**Two-sample Mendelian randomization using summary genetic data – practical in MR Base**

**A. Estimate the effect of genetically elevated BMI on CHD**

Estimate the effect of BMI on CHD using inverse-variance weighted (IVW) linear regression, simple median, weighted median and MR Egger regression. How do the results from each of the methods compare?

ANSWER: the point estimates and standard errors are fairly consistent between the IVW and median methods. MR Egger has a larger point estimate and standard error. All methods show a positive causal relationship between BMI and CHD.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| method | nsnp | b | se | pval |
| Inverse variance weighted | 95 | 0.3608317 | 0.06949419 | 2.077591e-07 |
| Simple median | 95 | 0.3378855 | 0.06937833 | 1.114926e-06 |
| Weighted median | 95 | 0.3945629 | 0.07121237 | 3.013496e-08 |
| MR Egger | 95 | 0.5368161 | 0.19883472 | 8.242343e-03 |

**B. Sensitivity analyses**

* + - 1. We discussed two tests for pleiotropy in the lectures; what were they? Is there any evidence for pleiotropy in our data?

ANSWER: The two tests of heterogeneity were, 1) using a Cochran’s Q test, and 2) testing the intercept from MR Egger which indicates directional pleiotropy. There is no evidence of directional pleiotropy using the intercept from MR Egger (intercept=-0.004733028, SE=0.005009926, P=0.3472448). The Q statistic is greater than K-1 (95-1 SNPs), indicating that there is heterogeneity in the causal estimates across the SNPs.

* + - 1. Can you think of any other way that we could test for pleiotropy?

ANSWER: we could see if any of the 95 SNPs are associated with other traits, such as smoking, cholesterol, blood pressure etc, using publically available GWAS summary statistics for those traits. If some of the SNPs were associated with other traits, you would need to think about whether they were on the causal pathway (e.g. SNP -> BMI -> SBP -> CHD) or whether they were a pathway that didn’t operate through BMI. However, biological knowledge is needed.

**C. Visualising the results**

1. Create a scatter plot of the SNP-CHD and SNP-BMI associations. Does the SNP-CHD association increase linearly as the SNP-BMI association increases? What could deviations from linearity mean? Are there any unusual data points?

ANSWER: Scatter plot of results for BMI SNPs. The plot shows the SNP-CHD effects (Y axis) plotted against the SNP-BMI effects (X axis), where effect refers to the log odds for CHD per SD change in BMI per copy of the effect allele. If BMI causes CHD we would expect a linear dose response relationship, i.e. the SNP-CHD effects should increase as the SNP-BMI effects increase. Notice that there are a few “outlier” SNPs. Some BMI-raising alleles seem to be associated with negative log odds for CHD. Some BMI-raising alleles also seem to have unusually strong, risk raising, effects on CHD, in comparison to the other SNPs. “Outlier” SNPs in a scatter plot will generally correspond to the SNPs at the top or bottom of a forest plot of SNP ratio estimates (figure 1). Could these “outliers” be indicative of pleiotropic pathways to CHD? Possibly. However, there may be other explanations for outliers (see answer on heterogeneity above).



1. Create a funnel plot of the results. Does the funnel plot look symmetric? What could asymmetry mean? Are there any outliers?

ANSWER: Funnel plot of MR results. The plot shows ratio estimates of causal effect (x axis) plotted against the inverse of the standard error of the ratio estimates (Y axis). We generally expect the distribution to look symmetric, so that as estimates get less precise (going from top to bottom on the Y axis), the causal estimates “fan” out randomly on either side of the overall effect (indicated by the vertical lines). Asymmetry (i.e an imbalance of estimates on one side of the overall effect) is indicative of small study bias in a meta-analysis context. In Mendelian randomization, asymmetry is suggestive of unbalanced pleiotropy.



1. Create a forest plot of the results. Is there heterogeneity in the effects amongst SNPs? What could that indicate?

ANSWER: Forest plot of Mendelian randomization results. Plot shows effect of body mass index on coronary heart disease due to each SNP separately (estimated by ratio method) as well as combined across all SNPs into an overall effect (estimated by various methods). Heterogeneity could be indicative of violations of MR assumptions. Notice how some SNPs at the bottom and top of the plots have unusually strong protective or harmful effects, in comparison to the bulk of the SNPs. This could be indicative of alternative pathways between the SNPs and CHD not mediated by BMI (horizontal pleiotropy). Alternatively, these “outliers” could reflect other violations of assumptions, data handling errors or chance sampling variation.



**C. Interpret the results**

1. Do you think BMI causes CHD? Consider these questions:

* 1. What is the odds ratio for coronary heart disease per unit increase in genetically elevated BMI?

ANSWER: The odds ratio for coronary heart disease per SD higher BMI due to genetic variation was 1.43 (logOR= 0.3608317) (95% confidence interval: 1.25 to 1.64) (estimated by the IVW method).

* 1. Are the genetic and observational effects directionally similar? Are they comparable?

ANSWER: the causal OR=1.43 (95%CI: 1.25-1.64), which is slightly higher than the observational estimate of 1.23 (95%CI: 1.17-1.29). However, the confidence intervals overlap slightly.

* 1. Are there any reasons we should be cautious with these results?

ANSWER:

* The two samples used in the BMI and CHD GWAS might not be completely independent (i.e. some individuals may have contributed to both analyses), which could bias our results.
* We could have some weak instruments which would bias our estimate towards the null (unlikely in this case because of the large BMI GWAS, but still needs to be tested)
* Potential impact of winner’s curse, whereby the effects of the SNPs on BMI may be overestimated in the discovery stage of the GWAS
* We haven’t tested whether there are other potentially pleiotropic pathways from the SNPs to CHD that don’t go through BMI (see response to B.2. for more detail).