Mendelian randomization: Practical 2

Does BMI causally affect Coronary Heart Disease?

Pearl St Mall

The Flatirons

“Ajax”- A Boulder Resident

Hotel St Julien

Inverse variance weighted (IVW) fixed effects method

- There is one underlying 'true' effect
- All deviations of sample effects from the 'true' effect are due to chance

$$w_i = \frac{1}{\mathrm{var}(\beta_i)}$$

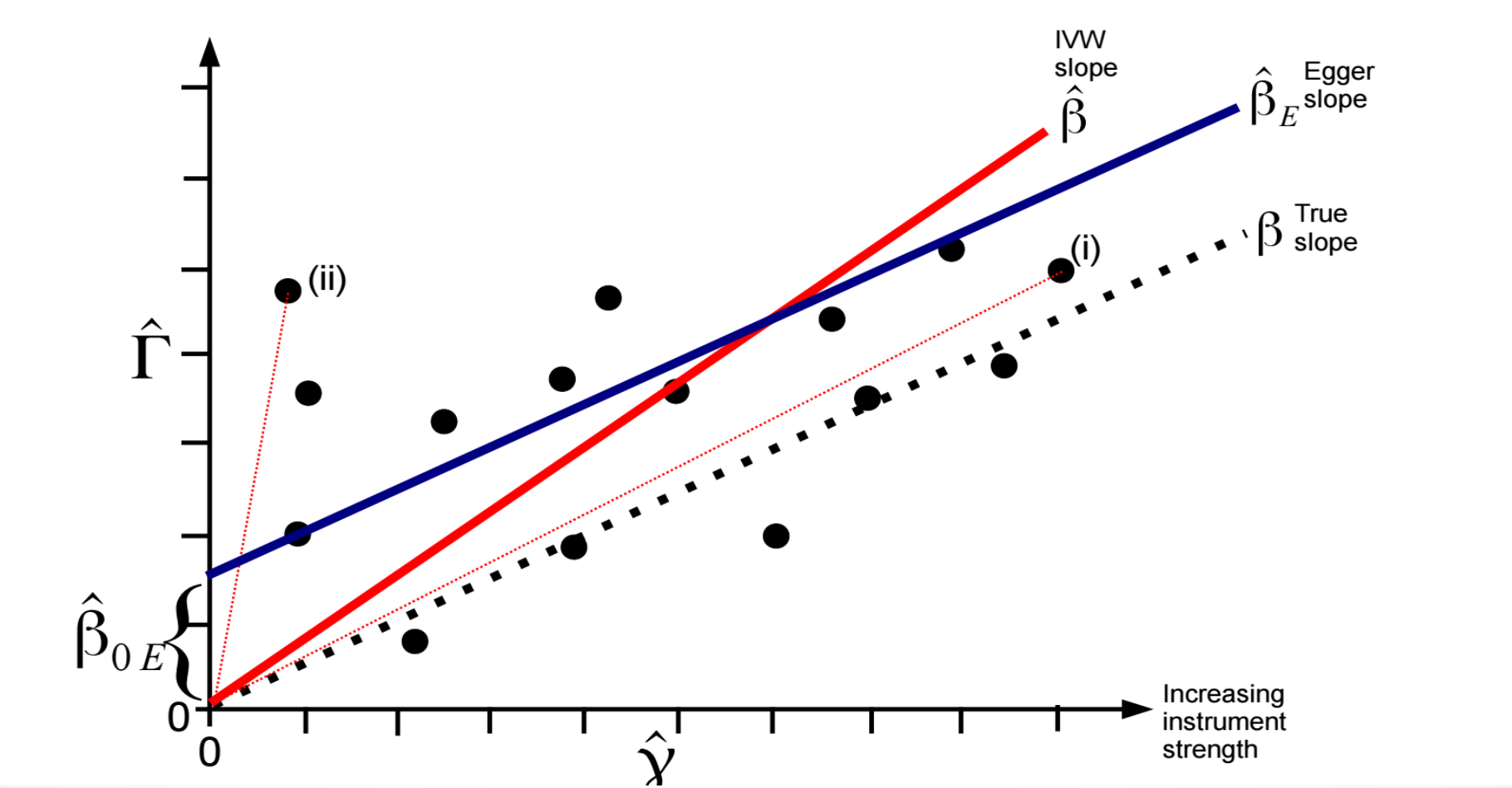
$$\beta_{pooled} = \frac{\sum_{i=1}^{N}(w_i * \beta_i)}{\sum_{i=1}^{N}(w_i)}$$

$$se_{pooled} = \sqrt{\frac{1}{\sum_{i=1}^{N}(w_i)}}$$

For N studies, each study *i* contributes more to the meta-analysis if its standard error is lower

MR Egger Regression

Egger regression: $\hat{\Gamma}_j = \beta_{0E} + \beta_E \hat{\gamma}_j.$



Simple and Weighted Median Method

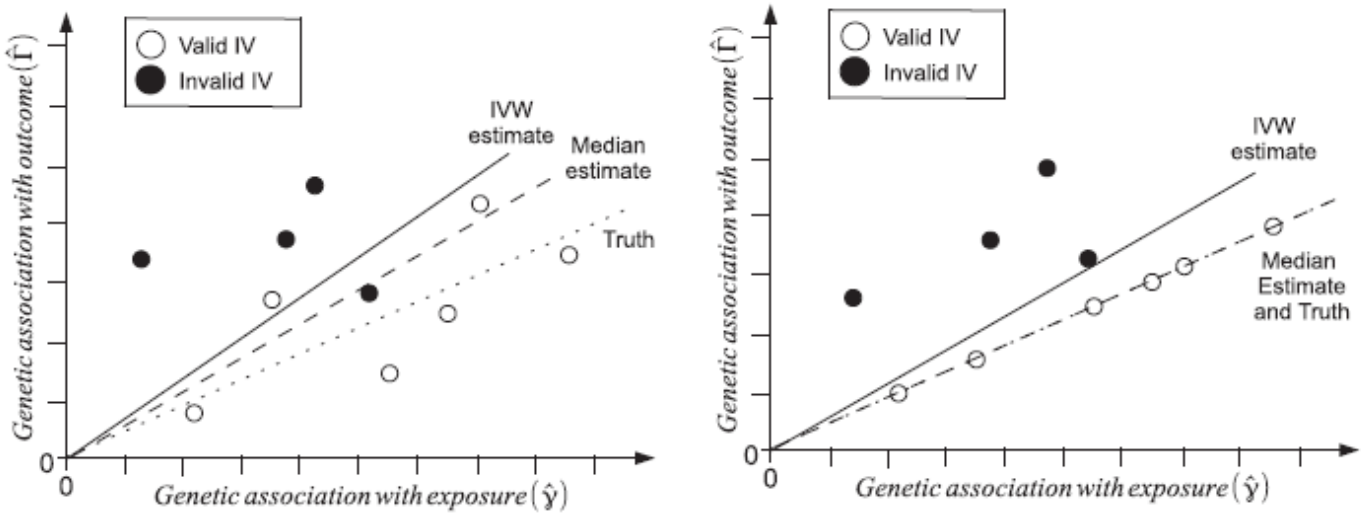
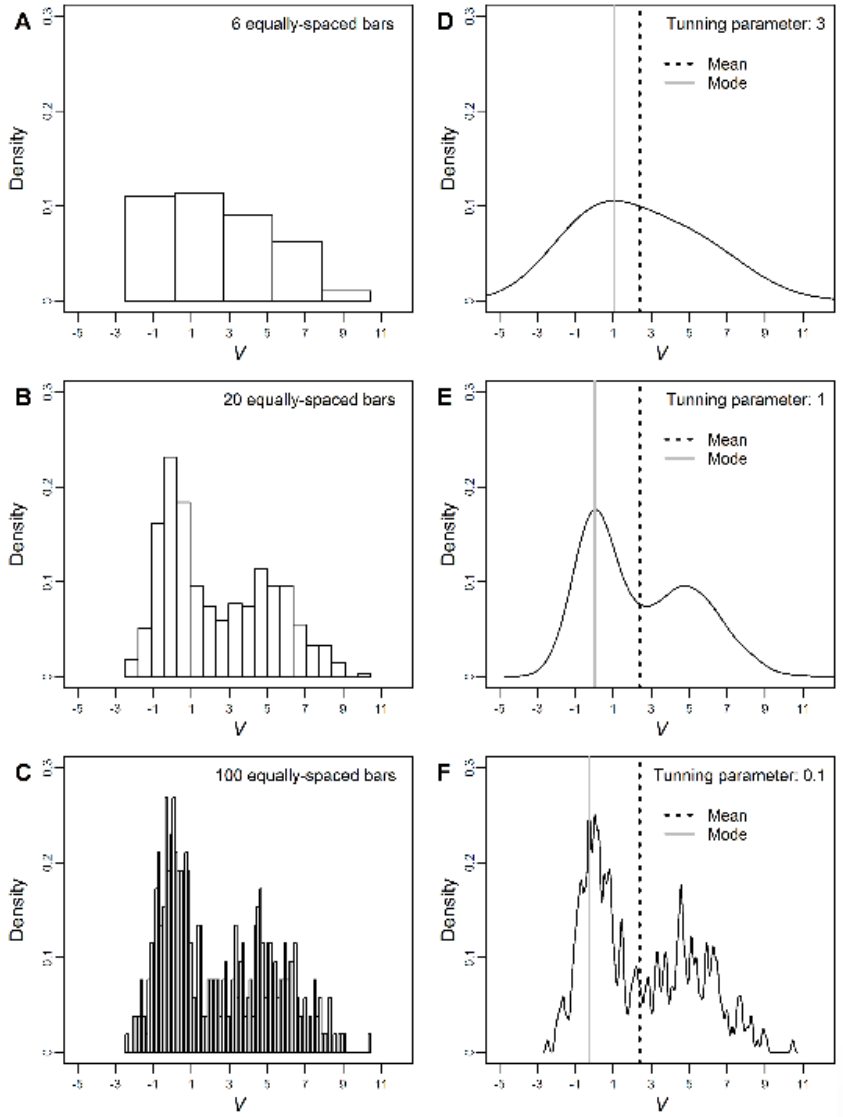


Figure 2. Fictional example of a Mendelian randomization analysis with 10 genetic variants—six valid instrumental variables (hollow circles) and four invalid instrumental variables (solid circles) for finite sample size (left) and infinite sample size (right) showing IVW (solid line) and simple median (dashed line) estimates compared with the true causal effect (dotted line). The ratio estimate for each genetic variant is the gradient of the line connecting the relevant datapoint for that variant to the origin; the simple median estimate is the median of these ratio estimates.

- Order instrumental variables estimates and take the median
- Like all subsequent estimators it enjoys a 50% breakdown limit

Modal Methods



Mendelian randomization: Practical 2

Does BMI causally affect Coronary Heart Disease?

- Web utilities and software packages exist to carry out two sample MR analyses quickly and efficiently

- This practical is not meant to illustrate the most efficient way to perform an MR study!

- This practical is designed to show you what goes on “under the hood” of these black boxes and to get you familiar with some of the data cleaning and interpretation issues when performing two sample Mendelian randomization. We will also get you to run some of the MR sensitivity analyses too.

- We have limited time!!!

- DO NOT FOCUS ON THE SYNTAX!!! (just accept that the code does what it says on the tin- go back later and check/run through)

- Remember to set your working directory at the beginning of the prac in R studio e.g.
  - setwd("/home/****YOURNAMEHERE****/MR/PRACTICAL2/")

- Run the code in 5 blocks (labelled PART ONE through PART 5)

<u>Timetable</u>

00-10 Minutes INTRO
10-20 minutes PART ONE
20-30 minutes PART TWO
30-40 minutes PART THREE
40-50 minutes PART FOUR
50-60 minutes PART FIVE
60-70 minutes Group Discussion

Mendelian randomization: Practical 2

Use your web browser to navigate to: https://workshop.colorado.edu/
#Login with your username and password
Click on the link "OnDemand"
Click on the R Studio pinned app and start a session

#Click on the "terminal" tab. This will take you to a UNIX like environment where you can copy the files over for this session's
#practical exercise

Now go to your home directory, and move to the directory "MR" you previously created:

```
cd
cd MR
```

Copy the PRACTICAL2 directory from David Evans' Faculty drive into this directory

```
cp -r /faculty/davide/BOULDER2021/PRACTICAL2 .
```

#Move into your newly created PRACTICAL2 directory and print the working directory here

```
cd PRACTICAL2
pwd
```

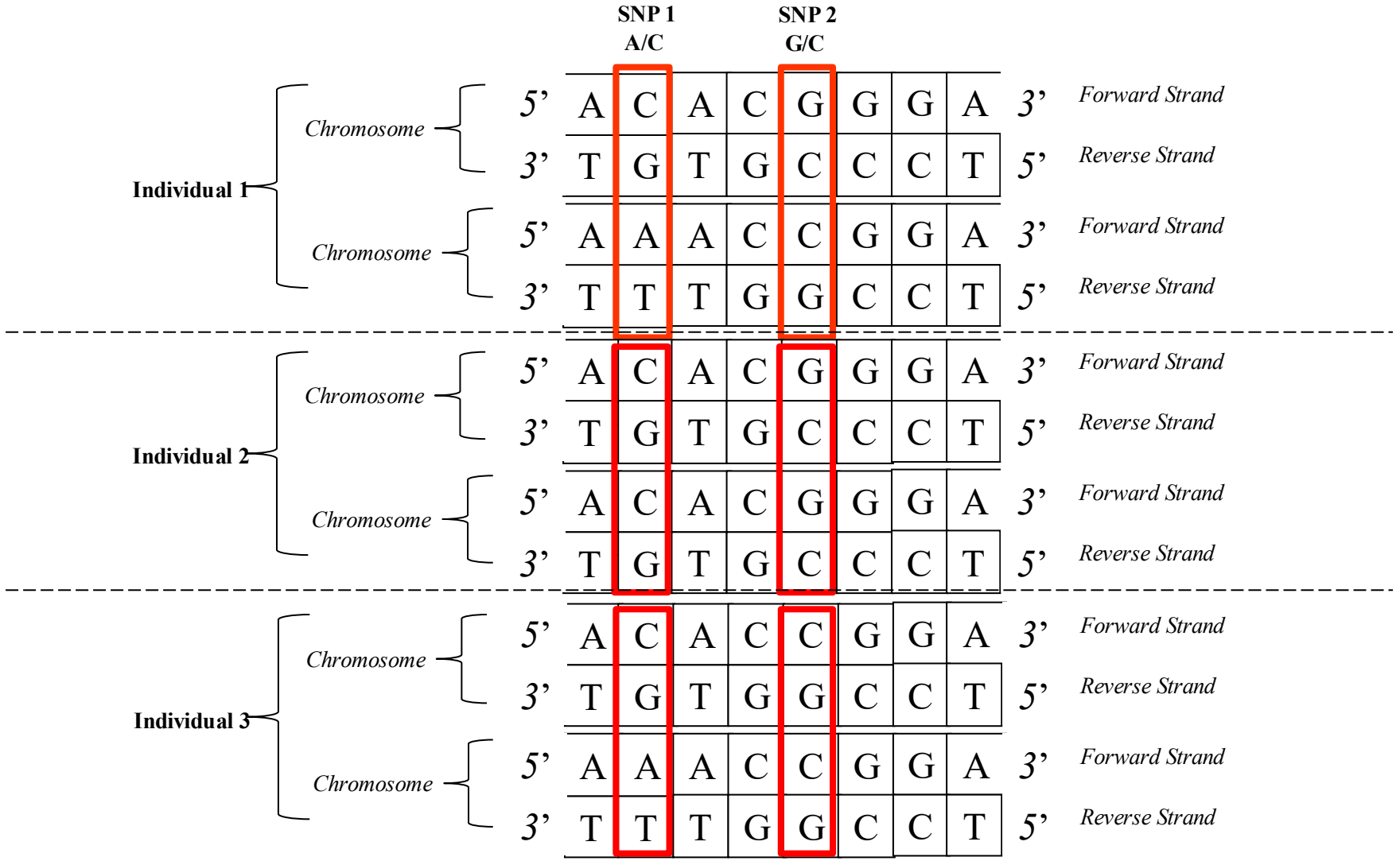
#Instructions, commands and questions for this practical are located in the file Practical2.R

Harmonise exposure and outcome effects

	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs12345	0.132	A	G	0.28	0.022	A	G	0.26
rs23456	-0.485	G	T	0.41	**0.056**	**T**	**G**	**0.61**
rs34567	0.203	G	C	0.11	**-0.046**	**G**	**C**	**0.88**

	Exposure GWAS				Outcome GWAS			
SNP	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*	*Effect*	*Effect allele*	*Other allele*	*Effect allele frequency*
rs12345	0.132	A	G	0.28	0.022	A	G	0.26
rs23456	-0.485	G	T	0.41	**-0.056**	**G**	**T**	**0.39**
rs34567	0.203	G	C	0.11	**0.046**	**G**	**C**	**0.12**

The Issue of Strand



Sensitivity Analyses- BMI and CHD

	parameter	estimate	se	lower_CI	upper_CI	p_value
1 IVW	beta	0.287658553	0.086237732	0.11590122	0.459415882	0.001318512
2 MR-Egger	beta	0.375935570	0.209803912	-0.04201525	0.793886395	0.077191677
3 MR-Egger	alpha	-0.002791481	0.006041587	-0.01482694	0.009243978	0.645387108
4 Weighted_median	beta	0.379652982	0.117839188	0.08299413	0.554455224	0.005920437
5 Weighted_mode	beta	0.311258179	0.129426222	0.05758745	0.564928912	0.018610428

- Results look very consistent!
- MR Egger lacks power- so look at coefficient rather than p value!
- MR Egger intercept is a test for directional horizontal pleiotropy