ESSAY

Mendelian randomisation for mediation analysis: current methods and challenges for implementation

Alice R. Carter[1,2] · Eleanor Sanderson[1,2] · Gemma Hammerton[1,2,3] · Rebecca C. Richmond[1,2] ·
George Davey Smith[1,2,4] · Jon Heron[1,2,3] · Amy E. Taylor[1,2,4] · Neil M. Davies[1,2,5] · Laura D. Howe[1,2]

Abstract
Mediation analysis seeks to explain the pathway(s) through which an exposure affects an outcome. Traditional, non-instrumental variable methods for mediation analysis experience a number of methodological difficulties, including bias due to confounding between an exposure, mediator and outcome and measurement error. Mendelian randomisation (MR) can be used to improve causal inference for mediation analysis. We describe two approaches that can be used for estimating mediation analysis with MR: multivariable MR (MVMR) and two-step MR. We outline the approaches and provide code to demonstrate how they can be used in mediation analysis. We review issues that can affect analyses, including confounding, measurement error, weak instrument bias, interactions between exposures and mediators and analysis of multiple mediators. Description of the methods is supplemented by simulated and real data examples. Although MR relies on large sample sizes and strong assumptions, such as having strong instruments and no horizontally pleiotropic pathways, our simulations demonstrate that these methods are unaffected by confounders of the exposure or mediator and the outcome and non-differential measurement error of the exposure or mediator. Both MVMR and two-step MR can be implemented in both individual-level MR and summary data MR. MR mediation methods require different assumptions to be made, compared with non-instrumental variable mediation methods. Where these assumptions are more plausible, MR can be used to improve causal inference in mediation analysis.

Keywords Mendelian randomisation · Mediation analysis · Multivariable Mendelian randomisation · Two-step Mendelian randomisation

Introduction

Mediation analysis can improve aetiological understanding and identify intermediate variables as potential intervention targets, when intervening on an exposure is not feasible. However, in order to make causal inference, non-instrumental variable (IV) regression based mediation analysis requires strong assumptions. Mendelian randomisation (MR) is an alternative causal inference approach using genetic variants as an IV for a phenotype [1]. In this paper we compare non-IV regression-based methods for mediation analysis with MR methods for mediation analysis, and describe the assumptions required for MR mediation methods to make valid causal inference.

✉ Alice R. Carter
  alice.carter@bristol.ac.uk

1 MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

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

3 Centre for Academic Mental Health, University of Bristol, Bristol, UK

4 National Institute for Health Research Biomedical Research Centre At the University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, UK

5 K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Mediation analysis

Methods for mediation analysis emerged in the early twentieth-century, although often not described as such at the time, with formal methods developed by Baron and Kenny in the 1980s [2, 3]. More recently, a large amount of research has built on and improved mediation methods for better causal inference [4].

Three parameters are typically estimated in a traditional mediation analysis i) the total effect (the effect of the exposure on the outcome through all potential pathways) ii) the direct effect, either controlled or natural (the remaining effect of the exposure on the outcome that acts through pathways other than the specified mediator or set of mediators) and iii) the natural indirect effect (the path from exposure to outcome that acts through the mediator(s)). In situations where the total effect, direct effect and indirect effect all act in the same direction, an estimate of the "proportion mediated" (i.e., proportion of the total effect explained by the mediator) can be calculated. Two common approaches to estimate the indirect effect are; the product of coefficients method and the difference in coefficients method [5] (see Fig. 1a).

Traditional non-IV mediation methods, such as Baron and Kenny methods, rely on several strong, untestable assumptions including, (i) no unmeasured confounding between the exposure, mediator and outcome (ii) no exposure-caused confounders of the mediator and outcome (intermediate confounders, see Fig. 2a) and (iii) no exposure-mediator interaction [4, 6, 7]. Furthermore, measurement error in either the exposure or mediator can introduce bias [8].

Baron and Kenny methods were introduced to estimate mediation with a continuous exposure, mediator and outcome, although they are also now often applied to binary variables. In the presence of a continuous or rare binary

Fig. 1 The decomposed effects in **a** non-IV regression-based mediation analysis where **c** represents the total effect, **c'** represents the direct effect and the indirect effect can be calculated by subtracting **c'** from **c** (difference method) or multiplying A times B (product of coefficients method) **b** multivariable Mendelian randomisation, using a combined genetic instrument for both the exposure and mediator of interest, to estimate the direct effect **c'** of the exposure and **c** two-step Mendelian randomisation, where the effect of the exposure on the mediator (A) and mediator on the outcome **b** are estimated separately, using separate genetic instrumental variables for both the exposure and mediator. These estimates are then multiplied together to estimate the indirect effect of the mediator (A*B)
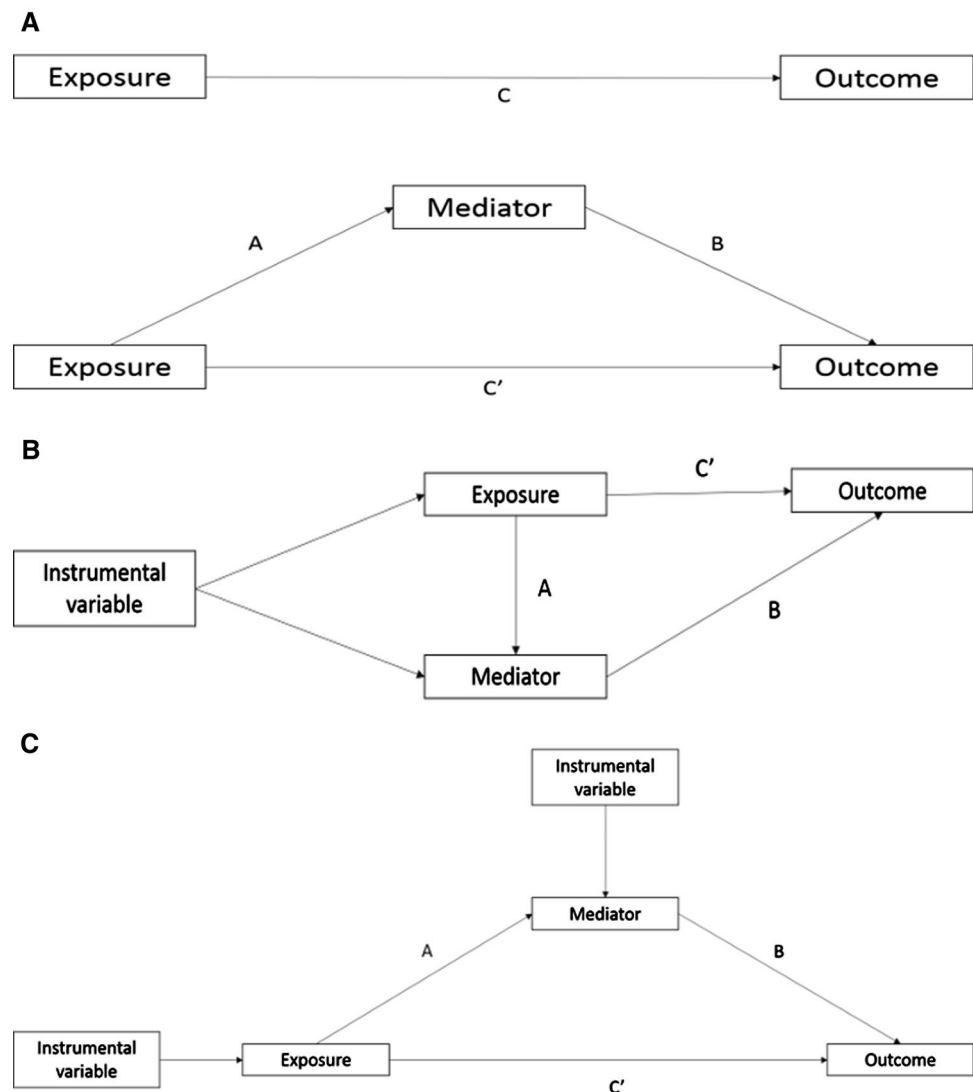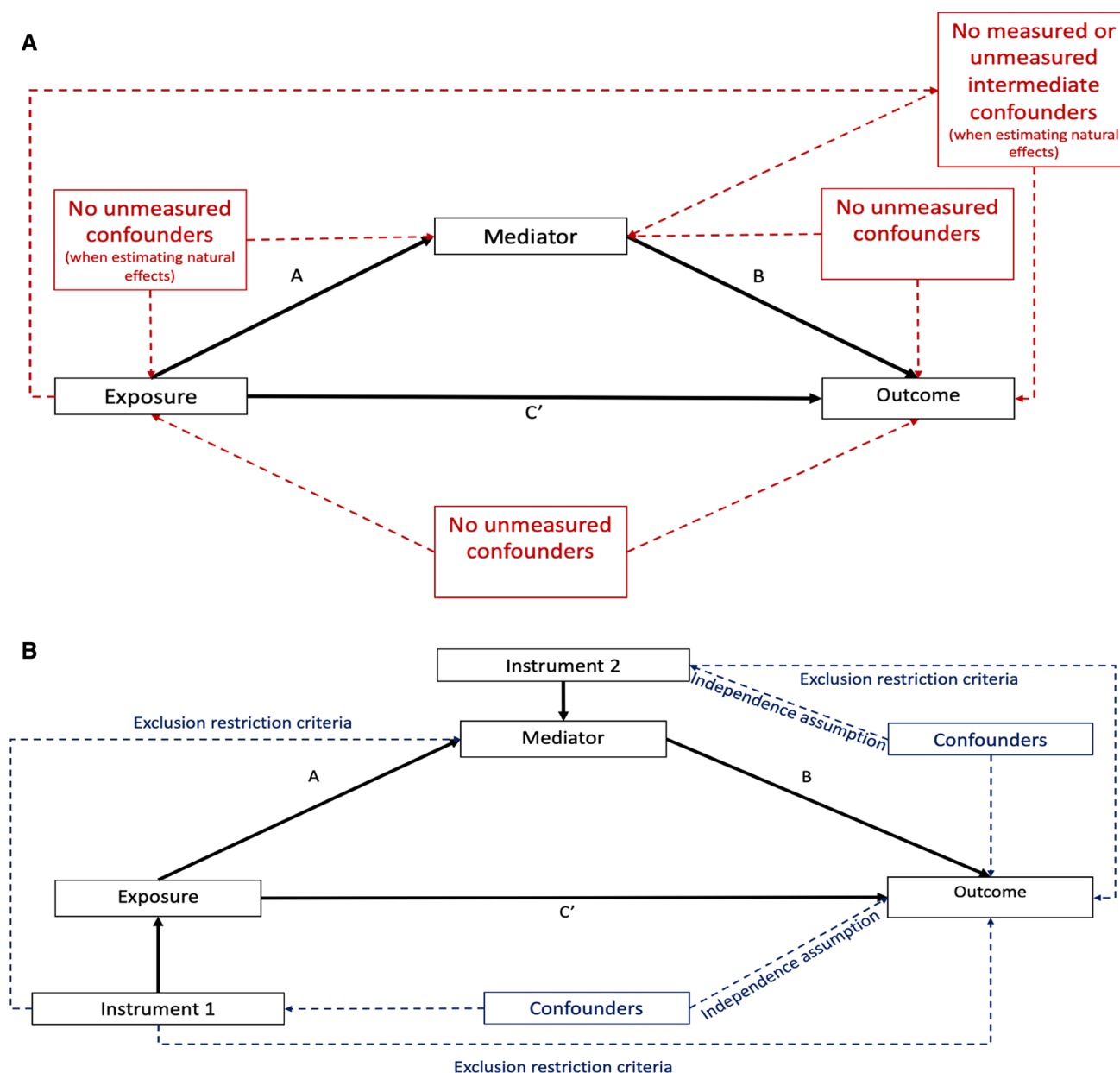
Fig. 2 Schematic diagram illustrating the causal assumptions (dashed lines) in **a** non-IV regression-based mediation methods and **b** Mendelian randomisation mediation analysis with the measured associations in solid black lines. Additional assumptions: in non-IV mediation there is no measurement error in the exposure or mediator; in Mendelian randomisation mediation there is no exposure-mediator interaction. In Mendelian randomisation, the exclusion restriction criteria mean there are no alternative pathways from the instrument to the outcome other than via the exposure (or mediator) of interest

outcome the estimates from the difference in coefficients and the product of coefficients method should coincide [4, 9].

Counterfactual reasoning has been used to develop confounder adjusted methods that can address some of the previously described strong assumptions in non-IV mediation methods [10–14]. The assumptions made by these counterfactual approaches, mean mediation can be estimated in the presence of exposure-mediator interactions and account for measured intermediate confounders. Additionally, these more flexible approaches can allow for binary mediators and outcomes. However, these methods remain biased in the presence of unmeasured confounding, measurement error in the exposure or mediator, or in a mis-specified model with reverse causality [4, 15]. Here, the estimated direct effect is described as being a "controlled direct effect" if the value of the mediator is controlled at a certain value for all individuals in the population, or a "natural direct effect", when the value of the mediator is allowed to take the value

for each person that it would have taken naturally had they been unexposed, in a counterfactual scenario. The "natural indirect effect" represents the average change in an outcome if the value of the exposure was fixed, but the value of the mediator changes from its natural value when exposed to its natural value when unexposed. If there is no interaction between the exposure and mediator, the estimate of the natural direct effect is equivalent to the controlled direct effect, and indeed would align with estimates from Baron and Kenny approaches to mediation [4, 9, 16].

Mendelian randomisation

In Mendelian randomisation (MR) randomly allocated genetic variants are used as instrumental variables (IV) for a phenotype [1, 17, 18]. Given the random allocation of genetic variants at conception, MR estimates are not biased by confounding between an exposure and outcome, reverse causation and measurement error [17]. Three core assumptions are required for a genetic variant to be a valid IV, these are (i) the genetic variants are associated with the exposure (the relevance assumption) (ii) genetic instruments are exchangeable with the outcome, across levels of the instrument (the independence assumption) and (iii) the genetic variants do not affect the outcome via any variable other than the exposure (the exclusion restriction criteria) (Online Resource 1: sFig. 1) [1]. Indeed, in the case of the independence assumption and exclusion restriction criteria, these are strong and unverifiable assumptions.

Rationale for using Mendelian randomisation in mediation analysis

MR can be used to overcome some of the previously described strong assumptions required for causal inference in mediation analysis. For example, estimates are not biased due to unmeasured confounding between an exposure, mediator or outcome.

In mediation terms, univariable MR estimates the total effect of the exposure on the outcome. Two differing MR approaches can then be used which broadly mirror traditional non-IV regression-based approaches to mediation to decompose the direct and indirect effects: multivariable MR (MVMR) [19, 20] and two-step MR [21–23].

In MVMR the controlled direct effect of the exposure on the outcome, controlling for the mediator, is estimated [19, 23]. The genetic instrument for both the primary exposure and the second exposure (mediator) are included as instruments in the analysis (Fig. 1b) [24, 25]. The indirect effect can then be estimated by subtracting the direct effect from the total effect (akin to the difference in coefficients method). MVMR assumes no interaction between the exposure and the mediator; therefore, the controlled direct effect estimated

is equivalent to the natural direct effect where this assumption holds true. As such, we refer to this as the direct effect, without further distinction, throughout this manuscript.

Two-step MR (also known as network MR) is akin to the product of coefficient methods. Two MR estimates are calculated i) the causal effect of the exposure on the mediator and ii) the causal effect of the mediator on the outcome (Fig. 1c) [21, 23, 26]. These two estimates can then be multiplied together to estimate the indirect effect. Two-step MR also assumes no interaction between the exposure and the mediator.

These MR methods are increasingly being used in mediation analysis [19, 27–30]. In this paper, we demonstrate how MVMR, and two-step MR can be used to estimate the direct effect, indirect effect and the proportion mediated, and which assumptions are required for the resulting estimates to be unbiased [23–25]. We provide guidance about how to carry out each method, with code provided, and illustrate each method using both simulated and real data (see Online Resource 2), applied to an individual level MR analysis.

Methods

Simulation study

We simulated data under the model illustrated in Fig. 1 with continuous, rare binary (5% prevalence) and common binary (25% prevalence) outcomes. We varied the total effect of our exposure and proportion mediated and obtained results using non-IV regression based mediation methods using both the difference and product of coefficients approaches, and MR methods using both MVMR and two-step MR. Additionally, we simulated results where the total effect of the exposure on the outcome is small, and where each of the exposure and mediator were subject to non-differential measurement error. Finally, we simulated how MR methods can estimate mediation in the presence of multiple mediators, these simulations are illustrated in Online Resource 1: sFig. 2. The full range of scenarios simulated are presented in sTable 1. Simulation analyses were carried out using R version 3.5.1 and the corresponding code for the simulation studies can be found at https://github.com/eleanorsanderson/MediationMR.

Applied example

Using data from UK Biobank (N = 184 778), we investigate the role of body mass index (BMI) and low-density lipoprotein cholesterol (LDL-C) in mediating the associations between education and systolic blood pressure, cardiovascular disease (CVD) and hypertension (continuous, rare binary and common binary outcomes, respectively). The effects on binary outcomes (hypertension and incident

CVD) were estimated on risk difference, log odds ratio, and odds ratio scales. Applied analyses were performed using Stata version 15 (StataCorp LP, Texas) and corresponding code is available at https://github.com/alicerosecarter/MediationMR. The full worked through example is available in Online Resource 2.

Statistical analysis

The following approaches were applied to both applied analyses and simulated data. Equations describing each of these analyses are given in Online Resource 1.

Difference in coefficients method

Each outcome was regressed on the exposure adjusting for the mediator to estimate the direct effect of the exposure. The direct effect was subtracted from the total effect, estimated using multivariable regression, to estimate the indirect effect. In all simulation scenarios the standard deviation of the regression coefficients was calculated across repeats to evaluate precision.

Product of coefficients method

Two regression models were estimated. Firstly, the mediator was regressed on the exposure. Secondly, the outcome was regressed on the mediator, adjusting for the exposure. These two estimates were multiplied together to estimate the indirect effect.

Multivariable Mendelian randomisation

Using MVMR to estimate the direct effect, in the first stage regression, the effect of the instrument for the exposure and the polygenic score for the mediator are used to predict each exposure respectively. In the second stage regression, the outcome was regressed on the predicted values of each exposure. The direct effect was then subtracted from the total effect, estimated using two-stage least squares regression, to estimate the indirect effect.

Two-step Mendelian randomisation

A univariable MR model was carried out to estimate the effect of the exposure on the mediator. A second model estimating the effect of the mediator on each outcome was carried out using MVMR. Both the genetic variants for the mediator and the exposure were included in the first and second stage regressions in MVMR. Previous approaches in the literature have not used MVMR for this second step [21, 23] and propose carrying out a univariable MR of the effect of the mediator on the outcome. However, using MVMR ensures any effect of the mediator on the outcome is independent of the exposure. Additionally, this method provides an estimate of the direct effect of the exposure on the outcome. The two regression estimates from the second stage regression are multiplied together to estimate the indirect effect.

Multiple mediators

In non-IV mediation analyses, to estimate the direct effect attributable to multiple mediators, the outcome was regressed on the exposure, controlling for all mediators, using multivariable regression. Here, the coefficient for the exposure reflects the direct effect [31]. This direct effect was then subtracted from the total effect to estimate the indirect effect. Secondly, the product of coefficients method was used to estimate the indirect effect of each mediator individually. The combined effect of all mediators was then estimated by summing together each individual effect.

In MR analyses, the direct effect attributable to multiple mediators was assessed using MVMR, controlling for all mediators. This direct effect was then subtracted from the total effect to estimate the combined indirect effect. Secondly two-step MR was used, as previously described, considering each mediator individually and summing the effects together to obtain the indirect effect of all mediators combined.

Proportion mediated

The proportion mediated is calculated by dividing the indirect effect by the total effect. In individual-level MR, the confidence intervals can be estimated via bootstrapping.

Testing the assumptions of mediation analysis

In this analysis, we have simulated a number of scenarios where non-IV regression based methods or MR methods for mediation analysis may provide biased answers. In this section we outline these results and any implications for analyses.

Unmeasured confounding between the exposure, mediator and outcome

Many of the key causal assumptions in non-IV mediation analysis relate to assumptions of no unmeasured confounding between all of the exposure, mediator and outcome, including where confounders of the mediator and outcome are descendants of the exposure (intermediate confounding). Controlling for confounders in multivariable regression

analyses often leads to residual confounding because it is generally impossible to measure all confounders, and frequently those that are measured are measured with error.

Indeed, in our simulations where residual covariance was simulated to reflect confounding, both the non-IV difference method and non-IV product of coefficients method were equally biased (Fig. 3 and Online Resource 1: sTables 2). Where no confounding was simulated in the case of no true total effect, estimates from non-IV approaches were free from bias (Online Resource 1: sTable 3). In simulations both with and without residual covariance to reflect confounding, MVMR and two-step MR estimated the direct effect, indirect effect and proportion mediated with no bias (Fig. 2 and Online Resource 1: sTables 4 and 5).

Collider bias can be introduced by adjusting for the mediator in the presence of un- or mis-measured mediator-outcome confounders, where a backdoor path opens up between the exposure and the confounder (Online Resource 1: sFig. 3) [6, 32, 33]. Given that MR estimates are unbiased by unmeasured confounding of the exposure-outcome and mediator-outcome relationships [1, 17], this means that within MR analyses, adjusting for the mediator does not result in collider bias.

Analysis of binary outcomes

Mediation analysis of binary outcome is challenging because of the non-collapsibility of odds ratio. This means the association between an exposure and outcome would not be constant on the odds-ratio scale by strata of categorical covariate [34, 35]. In mediation analysis, including the mediator in the model estimating the direct effect, means the model is no longer comparable with that for the total effect.

The mediation literature indicates that to estimate the direct and indirect effects of a binary outcome, the outcome must be rare (less than 10% prevalence), so the odds ratio approximates the risk ratio, and the product of coefficients method should be used [9]. In the presence of a common binary outcome, estimates from the product of coefficients method and difference method are unlikely to align (and indeed the literature suggests both are likely biased) [4].

In our simulations, both the difference in coefficients and the product of coefficients non-IV methods, with common and rare binary outcomes on a linear relative scale were biased as expected by unmeasured confounding (Fig. 3 and Online Resource 1: sTables 6–9). The size of bias was similar across the two non-IV methods. In simulated MR
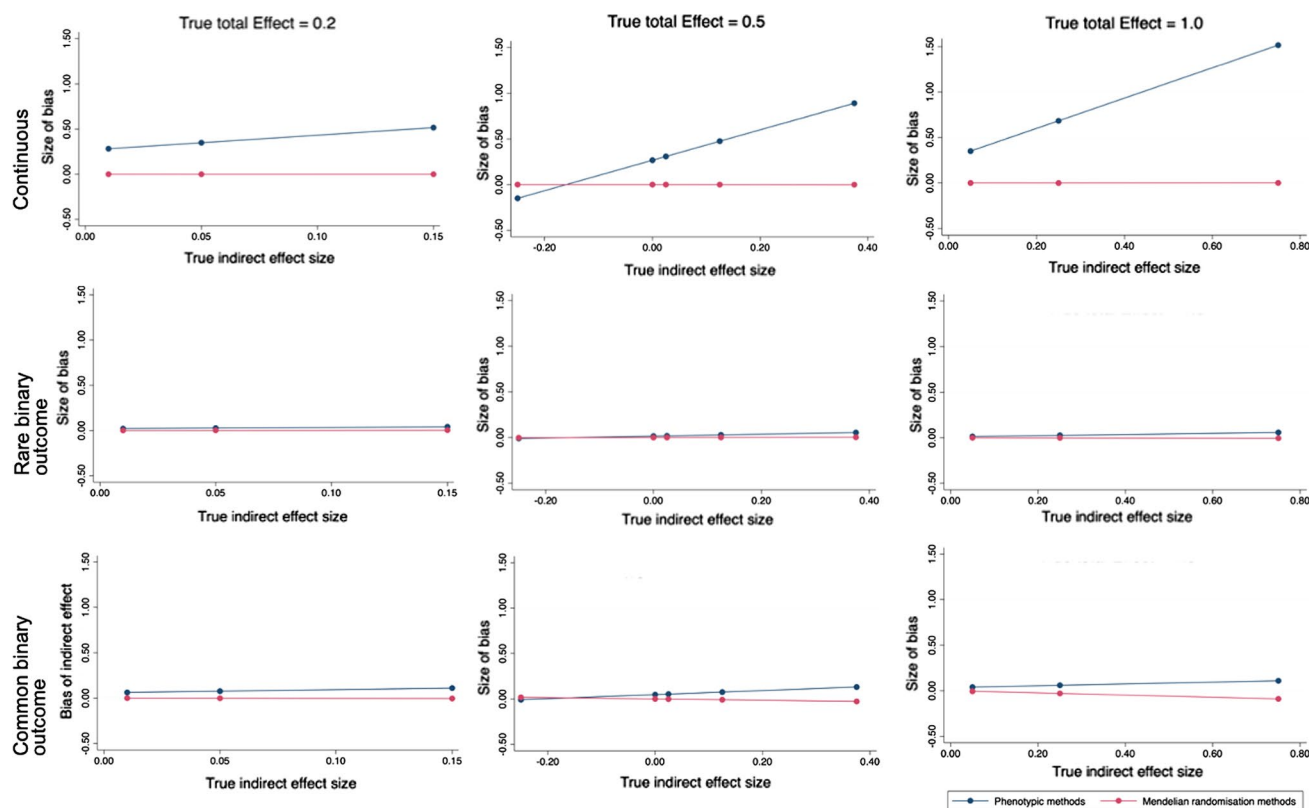

Fig. 3 Size of absolute bias for the indirect effect of an exposure on a continuous outcome, rare binary outcome and common binary outcome through a continuous mediator, for a range of fixed true total effect sizes (0.2, 0.5 and 1.0) and range of true indirect effect sizes using non-IV regression based mediation methods or Mendelian randomisation, on the relative scale (simulated N = 5000). In all scenarios, unmeasured confounding is simulated

scenarios with common and rare binary outcomes on a linear relative scale, estimated effects were concordant between MVMR and two-step MR, with little to no bias (Fig. 3 and Online Resource 1: sTables 10–13).

In the scenarios simulated, there was some bias when analysing binary outcomes on the log odds ratio scale using both MVMR and two-step MR, for both common and rare binary outcomes (Online Resource 1: sTables 14 and 15). This bias was small and typically would not alter conclusions made, although typically the size of absolute bias increased as the size of the true proportion mediated increased. However, the exact bias from non-collapsibility will be unique to each scenario, including depending on the strength of the mediators. Analyses in individual level MR can be conducted on the risk difference scale, which reduces bias due to non-collapsibility.

In simulation scenarios explored, neither MVMR nor two-step MR were able to estimate the mediated effects without bias when using the odds ratio scale (Online Resource 1: sTables 16 and 17).

Measurement error in the exposure or mediator

Our results show that in non-IV approaches, with a continuous exposure and mediator, non-differential measurement error in the mediator leads to an underestimate of the mediated effect. This is consistent with previous methodological and applied work [8]. Where non-differential measurement error was simulated in the exposure, the mediated effect was over estimated (Online Resource 1: sTable 18).

In MR simulations, both MVMR and two-step MR estimated the mediated effects with little bias when non-differential measurement error was simulated either in the exposure or the mediator (Online Resource 1: sTable 19). This is consistent with the previous literature demonstrating that MR estimates are less prone to bias by measurement error than conventional non-IV analyses [1, 17].

Weak instrument bias

In order to obtain valid causal inference for mediation, all standard MR assumptions must be met. This includes having strong instruments, typically determined through an F-statistic or conditional F-statistic of greater than 10. The conditional instrument strength in multivariable MR can be tested using the Sanderson-Windmeijer F-statistic [36]. When the instruments in the simulation were weakly associated with the exposure, both MVMR and two-step MR estimates of the indirect effect and proportion mediated were biased. The size of bias was greatest for a common binary outcome. When weak instruments were simulated for the mediator, estimates of the indirect effect and proportion mediated from both MVMR and two-step MR were biased

(Online Resource 1: sFig. 4 and sTable 19). Bias due to weak instruments have been discussed extensively in the literature [37–39], and methods are now available for testing for weak instrument bias in MVMR [40].

Pleiotropy

One of the core MR assumptions is that the genetic variants used as instruments do not affect the outcome other than via the exposure of interest, known as pleiotropy. Bias can be introduced to MR mediation analyses if any of the associations between the exposure and outcome, exposure and mediator or mediator and outcome are pleiotropic. In simulations with pleiotropy in the association between the exposure and outcome, estimates of the total effect and direct effect are biased (Online resource 1: sTable 20). In this scenario, no pleiotropy is present for the association between the exposure and mediator or mediator and outcome, therefore no bias is present for the indirect (mediated) effect. In simulations with pleiotropy in the association between the mediator and the outcome estimates of the direct effect and indirect effect are biased (Online resource 1: sTable 21).

Bias due to pleiotropy has been discussed extensively in the literature [41, 42]. Methods are available for testing for and assessing for pleiotropy, including in MVMR [40, 43–45]. Indeed, MVMR was developed as a method to account for pleiotropic variants [24, 25, 46].

Small total effects

In simulation studies with no true total effect the MR estimate of the proportion mediated is implausible (Online Resource 1: sTable 4). Where there is no evidence of a total effect, consideration should be given as to whether it is appropriate to continue with mediation analyses. Although an indirect effect can be estimated in the absence of a significant total effect, or absence of total effect when the indirect effect and direct effect act in opposing directions and cancel each other out, these estimates are prone to inflated type 1 errors (i.e. false positive results) [47].

Where the total effect is weak or estimated imprecisely, simulations show the indirect effect and the proportion mediated using MR can be estimated but have large standard deviations (Online Resource 1: sTables 22–25). In this case, results should be interpreted with caution, especially considering the bounds of error.

Interactions between the exposure and mediator

In simulation scenarios with an interaction between the exposure and mediator present, the estimate of the direct effect of the exposure from both the difference in coefficients method and MVMR method was biased. In our simulations,

as the size of the interaction increased, the size of both the absolute and relative bias of the direct effect increased (Online Resource 1: sTable 26). The size of the bias was typically larger in the non-IV analyses compared with MR analyses.

Analysis of multiple mediators

The direct effect of an exposure controlling for multiple mediators in a single model can be assessed using MVMR. Where all MR specific assumptions are satisfied, the direct effect of multiple mediators is estimated with no evidence of bias (Online Resource 1: sTable 27). Here, non-overlapping SNPs for all exposures and mediators are included in one set of instruments. The estimated direct effect attributable to multiple mediators is unbiased, even when one mediator causes another mediator, which in our simulations was demonstrated by M2 causing M3 (Online Resource 1: sFig. 2).

Where the mediators do not cause each other, estimates of the indirect effects and proportion mediated from both MVMR (mutually adjusting for all mediators) and two-step MR (considering each mediator individually and summing together) will coincide (Online Resource 1: sTable 27). In our simulations, both MR methods estimated the indirect effect of each mediator, and the three mediators jointly, with no bias (Online Resource 1: sTable 27). This is consistent with the existing literature on analyses with multiple mediators [31].

Where one mediator causes another mediator, the indirect effect estimated via two-step MR captures both the amount of the association explained by the mediator of interest, and the amount of the mediator-outcome association captured by related mediators. In our example, this means that the effect of M3 is estimated twice, once directly and once via M2. As such, the estimate for the proportion mediated summing all three mediators together will likely be an overestimate of the combined proportion mediated, but the estimated direct effects remain unbiased. In our simulations, the combined proportion mediated was over-estimated by 6% (Online Resource 1: sTable 27), which is equivalent to the proportion explained by M3 via M2. The indirect effect of M2 estimated using two-step MR is however unbiased and reflects both the direct effect of M2 on the outcome and the indirect effect via M3 (Online Resource 1: sFig. 2).

Limitations of Mendelian randomisation applied to mediation analysis

Instrument selection

When using MR for mediation, SNPs included in the instruments for the exposure and mediator should be independent. Contrastingly, when MVMR is being used to test for potential pleiotropic pathways, SNPs associated with the two exposures under consideration can be included [24, 25, 48]. This is not the case when MVMR is being used to test for mediation. Should non-independent SNPs be included as instruments it would not be possible to distinguish whether an attenuation in the direct effect, compared with the total effect, was due to mediation or pleiotropy.

In a two-step MR mediation analysis, the mediator is considered as both an exposure (of the outcome) and as an outcome (of the exposure) and therefore any instruments for the exposure that are also instruments for the mediator are pleiotropic in the estimation of the effects of the exposure on the mediator and should be excluded. Where there are no independent SNPs, or the SNPs had a perfectly proportional effect on both the exposure and the mediator, then it would not be possible to use MR methods to estimate mediation.

The exclusion restriction criteria assuming no pleiotropic pathway is an important assumption of standard univariable MR, which applies equally when MR is used for mediation analysis. Some methods are available to assess pleiotropy including for the use of MVMR [43–45].

Binary exposures and/or mediators

Very few binary exposures will be truly binary and are likely a dichotomization of an underlying liability, changing the interpretation of an MR analysis [49]. For example, smoking is often defined as ever versus never smokers, when the underlying exposure is a latent continuous variable reflecting smoking heaviness and duration. As a result, the exclusion restriction criteria are violated, where the genetic variant can influence the outcome via the latent continuous exposure, even if the binary exposure does not change [49]. In a mediation setting, the same would apply to a binary mediator. In these scenarios, two-step MR could be used to test whether there is evidence of a causal pathway between the binary exposure and/or mediator. However, the estimates of mediation would likely be biased.

Interactions between the exposure and mediators

Within non-IV methods based on counterfactual assumptions, exposure-mediator interactions can be accommodated

when estimating mediation parameters. This is not possible in the non-IV mediation methods assessed here (difference in coefficients and product of coefficients) or in the MR methods, MVMR or two-step MR.

Methods are available for estimating interactions in an MR framework with individual level data, but these do not currently extend to estimating mediation in the presence of exposure-mediator interactions [13, 50, 51]. Estimates of mediation from MR mediation methods will be assuming effect homogeneity of both the exposure on the mediator and outcome, and mediator on the outcome. This means that the effect of either the exposure or the mediator on the outcome is not modified by the genetic instrument [52]. For estimates of the direct effect and indirect effect from two-step MR to be unbiased, the homogeneity assumption must be satisfied between the exposure-mediator association and mediator-outcome association and there should be no interaction between the exposure and mediator [23, 53]. Similarly, where MVMR is used, the effects between the exposure, mediator and outcome should all be homogenous [45]. Where the homogeneity assumption cannot be satisfied, the causal estimates from MR analyses will provide a valid test of the causal null hypothesis, but not the local average treatment effect [45, 54]. Developing MR methods which can account for these interactions will be important areas of future research.

Non-linear effects of an exposure or mediator

A further limitation of MR methods are where the effects of the exposure or mediator are non-linear. Although some methods are emerging for carrying out MR analysis with non-linear exposures [55, 56], these methods have not yet been extended to MR mediation analyses. Current MR methods for mediation analysis will assume a linear association between the exposure and outcome. As above, where the linearity assumption cannot be satisfied, the causal estimates would be a valid test of the causal null hypothesis, but not the local average causal effect. Where non-linear effects are of interest in a mediation model, non-IV methods should be considered.

Time varying exposures or mediators

In non-IV mediation analyses, the effects of exposures and mediators throughout the life course can be investigated by analysing prospective longitudinal data. In a simple model, the exposure, mediator and outcome should all be measured at separate, sequential, time points. This approach is preferred even when time varying effects are not a focus of the analysis, as cross-sectional data does not reflect the implied temporality suggested by mediation analysis [57–59]. In a more complex model, the longitudinal time

varying relationships between exposures and mediators can be modelled using structural equation models [60]. However, all of these approaches assume the direction of the effect between the exposure, mediator and outcome has been correctly specified, and is not due to reverse causality. Where this is incorrectly specified, or there are bidirectional relationships over time, estimates of the mediated effects can be biased. Additionally, where prospective data are used, the length of the interval between the exposure and mediator measurement is typically not accounted for [61]. Therefore, these analyses assume all variables have been measured at the critical time point and interval lengths for the associations between the exposure, mediator and outcome to exist. Here, the indirect effect estimated will be dependent upon the timing of these three measurements.

Mediation estimates from MR represent lifetime effects of the exposure and mediator measured at a single point in time [23]. As with non-IV approaches for this model to be valid, the assumptions must be made that the temporal relationship between the exposure and mediator has been correctly specified, that these relationships do not change throughout the life course and that the exposure and mediator have been measured at the relevant time points. However, as the genetic variants used as IVs for the exposure and mediator used in MR represent lifetime effects cross-sectional data collection can be applied here. Where time varying effects are not accounted for, the direct effect from an MR mediation analysis can be thought of as the effects of a one unit change in the lifetime exposure, not explained by a change in the mediators.

Using simulations, Labrecque and Swanson have explored bias in MVMR and two-step MR due to time varying effects of exposures and mediators [62]. Where there are bidirectional relationships between an exposure and mediator at different timepoints, estimates of the total, direct and indirect effects can be biased. One key advantage to using MR is that evidence of reverse causality or causal bidirectional associations between the measured exposure, mediator and outcome can be tested. Where instruments are available for the effect of an exposure at different timepoints, the potential time-varying bidirectional relationships can be explored [63]. In some unique cases instruments may be available for an exposure at different time points (e.g., childhood and adulthood BMI), allowing for a longitudinal approach to MR mediation analysis to be carried out. However, using these instruments come with additional methodological challenges [63].

As GWAS methods develop and sample sizes increase, the potential opportunities for incorporating time varying effects into MR analyses will likely increase. Future methods developments should focus on methods to incorporate instruments reflecting changes in the exposure and mediator across the life course. Additionally, further research should consider the meaning and interpretation of these lifetime

effects, or indeed time varying effects, in the context of the direct and indirect effect.

Power

MR studies require very large sample sizes to achieve adequate statistical power. Conditional F-statistics in MVMR are typically weaker than standard F-statistics, and indeed are likely to become weaker with each additional mediator included, further decreasing the power of complex analyses. Therefore, to achieve adequate statistical power, or precision, sample sizes for mediation analysis likely need to be even larger than those needed in a univariable MR analyses.

In the absence of formal power calculators for complex MR scenarios, the power of these analyses can be considered by evaluating the precision of the confidence intervals for all of the total, direct and indirect effects, as well as assessing the conditional instrument strength.

Genetic confounding

Although assumptions about unmeasured confounding in MR can be relaxed compared with traditional non-IV analyses, confounding can be introduced through population stratification, assortative mating, and dynastic effects [64]. Here, the confounding is not between the exposure, the mediator and the outcome, but between the genetic instruments for the exposure (or mediator) and the outcome. Adjusting for genetic principal components and other explanatory variables that capture population structure or within family analyses can minimise bias [64].

Mediation analysis with summary sample Mendelian randomisation

Methods applied in this paper can be used with summary data MR. Similar considerations will apply for both individual level MR, as presented here, and summary data MR. Importantly, all sources of summary statistics for the exposure, mediator and outcome should be non-overlapping [65]. As the mediator is considered an outcome in the exposure-mediator model, sample overlap can introduce bias [65]. As individual level data is not available in summary data MR, bootstrapping cannot be used to estimate the confidence intervals for the indirect effect or proportion mediated, but the delta method can be used to approximate these confidence intervals if samples are independent [30]. Analyses will also be restricted to the scale reported by the GWAS used, so consideration will need to be given for binary outcomes where sensitivity analyses to test potential non-collapsibility are limited.

Which method and when

Although MR is not biased by many of the untestable causal assumptions in non-IV mediation methods, such as unmeasured confounding, there are instead a set of MR specific causal assumptions (Fig. 2), and careful consideration should be given to which assumptions are most plausible. Additionally, the data available, or research question of interest may not be suitable to test in an MR framework. For example, if the research question is interested in exposures and mediators with time varying effects, or where interactions are present between the exposure and mediator.

MR has specific advantages compared with non-IV mediation methods where causal assumptions are required. The causal effect of the exposure on the outcome, the exposure on the mediator and the mediator on the outcome can all be tested. Additionally, bi-directional MR could be used to determine which of two variables is the causal exposure and causal mediator, where this is not known.

Our results demonstrate that both MVMR (akin to the difference in coefficients method) and two-step MR (akin to the product of coefficients method) can estimate the mediating effects for both continuous and binary outcomes, with little evidence of bias. However, caution is required in some instances, for example where total effects are weak. Where all exposures, mediators and outcomes are continuous, MVMR may confer an advantage of power, where the standard deviations for the simulated effects estimated in MVMR were smaller compared with the same effects estimated using two-step MR.

If an analysis is interested in estimating the effects of multiple mediators, consideration should be given to the causal question of interest when deciding which method to use to analyse multiple mediators. Where the causal question specifically relates to identifying the combined effects of multiple mediators, MVMR is likely to be the most appropriate method. Where the causal question aims to estimate the effect of multiple mediators individually, and potentially any impact of intervening on a mediator, two-step MR may be most appropriate. However, it is important to note, that as the number of mediators included in an MVMR model increases, the power of the analysis would likely decrease. Additionally, future research should be carried out to determine if including increasing numbers of exposures in an MVMR model further violates any of the MR assumptions.

Although we have included a range of simulation scenarios, including both continuous and binary outcomes, this is not an exhaustive range of scenarios and there may be further scenarios where MR methods are biased.

The flow chart in Fig. 4 aims to help with the decision-making process, based on practical limitations of MR.
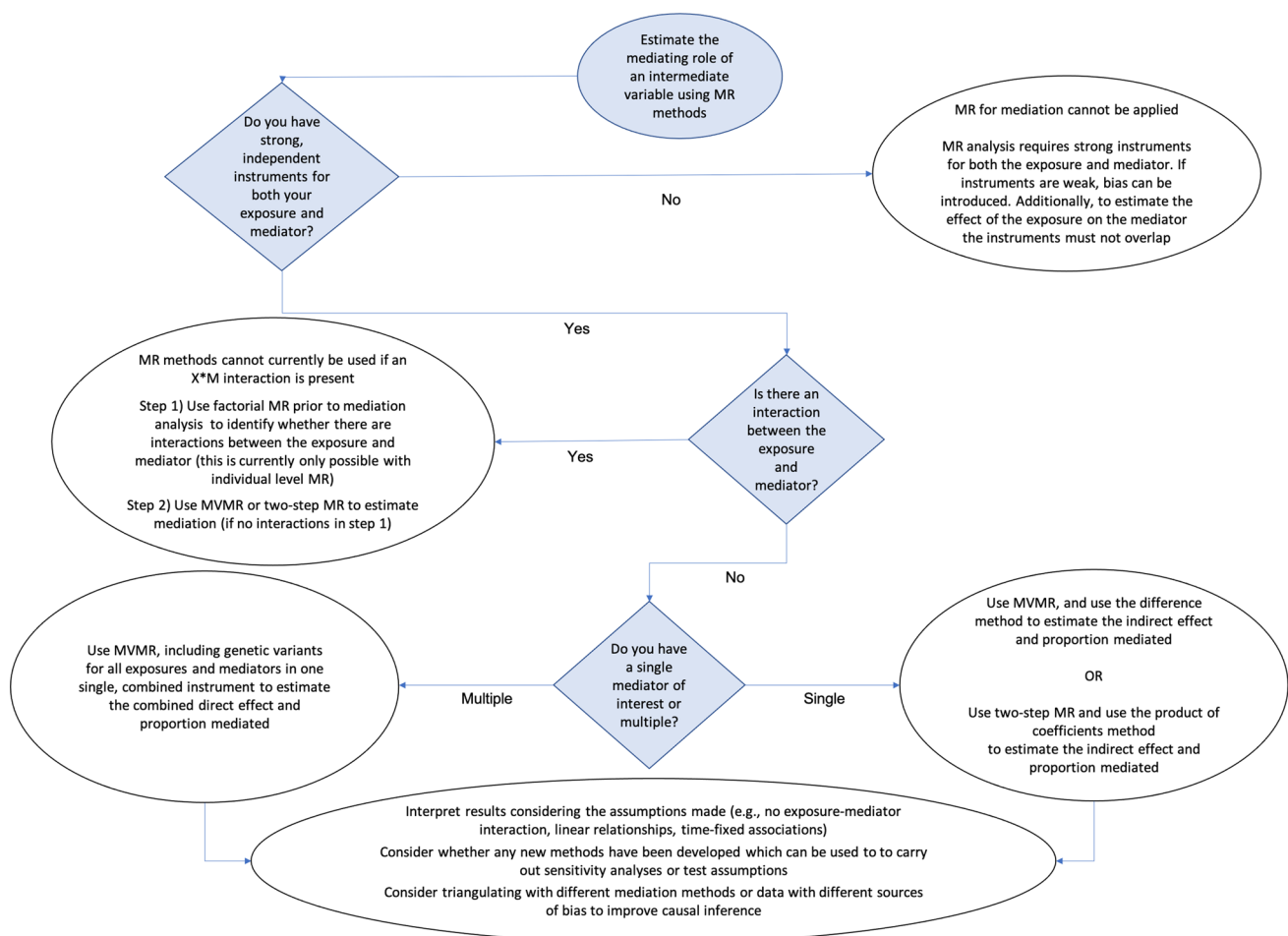
Fig. 4 Flow chart for analytical processes when carrying out mediation analyses using individual level Mendelian randomisation

However, best practice would always be to triangulate across non-IV and MR approaches, and across multiple data sources wherever possible [66].

Conclusions

MR can be extended to estimate direct effects, indirect effects and proportions mediated. MR estimates are not biased by violations of the often-untestable assumptions of non-IV mediation analysis, including unmeasured confounding and measurement error. MR analysis makes its own strong, but distinct assumptions, especially relating to instrument validity. To estimate mediation using MR, we require large sample sizes and strong instruments.

Data and code availability All code for simulation analyses, applied analyses and example code is available on GitHub (simulations: https://github.com/eleanorsanderson/MediationMR, applied analyses and example code: https://github.com/alicerosecarter/MediationMR). The cleaned dataset for UK Biobank analyses will be archived with UK Biobank. Please contact access@ukbiobank.ac.uk for further information.

Declarations

Conflict of interest The author declare that they have no conflict of interest.

Ethical approval This project was approved by UK Biobank under the study application 10953. No specific ethical approval or patient involvement was sought for this project.

References

1. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey SG. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–63. https://doi.org/10.1002/sim.3034.

2. Wright S. The method of path coefficients. Ann Math Stat. 1934;5:161–215. https://doi.org/10.1214/aoms/1177732676.

3. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51(6):1173–82.

4. VanderWeele TJ. Mediation analysis: a Practitioner's guide. Annu Rev Public Health. 2016;37:17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402.

5. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods. 2002;7(1):83–104. https://doi.org/10.1037/1082-989x.7.1.83.

6. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol. 2013;42(5):1511–9. https://doi.org/10.1093/ije/dyt127.

7. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. Annu Rev Psychol. 2007;58:593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542.

8. Blakely T, McKenzie S, Carter K. Misclassification of the mediator matters when estimating indirect effects. J Epidemiol Community Health. 2013;67(5):458–66. https://doi.org/10.1136/jech-2012-201813.

9. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. Am J Epidemiol. 2010;172(12):1339–48. https://doi.org/10.1093/aje/kwq332.

10. VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. Stat Interface. 2009;2(4):457–68.

11. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013;18(2):137–50. https://doi.org/10.1037/a0031034.

12. Vanderweele TJ. Controlled Direct and Mediated Effects: Definition. Identification and Bounds Scand J Stat. 2011;38(3):551–63. https://doi.org/10.1111/j.1467-9469.2010.00722.x.

13. VanderWeele TJ. A unification of mediation and interaction: a 4-way decomposition. Epidemiology. 2014;25(5):749–61. https://doi.org/10.1097/EDE.0000000000000121.

14. Vansteelandt S. Understanding counterfactual-based mediation analysis approaches and their differences. Epidemiology. 2012;23(6):889–91. https://doi.org/10.1097/EDE.0b013e31826d0f6f.

15. VanderWeele TJ, Valeri L, Ogburn EL. The role of measurement error and misclassification in mediation analysis: mediation and measurement error. Epidemiology. 2012;23(4):561–4. https://doi.org/10.1097/EDE.0b013e318258f5e4.

16. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. 1992;3(2):143–55. https://doi.org/10.1097/00001648-199203000-00013.

17. 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(12):e352. https://doi.org/10.1371/journal.pmed.0040352.

18. Davey Smith G, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1–22.

19. Burgess S, Thompson DJ, Rees JMB, Day FR, Perry JR, Ong KK. Dissecting causal pathways using mendelian randomization with summarized genetic data: application to age at menarche and risk of breast cancer. Genetics. 2017;207(2):481–7. https://doi.org/10.1534/genetics.117.300191.

20. Sanderson E. Multivariable Mendelian randomization and mediation. Cold Spring Harb Perspect Med. 2020. https://doi.org/10.1101/cshperspect.a038984.

21. Relton CL, Davey SG. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. Int J Epidemiol. 2012;41(1):161–76. https://doi.org/10.1093/ije/dyr233.

22. Richmond RC, Hemani G, Tilling K, Davey Smith G, Relton CL. Challenges and novel approaches for investigating molecular mediation. Hum Mol Genet. 2016;25(R2):R149–56. https://doi.org/10.1093/hmg/ddw197.

23. Burgess S, Daniel RM, Butterworth AS, Thompson SG, Consortium EP-I. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in

causal pathways. Int J Epidemiol. 2015;44(2):484–95. https://doi.org/10.1093/ije/dyu176.

24. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol. 2015;181(4):251–60. https://doi.org/10.1093/aje/kwu283.

25. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2018. https://doi.org/10.1093/ije/dyy262.

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

27. Varbo A, Benn M, Davey Smith G, Timpson NJ, Tybjaerg-Hansen A, Nordestgaard BG. Remnant cholesterol, low-density lipoprotein cholesterol, and blood pressure as mediators from obesity to ischemic heart disease. Circ Res. 2015;116(4):665–73. https://doi.org/10.1161/CIRCRESAHA.116.304846.

28. Xu L, Borges MC, Hemani G, Lawlor DA. The role of glycaemic and lipid risk factors in mediating the effect of BMI on coronary heart disease: a two-step, two-sample Mendelian randomisation study. Diabetologia. 2017. https://doi.org/10.1007/s00125-017-4396-y.

29. Marouli E, Del Greco MF, Astley CM, Yang J, Ahmad S, Berndt SI, et al. Mendelian randomisation analyses find pulmonary factors mediate the effect of height on coronary artery disease. Commun Biol. 2019;2:119. https://doi.org/10.1038/s42003-019-0361-2.

30. Carter AR, Gill D, Davies NM, Taylor AE, Tillmann T, Vaucher J, et al. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. BMJ. 2019;365:l1855. https://doi.org/10.1136/bmj.l1855.

31. VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. Epidemiol Methods. 2014;2(1):95–115. https://doi.org/10.1515/em-2012-0010.

32. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 2010;39(2):417–20. https://doi.org/10.1093/ije/dyp334.

33. Cole SR, Hernan MA. Fallibility in estimating direct effects. Int J Epidemiol. 2002;31(1):163–5. https://doi.org/10.1093/ije/31.1.163.

34. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. Stat Sci. 1999;14(1):29–46.

35. Pang M, Kaufman JS, Platt RW. Studying noncollapsibility of the odds ratio with marginal structural and logistic regression models. Stat Methods Med Res. 2016;25(5):1925–37. https://doi.org/10.1177/0962280213505804.

36. Sanderson E, Windmeijer F. A weak instrument [Formula: see text]-test in linear IV models with multiple endogenous variables. J Econom. 2016;190(2):212–21. https://doi.org/10.1016/j.jeconom.2015.06.004.

37. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med. 2011;30(11):1312–23. https://doi.org/10.1002/sim.4197.

38. Burgess S, Thompson SG. Improving bias and coverage in instrumental variable analysis with weak instruments for continuous and binary outcomes. Stat Med. 2012;31(15):1582–600. https://doi.org/10.1002/sim.4498.

39. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey SG. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr. 2016;103(4):965–78. https://doi.org/10.3945/ajcn.115.118216.

40. Sanderson E, Spiller W, Bowden J. Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation. bioRxiv. 2020. https://doi.org/10.1101/2020.04.02.021980.

41. Hemani G, Bowden J, Davey SG. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet. 2018;27(R2):R195–208. https://doi.org/10.1093/hmg/ddy163.

42. Davies NM, Holmes MV, Davey SG. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601. https://doi.org/10.1136/bmj.k601.

43. 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(2):512–25. https://doi.org/10.1093/ije/dyv080.

44. 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(4):304–14. https://doi.org/10.1002/gepi.21965.

45. Rees JMB, Wood AM, Burgess S. Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy. Stat Med. 2017;36(29):4705–18. https://doi.org/10.1002/sim.7492.

46. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658–65. https://doi.org/10.1002/gepi.21758.

47. Loeys T, Moerkerke B, Vansteelandt S. A cautionary note on the power of the test for the indirect effect in mediation analysis. Front Psychol. 2014;5:1549. https://doi.org/10.3389/fpsyg.2014.01549.

48. Zuber V, Colijn JM, Klaver C, Burgess S. Selecting likely causal risk factors from high-throughput experiments using multivariable Mendelian randomization. Nat Commun. 2020;11(1):29. https://doi.org/10.1038/s41467-019-13870-3.

49. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. Eur J Epidemiol. 2018;33(10):947–52. https://doi.org/10.1007/s10654-018-0424-6.

50. North TL, Davies NM, Harrison S, Carter AR, Hemani G, Sanderson E, et al. Using genetic instruments to estimate interactions in mendelian randomization studies. Epidemiology. 2019;30(6):e33–5. https://doi.org/10.1097/EDE.0000000000001096.

51. Rees JMB, Foley CN, Burgess S. Factorial Mendelian randomization: using genetic variants to assess interactions. Int J Epidemiol. 2019. https://doi.org/10.1093/ije/dyz161.

52. Labrecque J, Swanson SA. Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools. Curr Epidemiol Rep. 2018;5(3):214–20. https://doi.org/10.1007/s40471-018-0152-1.

53. Zheng J, Baird D, Borges MC, Bowden J, Hemani G, Haycock P, et al. Recent developments in Mendelian randomization studies. Curr Epidemiol Rep. 2017;4(4):330–45. https://doi.org/10.1007/s40471-017-0128-6.

54. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology. 2014;25(3):427–35. https://doi.org/10.1097/EDE.0000000000000081.

55. Staley JR, Burgess S. Semiparametric methods for estimation of a nonlinear exposure-outcome relationship using instrumental variables with application to Mendelian randomization. Genet Epidemiol. 2017;41(4):341–52. https://doi.org/10.1002/gepi.22041.

56. Burgess S, Davies NM, Thompson SG, Consortium EP-I. Instrumental variable analysis with a nonlinear exposure-outcome relationship. Epidemiology. 2014;25(6):877–85. https://doi.org/10.1097/EDE.0000000000000161.

57. Maxwell SE, Cole DA, Mitchell MA. Bias in cross-sectional analyses of longitudinal mediation: partial and complete mediation under an autoregressive model. Multivar Behav Res.

2011;46(5):816–41. https://doi.org/10.1080/00273171.2011.606716.

58. Goldsmith KA, MacKinnon DP, Chalder T, White PD, Sharpe M, Pickles A. Tutorial: the practical application of longitudinal structural equation mediation models in clinical trials. Psychol Methods. 2018;23(2):191–207. https://doi.org/10.1037/met0000154.

59. Jose PE. The merits of using longitudinal mediation. Educ Psychol. 2016;51(3–4):331–41. https://doi.org/10.1080/00461520.2016.1207175.

60. Cole DA, Maxwell SE. Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. J Abnorm Psychol. 2003;112(4):558–77. https://doi.org/10.1037/0021-843X.112.4.558.

61. Selig JP, Preacher KJ. Mediation models for longitudinal data in developmental research. Res Hum Dev. 2009;6(2–3):144–64. https://doi.org/10.1080/15427600902911247.

62. Labrecque JA, Swanson SA. Mendelian randomization with multiple exposures: the importance of thinking about time. Int J Epidemiol. 2019. https://doi.org/10.1093/ije/dyz234.

63. Richardson TG, Sanderson E, Elsworth B, Tilling K, Davey SG. Use of genetic variation to separate the effects of early and later

life adiposity on disease risk: mendelian randomisation study. BMJ. 2020;369:m1203. https://doi.org/10.1136/bmj.m1203.

64. Brumpton B, Sanderson E, Heilbron K, Hartwig FP, Harrison S, Vie GA, et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. Nat Commun. 2020;11(1):3519. https://doi.org/10.1038/s41467-020-17117-4.

65. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597–608. https://doi.org/10.1002/gepi.21998.

66. Lawlor DA, Tilling K, Davey SG. Triangulation in aetiological epidemiology. Int J Epidemiol. 2016;45(6):1866–86. https://doi.org/10.1093/ije/dyw314.