



Common Methods for Performing Mendelian Randomization

Alexander Teumer^{1,2*}

¹ Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany, ² Partner Site Greifswald, Deutsches Zentrum für Herz-Kreislaufforschung (DZHK), Greifswald, Germany

Mendelian randomization (MR) is a framework for assessing causal inference using cross-sectional data in combination with genetic information. This paper summarizes statistical methods commonly applied and strait forward to use for conducting MR analyses including those taking advantage of the rich dataset of SNP-trait associations that were revealed in the last decade through large-scale genome-wide association studies. Using these data, powerful MR studies are possible. However, the causal estimate may be biased in case the assumptions of MR are violated. The source and the type of this bias are described while providing a summary of the mathematical formulas that should help estimating the magnitude and direction of the potential bias depending on the specific research setting. Finally, methods for relaxing the assumptions and for conducting sensitivity analyses are discussed. Future researches in the field of MR include the assessment of non-linear causal effects, and automatic detection of invalid instruments.

OPEN ACCESS

Edited by:

Tanja Zeller,
Universität Hamburg, Germany

Reviewed by:

Bastiaan Geelhoed,
University Medical Center Groningen,
Netherlands

Joylene Elisabeth Siland,
University of Groningen, Netherlands

*Correspondence:

Alexander Teumer
ateumer@uni-greifswald.de

Specialty section:

This article was submitted to
Cardiovascular Genetics and
Systems Medicine,
a section of the journal
Frontiers in Cardiovascular Medicine

Received: 21 February 2018

Accepted: 04 May 2018

Published: 28 May 2018

Citation:

Teumer A
(2018) Common Methods for
Performing Mendelian
Randomization.
Front. Cardiovasc. Med. 5:51.
doi: 10.3389/fcvm.2018.00051

Keywords: mendelian randomization, causal inference, GWAS, bias, statistical methods

INTRODUCTION

Observational epidemiological studies made important contributions to our understanding of common diseases by identifying important risk factors. Although causal inference is of major interest as it builds a basis for intervention and prevention, it is difficult to perform using observational data from cross-sectional studies. Supposed causality was often revised e.g., by randomized controlled trials (RCTs) (1). Possible reasons for these contradicting findings include unobserved confounding, reverse causation and selection bias in the observational studies (2–4).

On the other hand, RCTs are often subject to long duration and ethical problems. Furthermore, confounding and selection bias is still a problem after the initiation of a RCT. This includes compliance problems or missing of follow-up information depending on treatment effect which may induce missing not at random problems.

During the last decade, huge efforts were undertaken searching for genetic risk factors underlying common traits and diseases. Genome-wide association studies (GWAS) revealed thousands of genetic associations predominantly based on single nucleotide polymorphisms (SNPs) including more than 950 related to cardiovascular diseases and measurements (by April 2018) and were made publically available (5). The effect sizes of these associations were often quite small (6–8), and thus their direct clinical relevance might be questioned. However, these genetic associations may help drawing causal inferences. This approach in which SNPs are used as instrumental variables (IVs) for specific exposures is called Mendelian randomization (MR) (9). By the Mendelian laws, alleles of SNPs segregate and are randomly inherited from parents to offspring. This principle can be seen analogously to the randomized treatment assignment in a RCT resulting in an unconfounded exposure-outcome relationship. Within an MR approach, the exposure represents a continuous or dichotomous risk factor of a disease, and

the outcome is the disease or a disease-related trait. These traits may e.g., be blood pressure defining hypertension, or estimated glomerular filtration rate (eGFR) defining the status of chronic kidney disease. Using the MR approach, causality between exposure and outcome can be tested. During recent years, the number of MR studies to assess causality increased substantially which includes also the field of cardiovascular diseases and nephrology (10–14). Furthermore, MR analyses revealed causal effects of blood lipids on coronary heart disease (15) as well as of alcohol consumption on cardiovascular traits (16). However, given the number of potential genetic instruments and statistical methods available nowadays, there is potential for assessing causality of many more traits by conducting successful MR analyses. Nevertheless, some important assumptions have to be fulfilled to be able to estimate an unconfounded and unbiased exposure-outcome relationship thus allowing drawing causal inference. This review describes the assumptions of MR and potential biases caused by violation of these assumptions, and provides an overview of commonly applied statistical methods for conducting MR analyses using individual level data as well as using GWAS meta-analyses results.

Estimation of the Causal Effect

The general aim of the MR approach is the estimation of a causal effect of an exposure X on an outcome Y using (one or more) genetic instruments Z for X (Figure 1). Basically, the causal effect will be obtained by two sequential steps. First, the exposure is estimated from its instruments. By using valid instruments, the estimated exposure will be independent of any confounders. In the second step, the outcome is regressed on this estimated exposure thus obtaining an unconfounded and therefore causal

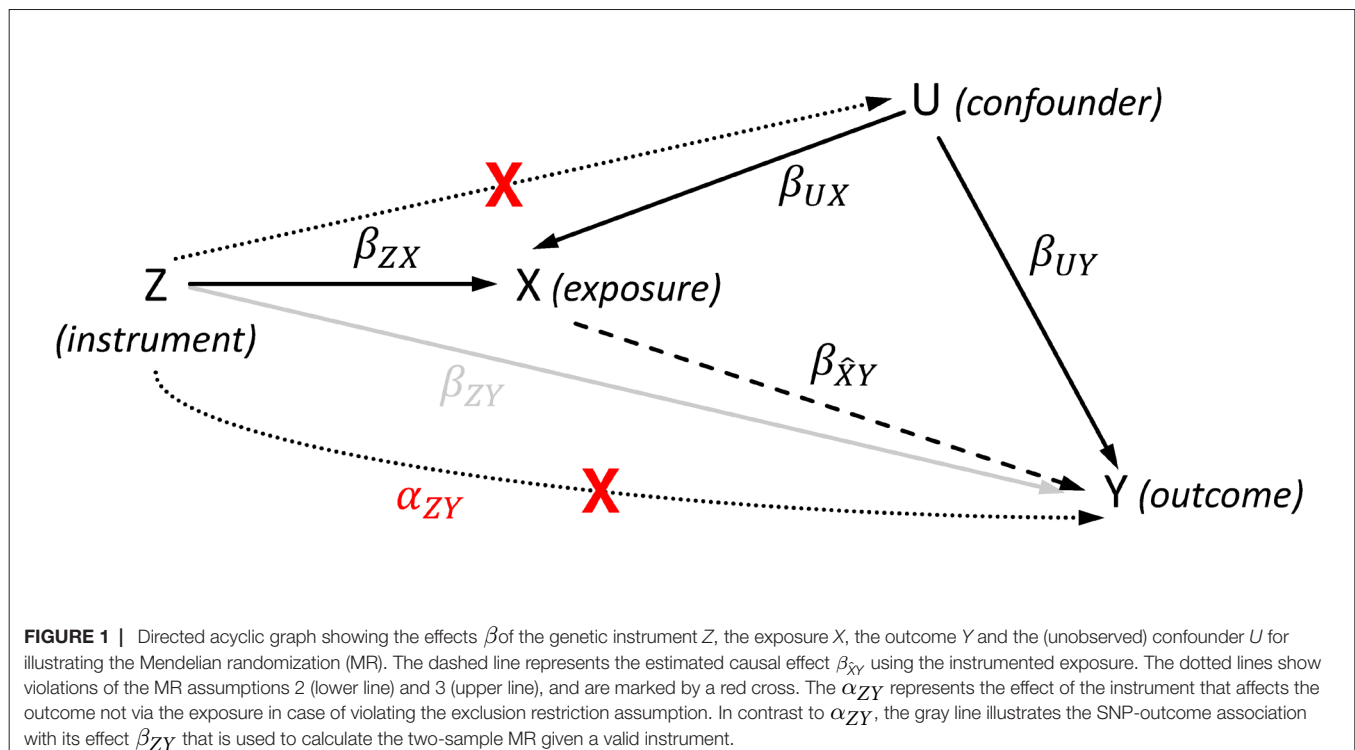
effect estimate. The instrument Z is usually coded by 0, 1 and 2 per individual according to its number of coding (e.g., exposure increasing) alleles.

2-Stage Least Squares Estimator

Given a continuous outcome Y and assuming linear effects between X and Y without interaction, the causal estimate of the exposure X on Y can be estimated through a 2-stage least squares (2SLS) regression. This method performs both steps described before implicitly. In the first step, the exposure \hat{X} which is independent of the confounders is estimated via the genotypes of the instruments by calculating the fitted values from the regression of X on Z . In the second step, the causal effect estimate $\beta_{\hat{X}Y}$ is obtained by regressing Y on \hat{X} . As both steps are performed in a single model instead of two separate regressions, the variation of both Z and \hat{X} is taken into account which is required for obtaining correct standard errors (SE) of $\beta_{\hat{X}Y}$ (17). The 2SLS regression can be calculated by standard methods in statistical software packages like R (18) using the function *tsls* of the package SEM, or by the STATA software (<https://www.stata.com/>) using the command *ivregress*. The 2SLS was included in an MR of testosterone with cardiometabolic risk factors, but the single study analysis limited the statistical power substantially (19).

Ratio Estimator

Alternatively, the causal effect can be estimated by triangulation without the need of calculating $\beta_{\hat{X}Y}$ from the exposure-outcome association directly. The principle of this method is illustrated through Figure 1: the standard approach (including 2SLS) for



obtaining the causal effect β_{XY} follows the path from the instrument Z via X to Y . In this case, the direct effect β_{ZY} of the instrument on the outcome Y equals the product of effects underlying the path mediated by the exposure, i.e., $\beta_{ZY} = \beta_{ZX} \cdot \beta_{XY}$. By rearranging this equation, the causal effect can be estimated through dividing the effect of the IV on the outcome (β_{ZY}) by the effect of the IV on the exposure (β_{ZX}): $\beta_{XY} = \beta_{ZY} / \beta_{ZX}$. As the triangulation approach calculates the causal effect (and its SE for testing significant deviation from null) by the ratio of the two IV based effect estimates, it is also known as ratio estimate or Wald estimate. It is important for the computation that both IV based effect estimates refer to the same allele of the IV. Furthermore, the same requirements as for the 2SLS apply. The SE of β_{XY} has to be estimated via the delta-method which is based on a Taylor series expansion, and can be approximated as (20):

$$\text{var}(\beta_{\hat{XY}}) = \text{var}\left(\frac{\beta_{ZY}}{\beta_{ZX}}\right) \cong \frac{\text{var}(\beta_{ZY})}{\beta_{ZX}^2} + \frac{\beta_{ZY}^2}{\beta_{ZX}^4} \text{var}(\beta_{ZX}) - 2 \frac{\beta_{ZY}}{\beta_{ZX}^3} \text{cov}(\beta_{ZY}, \beta_{ZX})$$

$SE(\beta_{\hat{XY}}) = \sqrt{\text{var}(\beta_{\hat{XY}})}$, where $\text{cov}(\beta_{ZY}, \beta_{ZX})$ is the covariance of the two effect estimates. This term will vanish if the effect estimates are obtained from distinct samples. That concise approximation can be easily implemented for significance testing in statistical software packages like R or STATA.

In contrast to the 2SLS which has to be performed using data of a single sample (one-sample MR), different sample sets can be used for conducting the triangulation: the effect estimates of the IV on exposure X and outcome Y can be obtained from genetic association studies with either disjunct or overlapping samples (two-sample MR). By this means, genetic associations revealed through large-scale GWAS meta-analyses can be used as β_{ZX} and β_{ZY} . These association results are often publically available for a variety of traits.

The triangulation method can also be applied if the outcome Y is dichotomous, i.e., an indicator of a disease status. In this case, log-linear effects without interaction on Y and an approximately normal distribution of X are required. Causal effect estimates on the odds ratio (OR) scale can be calculated by performing a logistic regression analysis using the disease as outcome. This model was also applied in most GWAS. To estimate causal OR using triangulation, the rare disease assumption (i.e., prevalence <10%) has to be fulfilled. Alternatively, estimates of a causal risk ratio may be calculated using a log-linear model instead of a logistic regression (21). The SE of the β_{XY} (i.e., the log causal OR) will be estimated by the same formula as applied in the case of a continuous outcome. An application of the ratio estimator is provided by the MR on statin c and cardiovascular disease (22).

Control Function Estimator

Another method for estimating the causal effect on a dichotomous outcome is provided through the control function estimator (21) which is a two-step approach. In the first step, the exposure X is regressed on the instruments Z . The residuals of the regression

correspond to the non-instrumented part of the exposure and may therefore correlate with a (unobserved) confounder U of the exposure-outcome association. In the second step, a logistic regression of the outcome Y on X is performed, adding the residuals of the first step as a covariate to the model. By adding the residuals of the first step into the model, the effects of U on Y will be controlled. Thus, the effect of X on Y of the second regression corresponds to the causal effect estimate. In case a linear regression is conducted in the second step (i.e., for a continuous outcome), the control function estimator is equivalent to the 2SLS estimator (21). This type of MR was conducted for assessing the causal effect of blood lipids on coronary heart disease (15).

Assumptions of the Instrumental Variables

SNPs have several properties predisposing them for instruments of the exposure. The inherited alleles are not changed by a disease or trait and thus also do not change over time. The random inheritance of the SNP alleles makes the genotype distribution mostly independent from socio-economic and lifestyle factors (1, 23). Nevertheless, specific assumptions still need to be fulfilled to ensure the validity of the genetic variant as an instrument. There are three core assumptions for MR (24–26):

1. The genetic variant is associated with the exposure
2. The genetic variant is independent of the outcome given the exposure and all confounders (measured and unmeasured) of the exposure-outcome association
3. The genetic variant is independent of factors (measured and unmeasured) that confound the exposure-outcome relationship

The first condition is required because within the MR the (unconfounded) exposure will be estimated using the allele distribution of the IVs. This assumption can be easily tested, and is considered as fulfilled if the SNP-exposure association has an F-statistic >10 (21, 27).

The second assumption, which is also known as exclusion restriction, is equivalent to the condition that an IV does not have an effect on the outcome when the exposure remains fixed. In general, this assumption is hard to validate as there may be pleiotropic effects of SNPs or SNPs in linkage disequilibrium correlated with genes that have effects on the outcome independently of the exposure. Even without considering the linkage disequilibrium, using SNPs of the pleiotropic gene *GCKR* exemplarily as instruments for kidney function to assess a causal effect on blood pressure would result in an invalid IV as there are effects of *GCKR* on blood pressure likely that are independent of kidney function, e.g., by the known associations of *GCKR* with serum lipid levels. Another violation would occur if the sample consists of a population substructure with different allele distributions, and which is also associated with the outcome. In this case, the substructure would be a common cause of both SNP and outcome opening a pathway from SNP to outcome not mediated by the exposure. Several examples of different scenarios violating the exclusion restriction are provided in the work of Glymour et al. (24).

The third assumption is also hard to validate. Similar problems due to pleiotropy and population substructure as described in the exclusion restriction may occur but affecting confounders of the

exposure-outcome relationship instead of the outcome directly. In an example of assessing causality of kidney function with heart disease, using *GCKR* as an instrument would violate the third assumption because these SNPs are also associated with blood pressure being a confounder of the association of kidney function and heart disease.

Weak Instrument Bias

Until today, more than 50,000 SNP-trait associations were revealed by GWAS and are usually accessible through public repositories like the GWASCatalog (5). These SNPs can be considered as potential instruments for MR analyses. Because the majority of these SNPs explain only a small proportion (i.e. <1%) of the phenotypic variance, GWAS with sample sizes of more than 10,000 or 100,000 individuals were required to unravel these associations at the level of genome-wide significance. However, the small effect sizes of the SNPs on the exposure result in weak instruments when using smaller sample sizes (28). Weak instruments tend to lead estimated causal effects towards the observational association (27). The reason for this bias is originated in using finite sample sizes. Although the IVs are asymptotically independent of confounders, there might be still an association by chance in finite samples. Increasing the sample size or the strength of the instruments will reduce the weak instrument bias. To illustrate the origin and the effect of the bias, let β_{UX} and β_{UY} be the effects of the confounder U on the exposure and the outcome, respectively (Figure 1). Furthermore, let Δ_U be the (by chance) difference in U depending on the instrument Z . The estimated causal effect $\beta_{\hat{XY}}$ can then be computed by the following sum of effects (27):

$$\beta_{\hat{XY}} = \beta_{causal} + \frac{\beta_{UY}\Delta_U}{\beta_{ZX} + \beta_{UX}\Delta_U}$$
 where β_{causal} is the true causal effect, and the $\text{mean}(\Delta_U) = 0$ because Z is an instrument (assumption 3). This leads the bias term towards zero with increasing sample size resulting in $\beta_{\hat{XY}} = \beta_{causal}$. The estimated causal effect is also close to the true causal effect in case the effect of the IV on the exposure β_{ZX} is relatively large compared to the by-chance difference in U on the exposure ($\beta_{UX}\Delta_U$). However, if β_{ZX} is small compared to $\beta_{UX}\Delta_U$ (in case of a weak instrument), the estimated causal effect will be biased towards the ratio of the effect of the confounder on the outcome and the effect of the confounder on the exposure, i.e., $\frac{\beta_{UY}}{\beta_{UX}}$.

Multiple Instruments Approach

Using multiple valid instruments will help to address the weak instrument bias. Adding multiple uncorrelated (linkage equilibrium) SNPs into a 2SLS model can increase the statistical power but might also increase the relative bias if weak instruments are added (28).

Alternatively, an allele score can be generated from the instruments and included as a single variable in the association model. This allele score Z is calculated per individual as the weighted or unweighted sum of the number of risk or trait increasing alleles Z_i of each SNP i , whereas the effect β_{Z_iX} of each SNP on the exposure X is used as weight: $Z = \beta_{Z_1X}Z_1 + \beta_{Z_2X}Z_2 + \dots + \beta_{Z_kX}Z_k$. In case of an

unweighted score where all β_{Z_iX} are set to 1, the allele score of an individual simplifies to the sum of its risk alleles. By using an allele score, the F-statistics increases because of the smaller degrees of freedom in the model. However, it has been shown that the unweighted score has lower power than adding multiple IVs into the 2SLS, but using an appropriately weighted allele score performs similarly. The causal effect is a little less biased when using a weighted allele score but might have a slightly lesser precision (and power) compared to the multiple IV 2SLS estimator. In general, effects obtained from external studies should be used as weights (28).

A third method for taking advantage of multiple IVs is to combine ratio estimates (triangulation) of single instruments using inverse variance weighting (29, 30). The method for combining the results is the same as used for meta-analyses, and is for example implemented in the R package *metafor*. Alternatively, the following simplification of this calculation can be used (31–33):

$$\beta_{\hat{XY}} = \frac{\sum \beta_{ZX}\beta_{ZY}\text{var}(\beta_{ZY})^{-1}}{\sum \beta_{ZX}^2\text{var}(\beta_{ZY})^{-1}}$$

with its approximated $SE(\beta_{\hat{XY}}) = \sqrt{\frac{1}{\sum \beta_{ZX}^2\text{var}(\beta_{ZY})^{-1}}}$, where the sum runs over the SNP specific estimates. This method is implemented in the R package *gtx*.

However, it is crucial that the effects of all IVs used in the calculation are corresponding to the allele referring to the same effect direction on the exposure (e.g., the trait increasing allele). In theory, problems of missing data may occur especially when using multiple IVs. Nowadays well established methods for imputing missing genotypes based on the linkage disequilibrium structure of the human genome are available to circumvent this problem (34–36).

Bias by Violation of the Assumptions 2 and 3

Importantly, valid instruments need to be included in the MR analyses. In case the assumptions are not fulfilled, different types of bias can occur leading to invalid causal effect estimates. Violation of the second assumption (the exclusion restriction) implies that there is at least a partial effect of the instrument on the outcome not mediated by the exposure, i.e. $\alpha_{ZY} \neq 0$ (Figure 1). Depending on the direction and strength of these pleiotropic effects, the causal effect will be over- or underestimated. As shown within the principle of triangulation, the estimated causal effect $\beta_{\hat{XY}}$ is the sum of the true causal effect β_{causal} and a bias term: $\beta_{\hat{XY}} = \beta_{causal} + \frac{\alpha_{ZY}}{\beta_{ZX}}$ (26). The bias increases due to larger pleiotropy (larger absolute α_{ZY} in the nominator) or weaker instruments (smaller absolute β_{ZX} in the denominator). Violation of assumption 3 leads to a bias similar to the weak instrument bias. In this case, the effect of the confounder U on exposure and outcome will not vary by chance but systematically because of the non-zero effect of the instrument Z on U . Thus, an increasing sample size will not remove the bias because $\text{mean}(\Delta_U) \neq 0$.

InSIDE Condition and Egger MR

Pleiotropic effects α_{ZY} of each IV will also be included in the model when applying the multiple instruments approach. However, in this scenario it is possible to substitute the exclusion restriction by a weaker assumption as explained below. If the ratio estimates of multiple instruments are combined via 2SLS or inverse variance weighting, equation (1) will result in $\beta_{\hat{X}Y} = \beta_{causal} + \frac{\sum \beta_{ZX}\alpha_{ZY}\text{var}(\beta_{ZY})^{-1}}{\sum \beta_{ZX}^2\text{var}(\beta_{ZY})^{-1}}$, where β_{causal} equals the right side of (1) and $\frac{\sum \beta_{ZX}\alpha_{ZY}\text{var}(\beta_{ZY})^{-1}}{\sum \beta_{ZX}^2\text{var}(\beta_{ZY})^{-1}}$ is a bias depending on

α_{ZY} and β_{ZX} . Thus, an unbiased causal effect will be obtained if the assumption 2 is true, i.e., all direct effects α_{ZY} of each IV on the outcome Y are zero. However, it will be sufficient for the bias term to equal zero if all pleiotropic effects α_{ZY} of all genetic IV cancel out. As shown below, this cancellation is sufficiently fulfilled if the correlation between direct genetic effects α_{ZY} on the outcome and their effects β_{ZX} on the exposure X (i.e., the strength of the IV) is zero. This independence between the genetic effects α_{ZY} and β_{ZX} is called InSIDE condition (Instrument Strength Independent of Direct Effect). If the InSIDE condition holds together with assumptions 1 and 3, an adaption of the Egger regression can be used to obtain a consistent causal estimate even for specific cases in which the exclusion restriction criteria is violated. The Egger regression for MR is an implementation of the meta-regression where the (total) SNP-outcome effect $\Gamma = \beta_{ZX}\beta_{causal} + \alpha_{ZY}$ for each SNP is regressed on the corresponding SNP-exposure effect β_{ZX} : $\Gamma \sim \beta_{0E} + \beta_E\beta_{ZX}$ where the slope β_E is the bias-reduced causal estimate (Figure 2). The principle behind this regression is that

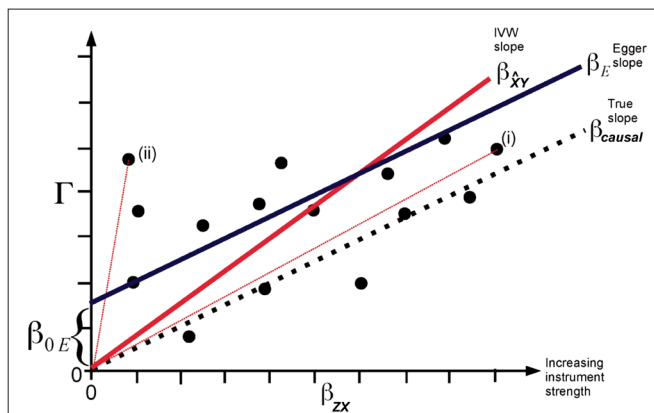


FIGURE 2 | Plot of the SNP-outcome (Γ) on the y-axis vs. the SNP-exposure (β_{ZX}) regression coefficients of potential genetic instruments (i.e., SNPs) of a Mendelian randomization analysis on the x-axis. The true causal effect represented by the slope β_{causal} is shown by a dotted line, the inverse variance weighted (IVW) causal estimate β_{IVW} by a red line, and the MR Egger regression estimate β_E by a dark blue line. The total SNP-outcome effect Γ is proportional to β_{ZX} for valid instruments. In case of invalid instruments but when the InSIDE assumption holds, stronger instruments are on average expected to be closer to the true causal effect (i) than weak instruments (ii). The intercept β_{0E} represents the overall directional pleiotropy of the instruments. The figure was adapted from the publication of Bowden et al., *Int J Epidemiol.* 2015;44(2):512–525 (26) (Creative Commons CC BY license).

Γ is proportional to the strength of the instrument β_{ZX} with the intercept $\beta_{0E} = 0$ for valid instruments, whereas under the InSIDE condition (i.e., α_{ZY} and β_{ZX} are uncorrelated) stronger instruments are expected to have a relatively small bias and thus are on average closer to the true causal effect than weak instruments. As the slope of the Egger MR can be calculated by the least squares estimator $\beta_E = \frac{\text{cov}(\Gamma, \beta_{ZX})}{\text{var}(\beta_{ZX})} = \beta_{causal} + \frac{\text{cov}(\beta_{ZX}, \alpha_{ZY})}{\text{var}(\beta_{ZX})}$, the bias term will be zero if α_{ZY} and β_{ZX} are uncorrelated, which is the case under the InSIDE assumption. A non-zero intercept β_{0E} indicates an overall directional pleiotropy of the IVs (26).

Considering Statistical Power

The statistical power of an MR strongly depends on the proportion of variance of the exposure that is explained by the IV. The use of multiple IVs, either by direct inclusion or as an allele score in the model, may therefore increase the power as more variance of the exposure is explained. However, the validity of these instruments has to be ensured (37). Two-sample MR additionally provide a possibility to increase statistical power if published GWAS meta-analyses of both the exposure and the outcome are available. In this case, effect estimates based on large sample sizes of independent studies can be used to estimate the causal effect. Formulas for performing power calculations of MR using single instruments or allele scores are provided in the study of Burgess (37). Brion et al. (38) discusses the statistical power in case of single IV and continuous outcomes for 2SLS MR, and provide an online power calculator for both continuous and binary outcomes which is available at <http://cns.genomics.com/shiny/mRnd/>. A tool for estimating statistical power of complex MR settings based on simulations is MR_predictor (39), whereas the PERL scripts required to run the estimator are available via GitHub.

Measurement Unit of the Causal Effect Estimates

When conducting two-sample MR, the causal effect corresponds to the unit of the outcome on a per unit change of the exposure that was used in the respective genetic association study of the IV with the corresponding trait (32). Some GWAS were meta-analyzed using the sample-size weighted z-score method (40) and thus do not provide effect estimates that can be directly included in a two-sample MR. However, it is possible to estimate the effect $\hat{\beta}$ for each SNP in Hardy-Weinberg equilibrium using its minor allele frequency MAF , its (large) GWAS sample size N , and its z-statistics z (which can be calculated from the inverse of the standard normal distribution using the association p-value and the corresponding effect direction) through the formula (41): $\hat{\beta} \approx z \cdot \frac{\sigma}{\sqrt{N \cdot 2 \cdot MAF \cdot (1 - MAF)}}$, whereas the corresponding

$SE(\hat{\beta}) = \frac{\hat{\beta}}{z}$. The SD σ of the trait can be set to 1 for standardizing the phenotype (i.e., the effect corresponds to a change of one SD of the trait unit). If the outcome is a binary trait, e.g., a disease with prevalence p in the sample, then $\sigma = \sqrt{p \cdot (1 - p)}$.

DISCUSSION

MR provides a method for testing causality of different traits using cross-sectional data and genetics. Although large sample sizes are required to achieve sufficient statistical power for revealing causal effects, it is often possible to overcome this limitation by using the publically available genetic association results of large GWAS meta-analyses conducted during the last decade.

The statistical methods needed for conducting MR analyses are implemented in common statistical software frameworks. Additionally, the MRbase platform provides a possibility to conduct two-sample MR analyses both online and via the R package *TwoSampleMR*, including the methods discussed in this article (42). A detailed overview of different statistical methods for calculating MR is provided in the review of Burgess et al. (17).

However, it is important that the genetic associations that are used as instruments fulfil the MR assumptions to avoid calculation of biased or spurious causal estimates resulting in false causal inferences. Other than the required strong association of the genetic variant with the exposure, the remaining two assumptions are in general hard to validate.

This review emphasizes the bias that may occur by using invalid instruments, whereas the presented formulas should help estimating the magnitude and direction of this bias depending on the specific MR study that needs to be conducted. Using multiple instruments can help to test the violation of the MR assumptions which may occur due to pleiotropy and via SNPs in linkage disequilibrium (but not for a violation due to population stratification) (28), or to conduct sensitivity analyses (25). A strategy for assessing pleiotropy and population substructure specifically to MR analyses is discussed for example in the work of Lawler et al. (9). The Egger regression can be used as a multiple IV approach to relax the exclusion restriction criteria, and as a sensitivity analysis to test the robustness of the causal association (26). However, if the Egger MR-specific InSIDE assumption is violated, a biased causal estimate and an increased Type I error rate may occur (43). Thus, seeking for genetic variants that are valid IV should be performed as far as possible. Knowledge of the physiology or the biological pathways of the SNPs and their causal genes might be useful for selecting instruments.

REFERENCES

- Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* (2005) 330(7499):1076–9. doi: 10.1136/bmj.330.7499.1076
- Lawlor DA, Davey Smith G, Kundu D, Bruckdorfer KR, Ebrahim S. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? *Lancet* (2004) 363(9422):1724–7. doi: 10.1016/S0140-6736(04)16260-0
- Lawlor DA, Davey Smith G, Ebrahim S. Commentary: the hormone replacement-coronary heart disease conundrum: is this the death of observational epidemiology? *Int J Epidemiol* (2004) 33(3):464–7. doi: 10.1093/ije/dyh124
- Pilz S, Verheyen N, Gröbler MR, Tomaschitz A, März W. Vitamin D and cardiovascular disease prevention. *Nat Rev Cardiol* (2016) 13(7):404–17. doi: 10.1038/nrcardio.2016.73

The methods summarized in this review assume linear effects between exposure and outcome (or log-linear in case of a binary outcome) without effect modifications by the variables. Addressing these limitations is subject to future research. A method for successfully revealing non-linear causal effects was provided in an example for alcohol intake on cardiovascular traits, but this approach is restricted to additional assumptions and limitations (16). With respect to the presence of effect modifications, other statistical methods for conducting binary outcome MR like structural mean models or generalized method of moments make weaker assumptions but still not solve this issue completely (21). Finally, methods for automatically detecting invalid instruments (i.e., due to pleiotropy) are under development (44). Selection of valid instruments still remains a main challenge for automated causal inference.

AUTHOR CONTRIBUTIONS

AT designed and wrote the review.

FUNDING

AT was supported by the SHIP study which is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. We acknowledge support for the Article Processing Charge from the DFG (German Research Foundation, 393148499) and the Open Access Publication Fund of the University of Greifswald.

ACKNOWLEDGMENTS

The author thanks the reviewers for helpful comments on the manuscript.

- MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* (2017) 45(D1):D896–D901. doi: 10.1093/nar/gkw1133
- Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med* (2010) 363(2):166–76. doi: 10.1056/NEJMra0905980
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* (2009) 461(7265):747–53. doi: 10.1038/nature08494
- Zuk O, Schaffner SF, Samocha K, do R, Hechter E, Kathiresan S, et al. Searching for missing heritability: designing rare variant association studies. *Proc Natl Acad Sci U S A* (2014) 111(4):E455–E464. doi: 10.1073/pnas.1322563111
- Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* (2008) 27(8):1133–63. doi: 10.1002/sim.3034
- Sekula P, del Greco M F, Pattaro C, Köttgen A. Mendelian Randomization as an Approach to Assess Causality Using Observational Data. *J Am Soc Nephrol* (2016) 27(11):3253–65. doi: 10.1681/ASN.2016010098

11. Pattaro C. Genome-wide association studies of albuminuria: towards genetic stratification in diabetes? *J Nephrol* (2017) 0(0). doi: 10.1007/s40620-017-0437-3
12. Lanktree MB, Thériault S, Walsh M, Paré G. Cholesterol HDL. HDL Cholesterol, LDL Cholesterol, and Triglycerides as Risk Factors for CKD: A Mendelian Randomization Study. *Am J Kidney Dis* (2018) 71(2):166–72. doi: 10.1053/j.ajkd.2017.06.011
13. Ding EL, Song Y, Manson JE, Hunter DJ, Lee CC, Rifai N, et al. Sex hormone-binding globulin and risk of type 2 diabetes in women and men. *N Engl J Med* (2009) 361(12):1152–63. doi: 10.1056/NEJMoa0804381
14. van der Laan SW, Fall T, Soumaré A, Teumer A, Sedaghat S, Baumert J, et al. Cystatin C and Cardiovascular Disease: A Mendelian Randomization Study. *J Am Coll Cardiol* (2016) 68(9):934–45. doi: 10.1016/j.jacc.2016.05.092
15. Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J* (2015) 36(9):539–50. doi: 10.1093/eurheartj/ehv571
16. Silverwood RJ, Holmes MV, Dale CE, Lawlor DA, Whittaker JC, Smith GD, et al. Testing for non-linear causal effects using a binary genotype in a Mendelian randomization study: application to alcohol and cardiovascular traits. *Int J Epidemiol* (2014) 43(6):1781–90. doi: 10.1093/ije/dyu187
17. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* (2017) 26(5):2333–55. doi: 10.1177/0962280215597579
18. Core Team R. R: A Language and Environment for Statistical Computing. <https://www.r-project.org> (2016).
19. Haring R, Teumer A, Völker U, Dörr M, Nauck M, Biffar R, et al. Mendelian randomization suggests non-causal associations of testosterone with cardiometabolic risk factors and mortality. *Andrology* (2013) 1(1):17–23. doi: 10.1111/j.2047-2927.2012.00002.x
20. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using "Mendelian triangulation" by Bautista et al. *Ann Epidemiol* (2007) 17(7):511–3. doi: 10.1016/j.annepidem.2006.12.005
21. Palmer TM, Sterne JA, Harbord RM, Lawlor DA, Sheehan NA, Meng S, et al. Instrumental variable estimation of causal risk ratios and causal odds ratios in Mendelian randomization analyses. *Am J Epidemiol* (2011) 173(12):1392–403. doi: 10.1093/aje/kwr026
22. van der Laan SW, Fall T, Soumaré A, Teumer A, Sedaghat S, Baumert J, et al. Cystatin C and Cardiovascular Disease. *J Am Coll Cardiol* (2016) 68(9):934–45. doi: 10.1016/j.jacc.2016.05.092
23. Hingorani A, Humphries S. Nature's randomised trials. *Lancet* (2005) 366(9501):1906–8. doi: 10.1016/S0140-6736(05)67767-7
24. Glymour MM, Tchetgen Tchetgen EJ, Robins JM. Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *Am J Epidemiol* (2012) 175(4):332–9. doi: 10.1093/aje/kwr323
25. Didelez V, Meng S, Sheehan NA. Assumptions of IV Methods for Observational Epidemiology. *Statistical Science* (2010) 25(1):22–40. doi: 10.1214/09-STS316
26. 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–. doi: 10.1093/ije/dyv080
27. Burgess S, Thompson SG CRP CHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol* (2011) 40(3):755–64. doi: 10.1093/ije/dyr036
28. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* (2012) 21(3):223–42. doi: 10.1177/0962280210394459
29. Schmidt AF, Swerdlow DI, Holmes M V, Patel RS, Fairhurst-Hunter Z, Lyall DM. CSK9 genetic variants and risk of type 2 diabetes: a mendelian randomisation study. *lancet Diabetes Endocrinol* (2016).
30. Østergaard SD, Mukherjee S, Sharp SJ, Proitsi P, Lotta LA, Day F, et al. Associations between Potentially Modifiable Risk Factors and Alzheimer Disease: A Mendelian Randomization Study. *PLoS Med* (2015) 12(6):e1001841. doi: 10.1371/journal.pmed.1001841
31. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* (2013) 37(7):658–65. doi: 10.1002/gepi.21758
32. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG EPIC-InterAct Consortium. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. *Eur J Epidemiol* (2015) 30(7):543–52. doi: 10.1007/s10654-015-0011-z
33. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* (2011) 478(7367):103–9. doi: 10.1038/nature10405
34. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature* (2015) 526(7571):68–74. doi: 10.1038/nature15393
35. das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet* (2016) 48(10):1284–7. doi: 10.1038/ng.3656
36. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* (2016) 48(10):1279–83. doi: 10.1038/ng.3643
37. 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. doi: 10.1093/ije/dyu005
38. Brion MJ, Shakhbuzov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol* (2013) 42(5):1497–501. doi: 10.1093/ije/dyt179
39. Voight BF. MR_predictor: a simulation engine for Mendelian Randomization studies. *Bioinformatics* (2014) 30(23):3432–4. doi: 10.1093/bioinformatics/btu564
40. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* (2010) 26(17):2190–1. doi: 10.1093/bioinformatics/btq340
41. Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* (2013) 340(6139):1467–71. doi: 10.1126/science.1235488
42. Hemani G, Zheng J, Wade KH, Laurin C, Elsworth B, Burgess S. MR-Base: a platform for systematic causal inference across the phenome using billions of genetic associations. *bioRxiv* (2016) 078972.
43. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* (2017) 32(5):377–89. doi: 10.1007/s10654-017-0255-x
44. Hemani G, Bowden J, Haycock PC, Zheng J, Davis O, Flach P. Automating Mendelian randomization through machine learning to construct a putative causal map of the human phenome. *bioRxiv* (2017) 173682. doi: 10.1101/173682

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Teumer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.