Lecture 7

Mendelian Randomisation: Principles, Applications, and Challenges

by Dr. Mustafa İsmail Özkaraca

Contents

Introduction to Mendelian Randomisation (MR)

Mendelian Randomisation (MR)

- **Causal Inference:** **assessing causal relationship** **using genetic variants**
- **Same concept with Randomised Trials (RT):** Corrects for confounding effects through **randomisation**
- **Practical Approach:** **Overcome limitations of RTs**

Introduction to Mendelian Randomisation (MR)

Randomised Trials (RT)

Gold standard for causal inference

Limitations of Randomised Trials (RT)

1. *Costly*
2. *Time-consuming*
3. *Ethical constraints*

Practical Solution: MR!

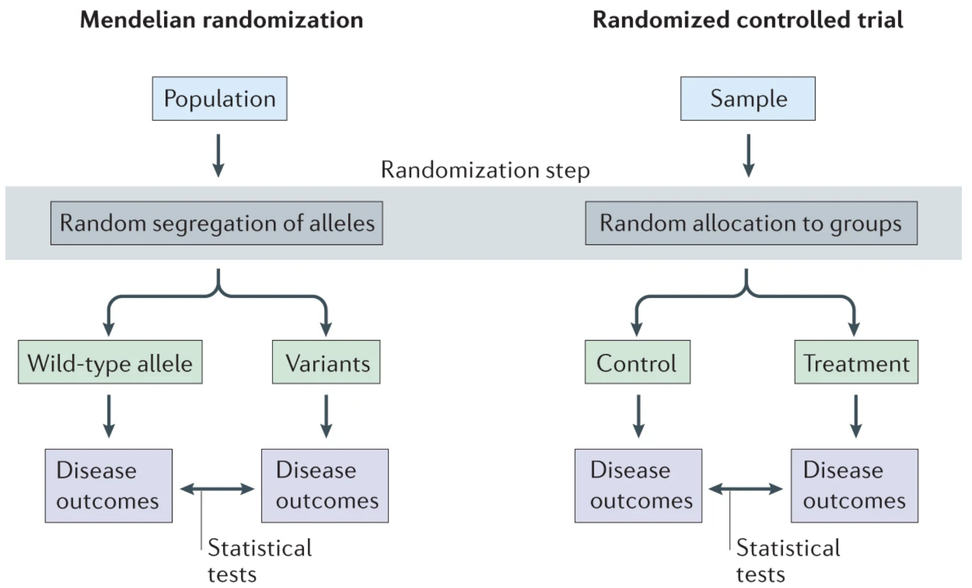
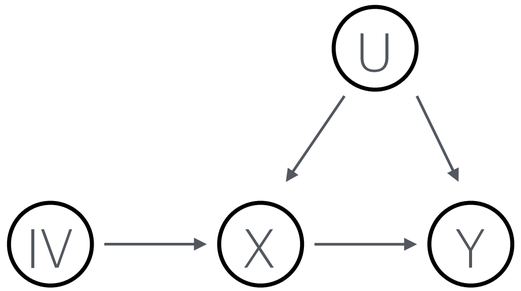


Figure is from:

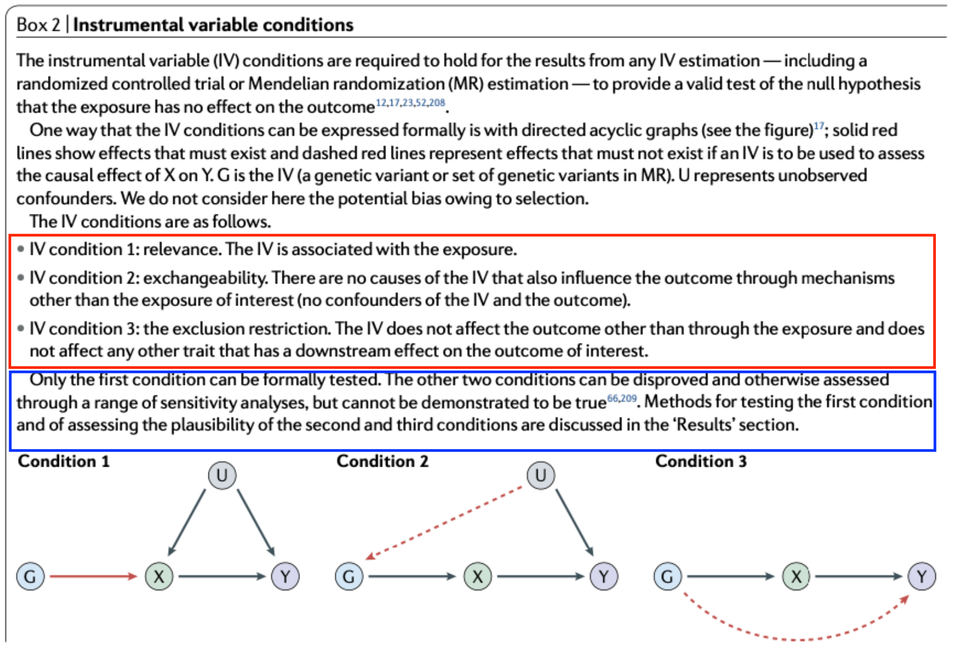
Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**, 6 (2022).

Core Logic and Theory of MR



U: Confounders
X: Exposure
Y: Outcome
IV: Instrumental Variable

Idea: Use genetic variants as IV [1]



Box 2 | Instrumental variable conditions

The instrumental variable (IV) conditions are required to hold for the results from any IV estimation — including a randomized controlled trial or Mendelian randomization (MR) estimation — to provide a valid test of the null hypothesis that the exposure has no effect on the outcome[12,17,23,52,208].

One way that the IV conditions can be expressed formally is with directed acyclic graphs (see the figure)[17]; solid red lines show effects that must exist and dashed red lines represent effects that must not exist if an IV is to be used to assess the causal effect of X on Y. G is the IV (a genetic variant or set of genetic variants in MR). U represents unobserved confounders. We do not consider here the potential bias owing to selection.

The IV conditions are as follows.

- IV condition 1: relevance. The IV is associated with the exposure.
- IV condition 2: exchangeability. There are no causes of the IV that also influence the outcome through mechanisms other than the exposure of interest (no confounders of the IV and the outcome).
- IV condition 3: the exclusion restriction. The IV does not affect the outcome other than through the exposure and does not affect any other trait that has a downstream effect on the outcome of interest.

Only the first condition can be formally tested. The other two conditions can be disproved and otherwise assessed through a range of sensitivity analyses, but cannot be demonstrated to be true[66,209]. Methods for testing the first condition and of assessing the plausibility of the second and third conditions are discussed in the 'Results' section.

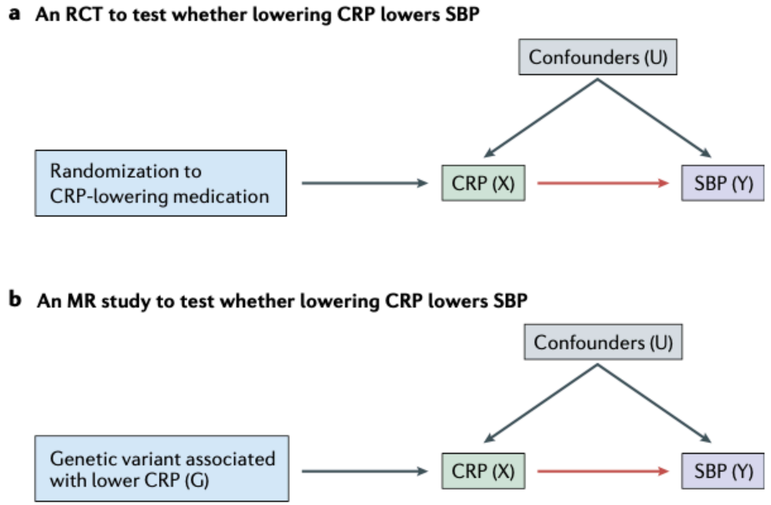


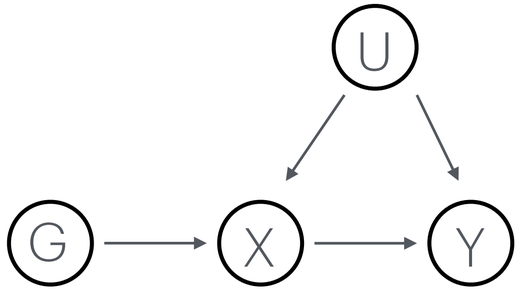
Fig. 2 | **Illustration of a randomized control study and instrumental variable estimation.** A randomized controlled trial (RCT) (panel **a**) and a Mendelian randomization (MR) study (panel **b**) to estimate the effect of lowering C-reactive protein (CRP) on systolic blood pressure (SBP). The arrows highlighted in red show the causal effect of interest.

Figure 2 and Box 2 are from:

Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**, 6 (2022).

[1] Davey Smith, George, and Shah Ebrahim. "'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?." *International journal of epidemiology* 32.1 (2003): 1-22.

Core Logic and Theory of MR



U: Confounders
X: Exposure
Y: Outcome
G: Instrumental Variable

Two-Stage Least Squares (2SLS) Estimation [2]

Step 1: Consider the model $X = \beta_{GX}G + \epsilon_X$. Compute $\widehat{X} = \beta_{GX}G$.

Step 2: Compute $\widehat{\beta_{XY}}$ using the model $Y = \beta_{XY}\widehat{X} + \epsilon_Y$.
NOTE: $\widehat{\beta_{XY}}$ is free from confounding because of conditions 2 and 3.

Wald Estimation (for single instrument) [2]

$$\widehat{\beta_{XY}} = \frac{\widehat{\beta_{GY}}}{\widehat{\beta_{GX}}}$$ where $\widehat{\beta_{GX}}$, $\widehat{\beta_{GY}}$ are associations between G and X and G and Y, respectively.

Wald Estimation (for multiple instruments) [2]

$$\widehat{\beta_{XY}} = \frac{\sum_i w_i \widehat{\beta_{GY,i}}\,\widehat{\beta_{GX,i}}}{\sum_i w_i \widehat{\beta_{GX,i}}^2}$$ where w_i are the inverse-variance weights of the genetic associations.

[2] Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**, 6 (2022).

Applications and Key Challenges

Applications

Table 1. Examples of MR

Type	Exposure/trait	Disease/outcome	Conclusion
Biomarkers	CRP	Coronary heart disease	Observational association between CRP and coronary heart disease is a result of confounding and/or reverse causation (18)
	Serum iron	Parkinson's disease	Higher serum iron levels lower the risk of Parkinson's disease (19)
	Uric acid	Coronary heart disease	Observational association between uric acid and coronary heart disease is, in part, due to confounding by BMI (20)
	Macrophage migration inhibitory factor (MIF)	Type 2 diabetes	Elevated MIF, amongst other factors, increases the risk of type 2 diabetes (21)
	Interleukin 6 (IL6)	Coronary heart disease	IL6 increases the risk of coronary heart disease (22,23)
Behaviours	Smoking	Anxiety/depression	Anxiety and depression amongst smokers does not appear to be a consequence of smoking (24,25)
	Alcohol consumption	Blood pressure	Alcohol use increases blood pressure (26)
Physiological measures	BMI	Symptomatic gallstone disease	Higher BMI increases the risk of symptomatic gall stone disease (27).
Maternal influences (corrected for genetic correlation between mother and child)	Alcohol consumption	Childhood school performance	The observational finding that moderate maternal alcohol intake is associated with more favourable school performance is due to confounding, and the casual association is in the opposite direction (28)
	Maternal BMI	Fat mass of offspring	Fat mass in children aged 9–11 is not strongly influenced by BMI of mothers during pregnancy (29)

Table is from:

Davey Smith, George, and Gibran Hemani. "Mendelian randomization: genetic anchors for causal inference in epidemiological studies." *Human molecular genetics* 23.R1 (2014): R89-R98.

Applications and Key Challenges

Key Challenges

1. **Assumptions (Condition 2 & 3):** cannot be proven; only disproven [3].
2. **Weak Instrument Bias:** low power due to weak associations [3,4]
3. **Pleiotropy:** genetic variants influencing multiple diseases or traits [3,5]
4. **Population Stratification:** Confounding due to ancestry differences [3,5]
5. **Causal Pathway Complexity:** Bias from multiple pathways [6]

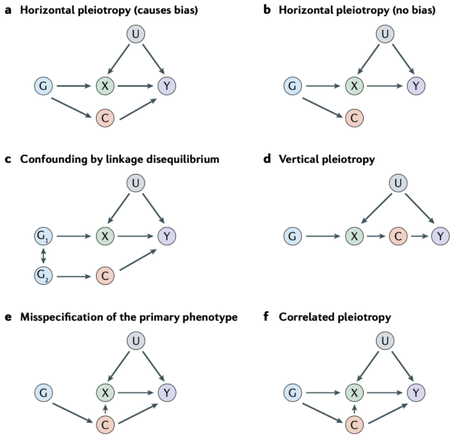


Fig. 3 | **Types of pleiotropy.** Figure showing different types of pleiotropy in Mendelian randomization (MR), where G is a genetic variant or set of genetic variants associated with the exposure, X is the exposure of interest, Y is the outcome of interest, U is an unmeasured confounder and C is another (potentially unmeasured) phenotype that is also associated with the genetic variants. **a**,**b** | Horizontal pleiotropy. Sometimes referred to as biological pleiotropy, this occurs where a genetic variant is associated with multiple phenotypes and these phenotypes lie on different pathways. In horizontal pleiotropy with bias (panel **a**), the third instrumental variable condition (IV3) is violated because there is a pathway from the genetic variant to the outcome that does not occur via the exposure. In horizontal pleiotropy with no bias (panel **b**), as the genetic variants are not associated with other phenotypes on the pathway to the outcome, MR estimates are not biased. **c** | Confounding by linkage disequilibrium. When G_2 has an effect on the outcome through a pathway that is not via the exposure, correlation between G_1 and G_2 creates a bias that is indistinguishable from that shown in panel **a**. **d** | Vertical pleiotropy. Another phenotype lies on the genetic variant–exposure–outcome pathway. This could occur either before or after the exposure of interest. Sometimes referred to as mediated pleiotropy, this form of pleiotropy does not bias MR studies and can even be used to elucidate causal intermediaries[41]. **e** | Misspecification of the primary phenotype. Vertical pleiotropy can bias MR estimates if the wrong phenotype is specified as the primary phenotype. Here the genetic variants are primarily associated with C. If X is misspecified as the primary phenotype, MR estimation of the effect of X on Y would be biased by the alternative pathways from C to Y[6,61]. **f** | In correlated pleiotropy, genetic variants for the exposure are also associated with a confounder of the exposure and outcome. In this setting, the size of the pleiotropic effect is correlated with the size of the association between the genetic variant and the exposure. This form of pleiotropy is particularly hard to detect and correct for. The scenarios in panels **b** and **d** produce settings where the pleiotropy will not bias the MR estimation. All other settings violate assumptions IV2 or IV3 and can cause meaningful bias in MR estimation.

Figure is from:
Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**, 6 (2022).

[3] Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**, 6 (2022).
[4] Burgess, Stephen, Simon G. Thompson, and Crp Chd Genetics Collaboration. "Avoiding bias from weak instruments in Mendelian randomization studies." *International journal of epidemiology* 40.3 (2011): 755-764..
[5] Bowden, Jack, George Davey Smith, and Stephen Burgess. "Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression." *International journal of epidemiology* 44.2 (2015): 512-525.
[6] Sanderson, Eleanor. "Multivariable Mendelian randomization and mediation." *Cold Spring Harbor perspectives in medicine* 11.2 (2021): a038984.

Thank You!

Thank You for Participating!

Contact


ozkaraca@duck.com


https://iozkaraca.github.io/