nature genetics

Perspective

https://doi.org/10.1038/s41588-026-02546-6

Challenges and future directions for Mendelian randomization

Received: 14 July 2025

Accepted: 12 February 2026

Published online: 15 April 2026

Check for updates

Eleanor Sanderson[1,2] ✉, Michael G. Levin[3,4], Venexia Walker[1,2,5], Shuai Yuan[4,5,6], Isabella Badini[7], Julia Dolce[3], Karina J. Mahida[3,8], Ju-Woo Nho[3,9], Jean-Baptiste Pingault[10], Scott M. Damrauer[4,5] ✉, Gibran Hemani[1,2] ✉ & Neil M. Davies[7,11,12] ✉

Mendelian randomization has evolved from a niche methodology to a widely adopted research approach. In this Perspective, we briefly present a bibliometric analysis of the Mendelian randomization literature to inform a discussion of how Mendelian randomization studies are conducted and how they do not fully realize the potential of the data and techniques available to empirically examine the reliability of assumptions. We propose that future progress will depend on integrating empirical evidence from molecular, cellular, animal and quasi-experimental studies to assess its assumptions and causal claims. We also highlight how the shifting landscape of genetic and genomic data presents new challenges and opportunities for the Mendelian randomization framework, providing a deeper understanding of causal mechanisms.

Mendelian randomization is a method for causal inference that leverages the random inheritance of alleles as instrumental variables to estimate the causal effects of exposures on outcomes[1–4] (Box 1). Exposures investigated in Mendelian randomization studies have included risk factors such as body mass index (BMI), the proteome, epigenome, microbiome and even environmental exposures such as air pollution. Outcomes have included most diseases, intermediate risk factors and social outcomes. The instrumental variables are genetic variants associated with the exposure of interest. Mendelian randomization can overcome bias arising from unmeasured confounding of exposures and outcomes, a key limitation of estimates from other observational study designs (Box 1). Mendelian randomization is an increasingly used and accessible approach that has substantially expanded our understanding of disease etiology. It has provided insights into a range of research questions, including many that are either very expensive or challenging to address through alternative methods for causal inference, such as randomized controlled trials or other observational designs. For example, Mendelian randomization has helped refute claims of detrimental effects of C-reactive protein[5] and protective effects of high-density lipoprotein cholesterol[6] on coronary heart disease, as well as clarified the direction of causation between educational attainment and myopia[7].

We reviewed the literature to characterize the nature of published Mendelian randomization papers, analyzing the abstracts of 10,126 studies that included the term ‘Mendelian randomization’ in the title or abstract (Supplementary Note 1). The number of articles using Mendelian randomization has increased exponentially, especially since 2019, primarily due to the increased availability of publicly available genome-wide association study (GWAS) summary statistics (Supplementary Fig. 1). The number of abstracts reporting

[1]Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK. [2]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. [3]Division of Cardiovascular Medicine, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. [4]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA. [5]Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. [6]Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [7]Division of Psychiatry, University College London, London, UK. [8]Harvard University, Cambridge, MA, USA. [9]Warren Alpert Medical School, Brown University, Providence, RI, USA. [10]Department of Clinical, Educational, and Health Psychology, Division of Psychology and Language Sciences, University College London, London, UK. [11]Department of Statistical Science, University College London, London, UK. [12]Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway. ✉e-mail: eleanor.sanderson@bristol.ac.uk; scott.damrauer@pennmedicine.upenn.edu; g.hemani@bristol.ac.uk; neil.m.davies@ucl.ac.uk

https://doi.org/10.1038/s41588-026-02546-6

BOX 1

Glossary of terms used in Mendelian randomization studies

Bias The difference between an effect estimate (for example, a phenotypic association of an exposure and outcome) and the true causal effect in the population (that is, the difference in outcomes that would occur if the exposure were changed).

Causal effect The effect of an exposure on an outcome.

Estimand The statistical parameter of interest in a study (for example, the average causal effect of the exposure on the outcome).

Exposure A measured phenotype that might affect the outcome.

Gene–environment equivalence The concept that the effect of changing the exposure genetically will have the same effect as changing the exposure by other environmental means. This is implied if the instrumental variable assumptions hold.

Individual-level analyses Analyses that use molecular genetic data structured at the individual level, where each row is an individual and genotypes are stored in columns.

Instrumental variables Must be associated with the exposure, share no common cause with the outcome, and only affect the outcome through the exposure. Mendelian randomization studies use genetic variants as instrumental variables.

One-sample analyses Estimate the association of the genetic variant, the exposure and the outcome in the same samples.

Outcome A measured phenotype that is thought to be affected by the exposure(s).

Summary-level analyses Analyses that use data structured at the genetic-variant level, typically involve associations of genetic variants and different phenotypes.

Two-sample analyses Estimate the associations of the genetic variants and the exposure and outcome in different samples. Note that one- and two-sample analyses can be conducted using either individual-level or summary-level data.

'statistically significant' effects increased more rapidly than those reporting effects consistent with the null, which may indicate selective reporting of study results, data-driven analytic decisions and p-hacking[8] (Supplementary Fig. 2). P-hacking occurs when researchers select which hypotheses to test and which results to report from among many possible hypotheses or analytic approaches solely on the basis of statistical significance. Summary-level Mendelian randomization is particularly susceptible to p-hacking because the data are readily available and the analyses are quick to run and easy to automate across a large set of exposures and/or outcomes. Although not specific to Mendelian randomization[9–12], there is concern that paper mills and AI-generated papers have driven a substantial proportion of this growth[13], leading some to suggest the field faces a credibility crisis[13–15]. Similar problems in other open-access health datasets, such as the National Health and Nutrition Examination Survey, have led to calls for preregistration of studies[9,16], which could also help reduce these issues in summary-level GWAS. This would require researchers to publish detailed analysis plans, specifying primary outcomes, datasets, estimators and an approach to statistical significance before analyzing the data.

Although the production of papers has accelerated across many study designs in biomedical research (Supplementary Fig. 1), the rapid growth of Mendelian randomization studies is relatively recent; hence, the absolute number of published studies remains very small compared with, for example, GWAS and systematic reviews. Nevertheless, this may indicate that the volume of Mendelian randomization studies will continue to increase in the coming years, underscoring the importance of developing appropriate analytical guidelines and publishing norms.

Summary-level analysis expands the scope of causal inference

There has been a rapid shift towards the use of summary statistics rather than individual-level data in Mendelian randomization studies (Supplementary Fig. 2). A key advantage of summary-level approaches is their ability to efficiently combine information from different studies, thereby increasing effective sample sizes and the scope of questions that can be addressed (Figs. 1 and 2). This approach is particularly powerful in cases where the exposure or outcome is difficult to measure, for example, rare disease outcomes or traits that are expensive to measure, including proteomic and metabolomic markers[17]. GWAS summary statistics are far easier to share than individual-level genotype data and are freely available in various forms[18]. Many software tools have been developed to automate most steps of summary-data analysis—for example, the OpenGWAS database can be directly integrated with the TwoSampleMR and MendelianRandomization analysis packages[18,19]. This framework also increases the equity in research by removing barriers, such as access to funding, data and computational power.

Our analysis of Mendelian randomization studies indicates that they are being conducted across a broad spectrum of health and social domains, with reports incorporating almost all possible combinations of trait categories, which reflects the value of the summary-level analytical framework being able to integrate data generated across disciplines (Fig. 2). There is a particular focus on molecular exposure traits, reflecting the important role of Mendelian randomization in drug-target prioritization[20]. Cardiometabolic, cancer and nervous system-related health conditions were substantially overrepresented among the outcome traits analyzed in Mendelian randomization studies.

Bias in Mendelian randomization

Mendelian randomization depends on the underlying instrumental variable assumptions—relevance, independence and the exclusion restriction[4,21] (Box 2). In addition, a fourth assumption is required to interpret the effect estimate obtained—typically the average causal effect of the exposure across the population—linking the statistical model to an interpretable estimand (Supplementary Note 2). Together, these instrumental variable assumptions are necessary for gene-environment equivalence, that is, that changing exposure through genetic variation will have the same effect as changing it through an environmental intervention (Box 1).

Mendelian randomization analyses generally provide estimates of the effect of a lifetime change in an exposure on an outcome. Thus, the effect estimate obtained from a Mendelian randomization study may not be equivalent to the effects of a real-world intervention that occurs at a specific age. The disconnect between the causal effect of potential interest (which could be described as 'what would happen if I changed this exposure in this way now') and what can be estimated with Mendelian randomization ('what would happen if this exposure had been changed for someone's entire lifetime') potentially shifts emphasis towards hypothesis testing, rather than interpretation of the effect estimate obtained (Box 3). However, by interpreting the magnitude of the effect of a lifetime change in an exposure,

https://doi.org/10.1038/s41588-026-02546-6

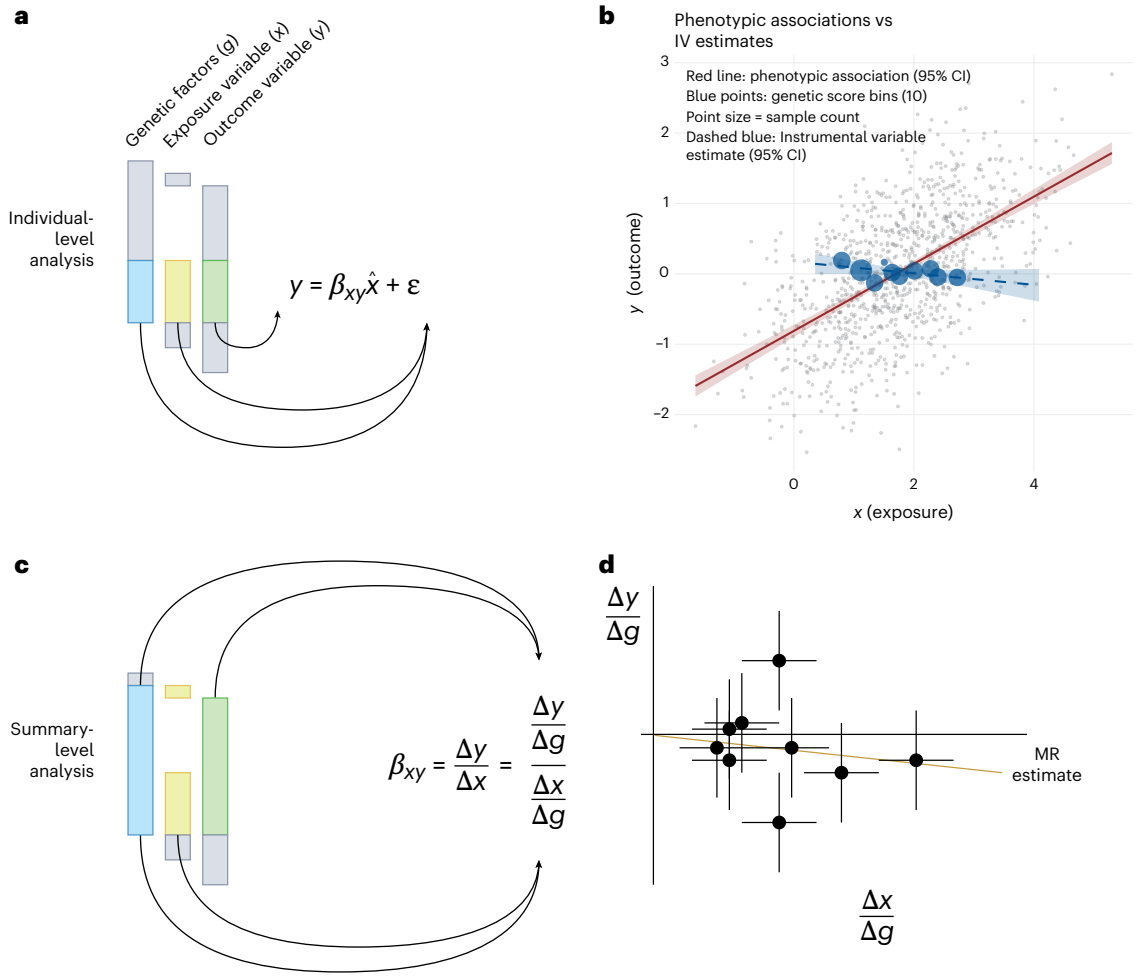


Fig. 1 | Individual- and summary-level Mendelian randomization. **a**–**d**, Mendelian randomization (MR) can be conducted using either individual-level data (**a**,**b**) or summary-level data (**c**,**d**). **a**, Analyses using individual-level data are restricted to those samples for which genetic (g), exposure (x) and outcome (y) measures (plus any covariates, not shown here) are available. **b**, In individual-level analyses, MR estimates (blue) can be contrasted against multivariable-adjusted phenotypic associations (red). **c**, Analyses using summary data are more flexible and do not require that the exposure and outcome variables be measured in the same samples. **d**, They use genetic associations between a variant and exposure, along with corresponding estimates of those variants on the outcome, which can be obtained from the same samples or from a different dataset. While the terms 'individual-level' MR and 'summary data' MR are clearer definitions[3], the terms 'one-sample' (to refer to individual-level data) and 'two-sample' (to refer to summary data) are still much more widely used in the literature. Both individual and summary data analyses typically select specific genetic variants associated with the exposure with $P < 5 \times 10^{-8}$ in GWAS summary data as instruments[83]. CI, confidence interval; IV, instrumental variable.

the relative importance of an exposure can be assessed. By clearly defining the causal effect of interest and the causal effect that can be estimated with Mendelian randomization, the distinction BETWEEN them can be made clear and the results obtained can be interpreted in this context.

Progress in assessing bias in Mendelian randomization studies

Several commentaries have argued that, given the ease with which Mendelian randomization can be implemented, a large fraction of these studies may be of low quality, owing to insufficiently rigorous study design or implementation[13,14,22]. A key question therefore is how should Mendelian randomization studies be conducted to provide reliable and impactful scientific insights.

There are two broad approaches for assessing the instrumental variable assumptions, which, for convenience, we categorize as 'statistical' and 'empirical'. Statistical approaches impose different assumptions on the same variant-exposure and variant-outcome associations or an individual-level dataset. In contrast, empirical approaches use additional data beyond the variant-exposure and variant-outcome associations to explore the robustness of the results and whether other mechanisms are likely to explain the findings[23,24].

A large proportion of methods for Mendelian randomization have concentrated on assessing horizontal pleiotropy using statistical methods, which occurs when a genetic variant directly affects two traits. Horizontal pleiotropy can violate the third instrumental-variable condition, the exclusion restriction[25] (defined in Box 2). Consequently, many statistical methods have been developed to assess the sensitivity of estimated effects to various assumptions about the presence of pleiotropic effects. Studies typically use several different statistical methods, all applied to the same underlying data. If multiple methods yield comparable effect estimates, it is often argued that these estimates are more likely to reflect the true causal effect of the exposure. Conversely, if estimates differ across methods, this suggests violations of the instrumental variable assumptions and potential bias (see Box 3 for further details on bias).

An advantage of this rationale is that it does not depend on a comprehensive measurement of the sources of bias. However, these

https://doi.org/10.1038/s41588-026-02546-6

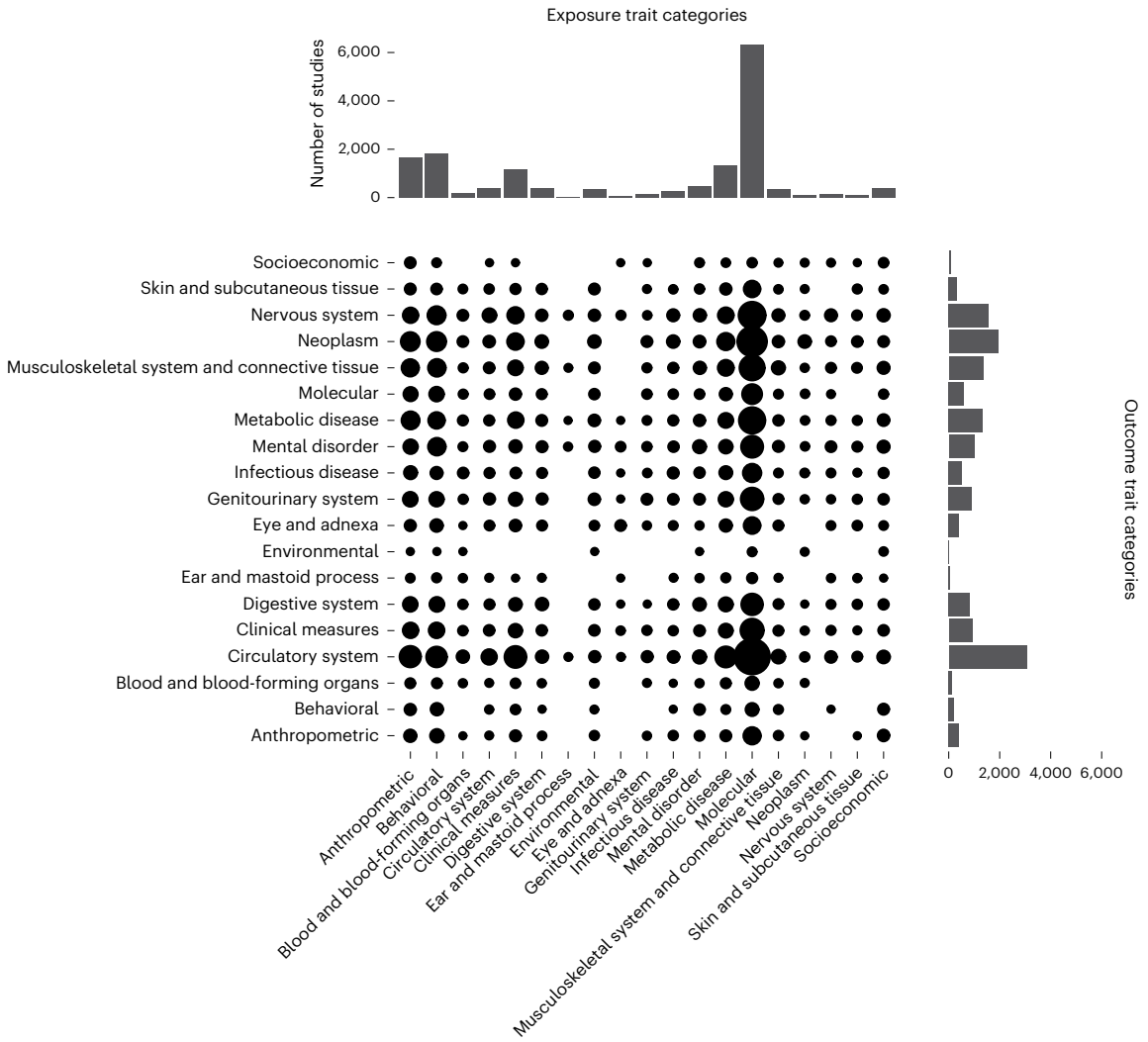


Fig. 2 | Categorizing exposures and outcomes used in MR studies. A bibliographic review of exposures and outcomes reported in MR studies 2003–2024 (n = 9,271 manuscripts; 24,298 exposure-outcome pairs). Point sizes represent the number of studies analyzing that exposure (x axis)–outcome (y axis) pair. Marginal bar plots represent the number of studies reporting on each exposure (top) and outcome (right). Studies have been conducted on most exposure–disease pairs, with particularly large numbers of studies on molecular exposures and circulatory outcomes.

methods still require additional assumptions about how the instrumental variable assumptions are violated, such as the structure of any horizontal pleiotropy. Although several Mendelian randomization methods allow estimation under different assumptions, they are not always clearly defined[25]. It can therefore be unclear how distinct the assumptions are across methods. This could lead to unwarranted confidence in the validity of a result if the assumptions are very similar and being applied to the same underlying data, especially across a class of estimators, for example, the simple, weighted and penalized-weighted-median approaches. Notably, these methods typically impose different assumptions on the same underlying data. As a result, the estimates are not orthogonal or independent; for example, the inverse-variance-weighted and weighted-median estimates using the same data will not be independent. It can also be challenging to assess heterogeneity across estimates or to determine whether inferences drawn from one estimator differ from those of another. Furthermore, there are limits to the range of assumptions that can be assessed and the settings in which these models can be applied. For example, pleiotropy is often assumed to affect different variants to different extents, with a majority or plurality assumed to be unaffected by any (where different estimators define plurality differently). This means analyses involving one or very few variants, for example, when restricting to the *cis* region of proteins, often cannot use many estimators and can struggle to assess pleiotropy.

By contrast, comparatively less attention has been given to empirical methods for assessing instrumental variable assumptions; however, these approaches offer a promising avenue for improving the robustness of Mendelian randomization studies. By incorporating evidence from a wider range of study designs regarding the plausibility of the instrumental variable assumptions in a particular analysis, the most appropriate Mendelian randomization methods can be assessed. Empirical approaches can include evaluating the association of variants used as instruments across a range of phenotypes to investigate whether these variants are likely to have horizontally pleiotropic effects. For example, the strength of associations between genetic variants used as instruments for an exposure and other phenotypes in published GWAS can help assess how many other traits the variants associate with, which may indicate horizontal pleiotropy.

Additional empirical approaches include negative-control outcomes or exposures, considering evidence from other studies to support the selection of variants as instruments or adding broader context from cell, animal and human studies[26]. Such approaches are currently not widely used in Mendelian randomization studies, with statistical approaches (described above) relied on to assess how sensitive an

https://doi.org/10.1038/s41588-026-02546-6

BOX 2

Assessing the instrumental variable assumptions

Instrumental variable 1: relevance
The genetic variants (*g*) must be associated with the exposure (*x*) (panel **a** of Box figure).

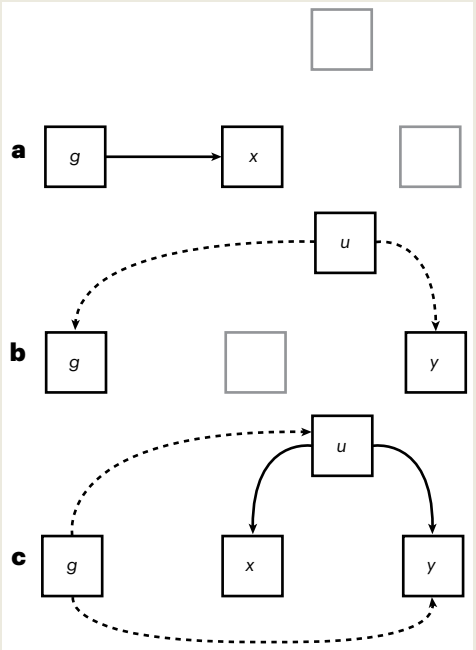
- This assumption can be assessed using *F*-statistics and conditional *F*-statistics for weak instruments in univariable and multivariable Mendelian randomization, respectively[68].
- Weak instrument bias and winner's curse can be addressed using statistical approaches[84–87].

Instrumental variable 2: independence
There must be no uncontrolled common cause (*u*) of the genetic variant and the outcome (*y*) or controlled common consequence (panel **b** of Box figure).

- Within-family studies can be used to assess independence by obtaining estimates of genetic associations that are less biased by population and familial processes[88,89].
- Phenome-wide approaches can help identify variants that capture sociodemographic confounding[23].
- Negative-control exposures and outcomes can be used to assess independence, such as geographically structured physical features that are unlikely to be causal[90].

Instrumental variable 3: exclusion restriction
The genetic variant's effect on the outcome must be mediated solely through the exposure (panel **c** of Box figure).

- The exclusion restriction can be assessed using estimators[67] that evaluate the consistency of effect estimates across multiple variants and different approaches to down-weighting or excluding outliers[25,91].
- Multivariable Mendelian randomization estimators can jointly estimate the effects of multiple exposures to assess potentially horizontal pleiotropic paths[66,92].
- Negative-control subpopulations can be used to assess variant-outcome associations across groups in which the variant-exposure effect is more or less attenuated, for example, in groups with minimal exposure level[93].
- Manual or automated inspection of instrument pleiotropy profiles using broad GWAS summary data can provide evidence of horizontal pleiotropy. Additionally, incorporating data from other study types that might indicate how likely an instrument is to be pleiotropic.



Box Fig. 2 | Directed acyclic graphs illustrating the three core instrumental variable assumptions—**a**, relevance; **b**, independence; and **c**, exclusion restriction.

effect estimate obtained is to the specific set of genetic variants used as instruments. The likelihood that those variants are valid instruments is often not assessed. Including evidence from a broader range of sources in assessing assumptions and combining it with existing statistical approaches could improve our understanding of the underlying biological mechanisms. This would strengthen our understanding of when the assumptions indeed hold and increase the reliability of the evidence from Mendelian randomization studies.

These approaches can help assess the plausibility of the assumptions required for a Mendelian randomization study. However, they do not yield a straightforward list of methods or approaches to apply to each study. Instead, the points raised above highlight the need for a careful case-by-case assessment, using all available information, to determine how likely the assumptions are to hold for each hypothesis. The most relevant approaches will depend on the nature of the application and the availability of data. For example, for more biological exposures—such as proteins (Box 4) or metabolites—there may be cell or animal studies that provide information on the relevance of particular genes[26]. For traits that are socially structured, such as educational attainment or exposure to air pollution (Box 5), examining the association of genetic variants with other traits or conducting negative-control analyses may help assess their validity.

One way to overcome instrument-outcome confounding—that is, violations of the second instrumental variable condition (independence; Box 2)—is to use family-based data, which can control for population structure, dynastic effects and assortative mating. However, there is relatively little familial data available[27], and we need more studies and individual- and summary-level data from large-scale consortia to support family-based analyses to assess the prevalence and impact of familial factors on estimates of the causal effects of exposures[28,29]. Family-based studies also provide insights into the mechanisms underlying the intergenerational transmission of traits. See Supplementary Note 3 for a discussion of how different analytical approaches can be used to examine the plausibility of the instrumental variable assumptions for anthropometric, gut microbiota and social and behavioral phenotypes.

Future challenges and opportunities

The tools for examining the assumptions required for reliable causal inference in Mendelian randomization are often developed within the context of the available data. Yet we are currently experiencing tremendous technological advances that are transforming the nature and scale of biobanks. In this context, we consider how the use of Mendelian randomization could be improved and highlight emerging challenges and opportunities arising from these changes in the data landscape.

BOX 3

What is bias in Mendelian randomization studies?

The bias of any estimate must be defined relative to the causal effect of interest[56,94]. Mendelian randomization estimates are based on a small difference in the liability of the exposure induced by the genetic variants across the entire life course. By contrast, interventions such as pharmacological treatments are more likely to induce large, temporally restricted perturbations on the level of the exposure. Hence, bias in Mendelian randomization studies can be defined as (1) the difference between the Mendelian randomization estimate obtained and the estimate if all assumptions are met (internal validity) or (2) the difference between the Mendelian randomization estimate and the estimand of interest (external validity)[95].

Internal validity can be compromised when the assumptions underlying instrumental variables are violated. For example, if the relevance condition does not hold because the genetic variants are only weakly associated with the exposure. In that case, the estimates can be biased toward the phenotypic association in individual-level data and towards the null in summary-data studies. Violations of the independence condition, for example, if population structure leads to differences in the frequencies of genetic variants associated with BMI and cardiovascular disease incidence, can also induce bias. Finally, if the exclusion restriction does not hold, for example, in case there are direct, horizontally pleiotropic paths from the genetic variants to the outcome, this can lead to biased estimates of the causal effect of the exposure on the outcome.

External validity depends on the extent to which a Mendelian randomization estimate is likely to reflect the real-world causal contrast of interest. By contrast, internal validity depends on ensuring—as far as is feasible—that the assumptions required for Mendelian randomization hold. For example, consider a Mendelian randomization study aiming to estimate the effects of bariatric surgery on cardiovascular disease risk using BMI as the exposure and cardiovascular disease as the outcome. It would not be possible to estimate the effect of interest using Mendelian randomization, as the estimates obtained reflect differences in BMI across the entire life course, from conception to outcome measurement. These Mendelian randomization estimates are therefore likely a biased estimate of the effect of a specific intervention at a specific point in time, such as bariatric surgery. The estimate is therefore not externally valid. However, the results obtained could provide an unbiased estimate of the causal lifetime effects of lifelong differences in BMI on the outcome of interest, if the instrumental variable assumptions hold. Therefore, if the estimand of interest is the lifetime effect of BMI on cardiovascular disease risk, the estimate would have external validity and by assessing the instrumental variable assumptions, we could determine how likely it is that the results are internally valid and provide a consistent estimate of that estimand.

Thus, assessing the potential bias of Mendelian randomization estimates requires stating both the real-world causal estimand of interest and assessing the plausibility of the assumptions required to estimate that effect.

Evidence triangulation

Integrating the diversity of approaches described above to reduce bias due to internal validity issues exploits the idea of evidence triangulation. The objective of triangulation is to integrate evidence with independent forms of bias to arrive at a more robust conclusion[30–38]. However, the extent to which biases across these methods are independent is not well understood, which may lead to overconfidence when multiple methods that rely on similar assumptions are applied. Further methodological work is needed to understand the independence of biasing mechanisms that underpin different Mendelian randomization methods and how those methods behave when their underlying assumptions do not hold. In parallel, more empirical research is required to assess the plausibility of the key assumptions for estimation across a wide range of settings.

Triangulation may also offer Mendelian randomization an opportunity to reduce bias in estimating the causal effect of interest, particularly when the mechanistic context of a study—such as an intended policy or health intervention—differs from the small, lifetime genetic perturbations leveraged in Mendelian randomization. This can be achieved by building triangulation frameworks that integrate results from independent experimental approaches, such as randomized controlled trials, cellular and animal models, other natural experiments and well-conducted observational studies[39]. Examples of such integration at the qualitative level are emerging and making a particular contribution to drug-target prioritization[20,40,41].

Diversity of samples

Historically, most molecular genetic data used in GWAS have come from populations of European ancestry living in Northern Europe or the USA[42]. This means that most Mendelian randomization studies are conducted on data from, or using genetic variants identified in, populations of European ancestry. If estimates vary across populations, this may imply that there are differing interactions or mediators that modify the effects of the exposure or that the third variable has a different distribution across populations. Understanding such relationships has two important implications. First, we gain a richer understanding of causal mechanisms that are likely relevant to all ancestral groups. Second, it may contribute to our understanding of health inequalities. However, most likely, health inequalities are driven by differential levels of disease risk factors rather than differential causal effects of those risk factors.

Data from diverse populations can also provide unique insights into disease risk and the mechanisms that mediate the effects of exposures on outcomes. Perhaps the most well-known example is the study of alcohol consumption and its effects on various outcomes, including blood pressure, which has been most compellingly made using variation in *ALDH2* (which encodes an enzyme that helps metabolize acetaldehyde, a harmful intermediate product of alcohol metabolism). A common missense variant in the *ALDH2* gene, encoding the p.Glu504Lys substitution, substantially impairs enzymatic function. This variant is common in East Asian populations, carried by approximately 20% of individuals, but has a very low frequency in European populations. Furthermore, for cultural reasons, some women in East Asian populations have very low alcohol consumption, which provides a credible negative-control population, in which, if the Mendelian randomization assumptions hold, the variants in *ALDH2* should not affect blood pressure[43]. Therefore, here genetic variation distinct to one population can provide evidence about the effects of an environmental exposure in all populations.

An important challenge with the increasing diversity of population-scale genetic association data is understanding the consistency of genetic associations across populations and the causal effects they suggest[44]. Numerous processes, in particular differential linkage disequilibrium, allele frequencies and sample sizes, can lead to spurious evidence for differential genetic or Mendelian randomization estimates across ancestral groups[44]. However, when appropriately accounted for, genetic effects appear largely consistent across ancestral groups, suggesting an opportunity to leverage well-powered European studies for instrument selection in other ancestral groups with smaller sample sizes. For such analyses to be rigorously performed, an extensive expansion of biobanks for underrepresented

BOX 4

Plausibility of instrumental variable assumptions for the proteome

Large GWAS have been conducted on the blood serum proteome[75,96]. Data from these GWAS have been used to assess the effects of the proteome on health outcomes. This approach is particularly interesting because the proteome represents the downstream product of genes and serves as the target of many pharmaceuticals. Studies are increasingly using these results in Mendelian randomization for drug-target associations. The instrumental variable assumptions can be quite plausible for specific proteomic mechanisms because several potential pitfalls are assessable and avoidable. Thus, Mendelian randomization can be potentially helpful for evaluating the causal influence of proteomic traits on various outcomes.

Relevance

Robust genetic associations for protein levels, known as protein quantitative trait loci (pQTLs), are being extensively cataloged, particularly for a majority of those proteins that are easily assayable[75,78]. However, pQTL effects may vary substantially across tissues, protein isoforms or cellular contexts. Detectable measurement errors may arise in several ways that are consequential for Mendelian randomization, for example, due to uncorrected batch effects[80], technical errors in the assay[97], cell count variation[80] or epitope effects[98]. A limitation to the relevance condition for summary data studies is that we cannot be sure that genetic variants associated with protein levels detected in one sample (for example, the UK Biobank) will also predict protein levels in other samples. However, the general consistency of pQTL effects across studies has been demonstrated[17].

Independence

Modest evidence from family-based designs suggests that population-based estimates of genetic effects on more widely available molecular measures, such as low-density lipoprotein (LDL) cholesterol, are unlikely to be biased by confounding processes, whereas estimates for proteins, for instance, C-reactive protein, are[21]. It is argued that biological traits such as protein levels are less likely to be affected by familial and social factors, including indirect effects of parents and assortative mating. However, assortative mating can induce spurious correlations between traits that are directly subject to assortment (for example, BMI) and other traits that are potentially less likely to be affected by assortative mating (for example, the proteome). Furthermore, most traits investigated show evidence of assortment[99].

Exclusion restriction

A key distinction is between *cis* (proximal to the functional gene) and *trans* (distal to the functional gene) variants. *Cis* variants typically have larger effects relative to other GWAS hits, explain a larger fraction of heritability and are likely more specific to the gene and protein of interest than *trans* instruments[80]. However, *cis* instruments are not guaranteed to be molecule-specific and empirical searches can be performed to examine this. Biological complexity can often be poorly captured as molecular phenotypes tend to reflect levels rather than function, which may lead to violations of the exclusion restriction.

BOX 5

Plausibility of instrumental variable assumptions for air pollution

A small but growing body of literature has applied Mendelian randomization to assess the effects of air pollution on health. The instrumental variable assumptions are unlikely to be plausible for assessing air pollution. It is unlikely that an individual's genetic variation directly influences their exposure to air pollution. A more plausible explanation is that these associations reflect the confounding of genetic associations by demographic factors. The instrumental variable assumptions provide a framework for assessing their plausibility and considering alternative explanations.

Relevance

GWAS in the UK Biobank has identified eight genetic variants strongly associated with air pollution (*F*-statistic for instrument strength > 10)[100]. These genetic variants are relevant because they are associated with the measured exposure. However, the measured exposure is not necessarily the exposure of interest and may therefore not capture an individual's total exposure to air pollution; for example, it may not reflect exposure at home or at work.

Independence

The most likely explanations for the associations of genetic variants and air pollution are demographic, familial and socioeconomic factors. The air pollution measures are based on regional summaries rather than individual-level data and therefore depend on an individual's place of residence. At the national or regional level, residual population stratification can lead to substantial genetic associations with geographic location. At the local level, residential location is strongly determined by socioeconomic status and more socioeconomically disadvantaged individuals typically experience higher levels of air pollution. Thus, genetic variants associated with the region (due to population stratification) or socioeconomic position (for example, educational attainment) could also be associated with air pollution. Family-based designs could help determine whether the reported associations between genetic variants and air pollution are attributable to individual-level or familial/demographic factors.

Exclusion

An important question is whether these genetic variants are likely to affect only the specific measure of air pollution of interest. The biological relevance of the variants for the measure can be assessed. Alternatively, the association between variants and air pollution should be assessed for potential mediating pathways, such as socioeconomic position or educational attainment.

communities is required. Furthermore, we need methods and larger samples to estimate gene-environment interactions or gene–gene interactions (epistasis)[45,46]. We also need to overcome significant challenges in analyzing diverse samples, such as differential statistical power[47], treating ancestry as a continuum[48] and differential linkage disequilibrium patterns[49].

Selection bias

Selection bias occurs when the samples available for analysis are nonrepresentative of the target population. Recent studies have

demonstrated that population-based samples, particularly large volunteer samples, are highly unrepresentative and may be affected by selection and collider bias[50–53]. This type of bias can affect Mendelian randomization studies, as with all epidemiological analyses of highly selected data[54], especially when the selection mechanism is related to the traits under analysis. Summary-data Mendelian randomization studies typically do not directly control for selection or collider bias, although sensitivity analyses and negative-control analyses can be performed. Individual-level studies can use statistical approaches, such as inverse propensity weights[55,56] or g-estimation[57], to correct for selection or collider bias, provided the selection mechanism is known. However, these methods are not frequently applied in genetic epidemiological studies[57,58]. This is partly for practical reasons, as many of these methods are not integrated into widely used statistical genetics packages. The effect of selection bias on GWAS associations and downstream Mendelian randomization studies remains unclear. Recent studies analyzing the UK Biobank found that inverse weighting by participation probability produces different effect estimates for the impact of education on BMI[58]. However, it remains challenging to determine whether these methods yield less biased estimates of the causal effects of exposures on outcomes because the underlying selection process in analytic samples is often unknown and imperfect modeling may exacerbate selection bias[58].

Another way to induce selection bias is by analyzing the progression of established disease (for example, to reduce the risk of recurrent stroke or slow the progression of cancer[59,60]), where selection based on cases may lead to factors associated with disease incidence becoming correlated with confounders of incidence and progression among cases[60]. Despite this difficulty, such studies are critically important, as the exposures that drive the progression and severity of an established disease may differ from those that lead to disease incidence and may be more relevant for treatment[61]. Statistical methods to address collider bias in studies of disease progression in genetic epidemiology have been proposed[62,63]. Recent GWAS have suggested that the genetic variation underlying disease incidence and progression may differ[61]. However, there have been relatively few applications of Mendelian randomization to progression phenotypes, largely because of a paucity of sufficiently large datasets.

Inclusion of multiple traits

Many hypotheses relate two or more exposures, such as the effects of educational attainment and cognition on Alzheimer's disease or the effects of multiple metabolites on age-related macular degeneration[64,65]. Widely used methods, such as multivariable Mendelian randomization[66,67], can effectively estimate the effect of a limited number of exposures on a single outcome. However, these estimators rely on sufficient genetic variation to estimate the effects of each exposure and so typically require hundreds of thousands of samples to obtain a precise estimate of the effects of correlated traits[68]. Alternative approaches, such as Bayesian model averaging, can determine which exposures to include; however, it is generally assumed that the true model includes relatively few exposures to make the analysis tractable and well powered[65]. Methods for reliably estimating the effects of multiple exposures on multiple outcomes in a single model have been developed[69]. However, new methodological approaches are still needed to identify specific causal pathways and accurately estimate the complex network of causal effects among phenotypes[65].

Effects across the lifecourse

Many exposures of interest will vary across an individual's lifetime and identifying whether the causal effect varies across the life course is vital to understanding how to intervene to improve outcomes. If there is sufficient variation in the effects of genetic variants across ages, it is possible to identify the effects of exposures at different ages[70–72]. One approach to estimating these effects is to exploit variation in the effects of genetic variants on the exposure (for example, BMI) by age (for example, childhood versus adulthood) to estimate the direct effects of the exposure in different ages (for example, the direct effects of childhood versus adulthood BMI on breast cancer)[70,73,74]. Approaches have also been proposed to estimate the effects of exposure trajectories. However, a limitation of these studies to date is that they rely on our limited understanding of how genetic variants associate with exposure throughout life. As a result, this approach has been applied so far only to limited exposures beyond BMI. Further GWAS of phenotypes stratified by age or exposure trajectories will help expand the application of these methods.

Molecular phenotypes

Heritability of molecular traits is typically concentrated in *cis* regions[75] (Box 4) and the routine use of *cis*-acting genetic variants for molecular exposures is intuitive, given that such variants are more likely to act directly on the molecular exposure of interest than *trans*-acting variants; therefore, they are less likely to act through pleiotropic pathways[76]. However, as molecular phenotypes are often highly correlated—both genetically and phenotypically—or can be grouped into classes that are difficult to separate, this approach has the limitation that there are often very few independent *cis* variants, which can make it challenging to use routine statistical tests for horizontal pleiotropy. Similarly, multivariable Mendelian randomization relies on a moderate level of genetic separation between the traits for reliable effect estimation; it cannot easily be applied to estimate the effects of highly correlated molecular traits[68]. One approach is to select a candidate trait from the set to consider as the reference exposure for a class of traits[77]. This strategy is likely to become increasingly important as the density of molecular phenotyping increases, for example, through single-cell assaying and the informational content of a single *cis* region becomes saturated. Further molecular quantitative trait locus studies using rare variants from sequencing studies may provide further insights[78].

Attempts to expand instruments to include *trans*-variants have been shown to induce substantial heterogeneity, consistent with the understanding that less-specific, smaller genetic effects on an exposure are more likely to capture heritable confounders[79]. Further work is also needed to integrate other genomic and biological information towards molecular instrument selection.

Several large studies have provided new evidence on the association between genetic variation and gene expression in specific tissues (for example, GTEx) and proteomics/metabolomics in large serum samples (for example, UK Biobank)[80]. These datasets have identified potential molecular mechanisms that mediate the effects of genetic variants, enabling Mendelian randomization studies of drug targets[81,82]. However, these studies are highly heterogeneous in their methods, for example, in the use of proxy exposures and tissues and there are no established standards or agreed-upon best practices for conducting drug-target studies. Furthermore, we have limited data from many of the tissues of interest; therefore, it is unclear how informative easily accessible tissues, such as serum, are regarding the effects of specific drugs across different tissues. Consequently, larger samples of tissue-specific proteomics and expression data could improve resolution, power and credibility of these studies.

Conclusions

Mendelian randomization is a powerful and efficient tool for testing causal relationships between exposures and outcomes in populations and has contributed to scientific understanding in several areas. However, as with all empirical research, the credibility of studies using these methods depends on the plausibility of their assumptions for each hypothesis. Further empirical studies are needed to assess the plausibility of the assumptions underlying different Mendelian randomization methods. Mendelian randomization studies also need to draw on a

wider range of evidence from other sources, in conjunction with new and existing statistical methods, to improve confidence in the conclusions drawn from the analysis.

References

1. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* **362**, k601 (2018).
2. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **32**, 1–22 (2003).
3. Sanderson, E. et al. Mendelian randomization. *Nat. Rev. Methods Primers* **2**, 6 (2022).
4. Walker, V. et al. Reading and conducting instrumental variable studies: guide, glossary, and checklist. *BMJ* **387**, e078093 (2024).
5. 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* **342**, d548 (2011).
6. Kingwell, B. A., Chapman, M. J., Kontush, A. & Miller, N. E. HDL-targeted therapies: progress, failures and future. *Nat. Rev. Drug. Discov.* **13**, 445–464 (2014).
7. Mountjoy, E. et al. Education and myopia: assessing the direction of causality by Mendelian randomisation. *BMJ* **361**, k2022 (2018).
8. Head, M. L., Holman, L., Lanfear, R., Kahn, A. T. & Jennions, M. D. The extent and consequences of p-hacking in science. *PLoS Biol.* **13**, e1002106 (2015).
9. O'Grady, C. Journals and publishers crack down on research from open health data sets. *Science* https://doi.org/10.1126/science.z6eukgo (2025).
10. Liverpool, L. AI intensifies fight against 'paper mills' that churn out fake research. *Nature* **618**, 222–223 (2023).
11. Candal-Pedreira, C. et al. Retracted papers originating from paper mills: cross sectional study. *BMJ* **379**, e071517 (2022).
12. Scancar, B., Byrne, J. A., Causeur, D. & Barnett, A. G. Machine learning based screening of potential paper mill publications in cancer research: methodological and cross sectional study. *BMJ* **392**, e087581 (2026).
13. Burgess, S., Woolf, B., Mason, A. M., Ala-Korpela, M. & Gill, D. Addressing the credibility crisis in Mendelian randomization. *BMC Med.* **22**, 374 (2024).
14. Davey Smith, G. & Ebrahim, S. Mendelian randomisation at 20 years: how can it avoid hubris, while achieving more? *Lancet Diabetes Endocrinol.* **12**, 14–17 (2024).
15. Chen, L., Guillot, A. & Schneider, C. V. Attention to the misuse of Mendelian randomisation in medical research. *eGastroenterology* **3**, e100187 (2025).
16. Baldwin, J. R., Pingault, J.-B., Schoeler, T., Sallis, H. M. & Munafò, M. R. Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur. J. Epidemiol.* **37**, 1–10 (2022).
17. Zheng, J. et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* **52**, 1122–1131 (2020).
18. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
19. Yavorska, O. O. & Burgess, S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* **46**, 1734–1739 (2017).
20. Minikel, E. V., Painter, J. L., Dong, C. C. & Nelson, M. R. Refining the impact of genetic evidence on clinical success. *Nature* **629**, 624–629 (2024).
21. Carter, A. R. & Anderson, E. L. Correct illustration of assumptions in Mendelian randomization. *Int. J. Epidemiol.* **53**, dyae050 (2024).
22. Tobias, J. H. et al. The proliferation of derivative and redundant studies in endocrinology due to the application of Mendelian randomisation and other methods to open databases. *Front. Endocrinol.(Lausanne)* **15**, 1400583 (2024).
23. Darrous, L., Hemani, G., Davey Smith, G. & Kutalik, Z. PheWAS-based clustering of Mendelian randomisation instruments reveals distinct mechanism-specific causal effects between obesity and educational attainment. *Nat. Commun.* **15**, 1420 (2024).
24. Davies, N. M. et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. *Int. J. Epidemiol.* **46**, 2067–2077 (2017).
25. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).
26. Howe, L. J. et al. Evaluating transportability of in vitro cellular models to in vivo human phenotypes using gene perturbation data. *Nat. Commun.* **17**, 513 (2026).
27. Davies, N. M. et al. The importance of family-based sampling for biobanks. *Nature* **634**, 795–803 (2024).
28. Assary, E. et al. Genetics of monozygotic twins reveals the impact of environmental sensitivity on psychiatric and neurodevelopmental phenotypes. *Nat. Hum. Behav.* **9**, 1683–1696 (2025).
29. Howe, L. J. et al. Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nat. Genet.* **54**, 581–592 (2022).
30. Davies, N. M., Dickson, M., Davey Smith, G., Windmeijer, F. & van den Berg, G. J. The causal effects of education on adult health, mortality and income: evidence from Mendelian randomization and the raising of the school leaving age. *Int. J. Epidemiol.* **52**, 1878–1886 (2023).
31. Hammerton, G. & Munafò, M. R. Causal inference with observational data: the need for triangulation of evidence. *Psychol. Med.* **51**, 563–578 (2021).
32. Treur, J. L., Lukas, E., Sallis, H. M. & Wootton, R. E. A guide for planning triangulation studies to investigate complex causal questions in behavioural and psychiatric research. *Epidemiol. Psychiatr. Sci.* **33**, e61 (2024).
33. Pearce, N., Vandenbroucke, J. P. & Lawlor, D. A. Causal inference in environmental epidemiology: old and new approaches. *Epidemiology* **30**, 311–316 (2019).
34. Bowden, J., Pilling, L. C., Türkmen, D., Kuo, C.-L. & Melzer, D. The Triangulation WIthin a STudy (TWIST) framework for causal inference within pharmacogenetic research. *PLoS Genet.* **17**, e1009783 (2021).
35. Dahabreh, I. J. & Bibbins-Domingo, K. Causal inference about the effects of interventions from observational studies in medical journals. *JAMA* **331**, 1845–1853 (2024).
36. Shi, X., Zhao, W., Chen, T., Yang, C. & Du, J. Evidence triangulator: using large language models to extract and synthesize causal evidence across study designs. *Nat. Commun.* **16**, 7355 (2025).
37. Van Hooijdonk, K. J. M. et al. Triangulated evidence provides no support for bidirectional causal pathways between diet/physical activity and depression/anxiety. *Psychol. Med.* **55**, e4 (2025).
38. Munafò, M. R., Higgins, J. P. T. & Davey Smith, G. Triangulating evidence through the inclusion of genetically informed designs. *Cold Spring Harb. Perspect. Med.* **11**, a040659 (2021).
39. Shapland, C. Y. et al. A quantitative approach to evidence triangulation: development of a framework to address rigour and relevance. Preprint at *medRxiv* https://doi.org/10.1101/2024.09.20.24314046 (2024).
40. Wang, D. et al. GDF15 promotes weight loss by enhancing energy expenditure in muscle. *Nature* **619**, 143–150 (2023).

41. Goudswaard, L. J. et al. Using trials of caloric restriction and bariatric surgery to explore the effects of body mass index on the circulating proteome. *Sci. Rep.* **13**, 21077 (2023).
42. Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* **2**, 9 (2019).
43. Chen, L., Davey Smith, G., Harbord, R. M. & Lewis, S. J. Alcohol intake and blood pressure: a systematic review implementing a Mendelian randomization approach. *PLoS Med.* **5**, e52 (2008).
44. Hemani, G. et al. Jointly modelling multiple ancestral populations using GWAS summary data improves causal inference. Preprint at *Res. Sq.* https://doi.org/10.21203/rs.3.rs-6091701/v1 (2025).
45. Hemani, G. et al. Phantom epistasis between unlinked loci. *Nature* **596**, E1–E3 (2021).
46. Spiller, W., Hartwig, F. P., Sanderson, E., Davey Smith, G. & Bowden, J. Interaction-based Mendelian randomization with measured and unmeasured gene-by-covariate interactions. *PLoS ONE* **17**, e0271933 (2022).
47. Meng, X. et al. Multi-ancestry genome-wide association study of major depression aids locus discovery, fine mapping, gene prioritization and causal inference. *Nat. Genet.* **56**, 222–233 (2024).
48. Saunders, G. R. B. et al. Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
49. Martin, A. R. et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
50. Fry, A. et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am. J. Epidemiol.* **186**, 1026–1034 (2017).
51. Mignogna, G. et al. Patterns of item nonresponse behaviour to survey questionnaires are systematic and associated with genetic loci. *Nat. Hum. Behav.* **7**, 1371–1387 (2023).
52. Benonisdottir, S. & Kong, A. Studying the genetics of participation using footprints left on the ascertained genotypes. *Nat. Genet.* **55**, 1413–1420 (2023).
53. Pirastu, N. et al. Genetic analyses identify widespread sex-differential participation bias. *Nat. Genet.* **53**, 663–671 (2021).
54. Gkatzionis, A. & Burgess, S. Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int. J. Epidemiol.* **48**, 691–701 (2019).
55. Robins, J. M., Hernán, M. A. & Brumback, B. Marginal structural models and causal inference in epidemiology. *Epidemiology* **11**, 550–560 (2000).
56. Hernan, M. A. & Robins, J. M. *Causal Inference: What If* (Taylor and Francis, 2024).
57. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A structural approach to selection bias. *Epidemiology* **15**, 615–625 (2004).
58. Schoeler, T. et al. Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nat. Hum. Behav.* **7**, 1216–1227 (2023).
59. Paternoster, L., Tilling, K. & Davey Smith, G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: conceptual and methodological challenges. *PLoS Genet.* **13**, e1006944 (2017).
60. Mitchell, R. E. et al. Strategies to investigate and mitigate collider bias in genetic and Mendelian randomisation studies of disease progression. *PLoS Genet.* **19**, e1010596 (2023).
61. Yang, Z. et al. Limited overlap between genetic effects on disease susceptibility and disease survival. *Nat. Genet.* **57**, 2418–2426 (2025).
62. Dudbridge, F. et al. Adjustment for index event bias in genome-wide association studies of subsequent events. *Nat. Commun.* **10**, 1561 (2019).
63. Dib, M.-J. et al. Proteomic associations of adverse outcomes in human heart failure. *J. Am. Heart Assoc.* **13**, e031154 (2024).
64. Anderson, E. L. et al. Education, intelligence and Alzheimer's disease: evidence from a multivariable two-sample Mendelian randomization study. *Int. J. Epidemiol.* **49**, 1163–1172 (2020).
65. Zuber, V., Colijn, J. M., Klaver, C. & Burgess, S. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. *Nat. Commun.* **11**, 29 (2020).
66. Sanderson, E., Davey Smith, G., Windmeijer, F. & Bowden, J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int. J. Epidemiol.* **48**, 713–727 (2019).
67. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**, 658–665 (2013).
68. Sanderson, E., Spiller, W. & Bowden, J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat. Med.* **40**, 5434–5452 (2021).
69. Zuber, V. et al. Multi-response Mendelian randomization: identification of shared and distinct exposures for multimorbidity and multiple related disease outcomes. *Am. J. Hum. Genet.* **110**, 1177–1199 (2023).
70. Sanderson, E., Richardson, T. G., Morris, T. T., Tilling, K. & Davey Smith, G. Estimation of causal effects of a time-varying exposure at multiple time points through multivariable mendelian randomization. *PLoS Genet.* **18**, e1010290 (2022).
71. Power, G. M. et al. Methodological approaches, challenges, and opportunities in the application of Mendelian randomisation to lifecourse epidemiology: a systematic literature review. *Eur. J. Epidemiol.* **39**, 501–520 (2024).
72. Shi, J. et al. Mendelian randomization with repeated measures of a time-varying exposure: an application of structural mean models. *Epidemiology* **33**, 84–94 (2022).
73. Vabistsevits, M. et al. Mammographic density mediates the protective effect of early-life body size on breast cancer risk. *Nat. Commun.* **15**, 4021 (2024).
74. Richardson, T. G., Sanderson, E., Elsworth, B., Tilling, K. & Davey Smith, G. Use of genetic variation to separate the effects of early and later life adiposity on disease risk: Mendelian randomisation study. *BMJ* **369**, m1203 (2020).
75. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
76. Neumeyer, S., Hemani, G. & Zeggini, E. Strengthening causal inference for complex disease using molecular quantitative trait loci. *Trends Mol. Med.* **26**, 232–241 (2020).
77. Bull, C. J. et al. Adiposity, metabolites, and colorectal cancer risk: Mendelian randomization study. *BMC Med.* **18**, 396 (2020).
78. Dhindsa, R. S. et al. Rare variant associations with plasma protein levels in the UK Biobank. *Nature* **622**, 339–347 (2023).
79. Min, J. L. et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat. Genet.* **53**, 1311–1321 (2021).
80. Sun, B. B. et al. Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
81. Schmidt, A. F. et al. Genetic drug target validation using Mendelian randomisation. *Nat. Commun.* **11**, 3255 (2020).
82. Trajanoska, K. et al. From target discovery to clinical drug development with human genetics. *Nature* **620**, 737–745 (2023).
83. Pierce, B. L. & Burgess, S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* **178**, 1177–1184 (2013).
84. Sadreev, I. I. et al. Navigating sample overlap, winner's curse and weak instrument bias in Mendelian randomization studies using the UK Biobank. Preprint at *medRxiv* https://doi.org/10.1101/2021.06.28.21259622 (2021).

85. Forde, A., Hemani, G. & Ferguson, J. Simulated sample splitting approach to address biases due to instrument selection and participant overlap in two-sample Mendelian randomization studies. Preprint at *bioRxiv* https://doi.org/10.1101/2025.11.04.686460 (2025).
86. Zhao, Q., Wang, J., Hemani, G., Bowden, J. & Small, D. S. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann. Stat.* **48**, 1742–1769 (2020).
87. Mounier, N. & Kutalik, Z. Bias correction for inverse variance weighting Mendelian randomization. *Genet. Epidemiol.* **47**, 314–331 (2023).
88. Davies, N. M. et al. Within family Mendelian randomization studies. *Hum. Mol. Genet.* **28**, R170–R179 (2019).
89. Brumpton, B. et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **11**, 3519 (2020).
90. Sanderson, E., Richardson, T. G., Hemani, G. & Davey Smith, G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. *Int. J. Epidemiol.* **50**, 1350–1361 (2021).
91. Zhang, W., Su, C.-Y., Yoshiji, S. & Lu, T. MR Corge: sensitivity analysis of Mendelian randomization based on the core gene hypothesis for polygenic exposures. *Bioinformatics* **40**, btae666 (2024).
92. Burgess, S. & Thompson, S. G. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am. J. Epidemiol.* **181**, 251–260 (2015).
93. Spiller, W., Slichter, D., Bowden, J. & Davey Smith, G. Detecting and correcting for bias in Mendelian randomization analyses using gene-by-environment interactions. *Int. J. Epidemiol.* **48**, 702–712 (2019).
94. Pearl, J. *Causality: Models, Reasoning, and Inference* (Cambridge Univ. Press, 2022).
95. Angrist, J. D., Imbens, G. W. & Rubin, D. B. Identification of causal effects using instrumental variables. *J. Am. Stat. Assoc.* **91**, 444–455 (1996).
96. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**, 1712–1721 (2021).
97. Haslam, D. E. et al. Stability and reproducibility of proteomic profiles in epidemiological studies: comparing the Olink and SOMAscan platforms. *Proteomics* **22**, e2100170 (2022).
98. Kuliesius, J. et al. Epitope effect prevalence in affinity-based pQTL studies. Preprint at *bioRxiv* https://doi.org/10.1101/2025.06.20.660695 (2025).
99. Border, R. et al. Cross-trait assortative mating is widespread and inflates genetic correlation estimates. *Science* **378**, 754–761 (2022).
100. Fan, H. et al. Linking ambient air pollution to mental health: evidence based on the two-sample Mendelian randomization and colocalization study. *Transl. Psychiatry* **14**, 489 (2024).

Acknowledgements

E.S. is supported by the Medical Research Council (MRC; grant UKRI077). N.M.D. is supported by the Research Council of Norway (grant 295989), MRC (grants MR/V002147/1 and UKRI1510), the National Institute for Mental Health (grant MH130448) and the Division of Psychiatry. G.H. is supported by the Wellcome Trust and the Royal Society (grant 208806/Z/17/Z) and by the MRC and the University of Bristol through the MRC Integrative Epidemiology Unit (grant MC_UU_00032/1). V.W. is supported by the UK Medical Research Council Integrative Epidemiology Unit (grant MC_UU_00032/3). M.G.L. is supported by the Doris Duke Foundation (award 2023–0224) and the US Department of Veterans Affairs Biomedical Research and Development (award IK2-BX006551). S.M.D. is supported by the US National Heart, Lung, and Blood Institute (grants R01HL181587, R01HL166991, HL141232) and the US Department of Veterans Affairs Biomedical Research and Development (award I01BX006159). This study does not represent the views of the Department of Veterans Affairs or the US Government.

Author contributions

G.H. conceived and performed the analysis. All authors contributed to writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-026-02546-6.

Correspondence should be addressed to Eleanor Sanderson, Scott M. Damrauer, Gibran Hemani or Neil M. Davies.

Peer review information *Nature Genetics* thanks Stephen Burgess and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.