Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2024;10(2): 132-149

Review Article

A Review of Mendelian Randomization in the Presence of Weak Instrumental Variables; Statistical Methods and Challenges

Danial Habibi[1], Maryam S Daneshpour[2], Marjan Mansourian[3], Mahdi Akbarzadeh[2*]

[1]Department of Epidemiology and Biostatistics, School of Public Health, Babol University of Medical Sciences, Babol, Iran.
[2]Cellular and Molecular Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[3]Department of Biostatistics and Epidemiology, Faculty of Health, Isfahan University of Medical Sciences, Isfahan, Iran.

ARTICLE INFO

Received 01.02.2024
Revised 20.02.2024
Accepted 09.03.2024
Published 15.06.2024

Key words:
Mendelian Randomization;
Instrumental variable;
Weak instrumental variables;
Statistical problem;
Statistical remedy

ABSTRACT

Introduction: Mendelian randomization (MR) assesses the causal effect of risk factors by using genetic variations as instrumental variables (IV) in nonexperimental data. IV strategies are one of a few available methods for determining causal effects in the absence of specific knowledge of all confounders in the exposure-outcome relationship. To use an IV as a legitimate instrumental variable, it must meet the following criteria: relevance, exchangeability, and exclusion restriction. A weak instrument is a circumstance in which there is a piece of weak statistical evidence for an association between IV and exposure. Weak instruments cause significant issues, including (i) insufficient statistical power to hypothesis testing, (ii) increasing bias with deviation from IV assumptions, and (iii) asymptotic estimation of standard errors and confidence intervals. Therefore, in this study, we intend to introduce the Mendelian randomization method, weak instrumental bias, and statistical remedy methods used in this bias.
Methods: Current study was conducted by using Medline/PubMed, Scopus, Web of Sciences and Google Scholar.
Results: This review provides a comprehensive description of the principles of MR, and a guide to basic MR methodology. To deal with these challenges, the bulk of the review considered statistical remedies. The review ends with a section that details the practical limitations, and recommendations regarding MR and the weak instruments.
Conclusion: Depending on the type of data, several solutions can be used in one and multiple IV. Moreover, it can be used in solutions in the design and analysis phase to minimize the effects of weak instruments.

Introduction

One of the difficulties for medical researchers is to determine the causal risk factors or exposures. Randomized controlled trials (RCTs) are used as a standard method. Random allocation in the RCTs removes the potential for bias in the allocation of participants; that is, the intervention and non-intervention groups will be balanced at unmeasured prognostic factors

*.Corresponding Author: akbarzadeh.ms@gmail.com, akbarzadehms@sbmu.ac.ir

and other characteristics of the participants at randomization.[1] Nevertheless, RCTs are always not feasible,[2] then the only remaining option is observational studies.[3] Here, confounding and reverse causation can hamper causal inference in observational epidemiological studies.[4] On the other hand, causal inference is challenged because of the absence of a random assignment for the exposure.

If confounders are known and measured, they can control by including them in regression models.[5] Due to the nature of the confounder variables, if the number of these confounders is large we cannot expect an unbiased estimate from the regression model and a true relationship of exposure and outcome, even if the study design and analysis method are appropriate.

The instrumental variable (IV) method has been proposed as an alternative statistical approach. The IV techniques are a few available ways to estimate the causal effects without fully knowing all the confounders of the exposure-outcome association.[6] The IV method was initially introduced by economists and later used in Mendelian randomization (MR) analysis in medical statistics.[7] The MR is an analytical method that uses genetic variants as IVs in nonexperimental data for assessing and estimating the causal effects of risk factors.[8,9] MR is a popular technique that eludes confounding variables in an observational study and evaluates causal factors for phenotypes that would not suit RCTs (figure 1).[10]

MR is an application of IV analysis. According to Mendel's first and second inheritance laws, MR is based on that genetic variants are randomly allocated during meiosis, thus mitigating concerns about reverse causation,[10–12] though this concern is not completely eliminated.[13] The term MR was coined due to its relevance to Mendel's laws.

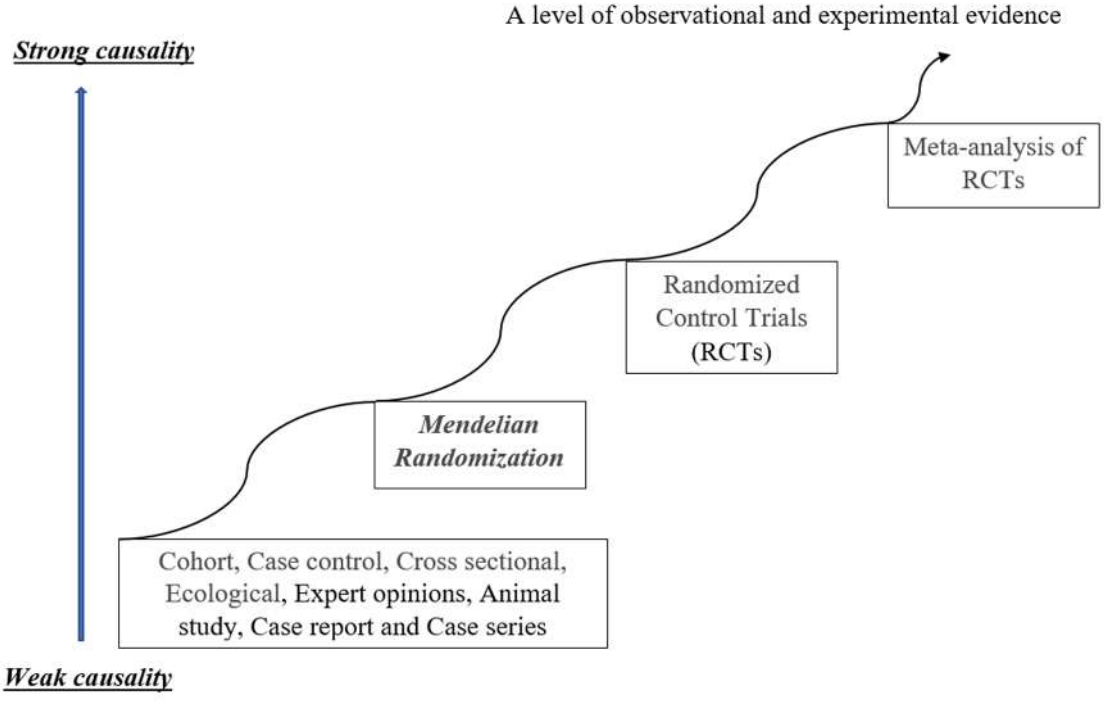

Figure 1. A hierarchy of causality from observational epidemiology to interventional studies.

As a result, inherited variants are free of the potentially confusing environment.[14] MR study's benefits are that it identifies differences genetically, remains constant if untreated, is not affected by selection bias, and reflects long-term differences.[15] In MR implementation, several steps must be performed in a row to obtain satisfactory results, summarized in Figure 2.

Mendelian deconfounding is another name for Mendelian randomization because of the effect of causality without biases due to confounding.[16] By finding a genetic variant that meets IV assumptions, we can estimate the unconfounded connection between exposure and outcome, summarized in table 1 and figure 3.[17]

A weak instrument is a phenomenon where the statistical association between risk factor and IV in the data set is weak and does not explain much variation in the risk factor. By and large, a weak instrument bias occurs when there is insufficient statistical evidence to support an association between IV and exposure.

Another assumption is required for estimating a causal effect via the IV analysis; all associations are linear and not affected by statistical interactions. This assumption becomes problematic when the outcome is binary. Because we want to estimate the odds ratio or risk ratio in this situation, and this association is exponential or non-linear.[18]

Although the instrument variable could be essentially any variable, single-nucleotide polymorphisms (SNPs) are widely utilized. Genome-wide association studies (GWAS) are observational studies and hypothesis-free methods to identify associations between SNPs and phenotypic traits (diseases).[19] This

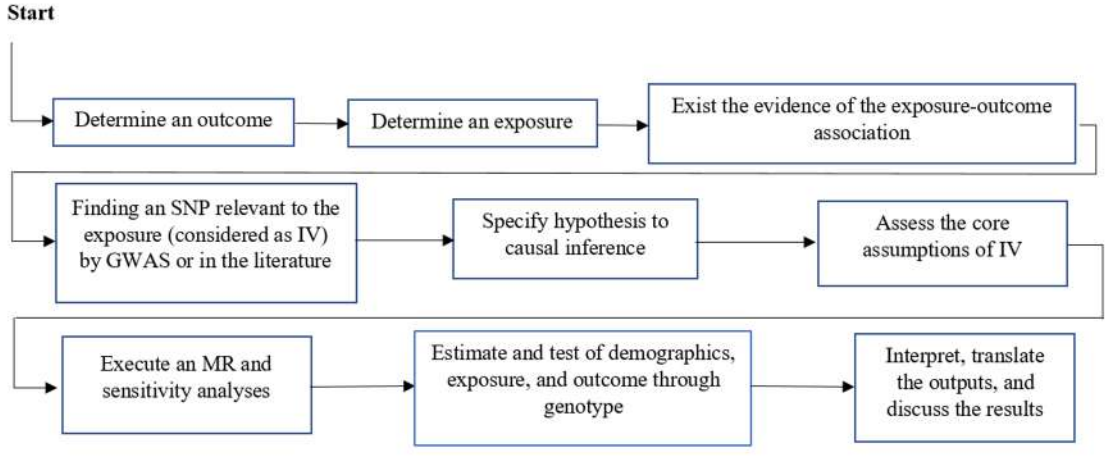

Figure 2. Flow chart of steps for MR analysis.

Table1. Core assumptions underlying MR

Assumptions	Terms	Description
(1)	Relevance	G has a causal effect on X
(2)	Exchangeability or independence	No confounding for the effect of G on Y
(3)	Exclusion restriction	G affects the outcome Y only via X

Demonstrated the main assumption of the instrument variable. If these assumptions are met, then the estimate has a free confounding effect of the exposure and outcome association.

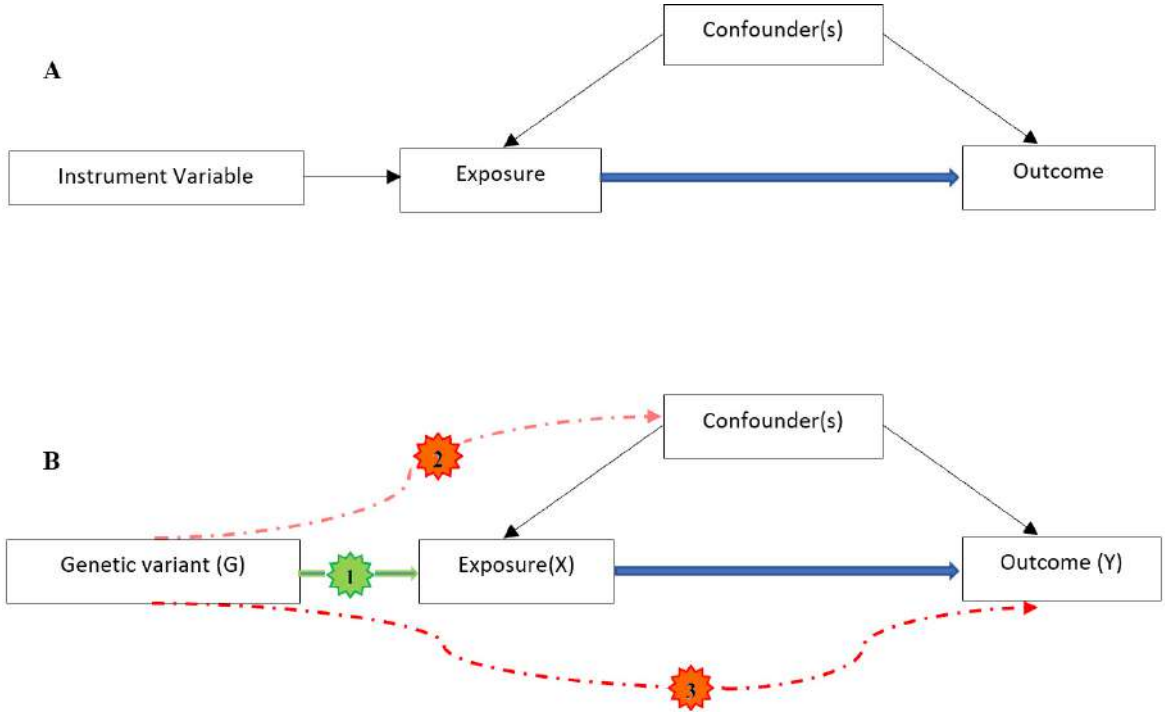

Figure 3. The A case represents the conceptual of Mendelian Randomization, and the B case represents the main assumptions of Mendelian randomization method. The main purpose is to estimate the exposure-outcome association without affecting the confounding variable. We try to estimate this association using an IV such as SNP.

means that many of the candidate genes have been published through GWAS so that MR studies can now be exploited without the need to attract new patients or additional design studies.

Herein, several types of information existed. Summary data refers to methods that use the summary level of the IV-exposure and the IV-outcome association, including beta-coefficients and standard errors from linear or logistic regression to get causal effect estimates. The effect directly represents the genetic instrument.[20,21] Each individual's information is accessible in individual data and can achieve regression estimates by logistic or linear regression.

Other data is available under meta-analysis. It involves (i) individual data meta-analysis (that is, each individual's information (exposure and outcome) are incorporated before regression analysis to obtain associations (IV-exposure and IV-outcome)), (ii) summary data meta-analysis (that is, the effect estimates can be pooled for identical or different relationships of the IV-exposure and IV-outcome), (iii) study data meta-analysis (that is, MR estimates for various SNPs are pooled straight to render the synthesized causal estimate).[22] The following figure 4 was indicated the summary of methods for summary and individual-level data information.

MR can be undertaken in a one-sample, a two-sample, or a subsample. In one-sample MR, the exposure and outcome data are obtained from a single dataset. In two-sample MR, the exposure and outcome data are obtained from different datasets.[8,23] A vital feature of a two-sample is that the analysis can be performed using summary-level data from a GWAS.[24] In a subsample, data on the exposure are achievable

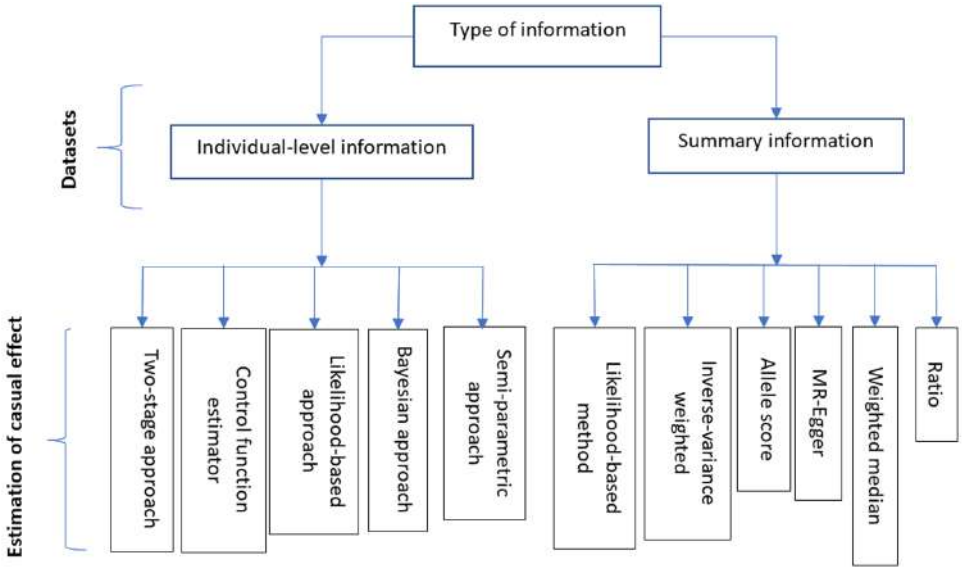

Figure 4. Methods for the estimation of causal effects. Summary data refers to methods that use the summary level of the IV-exposure and the IV-outcome association, including beta-coefficients and standard errors. In individual data, each individual's information is accessible.

only for a subset of participants, yet the outcome data are attainable for all participants, or data on outcome are accessible for a subset of participants. Nevertheless, data on exposure exist for all participants.[25]

Other designs are also available. In bidirectional MR, it can assess both exposure and outcome instruments whether the exposure variable causes the outcome or whether the outcome variable causes the exposure.[26] In two-step MR, the goal is to evaluate whether an intermediate trait works as a causal mediator role between the exposure-outcome association.[27] Multivariable MR is an extension of standard (univariable) MR that permits multiple exposures, and it can dominate the genetic variants that are pleiotropically associated with multiple correlated exposures.[28] Factorial MR is designed to answer questions on interactions.[29]

In the last two decades, the use of the Mendelian randomization approach has increased dramatically. Therefore, this study intends to overview of the Mendelian randomization method, weak instrumental bias, problem and challenges, and statistical remedy methods used in this bias.

Weak instrument

A weak instrument is a scenario that lacks strong statistical evidence of an association between IV and exposure. This phenomenon will occur if the first condition of IV is not met.[30] In other words, an IV is weak if it reports a little small amount of the variation of the exposure, and this weak is directly related to the sample size.[31]

The IV approach suffers from finite samples howbeit it can be used to render asymptotically unbiased estimates of causal effects in confounding attendance.[32] Even a weak IV is valid in infinite samples, so convinced assumptions and estimates will be unbiased. In finite samples, the IV estimator's average value will be biased, and its magnitude belongs to the strength of the IV-exposure association.[33]

A measure of the strength of a weak instrumental variable

A measure of the strength of a weak instrumental variable is the Cragg–Donald F statistic (F<10), which is the same in the IV-exposure regression.[34,35] The F-statistic is related to the proportion of variance in the phenotype explained by the genetic variants (R^2), sample size (n), and number of instruments (k) by the following formula

$$F = \left(\frac{n-k-1}{k}\right)\left(\frac{R^2}{1-R^2}\right)$$

The strength of selected single-IVs was also assessed in two sample MR by following formula

$$F = \left(\frac{Beta}{Standard\ Error}\right)^2$$

Since the F statistic depends on the sample size, the bias can be mitigated by increasing the sample size. Likewise, if there are instruments that do not contribute much to explaining the variation in the phenotype, removing these instruments will increase the F value.

However, such a cut-off point (<10) from the F statistic can be deceiving. This value is arbitrary and cannot be called strong or weak by looking at the binary to IV. In contrast, if we call an IV weak, it does not mean that it is an inherently weak IV because any IV can increase its power by increasing the sample size.[36] Also, regarding the sample size, the F statistic is not easily a measure of the coefficient of determination (R^2); hence, the F cannot trustworthy guide due to large sampling variability.[37] Besides, the post hoc selection of data according to measured F-statistics can intensify bias.[38] Indeed, the cut-off point was ascertained based on the Two-Stage least squares (2SLS) method and is not necessarily related to further IV methods. Hence, the F statistic may not even be a reliable measure of instrument strength for obtaining identification in a semi-parametric model.[39]

Weak instrumental variable in an one-sample design

In a one-sample design, estimates from the IV method are asymptotically unbiased. It can have a remarkable bias in finite samples, such as weak instrument bias, which leads to the confounded observational association between the exposure and outcome. The weak instrument's magnitude depends on the strength of association between the IV and the exposure[37] and inflates type 1 error rates.[34] The amplified bias in one-sample MR can prove by evidence through an assessment of the Wald ratio; hence, the IV-outcome association's coefficient remains constant, yet the coefficient of the IV-exposure association is reduced because of the weak instruments.

In the 2SLS method, weak instrument bias can be described as deriving from overfitting in the first stage, which happens at least in a portion resulting from chance correlations of the instrument variable with confounders.[37] In a one-sample design, the first stage's fitted values correlate with the outcome in the finite sample size, even in the lack of a causal effect. Thus, this phenomenon yields to the bias of finite sample 2SLS estimates.

Weak instrumental variable in a two-sample design

In a two-sample design, a weak instrument towards the null direction.[40] Thus, the bias in the way of the null can be lower earnest than bias in the way of the observational association. Besides, such a scenario is conservative, resulting in an inflated type 1 error; inversely, it may result in less power to identify a causal effect and raise the probability of a type 2 error.[25] Briefly, in the ratio method, if the outcome is continuous and regression analyses are a linear model, weak instrument bias in the one-sample design is derived from the correlation between the regression coefficients in the numerator and denominator. At the same time, the numerator and denominator will be uncorrelated in the two-sample design.[41]

In a two-sample design with individual-level data, estimates can be computed by getting estimates of the first stage in one dataset and creating fitted values of the second dataset exposure. In the second stage, the outcome and exposure are no longer correlated due to confounding. This scenario is referred to as a split-sample 2SLS.[42] Any bias resulting from weak instruments is in the way of the null.[25] Meanwhile, weak instruments assist in testing the effect of heterogeneity and distinguishing candidate IVs that do not meet the exclusion restriction.[43]

Problems and challenges

The bias is a term that refers to the difference between the expectation and the parameter's true value. The weak instrument and establishing bias can lead to underestimated confidence intervals and weak coverage properties.[37] In other words, a reasonable confidence interval should not be large or infinite if the data possess information about a parameter; hence, weak instruments lead to inconsistent estimates with wide confidence intervals, which can significantly influence the power of the analysis.[44]

Because of a weak instrument, 2SLS estimators have a finite sample bias; hence, they can grow bias by adding weak instruments to the first stage. Besides, weak instruments can enlarge any residual bias through a confounded instrument so that low violations of IV assumptions can make a direction to significant inconsistency in the IV estimator.[45] Moreover, a single weak instrument will possess small power to decline the null hypothesis and makes the point estimate hard to interpret. Any miniature violation in the exclusion restriction assumption can account for significant biases.[46]

In a two-sample design, if there is some overlap, it is uncertain whether bias leads to weak instruments in the way of the null (zero overlaps) or the observational association (complete overlap). Besides, if individuals were only utilized in the detection dataset in a binary outcome, this should not bias. However, a weak instrument bias will occur when controls and cases are used in the detection dataset.[20]

In general, weak instruments leads to the following problems. Firstly, they may provide small or no benefit information and prepare low statistical power to test hypotheses. Secondly, bias amplified when the core instrumental variable assumptions violation. Thirdly, make asymptotic approximations for standard errors and confidence intervals. Fourthly, even when available big samples, biased towards between the association of the outcome and exposure in the one-sample design and towards the direction of the null in the two-sample design.

A common problem of the weak instrument for semi-parametric approaches is that there is no warranty for estimating a unique parameter.[33]

Statistical method remedy

One IV in individual level data

Bias with one IV in medium-wide datasets is usually unimportant, yet bias may be discussed when several IVs.[47]

Multiple IVs in individual level data

In instrumental genetic variables, one debate is that many genetic variants are only weakly associated with exposure. To this end, several studies for inferential procedures provide better asymptotic approximations in the finite sample.[32] Although limited information maximum likelihood (LIMI) may not completely solve the problem, it is preferred over the 2SLS estimator.[48] Other partially robust estimators include Jackknife instrumental variables, Fuller-k estimator, or bias-adjusted 2SLS that provide relatively more reliability than 2SLS.[49]

On the other hand, it has been shown to be cautious for using "no moment" estimators such as LIML.[50] Hence, alternative approaches recommended including, the jackknife 2SLS (JK2SLS) estimator and Wayne Fuller's.[48] However, one of the merits of using the LIML over 2SLS is that it enables a limited confidence interval. Hence, LIML provides lower biased estimates in the attendance of weak instruments.[22]

Confidence intervals such as a Fieller's theorem and Bayesian posterior distribution drawn from Monte Carlo Markov chain sampling (MCMC) have better coverage properties. Moreover, inverting a test statistic, such as the Rubin test statistic or the conditional likelihood ratio test statistic, can be used as an alternative approach to decrease power with stronger instruments.[33]

Another way is allele score. The large numbers of instrumental variables, allele scores, genetic risk scores, gene scores, or genotype scores can reduce weak instruments' problems. A univariate allele score as one IV, rather than any single genetic variant as a discrete IV, helps solve IV assessment problems caused by the weak instruments. However, a weak instrument should not be entangled with an invalid instrument because a weak instrument can be amplified stronger by increasing more data.[36] Two approaches exist for incorporating information on multiple uncorrelated IVs into a single causal estimate; i) allele scores and ii) the summary statistic method. However, unweighted and externally weighted allele scores have been recommended for eluding bias depending on the weak instrument.[40] In 2SLS, when one or more genetic instruments are weak, one approach combines into a single genetic instrument through genetic risk scores (weighted or unweighted).[13]

The commonly used 2SLS estimator is biased.[51] To this end, two methods have been proposed for this situation. Building an allele score (weighted or unweighted) and utilizing the 2SLS with only one IV, and utilizing the complete collection of instruments simultaneously with a robust estimation method through the routine standard errors for limited information maximum likelihood (LIML) is wrong and result in false inference for testing hypotheses.[52]

Besides, robust methods such as merging the variants into a single allele score, using the

LIML and continuously updating estimator (CUE) estimator do not conflict with the many weak instruments. However, it has been shown that the LIML be biased with many weak instruments because of heteroskedastic errors.[53] Hence, The CUE estimator is one choice in this situation, though for estimating risk differences for binary outcomes and its standard errors again require to be corrected.[52] However, it is illustrated that if there are many weak instruments, then CUE is a better selection than LIML and 2SLS.[54]

To relieve weak instrument bias in the individual-level data, utilize the LIML or CUE method, also using a jackknife IV estimator or equivalently an allele score approach using leave-one-out cross-validated weights or an allele score approach using equal or externally specified weights.[20]

Two-sample summary data

Several tests were proposed in econometrics for individual-level data, such as the Anderson-Rubin test, the Kleibergen test, and the conditional likelihood ratio test. These tests have control on type I error regardless of the instrument's power(https://academic.oup.com/biometrics/article/78/4/1699/7460098?login=false). Wang et al. propose robust test statistics for two-sample summary data by developing these tests and demonstrating that these tests control type I error under weak instrument asymptotic. Besides, they show that the mrCLR (extending the conditional likelihood ratio test) has better performance.[55]

YE et al. proposed the debiased inverse variance weighted estimator in two-sample summary data for MR.[56] This approach is a simple adjustment of the inverse variance-weighted average method (IVW) estimator and doesn't need screening. Initially, meta-analysis methods have been used in MR as a tool for analyzing individual-level data and recently combining GWAS in two-sample summary data.[57] An attractive advantage of the two-sample design is a protection versus the weak instrument bias.[24] Also, sensitivity analyses can be conducted in this situation.[23]

Statistical remedy in the design and analysis phase

The measure of instrument strength for the MR-Egger method is based on the Bowden I^2 statistic, which gets values between 0 and 1. If I^2 obtains a value of one, then the MR-Egger estimate cannot suffer from the bias of the weak instrument.[58] In a one-stage meta-analysis, it is revealed that weak instrument bias can be mitigated.[59]

In addition to the above, the bias caused by the weak instruments can be minimized in the design and analysis phase. One way is to increase F statistics. As mentioned earlier, the F-statistic is used as a measure to assess the instrument's strength. As the F-statistic pertains to the sample size, the bias decreases as the sample size increases.[38] Excluding instruments that explain fewer variations in the phenotype leads to an increase in F-statistic. Parsimonious models typically increase F-statistics.[31] Another way is adjusted according to measures covariates. If the covariates can describe the variation in the exposure so that covariates are not in the path of causality between the exposure and the outcome, we can include these covariates in the model. Simulations have shown that may increase the genetic association with the exposure and reduce weak

instrument bias.[38] In this case, F-statistic is called partial F statistic.[60] Another way is based on the meta-analysis.[60] If an IV estimator is not biased, then the mean exposure's correct values under different genetic subgroups are known. Therefore, to obtain accurate estimates of exposure levels in each genetic subgroup, the estimates of exposure association could combine in different studies.[38]

Weak instrument bias will be low according to the P-value in a linear regression of the exposure for each IV lower than 1 × 10−5.[47] Strong instruments in polygenic scores are pleasant to forbid weak instrument bias.[61] Briefly, this section is summarised in figure 5.

Conclusion

The relationship between exposure and outcome in observational epidemiology studies can be disrupted by unmeasured or measured confounders, reverse causation, and potential biases.[62] Although these damaging effects can be reduced with the suitable study design and analysis approach, they cannot be eliminated. Moreover, the correlation between exposure and outcome cannot be considered causal.[63] We also have limitations in using the RCT as a standard method. Thus, the IV approach was found.[54]

The idea behind the MR approach was that in addition to overcoming the destructive

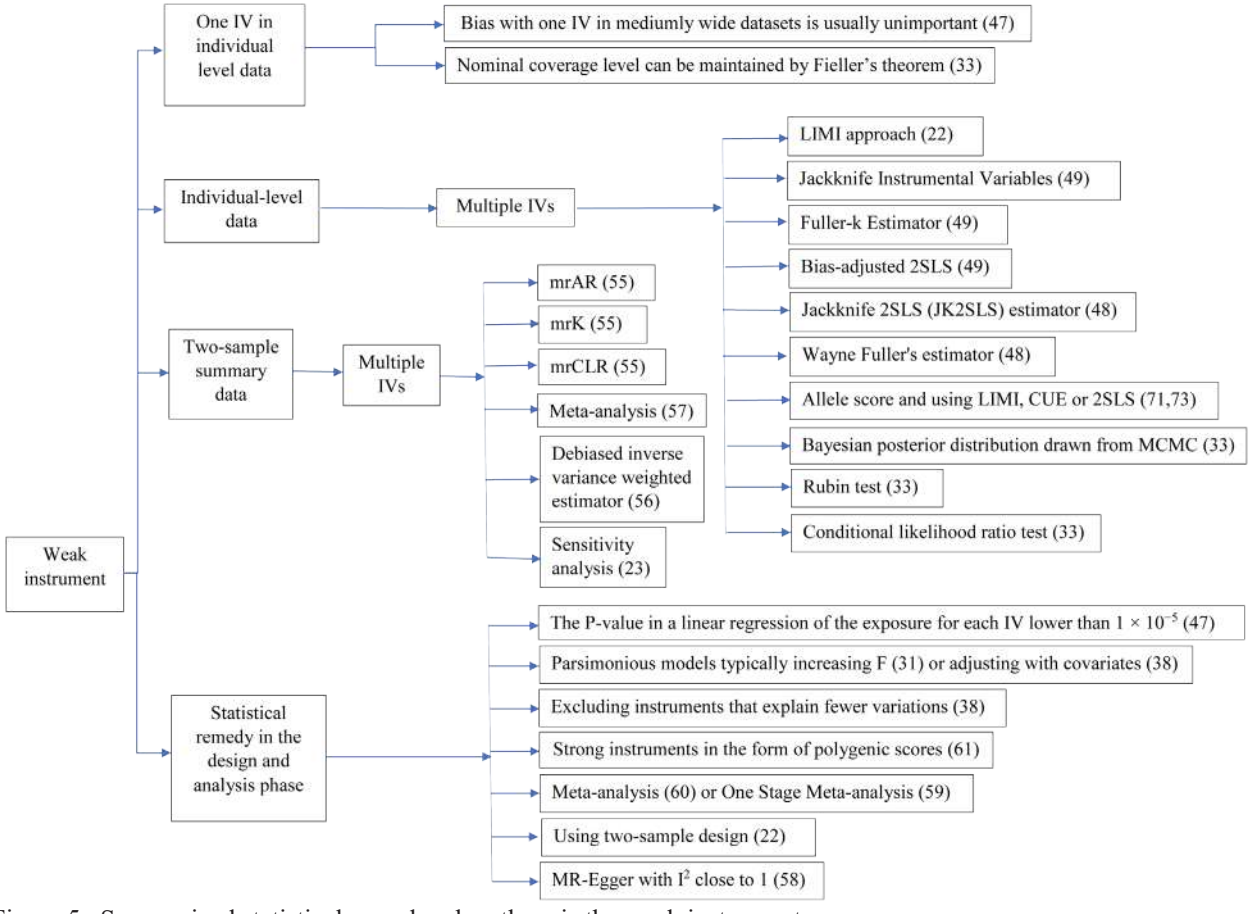

Figure 5. Summarized statistical remedy when there is the weak instrument.

effects in observational studies, a causal relationship could be found between exposure and outcome to simultaneously control the effect of confounding variables of exposure-outcome association, which does this using the IV.[55] The MR method behaved to mimic an RCT. Instead of assigning individuals to interventions, gene variants were used as a specific exposure because of the nature of conception.[56] The advantage of using genetic variants is that they are invariant and stay fixed throughout life. Today, great genetic variants are linked with various diseases and intricate traits found by GWAS;[64,65] that is, the MR can now easily leverage and launch through multiple observational study contexts.[60]

However, there are also important differences between MR and RCT. Alleles were used as IVs in MR usually produce minimal changes in the variable exposure level, while in RCTs, the intervention is much more effective. Another fundamental difference is that in RCT, intervention is introduced at a certain period in life. In contrast, in MR, there will be a change in exposure passed on by inheritance from the time of fertilization. In this way, the MR approach provides beneficial information before conducting a trial. For this reason, it can be argued that even if all MR assumptions are fully valid, an MR study can never prove the success of an environmental intervention.[57–59]

In general, there are two types of IV. First, those that are following the control and randomized by the investigator. Second, those are randomized through inherent. However, the second cause lies in the MR.[17] To be a valid IV in a univariable MR, each genetic variant must meet the core assumption. The IV is associated with the risk factor. The IV is independent of all confounders of the risk factor-outcome association.[66] The IV does not affect the outcome, except possibly through its association with the exposure.[67]

There are multiple accessible MR methods to estimate a causal effect. In general, the selection of MR method relevant to the accessibility of individual or summary data, number of SNPs in an IV, attendance and proportion of pleiotropic variants, strength of an IV, attendance of correlations between SNPs establishing an IV, and the binary or continuous outcome variable.[22,68] Methods can be classified according to whether the data is at the individual level or in summary data. In uncorrelated variants, summarized data are the same, similarly efficient as the individual-level data.[47] The 2SLS, control function estimator, and LIMI methods are at the individual level, whereas the ratio, IVW, weighted median, and MR-Egger methods are used in the summary data. However, it is reported that the 2SLS regression and ratio approach was the most popularly used method.[69]

A weak instrument is a phenomenon where the statistical association between risk factor and IV in the data set is weak and does not explain a large proportion of variation in the risk factor so that IV estimates are biased and the distribution of the IV estimate is weakly approximated through normal distribution.[70]

In general, weak instruments leads to the following problems. Firstly, they may provide small or no benefit information and prepare low statistical power to test hypotheses. Secondly, bias is amplified when the core instrumental variable assumptions violation. Thirdly, create asymptotic approximations for standard errors and confidence intervals; hence, they propose an unstable estimate instead of very stable. Fourthly, large samples are biased towards

between the association of the outcome and exposure in the single sample design and towards the direction of the null in the two sample design.[43,71]

Using the 2SLS method for a constant mean F statistic, it has been shown that the accuracy of the IV estimator proliferates as the number of instruments proliferates. Then the bias proliferates (Bias-variance trade-off). Nevertheless, with the LIML method, the bias did not increase with the number of instruments, but the accuracy is slightly less than 2SLS.[20] In practice, to ensure the large expected F statistic values, the subject of weak instrument bias should be considered before data collection through determining sample sizes, instruments, and genetic model.[20]

Weak instrument bias can be lessened through LIML, bayesian, and allele score methods. Besides, confidence intervals by the Rubin test statistic or Bayesian MCMC methods can be maintained. On the other hand, it is demonstrated that LIML and CUE methods had an unbiased estimate when standard errors were accurate for the attendance of many weak instruments.[72]

Finally, it is demonstrated that the allele score, CUE, and LIML provide accurate inferences under homoscedasticity. The LIML is to be a little more efficient than the CUE when homoscedasticity maintains. In small samples, LIML and CUE are probable to be lower efficient than allele scores. The CUE should use instead of the LIML when conditional heteroscedasticity holds. Moreover, CUE and LIML are less efficient when many instruments and modest samples are available.[52]

Abbreviations

List of common abbreviation terms used in Mendelian randomization studies

Instrumental variable (IV): The IV techniques are a few available ways to estimate the causal effects without fully knowing all the confounders of the exposure-outcome association. Briefly, The IV is referred to as an external variable that is associated with exposure, and it is independent of the outcome as well as any factor linked to the outcome, other than by exposure.

Mendelian randomization (MR): The MR is an analytical method that uses genetic variants as IVs in nonexperimental data for assessing and estimating the causal effects of risk factors. Mendelian randomization is a popular technique that eludes confounding variables in randomize clinical trials (RCTs) and evaluates causal factors for phenotypes that would not suit RCTs. In fact, Mendelian randomization is an application of instrumental variable.

Main assumption in mendelian randomization: To utilize a genetic variant to be a valid instrumental variable, several main assumptions must be satisfied. (i) The IVs are strongly associated with exposure(s) and should be clear quantifiably. (ii) The IVs are not linked with any confounder of the exposure-outcome association. (iii) The IVs do not affect the outcome, except possibly through its association with the exposure(s).

Weak instrument bias: A weak instrument is a scenario that lacks strong statistical evidence of an association between IV and exposure. This phenomenon will occur if the first condition is not met. An IV is weak if it reports a small amount of the variation of the exposure, which is directly related to the sample size.

Pleiotropic effect: Two types were existed. Horizontal pleiotropy is that the outcome

affected through another trait or biological pathway. Vertical pleiotropy is that affect other traits through the risk factor of interest (exposure). Horizontal pleiotropy will occur if the third condition is not met, because of not affect only the outcome by exposure. Horizontal pleiotropy is problematic for Mendelian randomization studies, but vertical pleiotropy is in general not problematic.

One sample Mendelian randomization: The data on the exposure and the outcome are obtained from a single dataset or genetic variants, risk factor, and the outcome are obtained in the same participants.

Two sample mendelian randomization: The exposure and outcome data are obtained from different datasets. The advantages are including, (i) measuring the risk factor or outcome, or both are expensive, (ii) increasing the statistical power, because of synthesizing data from multiple data sources.

Summary statistical data: summary-level data of the IV-exposure and the IV-outcome association, including beta-coefficients and standard errors from linear or logistic regression to get causal effect estimates.

Individual-level data: Each individual's information is accessible.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

None.

Conflicts of interest

The authors declare that they have no competing interests.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Code availability

Not applicable.

Authors' contributions

DH and MA participated in the sequence content and drafted the manuscript. MA conceived and contributed to the rationale for the manuscript. DH, MM, MD and MA contributed to the design of the manuscript. MM, MD and MA read and critically revised the draft manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

We would like to express our gratitude to all of our colleagues at the Research Institute for Endorcrine Research (RIES) and the GEMIRAN research group for their constructive suggestions that helped us enhance the work.

References

1. Dettori J. The random allocation process: two things you need to know. Evid Based Spine Care J. 2010;1(03):7–9.

2. Cornish AJ, Tomlinson IPM, Houlston RS. Mendelian randomisation: a powerful and inexpensive method for identifying and excluding non-genetic risk factors for colorectal cancer. Mol Aspects Med. 2019;69:41–7.

3. Berzuini C, Guo H, Burgess S, Bernardinelli L. A Bayesian approach to Mendelian randomization with multiple pleiotropic variants. Biostatistics. 2020;21(1):86–101.

4. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet. 2014;23(R1):R89–98.

5. Jepsen P, Johnsen SP, Gillman MW, Sørensen HT. Interpretation of observational studies. Heart [Internet]. 2004 Aug;90(8):956–60. Available from: https://pubmed.ncbi.nlm.nih.gov/15253985

6. DiPrete TA, Burik CAP, Koellinger PD. Genetic instrumental variable regression: Explaining socioeconomic and health outcomes in nonexperimental data. Proc Natl Acad Sci. 2018;115(22):E4970–9.

7. Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2000;29(4):722–9.

8. Habibi D, Daneshpour MS, Asgarian S, Kohansal K, Hadaegh F, Mansourian M, et al. Effect of C-reactive protein on the risk of Heart failure: a mendelian randomization study. BMC Cardiovasc Disord. 2023;23(1):1–7.

9. Habibi D, Teymoori F, Ebrahimi N, Fateh ST, Najd-Hassan-Bonab L, Saeidian AH, et al. Causal effect of serum 25 hydroxyvitamin D concentration on cardioembolic stroke: Evidence from Two-sample Mendelian randomization. Nutr Metab Cardiovasc Dis. 2024;

10. Gala H, Tomlinson I. The use of Mendelian randomisation to identify causal cancer risk factors: promise and limitations. J Pathol. 2020;250(5):541–54.

11. Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, et al. Recent developments in Mendelian randomization studies. Curr Epidemiol reports. 2017;4(4):330–45.

12. Evans DM, Davey Smith G. Mendelian randomization: new applications in the coming age of hypothesis-free causality. Annu Rev Genomics Hum Genet. 2015;16:327–50.

13. Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation?

Eur J Epidemiol. 2021;36(3):253–7.

14. Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. PLoS Genet. 2017;13(10):e1006944.

15. Bennett DA. An introduction to instrumental variables–part 2: Mendelian randomisation. Neuroepidemiology. 2010;35(4):307–10.

16. Lee K, Lim C-Y. Mendelian Randomization Analysis in Observational Epidemiology. J Lipid Atheroscler. 2019;8(2):67.

17. Sheehan NA, Didelez V. Epidemiology, genetic epidemiology and Mendelian randomisation: more need than ever to attend to detail. Hum Genet. 2020;139(1):121–36.

18. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–63.

19. Cano-Gamez E, Trynka G. From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. Front Genet. 2020;11:424.

20. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. Genet Epidemiol. 2016;40(7):597–608.

21. Hartwig FP, Davies NM, Hemani G, Davey Smith G. Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique. Int J Epidemiol. 2016;45(6):1717–1726.

22. Elston RC, Satagopan JM, Sun S. Statistical human genetics. Springer; 2012.

23. Rasooly D, Patel CJ. Conducting a Reproducible Mendelian Randomization Analysis using the R analytic statistical environment. Curr Protoc Hum Genet. 2019;101(1):e82.

24. Plotnikov D, Guggenheim JA. Mendelian randomisation and the goal of inferring causation from observational studies in the vision sciences. Ophthalmic Physiol Opt. 2019;39(1):11–25.

25. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. Am J Epidemiol. 2013;178(7):1177–84.

26. Go T-H, Kang DR. Basic Concepts of a Mendelian Randomization Approach. Cardiovasc Prev Pharmacother. 2020;2(1):24–30.

27. Burgess S, Daniel RM, Butterworth AS, Thompson SG, Consortium E-I. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. Int J Epidemiol. 2015;44(2):484–95.

28. 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. 2019;48(3):713–27.

29. Rees JMB, Foley CN, Burgess S. Factorial Mendelian randomization: using genetic variants to assess interactions. Int J Epidemiol. 2020;49(4):1147–58.

30. VanderWeele TJ, Tchetgen EJT, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology. 2014;25(3):427.

31. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. Int J Epidemiol. 2011;40(3):740–52.

32. Ebbes P, Wedel M, Böckenholt U. Frugal IV alternatives to identify the parameter for an endogenous regressor. J Appl Econom. 2009;24(3):446–68.

33. Burgess S, Thompson SG. Mendelian randomization: methods for using genetic variants in causal estimation. CRC Press; 2015.

34. Stock JH, Yogo M. Testing for weak instruments in linear IV regression. Identif inference Econom Model Essays Honor Thomas Rothenb. 2005;80(4.2):1.

35. Cragg JG, Donald SG. Testing identifiability and specification in instrumental variable models. Econom Theory. 1993;222–40.

36. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. Int J Epidemiol. 2013;42(4):1134–44.

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

38. Burgess S, Thompson SG, Collaboration CRPCHDG. Avoiding bias from weak instruments in Mendelian randomization studies. Int J Epidemiol. 2011;40(3):755–64.

39. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. Int J Epidemiol. 2014;43(3):922–9.

40. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. Stat Med. 2016;35(11):1880–906.

41. Nelson CR, Startz R. The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. J Bus. 1990;S125–40.

42. Angrist JD, Krueger AB. Split-sample instrumental variables estimates of the return to schooling. J Bus Econ Stat. 1995;13(2):225–35.

43. Zhao Q, Chen Y, Wang J, Small DS. Powerful three-sample genome-wide design and robust statistical inference in summary-data Mendelian randomization. Int J Epidemiol. 2019;48(5):1478–92.

44. Mikusheva A. Robust confidence sets in the presence of weak instruments. J Econom. 2010;157(2):236–47.

45. Brookhart MA, Rassen JA, Schneeweiss S. Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 2010;19(6):537–54.

46. Ertefaie A, Small DS, Flory JH, Hennessy S. A tutorial on the use of instrumental variables in pharmacoepidemiology. Pharmacoepidemiol Drug Saf. 2017;26(4):357–67.

47. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658–65.

48. Hahn J, Hausman J. Weak instruments: Diagnosis and cures in empirical econometrics. Am Econ Rev. 2003;93(2):118–25.

49. Stock JH, Wright JH, Yogo M. A survey of weak instruments and weak identification in generalized method of moments. J Bus Econ Stat. 2002;20(4):518–29.

50. Hahn J, Hausman J. A new specification test for the validity of instrumental variables. Econometrica. 2002;70(1):163–89.

51. Hansen C, Hausman J, Newey W. Estimation with many instrumental variables. J Bus Econ Stat. 2008;26(4):398–422.

52. Davies NM, von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Smith GD. The many weak instruments problem and Mendelian randomization. Stat Med. 2015;34(3):454–68.

53. Hausman JA, Newey WK, Woutersen T, Chao JC, Swanson NR. Instrumental variable estimation with heteroskedasticity and many instruments. Quant Econom. 2012;3(2):211–55.

54. Jiang L, Oualkacha K, Didelez V, Ciampi A, Rosa-Neto P, Benedet AL, et al. Constrained instruments and their application to Mendelian randomization with pleiotropy. Genet Epidemiol. 2019;43(4):373–401.

55. Wang S, Kang H. Weak-instrument robust tests in two-sample summary-data Mendelian randomization. Biometrics. 2021 Jul;

56. Ye T, Shao J, Kang H. Debiased inverse-variance weighted estimator in two-sample summary-data mendelian randomization. Ann Stat. 2021;49(4):2079–100.

57. Bowden J, Holmes M V. Meta-analysis and Mendelian randomization: A review. Res Synth Methods. 2019;10(4):486–96.

58. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I 2 statistic. Int J Epidemiol. 2016;45(6):1961–74.

59. Schmidt AF, Hingorani AD, Jefferis BJ, White J, Groenwold RHH, Dudbridge F. Comparison of variance estimators for meta-

analysis of instrumental variable estimates. Int J Epidemiol. 2016;45(6):1975–86.

60. Shea J. Instrument relevance in multivariate linear models: A simple measure. Rev Econ Stat. 1997;79(2):348–52.

61. Minică CC, Dolan C V, Boomsma DI, de Geus E, Neale MC. Extending causality tests with genetic instruments: an integration of Mendelian randomization with the classical twin design. Behav Genet. 2018;48(4):337–49.

62. Zou L, Guo H, Berzuini C. Overlapping-sample Mendelian randomisation with multiple exposures: a Bayesian approach. BMC Med Res Methodol. 2020;20(1):1–15.

63. Byrne BE, Rooshenas L, Lambert HS, Blazeby JM. A mixed methods case study investigating how randomised controlled trials (RCTs) are reported, understood and interpreted in practice. BMC Med Res Methodol. 2020;20:1–12.

64. Hosseini-Esfahani F, Zahedi AS, Akbarzadeh M, Seyedhamzehzadeh A, Daneshpour MS, Mirmiran P, et al. The resemblance of dietary intakes in three generations of parent-offspring pairs: Tehran lipid and glucose study. Appetite. 2021;105794.

65. Akbarzadeh M, Dehkordi SR, Roudbar MA, Sargolzaei M, Guity K, Sedaghati-Khayat B, et al. GWAS findings improved genomic prediction accuracy of lipid profile traits: Tehran Cardiometabolic Genetic Study. Sci Rep. 2021;11(1):1–9.

66. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nat Commun. 2020;11(1):1–11.

67. 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.

68. Akbarzadeh M, Moghimbeigi A, Morris N, Daneshpour MS, Mahjub H, Soltanian AR. A Bayesian structural equation model in general pedigree data analysis. Stat Anal Data Min ASA Data Sci J. 2019;12(5):404–11.

69. Boef AGC, Dekkers OM, Le Cessie S. Mendelian randomization studies: a review of the approaches used and the quality of reporting. Int J Epidemiol. 2015;44(2):496–511.

70. Myers JA, Rassen JA, Gagne JJ, Huybrechts KF, Schneeweiss S, Rothman KJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. Am J Epidemiol. 2011;174(11):1213–22.

71. Davies NM, Holmes M V, Smith GD. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. Bmj. 2018;362.

72. Hartwig FP, Davies NM. Why internal weights should be avoided (not only) in MR-Egger regression. Int J Epidemiol. 2016;45(5):1676–8.