2. Introduction and Background

B233241

2025-01-07

Word count: 766

Introduction and Background

Mendelian randomisation (MR) is a methodology intended to support causal inference from observational data. It applies the principles of instrumental variable (IV) analysis to genetic data: naturally occurring genetic variants - "instruments" – with a known association to an exposure of interest are chosen, and by comparing the association of those same genetic instruments to an outcome of interest, genetic data can be used to investigate causal links between exposures and outcomes[1]. In theory, provided that the assumptions of IV analysis are met, random assignment of genetic variants from parents to offspring during meiosis can create a form of natural experiment, analogous to randomisation during a clinical trial – both measured and unmeasured confounders should be distributed evenly between the groups created, allowing valid causal inference after other sources of bias and random variation are accounted for[2].

Three key assumptions of IV analysis must be met [@ ***REF***]: 1. Relevance – the genetic variant must be associated with the exposure of interest 2. Independence – the genetic variant is independent of confounders of the relationship between exposure and outcome 3. Exclusion restriction – the genetic variant must not be associated with the outcome except via the exposure

If these assumptions are satisfied, the "causal effect estimate" of the exposure on the outcome can be estimated by the Wald ratio, i.e. by dividing the co-efficient of gene-outcome association by the co-efficient of gene-exposure, giving a numerical measure of strength of causal exposure-outcome association[3].

Figure X. Taken from Burgess et al 2016 (DAG)[4]

In practice, only the relevance assumption can be directly tested and proven, as independence and exclusion restriction depend on all possible confounders of the exposure-outcome association, both measured and unmeasured[@ ***REF***]. Threats to the independence assumption will vary depending on the population, exposure and outcome being studied[@ ***REF***]. Exclusion restriction is a particularly universal issue in MR, due to so-called (horizontal) genetic pleiotropy, where a single genetic variant may have multiple "pleiotropic" effects – i.e. it may influence several traits simultaneously. Such pleiotropic effects may be unknown and open unmeasured causal pathways between a genetic instrument and the outcome, separate to the path involving the exposure of interest, thus potentially biasing MR estimates of the association between exposure and outcome[5].

Although not possible to prove exclusion restriction for any MR study, several methods attempt to produce exposure-outcome causal effect estimates which are robust to violations of this assumption. A common approach is the Weighted Median Estimator (WME) method, proposed by Bowden et al[6].

In WME analysis, several genetic instruments are used to estimate the exposure-outcome causal effect; each instrument is known to be associated with the exposure of interest, but a proportion of these instruments may be invalid due to unknown pleiotropic genetic effects. Any genetic instrument causally linked an outcome via multiple pleiotropic causal pathways would be expected to exhibit a less consistent gene-outcome association

than if only a single pathway mediated the gene-outcome relationship; this would be reflected in larger variance in causal effect estimates derived from invalid/pleiotropic genetic instruments. WME therefore assigns a weight to each genetic instrument's estimate of the causal effect according to the inverse of the variance of the estimate; these weighted effect estimates are used to construct a cumulative distribution function for probability of true causal effect size across the range of estimated values.

Causal effect estimates from each instrument are ordered by size, then used to create a cumulative distribution function for probability of true causal effect size. The relative contribution of each instrument's effect estimate to the probability distribution is weighted according to the inverse of the variance of the estimate. Genetic instruments whose causal effect estimates exhibit a large variance, which would be expected would therefore contribute less to]

In WME analysis, several genetic instruments are used to estimate the exposure-outcome causal effect. Any instrument linked to an outcome via multiple pleiotropic causal pathways will exhibit a less consistent gene-outcome association than a relationship mediated by a single pathway; this results in larger variance in causal estimates derived from invalid/pleiotropic genetic instruments. WME therefore assigns a weight to each genetic instrument's causal estimate according to the inverse of its variance, then constructs a cumulative distribution function for probability of true causal effect size across the range of estimated values, before taking the 50th percentile of this distribution as a "weighted median estimate" of the true causal effect, theoretically producing consistent causal estimates even if up to 50% of the included information comes from invalid instruments.

1. Richmond RC, Smith GD. Mendelian Randomization: Concepts and Scope. Cold Spring Harbor Perspectives in Medicine [Internet]. 2022 Jan [cited 2024 Oct 22];12(1):a040501. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8725623/

2. Davies NM, Holmes MV, Smith GD. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. BMJ [Internet]. 2018 Jul [cited 2025 Jan 7];362:k601. Available from: https://www.bmj.com/content/362/bmj.k601

3. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. Nature Communications [Internet]. 2020 Jan [cited 2025 Jan 7];11(1):376. Available from: https://www.nature.com/articles/s41467-019-14156-4

4. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. Epidemiology (Cambridge, Mass) [Internet]. 2016 Nov [cited 2024 Oct 22];28(1):30. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC5133381/

5. Hemani G, Bowden J, Smith GD. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Human Molecular Genetics [Internet]. 2018 May [cited 2024 Oct 23];27(R2):R195. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6061876/

6. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genetic Epidemiology [Internet]. 2016 Apr [cited 2024 Oct 22];40(4):304. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC4849733/