. BMC Bioinformatics. 2013 Dec 4;14:353 [↑](#endnote-ref-48)