Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians

Neil M Davies,[1,2] Michael V Holmes,[1,3,4,5] George Davey Smith[1,2,6]

[1]Medical Research Council Integrative Epidemiology Unit, University of Bristol, BS8 2BN, UK

[2]Population Health Sciences, Bristol Medical School, University of Bristol, Barley House, Oakfield Grove, Bristol, BS8 2BN, UK

[3]Medical Research Council Population Health Research Unit, University of Oxford, UK

[4]Clinical Trial Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Old Road Campus, Roosevelt Drive, Oxford OX3 7LF, UK.

[5]National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospital, Oxford, UK

[6]National Institute for Health Research Bristol Biomedical Research Centre, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK

Correspondence to: G Davey Smith pa-ieudirector@bristol.ac.uk

Additional material is published online only. To view please visit the journal online.

Mendelian randomisation uses genetic variation as a natural experiment to investigate the causal relations between potentially modifiable risk factors and health outcomes in observational data. As with all epidemiological approaches, findings from Mendelian randomisation studies depend on specific assumptions. We provide explanations of the information typically reported in Mendelian randomisation studies that can be used to assess the plausibility of these assumptions and guidance on how to interpret findings from Mendelian randomisation studies in the context of other sources of evidence

Understanding whether a biomarker or behaviour causes ill health is central to evidence based medicine, drug development, and better informed clinical decision making. Ideally, evidence of causal effects comes from well conducted randomised trials. Clinicians are well versed in the strengths and limitations of such trials and have an increasingly sophisticated understanding of traditional analyses of observational studies. But they may be less aware of the strengths and limitations of a more recently developed approach to analysing observational data known as Mendelian randomisation. Although numerous guides exist for conducting[1-3] and reporting Mendelian randomisation studies and related methods,[4 5] here we focus on helping clinicians and practitioners read and interpret them. Our goal is to provide explanations of

core concepts and recent developments in Mendelian randomisation methods.

A method to overcome confounding

Mendelian randomisation is an analytical method that uses genetic variants as instrumental variables for modifiable risk factors that affect population health.[1 6-8] It is increasingly being used because it can overcome a major limitation of evidence from observational studies: unmeasured confounding.[6 9] Suppose, for example, we wanted to investigate the effects of alcohol consumption on blood pressure with a view to understanding the overall relationship of alcohol with risk of coronary heart disease. One source of evidence is the association between alcohol and blood pressure in observational studies. This association may be a poor indicator of the causal effects of alcohol if there are other factors—"confounders"—that influence both alcohol intake and blood pressure. Many epidemiological methods attempt to correct for, or minimise, observed differences in confounders between study participants. These methods can give useful evidence about causal relations if we measure enough confounders so that, after adjustment or matching, study participants who consume different amounts of alcohol are otherwise comparable. But this assumption is unverifiable; if it does not hold, then findings from observational studies will be biased estimates of causal effects.

People who consume more alcohol may also have other risk factors for cardiovascular disease, such as smoking more heavily than those with lower alcohol consumption. The confounding factor (smoking) induces a positive association between the risk factor (alcohol) and an outcome (blood pressure); interpreting this as causal would be misleading. Measuring a confounder does not perfectly characterise it, so measurement error leads to residual confounding, even after apparent statistical adjustment. Reverse causality is a form of confounding that is difficult to account for. It arises if the outcome or preclinical aspects of the disease that lead to the outcome affect the risk factor. People with symptoms of cardiovascular disease, for example, may consume less alcohol than those without symptoms. This would lead to a negative association between a risk factor (alcohol) and an outcome (cardiovascular disease); interpreting this as being because alcohol consumption decreases the risk of cardiovascular disease would be misleading.

Mendelian randomisation uses genetic variants, which are fixed at conception, to support causal inferences about the effects of modifiable risk factors, which can overcome some types of confounding. In the case of alcohol and blood pressure, a variant in the *ALDH2* gene (specifically the minor A allele of

SUMMARY POINTS

- Mendelian randomisation is a research method that provides evidence about putative causal relations between modifiable risk factors and disease, using genetic variants as natural experiments

- Mendelian randomisation is less likely to be affected by confounding or reverse causation than conventional observational studies

- Like all analytical approaches, however, Mendelian randomisation depends on assumptions, and the plausibility of these assumptions must be assessed

- Moreover, the relevance of the results for clinical decisions should be interpreted in light of other sources of evidence

- We provide a critical appraisal checklist that can be used to assess and interpret Mendelian randomisation studies

---

Box 1: Glossary of common terms used in Mendelian randomisation studies

Readers interested in more detailed discussions of these terms are encouraged to read the references.[1 9 12 13]

Concepts

- *Instrumental variables*—variables that are associated with the risk factor of interest, that are not related to confounders, and that affect the outcome only through the risk factor.[14] An instrumental variable can be any trait (not necessarily a genetic variant) that meets these criteria, but the nature of genetic inheritance means that genetic variants are often plausible instrumental variables

- *Mendelian randomisation*—the use of genetic variants as instrumental variables to investigate the effects of modifiable risk factors for disease[6 15]

- *Multiple instruments*—the use of more than one genetic variant in a Mendelian randomisation analysis

- *Allele score*—the number of alleles associated with an increase in the risk factor of interest. These genetic variants are normally identified in large genome-wide association studies. The statistical efficiency (power) of allele scores can be increased by weighting each variant by the size of its association with the risk factor[16]

- *Weak instrument bias*—can occur in Mendelian randomisation studies when using one or more genetic variants that only explain a small proportion of the variation in the risk factor, coupled with a small sample size

- *Pleiotropic effects*—the effects of a genetic variant on multiple biological pathways. These can either affect the outcome through another trait or pathway to the one under investigation, known as horizontal pleiotropy, or affect other traits through the risk factor of interest, known as vertical pleiotropy.[1] Horizontal pleiotropy is a violation of the instrumental variable assumptions because the effects of the genetic variant on the outcome are not exclusively through the risk factor; this is problematic for Mendelian randomisation studies. Vertical pleiotropy is in fact the essence of Mendelian randomisation—showing that one factor influences a downstream outcome—and is in general not problematic[1]

Statistical methods

- *Single sample Mendelian randomisation*—using one dataset in the instrumental variable analysis to yield the causal estimate of the risk factor on the outcome

- *Two sample Mendelian randomisation*—using two different study samples to estimate the instrument-risk factor and instrument-outcome associations to estimate a causal effect of the risk factor on the outcome. This can be useful when the risk factor or outcome, or both, are expensive to measure.[17 18] It also provides an opportunity to substantially increase the statistical power, by incorporating data from multiple sources, including large consortia[17]

- *MR Egger regression*—a statistical technique that allows one or more genetic variants to have pleiotropic effects, as long as the size of these pleiotropic effects is independent of the size of the genetic variants' effects on the risk factor of interest[19]

Properties of the genetic instrument

- *Testing between observational and Mendelian randomisation estimate*—tests include the Hausman test (for continuous outcomes in single sample Mendelian randomisation) and tests for difference in estimates (for binary outcomes in single sample and two sample Mendelian randomisation[20 21]

- *Tests of instrument strength*—to evaluate the strength of the association between the instrument (the genetic variant) and the risk factor (for example, the partial F statistic and R^2)[22]

---

rs671, rather than the wild type or major allele G) found in east Asian populations slows the metabolism of acetaldehyde, which causes a flush response and other adverse responses to alcohol consumption. In a study of 4057 people selected from the general population, 170 of 1919 men carried two copies of the A allele and drank an average of 1.1 g of alcohol a day, whereas those with no copies drank 23.7 g.[10 11] If men with one or more copies of the A allele have lower blood pressure, then this implies that lower alcohol consumption decreases blood pressure.[11] But this inference relies on several assumptions. The two key challenges when reading Mendelian randomisation studies are evaluating the plausibility of the underlying assumptions and interpreting the results. We discuss these challenges below, using terms defined in the glossary (box 1).

What assumptions does Mendelian randomisation depend on?

Valid instrumental variables are defined by three key assumptions (table 1, fig 1): that they associate with the risk factor of interest (the relevance assumption); that they share no common cause with the outcome (the independence assumption); and that they do not affect the outcome except through the risk factor (the exclusion restriction assumption). A single genetic variant could plausibly meet these conditions if the biological process linking the variant with the risk factor is well understood. But in many cases Mendelian

---

Table 1 | Three key assumptions that must hold for a Mendelian randomisation study to be valid

| Assumption	Description	Tools to assess plausibility	
		Single sample	Two sample
Relevance assumption	The genetic variants associate with the risk factor of interest	The partial F statistic and partial r squared, or risk difference	Variants are associated with the risk factor in a large genome-wide study
Independence assumption	There are no unmeasured confounders of the associations between genetic variants and outcome	Covariate balance tests and bias component plots. Adjusting for principal components of population stratification	Evidence from large genome-wide association studies on the association of the genetic variants used as instruments with other baseline covariates
Exclusion restriction	The genetic variants affect the outcome only through their effect on the risk factor of interest	Biological knowledge, tests of association of the genetic variants and potential alternative mediating pathways	Evidence from large genome-wide association studies that the genetic variants associate with alternative pathways. MR Egger test for pleiotropy, Cook's distance evaluation of outliers
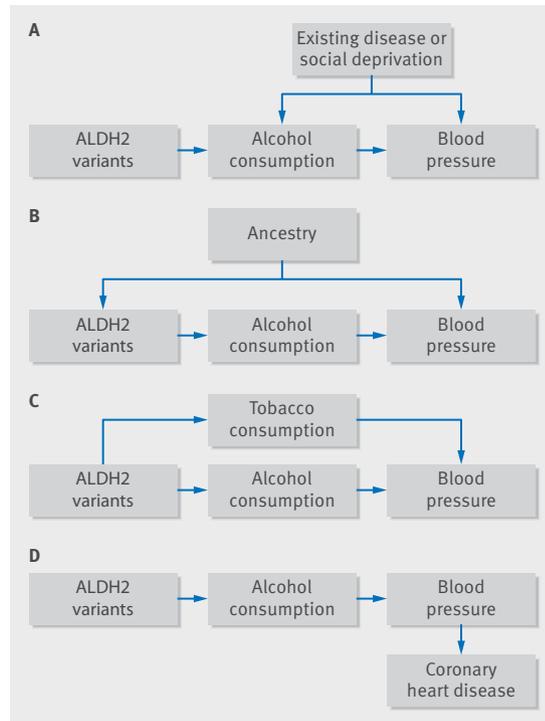
Fig 1 | Examples of Mendelian randomisation and potential violations of assumptions. (A) A simplified causal diagram depicting confounding of the association of alcohol consumption and blood pressure by existing disease or social deprivation. The instrumental variable assumptions are that the genetic variants are associated with the risk factor, that they have no other influence on the outcome, except through alcohol, and that there are no confounders of the genetic variants-outcome association. (B) Confounding by ancestry could occur if variants associated with alcohol consumption had different frequencies in different ethnic groups in the population sampled and if cultural differences affected blood pressure between ethnic groups. This would violate the second instrumental variable assumption—the independence assumption. (C) An example of horizontal pleiotropy, in which the genetic variants associated with alcohol consumption also affect tobacco consumption (violating the third assumption—the exclusion restriction assumption). (D) An example of vertical pleiotropy, in which the effect of *ALDH2* on coronary heart disease is mediated by blood pressure. This example does not violate the Mendelian randomisation assumptions and does not cause bias.

randomisation studies include multiple genetic variants, which can be used in sensitivity analyses to evaluate the underlying assumptions. Generally, the three key assumptions must hold for each of the genetic variants. We describe common strategies for assessing the plausibility of these assumptions and give examples, where possible, from published studies.

Sources of data in Mendelian randomisation

Historically, a typical Mendelian randomisation study required measures of genotypes (variants in *ALDH2* gene), risk factor (alcohol consumption), and outcome (blood pressure) from the same sample of people. This approach is known as single sample Mendelian randomisation. Two sample Mendelian randomisation involves two separate study populations; for example, data on the *ALDH2* genotype and alcohol consumption are measured in one sample, and *ALDH2* genotype and blood pressure in the other.[17 23] This design has two advantages. Firstly, neither the risk factor nor the outcome needs to be measured in all studies, which is particularly useful if they are difficult or expensive to measure. Secondly, it allows the summary results from genome-wide association studies to be used, which can be very large (often >50000) and thus highly precise (table 2 and table 3).

Statistical power is usually much higher in two sample studies. These advantages come with two additional assumptions: the two samples are assumed to represent the same underlying population, and overlap in participants between the two samples can cause bias towards the risk factor-outcome association.[17 39]

Single or multiple genetic variants

The simplest applications of Mendelian randomisation use a single genetic variant as an instrument for the risk factor. Some of the most persuasive published studies use a single genetic variant with a relatively well understood function, so the core assumptions can be supported by biological knowledge. One such example is the use of a missense variant in the *PCSK9* gene, which modifies the function of proprotein convertase subtilisin/kexin type 9, an enzyme that degrades the low density lipoprotein (LDL) receptor on hepatocytes. *PCSK9* variants are associated with

Table 2 | Publicly available data sources for two sample Mendelian randomisation studies		
Consortium name	**Description**	**Most recent sample size**
BCAC[24]	Breast cancer	256 123
CARDIoGRAMplusC4D[25]	Coronary artery disease and myocardial infarction	184 305
CKDGen[26]	Chronic kidney disease	111 666
DIAGRAM[27]	Diabetes	159 208
EAGLE[28]	Antenatal and early life and childhood phenotypes	47 541
EGG[29]	Early growth	153 781
GIANT[30]	Height, BMI, and other adiposity traits	693 529
GLGC[31]	Global lipids genetics consortium	331 368
ISGC[32]	Stroke	84 961
MAGIC[33]	Glucose and insulin related traits	224 459
PGC[34 35]	Psychiatric genetics, alcohol and tobacco, and other related traits	>500 000
SSGAC[36]	Educational attainment and wellbeing	293 723

Table 3 | Databases of genome-wide association study results

Data source	Description	Number of traits	Integrated with statistics package?
MR-Base	A curated database of genome-wide association study results with integrated R package for MR[23]	Over 1000	Yes
PhenoScanner	A curated database of genome-wide association study results with integrated R package for MR[37]	Over 500	Yes
GWAS catalog	Searchable database of genome-wide association study results[38]	Over 24 000	No

altered blood concentrations of LDL cholesterol and risk of heart disease.[40 41] This provides further evidence that LDL cholesterol causes heart disease, and also indicates that drugs that inhibit PCSK9 may have cardiovascular benefits. This was confirmed in phase III clinical trials.[42 43]

In most circumstances, however, single genetic variants individually typically explain only a very small proportion of the variation in a phenotype; investigators may refer to these as "weak instruments," particularly in modest sample sizes. Studies using variants with modest effects in small samples are likely to have very low statistical power and can be biased. To overcome this, investigators have developed methods that use multiple genetic variants that collectively explain more of the variation in a risk factor than a single variant and thus have more statistical power. One way of using multiple genetic variants is to use single nucleotide polymorphisms (SNPs, the most common form of DNA variation among people) as individual instruments in a statistical regression model in the setting of a single or two sample dataset (multiple instruments approach). The second way is to aggregate the variants into an allele score (also called a genetic risk score, gene score, or a multilocus allele score).[16] The allele score is effectively a single

instrumental variable that can be used to predict the risk factor in a Mendelian randomisation analysis.

Genetic variants in these scores are often weighted by their associations with the risk factor to maximise statistical power. Many recent Mendelian randomisation investigations of complex traits (such as blood pressure, BMI, or blood lipids)[44] have used multiple variants because they can help identify pleiotropy or other violations of the underlying assumptions.

Genetic pleiotropy

Genetic variants may affect the outcome through pathways other than through the risk factor of interest (so called horizontal pleiotropic effects, fig 1C). Genetic variants associated with alcohol consumption, for example, may affect other behaviours such as smoking, which would invalidate the exclusion restriction. When using single or multiple genetic variants, Mendelian randomisation estimates require that genetic variants do not have such horizontally pleiotropic effects.[16] This means the results could be biased if a genetic variant or an allele score has pleiotropic effects on the outcome that are not mediated through the risk factor of interest (fig 2). Genetic variants can also affect the outcome through a pathway affected by the risk factor of interest (vertical pleiotropic effects, fig 1D). This does not invalidate the instrumental variable assumptions and does not result in bias.

Various methods have been developed that allow for genetic pleiotropy.[19 46 47] These provide useful sensitivity analyses to explore whether a finding depends on the assumption that all the variants have no pleiotropic effects. One such approach, known as the median estimator, can provide reliable evidence as long as at least half the genetic variants have no pleiotropic effects.[47 49] A second method, known as MR Egger regression, allows all variants to have pleiotropic effects, provided they are not proportional to the variants' effects on the risk factor of interest.[19] A comprehensive analysis of cross trait genetic effects indicates that the magnitude of genome-wide associations are generally independent across traits that are not causally related, so the assumption may be plausible for many potential relationships.[50] MR Egger regression yields less precise estimates than other methods, owing to a power penalty. Most of these methods assume that the risk factor has the same effect on everyone.

How to assess the key assumptions

Tests of instrument strength—the relevance assumption
The power of a Mendelian randomisation study is determined by sample size and strength of the association between the proposed instrument and risk factor. Weak instruments that poorly predict the risk factor cause three problems.[12] Firstly, they provide very little statistical power to test hypotheses. Secondly, bias due to violations of the core instrumental variable assumptions, such as horizontally pleiotropic effects of variants, will be amplified. Thirdly, even when using
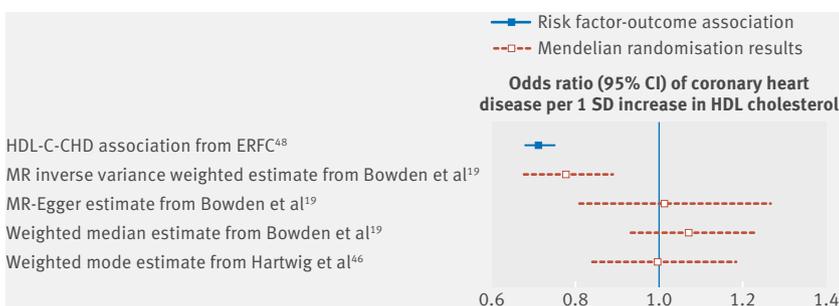
Fig 2 | Example of genetic pleiotropy in Mendelian randomisation: HDL cholesterol and risk of heart disease. Variants associated with HDL cholesterol are likely to have pleiotropic effects on risk of heart disease because they also associate with LDL cholesterol and triglycerides.[2 45] Thus the inverse variance weighted Mendelian randomisation estimate, which assumes no pleiotropy, provides (biased) evidence of a protective role for HDL cholesterol in coronary heart disease. But the estimates using MR Egger, weighted median, and weighted mode, which allow for genetic pleiotropy, are attenuated towards the null. The MR Egger estimator assumes that for the variants with pleiotropic effects on coronary heart disease the magnitude of these effects do not correlate with the magnitude of the variants' effects on HDL cholesterol. These results suggest that the inverse variance weighted estimate is driven by genetic pleiotropy[19] [46-48] and that HDL cholesterol is unlikely to have a major causal role in the development of coronary heart disease. CHD=coronary heart disease; ERFC=Emerging Risk Factors Collaboration; HDL-C=high density lipoprotein cholesterol; SD=standard deviation

very large samples, results using weak instruments are biased towards the outcome-risk factor association in the single sample setting and towards the null in the two sample setting. Precision (assessed with confidence intervals) is underestimated. Weak instruments can be detected using the F statistic for single sample settings.[51] A rule of thumb is that the F statistic should be greater than 10. Exceeding this threshold indicates that a result based on a valid instrumental variable ought not to suffer substantially from weak instrument bias but does not guarantee sufficient statistical power to test a specific hypothesis.

Independence and exclusion restriction assumptions

Confounding by violations of the independence assumption (no confounders) and exclusion restriction (genetic instruments work only through the risk factor) can be investigated by estimating the relation between the genetic instrument and a wide range of characteristics, analogous to the balance of baseline characteristics between treatment arms in a randomised trial.[52-54] Although these assessments cannot prove that the independence and exclusion restriction assumptions hold, they can provide evidence regarding their (lack of) plausibility.

One promising way to evaluate the plausibility of the assumptions is using "negative control" populations. Genetic variants that are known to affect alcohol consumption, for example, should not be associated with outcomes in populations that rarely or never drink, such as children or women in some societies.[11] This tests the exclusion restriction and independence assumptions. If investigators found an association between *ALDH2* variants and blood pressure in a non-drinking population, the instrumental variable assumptions may not hold in the original population either. The usefulness of negative controls depends on them being well reasoned and, of course, available.

How are Mendelian randomisation studies analysed?

The simplest approach is to report the association of the genetic variants and the outcome. This test is relatively robust and should be reported in all studies. It provides evidence as to whether the risk factor causes the outcome, but it is not informative about the size of the effect. The magnitude of the causal effect can be estimated by dividing the genetic variant-outcome association by the genetic variant-risk factor association. This ratio is known as the instrumental variable or Wald estimate.[55]

Inclusion of, or stratifying by, the risk factor of interest

To assess whether a genetic variant (or several in combination) mediates its effect on an outcome through the primary risk factor (and thus to test for exclusion restriction), some studies adjust or stratify for the primary risk factor of interest.[56] But this is problematic for several reasons. Firstly, a residual association of the genetic instrument with the disease after adjustment for the risk factor does not imply that the variant affects the disease through other pathways. The measured risk factor almost certainly does not completely account for the lifetime variation in the trait instrumented by the gene. A residual effect of the gene (on the outcome after adjusting for the measured risk factor) might therefore not indicate a violation of the exclusion restriction assumption. Secondly, adjusting or stratifying for the risk factor can lead to collider bias (see supplementary figure 1), in the same way that conducting a subgroup analysis by "on-treatment" response breaks the randomisation in a trial.[57] To see why, consider a study investigating the effect of statins on coronary heart disease.[58] The on-treatment (or "achieved") LDL cholesterol concentration will break the randomisation and is likely to be biased by other characteristics that influence lipid levels (such as age, sex, diet, physical activity, and so on). Although this is a seemingly common retrospective approach to try and elucidate dose-response relationships in a randomised trial, such analyses can be as vulnerable to confounding as observational analyses. Similarly, adjusting (or stratifying) a Mendelian randomisation analysis for the risk factor can induce spurious correlations between the genetic variants and the outcome.

Why do some Mendelian randomisation studies adjust for other traits?

If a genetic variant is a valid instrument, inclusion of other covariates is not necessary, but they can increase statistical efficiency. Some Mendelian randomisation studies adjust for other characteristics, including directly measured traits (such as age, sex, or smoking) or associations between the genetic instrument and traits other than the risk factor of interest. The genetic variant-risk factor and genetic variant-outcome associations should be adjusted for the same covariates. If not, the instrumental variable estimate may be biased. For example, if a two sample Mendelian randomisation study of the effect of BMI on blood pressure used a blood pressure genome-wide association study that adjusted for BMI, then the estimate would be unreliable and the direction of effect could even be reversed. These problems are an area of active methodological research.

In some samples, the association between a genetic variant and outcome may be confounded by hidden population structure. This can be tackled by adjusting for genetic ancestry or restricting to ethnically homogenous samples. Suppose a study sampled data from an ethnically mixed east Asian and European population. The minor allele of the alcohol variant rs671 is extremely rare in European populations, so an association between this variant and an outcome in this sample could be due to differences in ethnicity (fig 1B). This problem could be mitigated by stratifying by ethnicity or adjusting for genetic principal components (which are calculated in datasets with genome-wide arrays and provide proxy measures of genetic ancestry). The UK Biobank, for

example, provides up to 40 principal components for adjusting genetic analyses for differences in ethnicity. This improves reliability, as has been done in a recent Mendelian randomisation study of adiposity and risk of heart disease and diabetes.[59] If a study depends on these adjustments—that is, there are major differences between the unadjusted and adjusted findings—the results should be treated with caution. At the very least the source of potential bias must be investigated and justified, just as for residual confounding in a conventional epidemiology study.

Why do some Mendelian randomisation studies remove variants from the genetic instrument?

Some studies manually remove (or "prune") genetic variants thought to be pleiotropic from a genetic instrument to provide a more reliable estimate of the association. Voight and colleagues used genetic variants that were exclusively associated with HDL cholesterol (not LDL cholesterol or triglycerides) to investigate the relationship between HDL cholesterol and heart disease.[60] Including or excluding a genetic variant in this way can be arbitrary and limited by available data; such pruning should be done with caution (if at all) and for well justified and transparent reasons. Furthermore, as the sample size increases, genetic variants will be identified as showing associations with multiple traits owing to vertical or horizontal pleiotropy.[61] Excluding variants on the basis of vertical pleiotropic effects will bias findings, since these effects would be ones seen when manipulating the risk factor of interest.[62] One or more variants might be excluded from sensitivity analyses if they are obvious outliers based on visual assessment of the data or using more formal approaches such as Cook's distance.[63]

How are Mendelian randomisation studies reported?

Having checked whether the assumptions underpinning a Mendelian randomisation analysis are valid (table 1), readers should be familiar with how such studies are reported (box 2). They may report the association of one or more genetic variants with the outcome or provide an estimate of the causal (instrumental variable) effects of the risk factor on the outcome, or both. Ference and colleagues report the association between SNPs that alter LDL cholesterol and risk of heart disease, showing compelling genetic evidence of a dose-response association across multiple independent genetic loci. These SNPs can also be used to provide an "overall" causal effect of LDL cholesterol on the risk of coronary heart disease (odds ratio 0.46, 95% confidence ratio 0.41 to 0.51 per 1mmol/L lower LDL cholesterol).[64] These associations leverage genetic differences that occur at conception and can detect whether differences in the risk factor at any point over the life course affect the outcome. As a guide to future research and drug development, it is important to remember that findings of Mendelian randomisation potentially reflect lifetime differences in risk factors.[65 66]

Special cases of reporting findings
Testing for differences between Mendelian randomisation estimates from different SNPs

If multiple genetic variants and biological pathways influence the risk factor, we can test whether the effect of the risk factor on the outcome is similar when using different variants using Hansen tests or Cochrane's Q test for individual level data and summary data, respectively.[67 68] If these test statistics are large (yielding correspondingly small P values), the estimated causal effects of the risk factor may vary across the population or between variants. This might be due to multiple pathways causing the outcome. If these tests cannot be rejected, sample size and statistical power may be insufficient to detect differences that do exist and, therefore, may be falsely reassuring.

Comparing findings from observational and Mendelian randomisation analyses in the same dataset

For continuous outcomes in single sample analyses, the Hausman test can be used to assess whether the Mendelian randomisation and linear regression results (obtained from the same dataset) are systematically different.[20] Differences could occur because the linear regression results are biased by residual confounding, the independence or exclusion restriction assumptions underlying the instrumental variable regression approach are invalid, or both. Finally, the effects being estimated by the two methods may not be the same—the Mendelian randomisation estimate reflects the effects of lifelong perturbations in the risk factor, whereas linear regression results may reflect more acute effects. Mendelian randomisation estimates are almost always less precise and have wider confidence intervals than linear regression (fig 3), so tests for difference often have low statistical power.[74]

How should findings from Mendelian randomisation be interpreted?

The findings from Mendelian randomisation studies, which are less susceptible to confounding and reverse causality bias, sit at the interface between traditional observational epidemiology and interventional trials (fig 4). A well conducted Mendelian randomisation study that reasonably satisfies the above assumptions often provides more reliable evidence than a conventional observational study. But the findings should be interpreted in the context of existing evidence from other sources, using different study designs,[75] and clinical guidelines should not be rewritten solely on the basis of Mendelian randomisation results.

Triangulating the evidence

Some studies report the Mendelian randomisation estimate in the context of traditional observational epidemiology (fig 3; supplementary table 1). When there is clear evidence that the associations differ, the Mendelian randomisation estimate can provide strong evidence that the traditional observational estimate arises from confounding and/or reverse causality.

Box 2: Critical appraisal checklist for evaluating Mendelian randomisation studies

Some key questions readers can ask below.

Core Mendelian randomisation assumptions

- Is there sufficient evidence that the genetic variants are robustly associated with the risk factor of interest?
- Are the genetic variants associated with potential confounders? Do the authors present this relationship?
- Is there any way for the genetic variants to affect the outcome through alternative pathways (horizontal pleiotropy)? Do the authors present alternative Mendelian randomisation approaches (such as MR Egger, median, and mode estimators, or use of "negative control" populations) to investigate this more fully?

Methods reporting

All studies

- Are the effect and other alleles coded in the same direction for the exposure and outcome?

Two sample studies

- Were the two samples drawn from the same population?
- Were the two samples independent?
- Was the analysis restricted to independent variants (that is, pruned of SNPs in linkage disequilibrium) or did the analysis allow for the correlation between variants?

Data presentation

- Do the authors present the results as a genetic association, an instrumental variable estimate, or both?
- If they provide an instrumental variable estimate, do they compare it with the conventional observational estimate?
- Do the authors provide sensitivity analyses such as MR Egger, weighted median, and mode Mendelian randomisation, or use negative control populations?
- Do the authors manually pick and choose which SNPs go into the instrument to tackle pleiotropy? If so, is the approach and justification clear?
- Do the authors provide the data that they used (especially for Mendelian randomisation analyses conducted at the summary level) in a supplement to allow researchers to reproduce their findings?

Interpretation

- If the Mendelian randomisation estimate is similar to the observational estimate and provides evidence in support of a causal effect, could it be due to weak instrument bias in a single study or confounding through, for example, horizontal pleiotropy?
- If the Mendelian randomisation estimate differs from the observational estimate and provides little evidence of a causal effect, could this be due to weak instrument bias when using two different samples or negative confounding due to pleiotropy?
- Mendelian randomisation provides estimates of the effects of the risk factor over a lifetime, and the numerical effect estimates may not be clinically meaningful. Will interventions at a specific age have the same sized effects?
- Are the 95% confidence intervals of the Mendelian randomisation estimate sufficiently precise to identify the observational estimate and a clinically meaningful difference?

Clinical implications

- Do the results triangulate with other forms of evidence? Could a clinical trial be conducted to provide definitive evidence, as in the case of PCSK9 inhibitors?
- If a randomised clinical trial is not feasible (such as in the case of alcohol consumption and risk of heart disease) or unlikely to be conducted in the short term (such as the case of lifestyle interventions to lower BMI and risk of heart disease), and there is existing evidence from multiple Mendelian randomisation studies and other robust study designs that converge on a similar result and show consistency of association, this information can be used to guide patient care; for example, advising weight loss to prevent heart disease or advising against moderate alcohol consumption to prevent cardiovascular disease

---

In the case of C-reactive protein and coronary heart disease, the Mendelian randomisation estimates suggest that targeting CRP is unlikely to be a viable therapeutic target for the prevention of coronary heart disease.[72] In contrast, when the Mendelian randomisation estimate has very wide confidence intervals that overlap the observational estimate and include the null (as in the case of dairy consumption and blood pressure),[69] very little can be inferred from the Mendelian randomisation results.

Summary estimates from observational and Mendelian randomisation studies can be compared formally using tests for heterogeneity, such as Cochrane's Q test and tests for differences in estimates.[21] This can provide statistical evidence of a difference, but whether this should be interpreted as clinically meaningful is further outlined below.

Clinical and public health implications

Clinical and public health decisions about potential interventions ideally require evidence about the size and direction of an effect for a specific population. We might want to know, for example, how reducing alcohol consumption by an average of one serving a day beginning at age 45 affects blood pressure five years later. But this must be approached with care because genetic variants generally relate to lifelong differences in a risk factor, not the effects at a specific time. *ALDH2* is associated with alcohol consumption from adolescence through middle age,[76] so the Mendelian randomisation results (coupled with biological knowledge) indicate that consuming less alcohol would lead to lower blood pressure on average, but do not necessarily inform the effect size of an intervention at a specific time in life for a specific duration. Another example is the relation
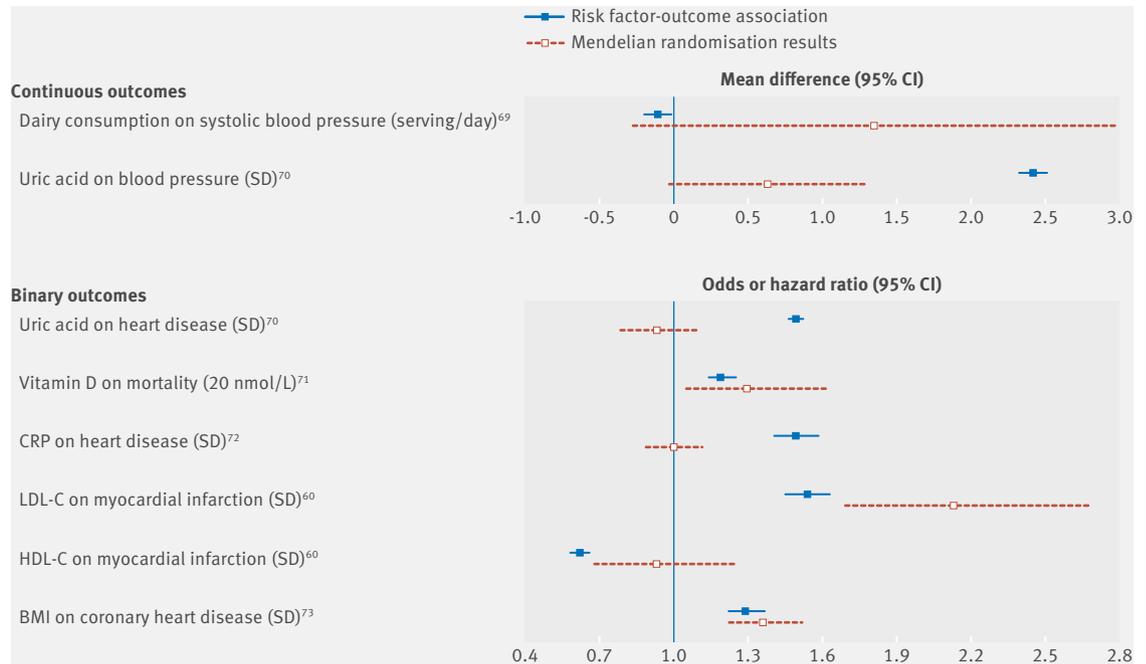
Fig 3 | Example associations between risk factors and outcomes from traditional observational epidemiology and Mendelian randomisation instrumental variable estimates.[60, 69-73] For some associations—such as vitamin D and mortality—the Mendelian randomisation results potentially confirm some causal relation. For other associations—CRP and heart disease—the Mendelian randomisation results are consistent with there being no causal effect. BMI=body mass index; CRP=C reactive protein; HDL-C=high density lipoprotein cholesterol; LDL-C=low density lipoprotein cholesterol

between vitamin D and multiple sclerosis, where data from Mendelian randomisation studies indicate a protective effect, but such protection may be limited to childhood and adolescence and not later adult life.[66 77] Biological and methodological knowledge about the relations between risk factors and outcomes is critical to interpreting Mendelian randomisation studies. Does risk accumulate over time? Does the risk factor have an acute effect? Is there a particular time in life where this risk factor has an effect on this disease?[6 66 78 79] This is why some studies focus on the association of the genetic variants and outcome,[56] rather than providing

a direct quantification of the causal effect scaled to a difference in the exposure of interest.

Consider another example, where the biology and context are important for the appraisal of Mendelian randomisation findings. Variants in the *CHRNA5* gene that associate with increased heaviness of smoking are (through smoking behaviour) associated with increased risk of lung cancer.[80] People with these variants are also likely to find it more difficult to reduce smoking after diagnosis of lung cancer. But variants in this gene cannot be used to assess the effect of quitting or reducing smoking on mortality after diagnosis with
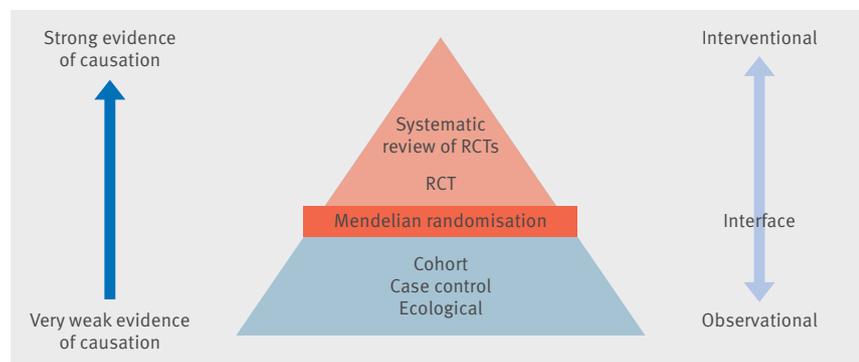

Fig 4 | A hierarchy of observational and experimental data. Mendelian randomisation studies sit at the interface of experimental and observational studies. Their findings can be used to provide more reliable evidence to guide interventional research and provide information about potential public health interventions when a randomised controlled trial may not be feasible. Although we adapt the conventional pyramid of evidence for presentation purposes, we consider that triangulation of findings from different study designs should be used.[75]

lung cancer, because people with the risk variants will have smoked more heavily before diagnosis. The association between *CHRNA5* risk variants and mortality in people with lung cancer reflects an individual's lifetime exposure to tobacco smoke, and we would not expect an instrumental variable estimate using *CHRNA5* to reflect the effect of current smoking on prognosis. Most importantly, the triggers of disease onset may be entirely different to the factors leading to progression.[81] This means that Mendelian randomisation studies need to be performed using data from studies of disease progression and prognosis, rather than disease occurrence, to generate findings of direct relevance to treating disease.[81] Finally, when an exposure is seemingly non-modifiable (such as height[82 83]), a causal relation can point towards potential mechanisms or pathways that lead to disease (even if the actual risk factor under investigation is non-modifiable); such pathways are likely to be modifiable, providing new insights in disease aetiology.

Summary

Mendelian randomisation studies can provide reliable evidence on the effect of modifiable risk factors for disease or ill health and can overcome some limitations of traditional observational epidemiology. In settings for which the instrumental variable assumptions are well justified (assessed as described above and using biological knowledge), the findings could help prioritise clinical trials or drug development and inform clinical or public health decision making.[84-86]

The conversation on how to conduct, report, and interpret Mendelian randomisation studies is still ongoing. Other areas of medical research, such as randomised controlled trials, have been greatly strengthened by close collaborations between methodologists, empirical researchers, and clinicians. Similar collaborations are needed to ensure the strengths and limitations of Mendelian randomisation are fully appreciated and realised.

1 Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 2014;23(R1):R89-98. doi:10.1093/hmg/ddu328

2 Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res Published Online First* 2015;17. doi:10.1177/0962280215597579.

3 Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr* 2016;103:965-78. doi:10.3945/ajcn.115.118216

4 Davies NM, Davey Smith G, Windmeijer F, Martin RM. Issues in the reporting and conduct of instrumental variable studies: a systematic review. *Epidemiology* 2013;24:363-9. doi:10.1097/EDE.0b013e31828abafb

5 Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology* 2013;24:370-4. doi:10.1097/EDE.0b013e31828d0590

6 Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003;32:1-22. doi:10.1093/ije/dyg070

7 Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* 2005;330:1076-9. doi:10.1136/bmj.330.7499.1076

8 Evans DM, Davey Smith G. Mendelian randomization: new applications in the coming age of hypothesis-free causality. *Annu Rev Genomics Hum Genet* 2015;16:327-50. doi:10.1146/annurev-genom-090314-050016

9 Burgess S, Butterworth A, Malarstig A, Thompson SG. Use of Mendelian randomisation to assess potential benefit of clinical intervention. *BMJ* 2012;345:e7325-7325. doi:10.1136/bmj.e7325

10 Takagi S, Baba S, Iwai N, et al. The aldehyde dehydrogenase 2 gene is a risk factor for hypertension in Japanese but does not alter the sensitivity to pressor effects of alcohol: the Suita study. *Hypertens Res* 2001;24:365-70. doi:10.1291/hypres.24.365

11 Chen L, Davey Smith G, Harbord RM, Lewis SJ. Alcohol intake and blood pressure: a systematic review implementing a Mendelian randomization approach. *PLoS Med* 2008;5:e52. doi:10.1371/journal.pmed.0050052

12 Hernán M, Robins J. *Causal Inference.* Chapman & Hall/CRC, 2018, (forthcoming).

13 Burgess S, Thompson SG. *Mendelian randomization: methods for using genetic variants in causal estimation*. Boca Raton: CRC Press, Taylor & Francis Group 2015.

14 Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica* 1994;62:467. doi:10.2307/2951620.

15 Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 2008;27:1133-63. doi:10.1002/sim.3034

16 Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 2013;42:1134-44. doi:10.1093/ije/dyt093

17 Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 2013;178:1177-84. doi:10.1093/aje/kwt084

18 Inoue A, Solon G. Two-sample instrumental variables estimators. *Rev Econ Stat* 2010;92:557-61. doi:10.1162/REST_a_00011.

19 Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol* 2015;44:512-25. doi:10.1093/ije/dyv080

20 Hausman JA. Specification tests in econometrics. *Econometrica* 1978;46:1251-71. doi:10.2307/1913827.

21 Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;326:219. doi:10.1136/bmj.326.7382.219

22 Angrist JD, Pischke JS. *Mostly harmless econometrics: an empiricist's companion.* Princeton Univ Pr, 2009.

23 Hemani G, Zheng J, Wade KH, et al. The MR-base enables causal inference across the phenome. *eLife* [forthcoming].

24 Michailidou K, Lindström S, Dennis J, et al, NBCS Collaborators, ABCTB Investigators, ConFab/AOCS Investigators. Association analysis identifies 65 new breast cancer risk loci. *Nature* 2017;551:92-4. doi:10.1038/nature24284

25 Nikpay M, Goel A, Won HH, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;47:1121-30. doi: 0.1038/ng.3396

26 Li M, Li Y, Weeks O, et al, CHARGE Glycemic-T2D Working Group, CHARGE Blood Pressure Working Group. SOS2 and acp1 loci identified through large-scale exome chip analysis regulate kidney development and function. *J Am Soc Nephrol* 2017;28:981-94. doi:10.1681/ASN.2016020131

27 Scott RA, Scott LJ, Mägi R, et al. DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. an expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes* 2017;66:2888-902. doi:10.2337/db16-1253

28 Felix JF, Bradfield JP, Monnereau C, et al, Bone Mineral Density in Childhood Study (BMDCS), Early Genetics and Lifecourse Epidemiology (EAGLE) consortium, Early Growth Genetics (EGG) Consortium, Bone Mineral Density in Childhood Study BMDCS. Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. *Hum Mol Genet* 2016;25:389-403. doi:10.1093/hmg/ddv472

29 Horikoshi M, Beaumont RN, Day FR, et al, CHARGE Consortium Hematology Working Group. Genome-wide associations for birth weight and correlations with adult disease. *Nature* 2016;538:248-52. doi:10.1038/nature19806

30 Yengo L, Sidorenko J, Kemper KE, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700,000 individuals of European ancestry. *bioRxiv* 2018, doi:10.1101/274654.

31 Liu DJ, Peloso GM, Yu H, et al, Charge Diabetes Working Group, EPIC-InterAct Consortium, EPIC-CVD Consortium, GOLD Consortium, VA Million Veteran Program. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* 2017;49:1758-66. doi:10.1038/ng.3977

32 Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium, the Stroke Genetics Network (SiGN), and the International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* 2016;15:695-707. doi:10.1016/S1474-4422(16)00102-2

33 Shungin D, Winkler TW, Croteau-Chonka DC, et al, ADIPOGen Consortium, CARDIOGRAMplusC4D Consortium, CKDGen Consortium, GEFOS Consortium, GENIE Consortium, GLGC, ICBP, International Endogene Consortium, LifeLines Cohort Study, MAGIC Investigators, MuTHER Consortium, PAGE Consortium, ReproGen Consortium. New genetic loci link adipose and insulin biology to body fat distribution. *Nature* 2015;518:187-96. doi:10.1038/nature14132

34 Wray NR, Ripke S, Mattheisen M, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 2018;50:668-81. doi:10.1038/s41588-018-0090-3

35 Sullivan PF, Agrawal A, Bulik CM, et al, Psychiatric Genomics Consortium. Psychiatric genomics: an update and an agenda. *Am J Psychiatry* 2018;175:15-27. doi:10.1176/appi.ajp.2017.17030283

36 Okbay A, Beauchamp JP, Fontana MA, et al, LifeLines Cohort Study. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 2016;533:539-42. doi:10.1038/nature17671

37 Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;32:3207-9. doi:10.1093/bioinformatics/btw373

38 MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017;45(D1):D896-901. doi:10.1093/nar/gkw1133

39 Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol* 2016;40:597-608. doi:10.1002/gepi.21998

40 Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 2006;354:1264-72. doi:10.1056/NEJMoa054013

41 Abifadel M, Varret M, Rabès J-P, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 2003;34:154-6. doi:10.1038/ng1161

42 Sabatine MS, Giugliano RP, Keech AC, et al, FOURIER Steering Committee and Investigators. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med* 2017;376:1713-22. doi:10.1056/NEJMoa1615664

43 American College of Cardiology. ODYSSEY outcomes: results suggest use of PCSK9 inhibitor reduces CV events, LDL-C in ACS patients. 10 Mar 2018. http://www.acc.org/latest-in-cardiology/articles/2018/03/05/15/53/sat-9am-odyssey-outcomes-cv-outcomes-with-alirocumab-after-acs-acc-2018

44 Benn M, Nordestgaard BG. From genome-wide association studies to Mendelian randomization: Novel opportunities for understanding cardiovascular disease causality, pathogenesis, prevention, and treatment. *Cardiovasc Res* 2018. doi:10.1093/cvr/cvy045

45 Holmes MV, Asselbergs FW, Palmer TM, et al, UCLEB consortium. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* 2015;36:539-50. doi:10.1093/eurheartj/eht571

46 Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 2017;46:1985-98. doi:10.1093/ije/dyx102

47 Bowden J, Davey Smith G, Haycock PC, Burgess S. consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 2016;40:304-14. doi:10.1002/gepi.21965

48 Di Angelantonio E, Sarwar N, Perry P, et al, Emerging Risk Factors Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. *JAMA* 2009;302:1993-2000. doi:10.1001/jama.2009.1619

49 Kang H, Zhang A, Cai TT, et al. Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. *J Am Stat Assoc* 2016;111:132-44. doi:10.1080/01621459.2014.994705.

50 Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* 2016;48:709-17. doi:10.1038/ng.3570

51 Staiger D, Stock J. Instrumental variables regression with weak instruments. *Econometrica* 1997;65:557-86. doi:10.2307/2171753.

52 Davey Smith G, 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:e352. doi:10.1371/journal.pmed.0040352

53 Tchetgen Tchetgen EJ, Walter S, Glymour MM. Commentary: building an evidence base for mendelian randomization studies: assessing the validity and strength of proposed genetic instrumental variables. *Int J Epidemiol* 2013;42:328-31. doi:10.1093/ije/dyt023

54 Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* 2012;175:332-9. doi:10.1093/aje/kwr323

55 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. doi:10.1093/ije/dyv071

56 Holmes MV, Dale CE, Zuccolo L, et al, InterAct Consortium. Association between alcohol and cardiovascular disease: Mendelian randomisation analysis based on individual participant data. *BMJ* 2014;349:g4164-4164. doi:10.1136/bmj.g4164

57 Rosenbaum PR. The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J R Stat Soc Ser Gen* 1984;147:656. doi:10.2307/2981697.

58 Giugliano RP, Pedersen TR, Park J-G, et al, FOURIER Investigators. Clinical efficacy and safety of achieving very low LDL-cholesterol concentrations with the PCSK9 inhibitor evolocumab: a prespecified secondary analysis of the FOURIER trial. *Lancet* 2017;390:1962-71. doi:10.1016/S0140-6736(17)32290-0

59 Lyall DM, Celis-Morales C, Ward J, et al. Association of body mass index with cardiometabolic disease in the uk biobank: a mendelian randomization study. *JAMA Cardiol* 2017;2:882-9. doi:10.1001/jamacardio.2016.5804

60 Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* 2012;380:572-80. doi:10.1016/S0140-6736(12)60312-2

61 Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell* 2017;169:1177-86. doi:10.1016/j.cell.2017.05.038

62 Merino J, Leong A, Posner DC, et al. Genetically driven hyperglycemia increases risk of coronary artery disease separately from type 2 diabetes. *Diabetes Care* 2017;40:687-93. doi:10.2337/dc16-2625

63 Corbin LJ, Richmond RC, Wade KH, et al. BMI as a modifiable risk factor for type 2 diabetes: refining and understanding causal estimates using Mendelian randomization. *Diabetes* 2016;65:3002-7. doi:10.2337/db16-0418

64 Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease: a Mendelian randomization analysis. *J Am Coll Cardiol* 2012;60:2631-9. doi:10.1016/j.jacc.2012.09.017

65 Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet* 2015;47:856-60. doi:10.1038/ng.3314

66 Holmes MV, Ala-Korpela M, Davey Smith G. Mendelian randomization in cardiometabolic disease: challenges in evaluating causality. *Nat Rev Cardiol* 2017;14:577-90. doi:10.1038/nrcardio.2017.78

67 Hansen LP. Large sample properties of generalized method of moments estimators. *Econom J Econom Soc* 1982;50:1029-54. doi:10.2307/1912775.

68 Greco M FD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* 2015;34:2926-40. doi:10.1002/sim.6522

69 Ding M, Huang T, Bergholdt HK, Nordestgaard BG, Ellervik C, Qi L, CHARGE Consortium. Dairy consumption, systolic blood pressure, and risk of hypertension: Mendelian randomization study. *BMJ* 2017;356:j1000. doi:10.1136/bmj.j1000

70 Palmer TM, Nordestgaard BG, Benn M, et al. Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. *BMJ* 2013;347:f4262-4262. doi:10.1136/bmj.f4262

71 Afzal S, Brøndum-Jacobsen P, Bojesen SE, Nordestgaard BG. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. *BMJ* 2014;349:g6330-6330. doi:10.1136/bmj.g6330

72 Wensley F, Gao P, Burgess S, et al, C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC). Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ* 2011;342:d548-548. doi:10.1136/bmj.d548

73 Dale CE, Fatemifar G, Palmer TM, et al, UCLEB Consortium; METASTROKE Consortium. 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;135:2373-88. doi:10.1161/CIRCULATIONAHA.116.026560

74 Sedgwick P. Understanding P values. *BMJ* 2014;349:g4550. doi:10.1136/bmj.g4550

75 Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016;45:1866-86.

76 Irons DE, Iacono WG, Oetting WS, McGue M. Developmental trajectory and environmental moderation of the effect of *ALDH2* polymorphism on alcohol use. *Alcohol Clin Exp Res* 2012;36:1882-91. doi:10.1111/j.1530-0277.2012.01809.x

77 Mokry LE, Ross S, Ahmad OS, et al. Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. *PLoS Med* 2015;12:e1001866. doi:10.1371/journal.pmed.1001866

78 VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. *Epidemiology* 2014;25:427-35. doi:10.1097/EDE.0000000000000081

79 Swanson SA, Tiemeier H, Ikram MA, Hernán MA. Nature as a trialist? Deconstructing the analogy between Mendelian Randomization and randomized trials. *Epidemiology* 2017;28:653-9. doi:10.1097/EDE.0000000000000699

80 Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *J Natl Cancer Inst* 2008;100:1552-6. doi:10.1093/jnci/djn363

81 Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. *PLoS Genet* 2017;13:e1006944. doi:10.1371/journal.pgen.1006944

82 Nelson CP, Hamby SE, Saleheen D, et al, CARDIoGRAM+C4D Consortium. Genetically determined height and coronary artery disease. *N Engl J Med* 2015;372:1608-18. doi:10.1056/NEJMoa1404881

83 Nüesch E, Dale C, Palmer TM, et al, Adult height, coronary heart disease and stroke: a multi-locus Mendelian randomization meta-analysis. *Int J Epidemiol* 2016;45:1927-37. doi:10.1093/ije/dyv074

84 Davey Smith G, Paternoster L, Relton C. When will Mendelian randomization become relevant for clinical practice and public health? *JAMA* 2017;317:589-91. doi:10.1001/jama.2016.21189

85 Walker VM, Davey Smith G, Davies NM, Martin RM. Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities. *Int J Epidemiol* 2017;46:2078-89. doi:10.1093/ije/dyx207

86 Fordyce CB, Roe MT, Ahmad T, et al. Cardiovascular drug development: is it dead or just hibernating? *J Am Coll Cardiol* 2015;65:1567-82. doi:10.1016/j.jacc.2015.03.016

Supplementary table: Strength of genetic instruments
Supplementary figure: An illustration of collider bias