Designing Your MR Study Practical

Jean Morrison

2024-02-23

Introduction

In this practical, we will walk through choices and considerations we might encounter in an MR analysis of the effect of blood pressure on risk of stroke. Some of the content of this practical is based on the two papers, Georgakis et al. (2020) and Georgakis and Gill (2021).

Locating Data Sources

The first step is to identify GWAS summary statistics for exposure and outcome. Two catalogs of GWAS summary statistics:

- NHGRI GWAS Catalog
- MRC-IEU GWAS Catalog

I find the NGHRI GWAS Catalog slightly easier to use for discovering data sources, but the MRC-IEU Catalog is conveniently linked with the R package ieugwasr which makes querying and extracting data very easy. The gwasrapidd package can be used for interfacing with the GWAS Catalog, but this package is lacking some of the functionalities of `ieugwasr`.

Outcome Data

If you go to the NHGRI Catalog and search for "Stroke" you will find the trait entry for stroke, which looks like this:

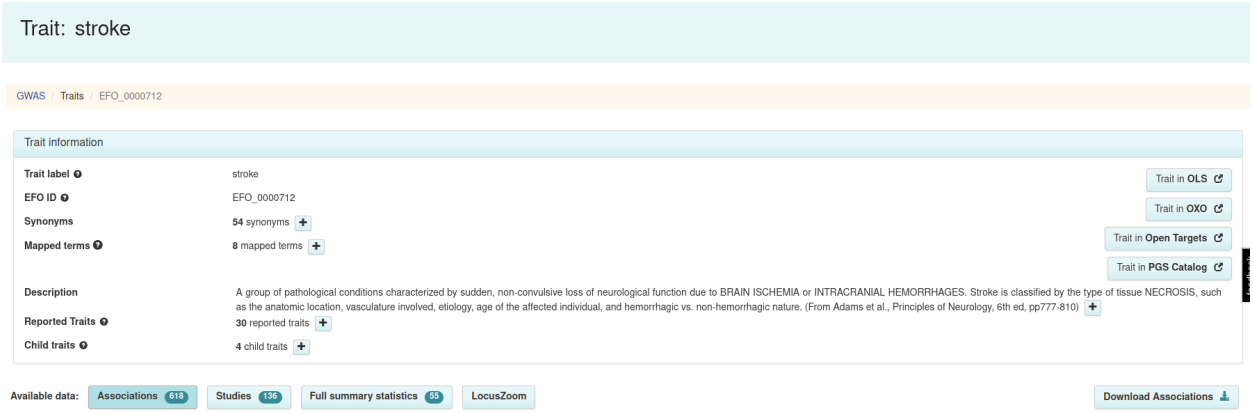
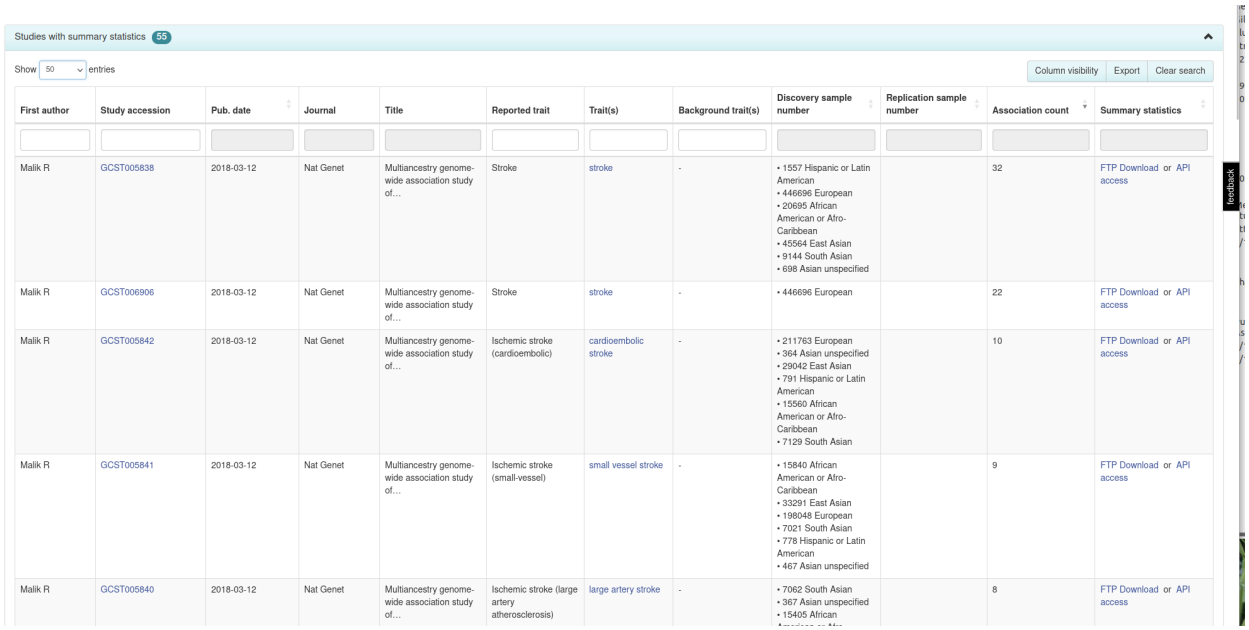


Figure 1: Entry for Stroke on GWAS Catalog

As of the time of this writing, there are 55 sets of full summary statistics for stroke.

Clicking on the "Full summary statistics" tab gives a list of all available studies. Sort the table largest to smallest by number of associations. This should look something like



Studies with summary statistics 55

Show 50 entries

Column visibility | Export | Clear search

First author	Study accession	Pub. date	Journal	Title	Reported trait	Trait(s)	Background trait(s)	Discovery sample number	Replication sample number	Association count	Summary statistics
Malik R	GCST005838	2018-03-12	Nat Genet	Multiancestry genome-wide association study of…	Stroke	stroke	-	• 1557 Hispanic or Latin American • 446696 European • 20695 African American or Afro-Caribbean • 45564 East Asian • 9144 South Asian • 698 Asian unspecified		32	FTP Download or API access
Malik R	GCST006906	2018-03-12	Nat Genet	Multiancestry genome-wide association study of…	Stroke	stroke	-	• 446696 European		22	FTP Download or API access
Malik R	GCST005842	2018-03-12	Nat Genet	Multiancestry genome-wide association study of…	Ischemic stroke (cardioembolic)	cardioembolic stroke	-	• 211763 European • 364 Asian unspecified • 29042 East Asian • 791 Hispanic or Latin American • 15560 African American or Afro-Caribbean • 7129 South Asian		10	FTP Download or API access
Malik R	GCST005841	2018-03-12	Nat Genet	Multiancestry genome-wide association study of…	Ischemic stroke (small-vessel)	small vessel stroke	-	• 15840 African American or Afro-Caribbean • 33291 East Asian • 198048 European • 7021 South Asian • 778 Hispanic or Latin American • 467 Asian unspecified		9	FTP Download or API access
Malik R	GCST005840	2018-03-12	Nat Genet	Multiancestry genome-wide association study of…	Ischemic stroke (large artery atherosclerosis)	large artery stroke	-	• 7062 South Asian • 367 Asian unspecified • 15405 African		8	FTP Download or API access

The most well-powered study is a multi-ancestry meta analysis. However, using multi-ancestry data in MR could cause problems due to mismatching populations in exposure and outcome data. Therefore, we will go with the second result which is from the same paper and uses only the European ancestry subset.

Exposure Data

Next, we need to find GWAS results for blood pressure. There are many blood pressure related phenotypes we could look for including hypertension status, systolic, and diastolic blood pressure. It is better for us to select a continuous exposure variable (so not hypertension status).

If we examine the available studies for systolic blood pressure, we find that there are two large European ancestry studies available, one by Surendran et al(Surendran et al. 2020), and one by Evangelou et al(Evangelou et al. 2018). The Surendran et al. study uses exome chip and exome sequencing data while the Evangelou et al. study uses a genome-wide genotyping array, imputed to around 8 million variants.

Unfortunately for us, both studies use blood pressure adjusted for blood pressure medication and both studies adjust for BMI in the GWAS. Because BMI is heritable, this could introduce some collider bias if there are common causes of BMI and stroke. Georgakis et al. (2020) use the Evangelou et al. data and address the risk of collider bias by performing a sensitivity analysis using results only from UK Biobank that do not adjust for BMI or medication use. A recently proposed alternative solution is to perform multivariable MR adjusting for BMI (Gilbody et al. 2022). If all effects are linear, this strategy will recover the effect of SBP on stroke but will not provide a valid estimate of the effect of BMI on stroke.

Another thing we need to pay attention to in the exposure data is the unit of the measurement. In the Evangelou data, SBP and DBP are untransformed, so our effects will be interpretable as effect per unit increase in SBP or DBP. In the Surendran et al. paper, each constituent study inverse normalized the phenotype before performing GWAS. Then all of the studies were meta analyzed together. If all of the studies had the same standard deviation of SBP, the resulting unit would be whatever that standard deviation is. However, each study likely has a somewhat different distribution of blood pressures, so it is not exactly clear what the resulting unit of the estimate is.

For our example, we will move on using the Evangelou data.

Instrument Selection

Conveniently, both the Malik et al. stroke study and the Evangelou et al. blood pressure study are contained in the MRC-IEU database, so we can use the `ieugwasr` package to query these datasets. You can also find some of the processed data sets below on the course website.

Genome-Wide Instrument Selection

It would be nice to have a separate exposure study for instrument selection, but the Evangelou et al. study is a meta analysis and includes most other studies measuring SBP. This means that we will need to do in-sample selection and accept that we will have some bias from winner's curse. Fortunately, there is not very much overlap in samples between the exposure and outcome studies, so this bias will be towards zero making the estimate a little bit conservative. We will pull data for both systolic and diastolic blood pressure, for now we will simply analyze these two phenotypes in parallel. In a later section, we will look at them together.

```
library(ieugwasr)
library(TwoSampleMR)
library(dplyr)
library(qqman)
library(ggplot2)
library(GRAPPLE)
library(MVMR)
```

```
sbp_id <- "ieu-b-38" # Systolic blood pressure
dbp_id <- "ieu-b-39" # Diastolic blood pressure
stroke_id <- "ebi-a-GCST006906" # Stroke

# extract top hits using default LD pruning settings
# this uses extract_instruments from
# TwoSampleMR but tophits from ieugwasr would also work
sbp_dat <- extract_instruments(outcomes = sbp_id, p1 = 5e-8)
dbp_dat <- extract_instruments(outcomes = dbp_id, p1 = 5e-8)

dim(sbp_dat)
#[1] 461  15
dim(dbp_dat)
#[1] 460  15

head(sbp_dat)

# extract outcome data and merge
stroke_sbp_dat <- extract_outcome_data(outcomes = stroke_id, snps = sbp_dat$SNP)
sbp_stroke_full_data <- harmonise_data(exposure_dat = sbp_dat,
                                       outcome_dat = stroke_sbp_dat)


stroke_dbp_dat <- extract_outcome_data(outcomes = stroke_id, snps = dbp_dat$SNP)
dbp_stroke_full_data <- harmonise_data(exposure_dat = dbp_dat,
                                       outcome_dat = stroke_dbp_dat)
```

The data frames created in the previous code are available on the course website. You can download them and load them in using

```
sbp_stroke_full_data <- readRDS("sbp_stroke_full_data.RDS")
dbp_stroke_full_data <- readRDS("dbp_stroke_full_data.RDS")
```

Georgakis et al. (2020) also look at a set of instruments in promoters or enhanchers for genes targeted by beta blockers or calcium channel blockers, but we will not include this in today's tutorial.

Per-Variant Screening

Two common variant screening strategies are Steiger filtering, which removes variants that are more strongly associated with the outcome than with the exposure and filtering on contribution to the Q-statistic, which measures heterogeneity.

Steiger filtering is usually a good idea but could give misleading results if there is differential measurement error or certain arrangements of strong confounding. I believe that the supplied prevalence does not have a very big effect on the results, but it may be worth looking at multiple values.

```
sbp_stroke_full_data <- sbp_stroke_full_data %>%
                        mutate(units.outcome = "log odds",
                               ncase.outcome = 40585,
                               ncontrol.outcome =406111,
                               prevalence.outcome = 0.01) %>%
                        steiger_filtering()
table(sbp_stroke_full_data$steiger_dir)
```

```
##
## FALSE  TRUE
##    13   446
```

```
dbp_stroke_full_data <- dbp_stroke_full_data %>%
                        mutate(units.outcome = "log odds",
                               ncase.outcome = 40585,
                               ncontrol.outcome =406111,
                               prevalence.outcome = 0.01) %>%
                        steiger_filtering()
table(dbp_stroke_full_data$steiger_dir)
```

```
##
## FALSE  TRUE
##     9   446
```

Outlier removal by filtering on contribution to the Q-statistic can be useful when most of the data are homogeneous and there are a few outliers. However, if there is complex horizontal pleiotropy, outlier removal could remove the wrong variants. It can also give a false impression of confidence, as definitionally, after outlier removal, all of the remaining variants are in agreement. It is a good idea to look at the data and the distribution of the Q-statistic contributions before removing outliers. It is also a good idea to compare outlier filtered results (if you are using them) to alternative methods that retain variants but model heterogeneity such such as MR-RAPS/GRAPPLE or mixture methods like CAUSE or MR-MIX. Outlier filtering should be performed after Steiger filtering.

The easiest way I know to extract the Q-statistic is using the (non-exported) `mr_all` function in `TwoSampleMR`. This function will also calculate a variety of 2SMR estimators. A common threshold for filtering on the contribution to the Q-statistic is a p-value threshold of 0.05. Removing all variants above this threshold will guarantee that the remaining variants have a non-sgnificant total Q-statistic.

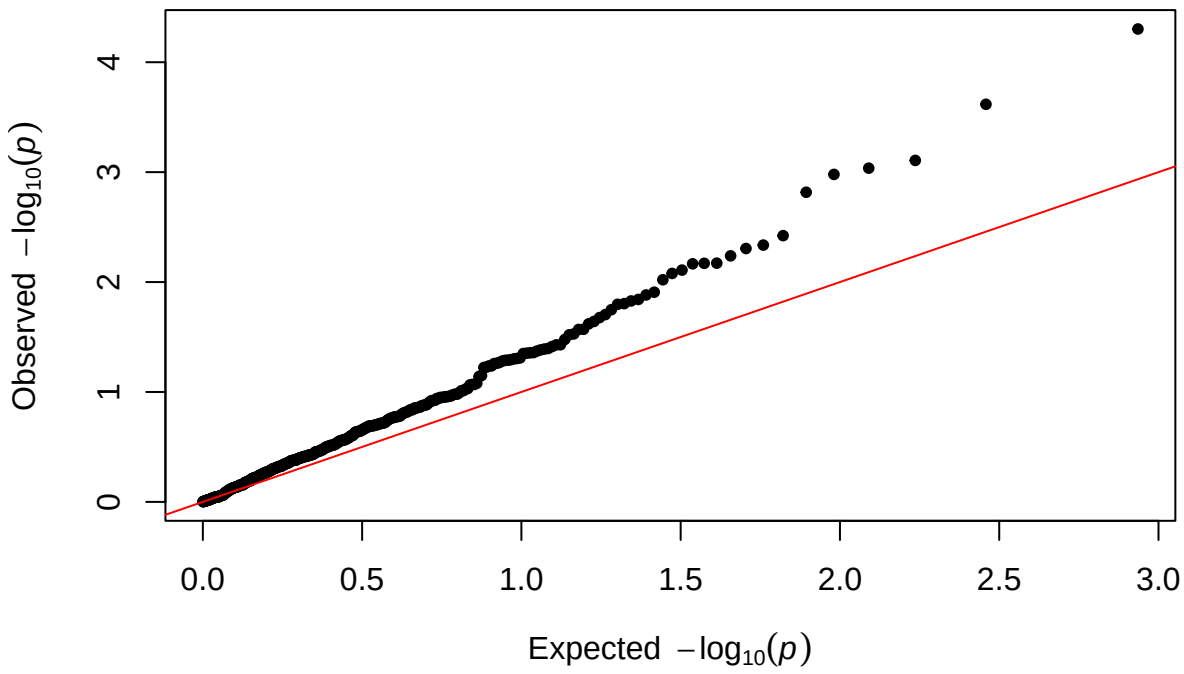
```
mr_all_sbp <- sbp_stroke_full_data %>%
              filter(steiger_dir == TRUE) %>%
              TwoSampleMR:::mr_all()
```

```
## systolic blood pressure || id:ieu-b-38 - Stroke || id:ebi-a-GCST006906
```

```
head(mr_all_sbp$outliers)
```

```
## # A tibble: 6 x 5
##   id.exposure id.outcome       SNP              Qj    Qpval
##   <chr>       <chr>            <chr>         <dbl>    <dbl>
## 1 ieu-b-38    ebi-a-GCST006906 rs1000423   0.147   0.701
## 2 ieu-b-38    ebi-a-GCST006906 rs10008637  0.331   0.565
## 3 ieu-b-38    ebi-a-GCST006906 rs10028284  0.341   0.559
## 4 ieu-b-38    ebi-a-GCST006906 rs10045307  0.00113 0.973
## 5 ieu-b-38    ebi-a-GCST006906 rs10048404  0.627   0.428
## 6 ieu-b-38    ebi-a-GCST006906 rs1006545  11.3     0.000782
```

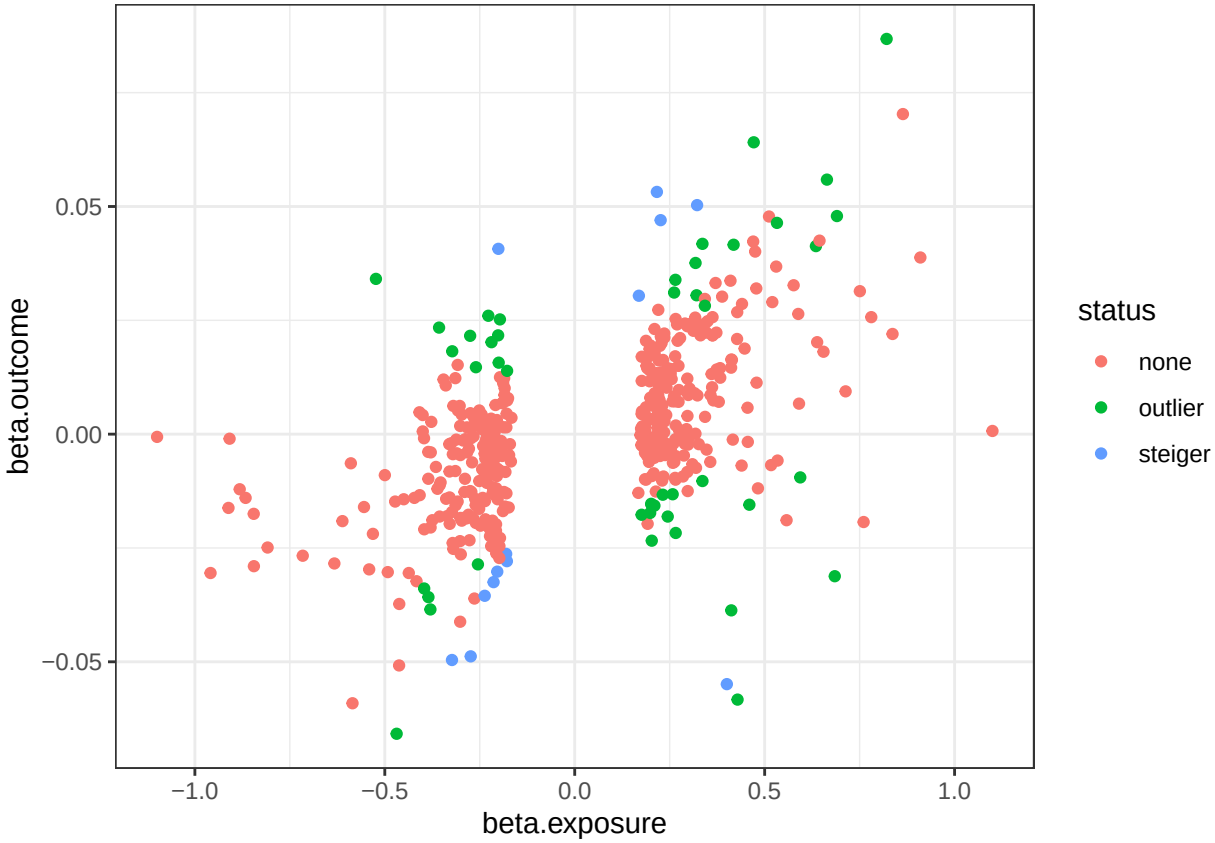
```
qq(mr_all_sbp$outliers$Qpval)
```



Looking at the distribution of Q-statistic contributions shows that we have overdispersion in all variants and not just a few large outliers, so this application may not be a good candidate for outlier removal. We can also look at the data colored by whether a variant is removed by Steiger filtering or by Q-statistic based outlier removal.

```r
left_join( sbp_stroke_full_data, mr_all_sbp$outliers) %>%
  mutate(status = case_when(steiger_dir == FALSE ~ "steiger", Qpval < 0.05 ~ "outlier", TRUE ~ "none")) +
  ggplot() +
    geom_point(aes(x = beta.exposure, y = beta.outcome, color = status)) +
    theme_bw()
```

```
## Joining with `by = join_by(SNP, id.outcome, id.exposure)`
```



Assessing Sample Overlap

If there is sample overlap, it is a good idea to use a method that can account for it. There are two ways to determine if sample overlap is an issue. The first is to read the source papers and attempt to determine if the exposure and outcome data sets contain overlapping individuals. The second is to estimate the residual correlation due to sample overlap. If this value is statistically different from zero, there is substantial residual correlation that needs to be accounted for. There are two methods to estimate this value. Both require full genome-wide data and are a little time consuming so we won't do them here. However, we will describe how to do them.

P-value thresholding

For this method, we need to perform these steps:

1. Download, merge, and harmonise genome-wide data for exposure and outcome.

2. Remove variants that have $p < 0.05$ for either exposure or outcome.
3. Prune remaining variants for LD.
4. Compute the correlation in z-scores for exposure and outcome in the LD-pruned, non-significant variant set.

Cross-trait LD score regression

For this method, we need to perform these steps:

1. Download, merge, and harmonise genome-wide data for exposure and outcome.
2. Download LD scores for the appropriate continental population.
3. Harmonise summary statistics with LD scores.
4. Perform cross-trait LD-score regression.

The residual correlation parameter is estimated by the intercept of cross-trait LD-score regression.

I performed this method for our data sets and found no significant residual correlation between SBP and stroke or between DBP and stroke, suggesting either no or very minimal sample overlap. Since SBP and DBP are from the same individuals, there is residual correlation, estimated at 0.154 (we will use this later in the MVMR section).

Univariable MR Analysis

It is a good idea to compare results from methods that make different assumptions about horizontal pleiotropy. Since accounting for horizontal pleiotropy was covered in Lecture 2, here we will only use random effects IVW regression and GRAPPLE with robust loss function. GRAPPLE should give very similar results to MR-RAPS with overdispersion, but GRAPPLE provides a plotting function to examine the distribution of residuals. Below, I am using data after Steiger filtering but did not filter on the Q-statistic.

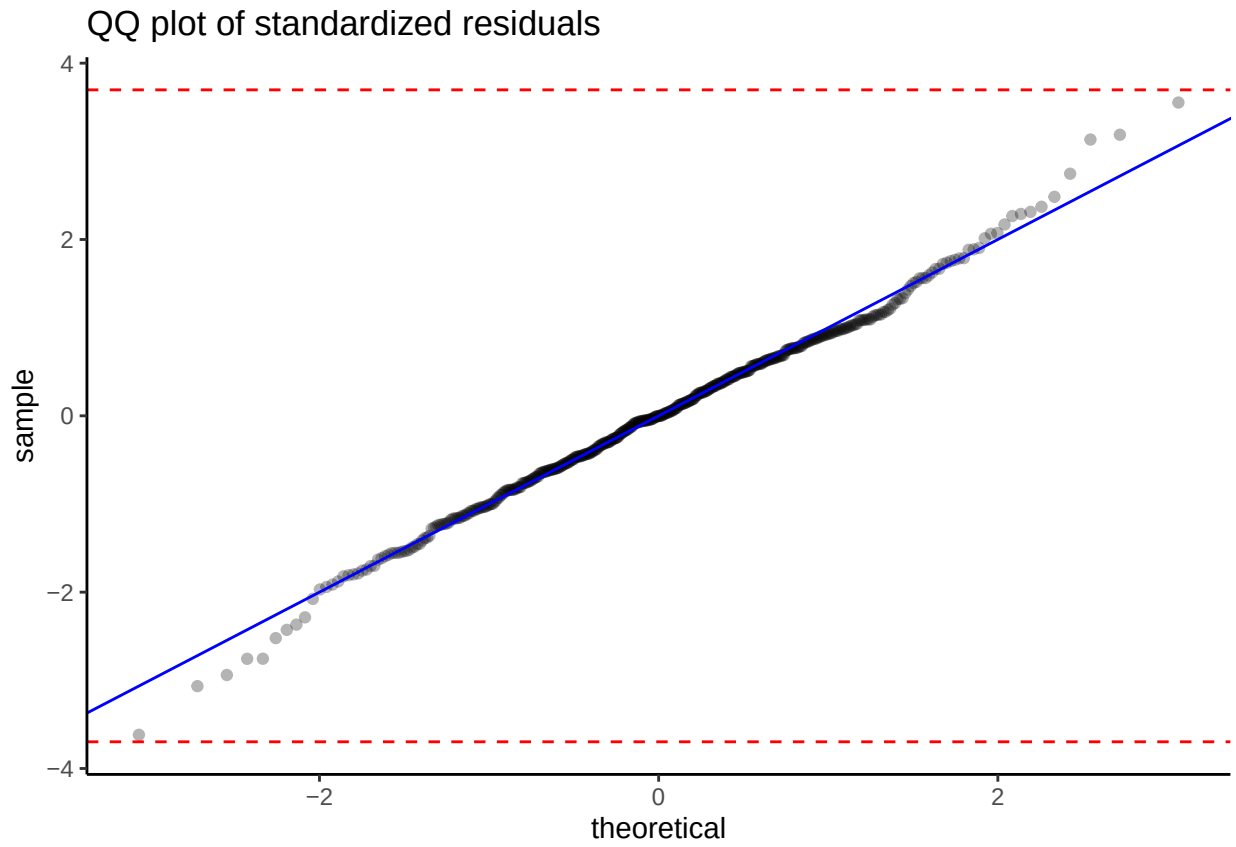
Note that functions like `mr_wrapper` and `mr` in `TwoSampleMR` will conduct many different MR tests. These results can be interesting. However, it is good to keep in mind that high concordance between methods that make the same/similar assumptions does not strengthen our belief in the estimate. Additionally, these tests are not all equally good/reliable/robust. For example, MR Egger is notoriously imprecise (very wide confidence intervals) while remaining sensitive to many types of horizontal pleiotropy.

```
ivw_res_sbp <- sbp_stroke_full_data %>%
           filter(steiger_dir == TRUE) %>%
           with(., mr_ivw(b_exp = beta.exposure,
                          b_out = beta.outcome,
                          se_exp= se.exposure,
                          se_out = se.outcome))

grapple_res_sbp <- sbp_stroke_full_data %>%
               rename(gamma_out1 = beta.outcome,
                      se_out1 = se.outcome,
                      gamma_exp1 = beta.exposure,
                      se_exp1 = se.exposure) %>%
           grappleRobustEst(data = .,
                          loss.function = "tukey",
                          plot.it = TRUE)
```

```
## [1] "Test of normality for the standardized residuals:"
```

```
## [1] "Anderson-Darling test: p-value = 0.189"
## [1] "Shapiro-Wilk test: p-value = 0.0971"
## [1] "0 outliers detected!"
```




```
sbp_results_tab <- data.frame(method = c("IVW", "GRAPPLE"),
                          beta_hat = c(ivw_res_sbp$b,  grapple_res_sbp$beta.hat),
                          se = c(ivw_res_sbp$se, sqrt(grapple_res_sbp$beta.var)),
                          pval = c(ivw_res_sbp$pval, grapple_res_sbp$beta.p.value))
sbp_results_tab
```

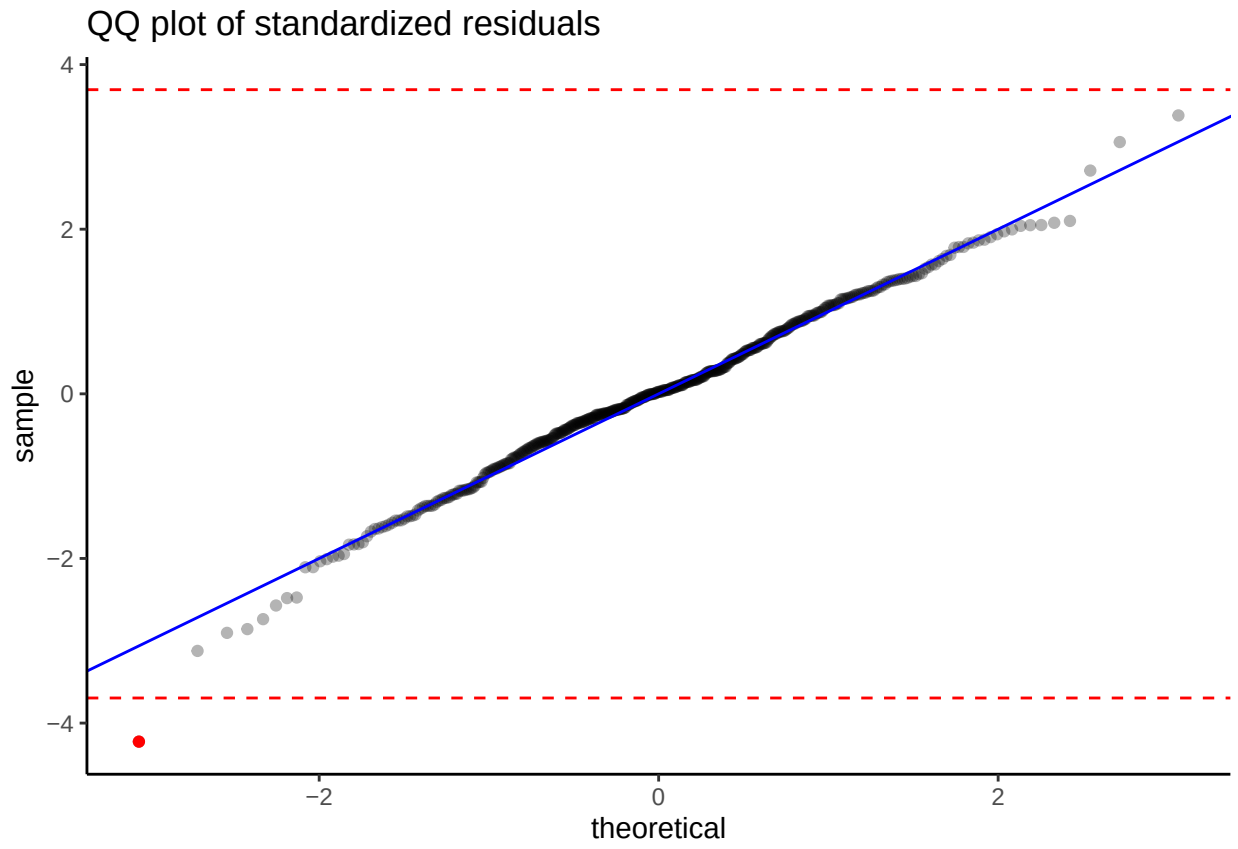
```
##    method   beta_hat          se         pval
## 1     IVW 0.02982745 0.002068623 3.930806e-47
## 2 GRAPPLE 0.03133704 0.002272708 2.992669e-43
```

```
ivw_res_dbp <- dbp_stroke_full_data %>%
           filter(steiger_dir == TRUE) %>%
           with(., mr_ivw(b_exp = beta.exposure,
                          b_out = beta.outcome,
                          se_exp= se.exposure,
                          se_out = se.outcome))

grapple_res_dbp <- dbp_stroke_full_data %>%
                rename(gamma_out1 = beta.outcome,
                       se_out1 = se.outcome,
                       gamma_exp1 = beta.exposure,
```

```
                      se_exp1 = se.exposure) %>%
            grappleRobustEst(data = .,
                             loss.function = "tukey",
                             plot.it = TRUE)
```

```
## [1] "Test of normality for the standardized residuals:"
## [1] "Anderson-Darling test: p-value = 0.0376"
## [1] "Shapiro-Wilk test: p-value = 0.0172"
## [1] "1 outliers detected!"
```




```
dbp_results_tab <- data.frame(method = c("IVW", "GRAPPLE"),
                              beta_hat = c(ivw_res_dbp$b,  grapple_res_dbp$beta.hat),
                              se = c(ivw_res_dbp$se, sqrt(grapple_res_dbp$beta.var)),
                              pval = c(ivw_res_dbp$pval, grapple_res_dbp$beta.p.value))
dbp_results_tab
```

```
##    method   beta_hat          se         pval
## 1     IVW 0.04453495 0.003554744 5.226219e-36
## 2 GRAPPLE 0.04731124 0.003792042 1.003553e-35
```

The residual qq-plots for GRAPPLE look pretty good (with one outlier for DBP), so we don't have strong evidence that the GRAPPLE model fits the data poorly. To interpret these results, we should first verify the units of the original exposure study (in this case SBP and DBP were untransformed). Since stroke is binary with association statistics measured by logistic regression, our MR estimate is the expected increase

in log odds for stroke caused by an increase in SBP/DBP of one unit. Georgakis et al. (2020) give results per 10 systolic blood pressure units and per 5 diastolic blood pressure units, so we will do the same thing. Using the GRAPPLE results, we find an estimated causal odds ratio of 1.37 (95% CI: 1.34 to 1.40) per 10 unit increase in SBP and an estimated causal odds ratio of 1.26 (95% CI: 1.24 to 1.29) per 5 unit increase in DBP

Multivariable MR Analysis

If there are known heritable confounders, we can look at an MVMR analysis. In this case, we know that SBP and DBP share a large number of genetic variants, so it is more sensible to look at these together in a multivariable analysis than separately. We will also look at an analysis adjusting for BMI to resolve bias induced by adjusting for BMI in the original blood pressure GWAS.

The simplest thing to do is multivariable IVW. We need to re-extract instruments for the two exposures together so that we have a single set of LD-pruned exposures.

```
## extract data
sbp_dbp_dat <- mv_extract_exposures(id_exposure = c(sbp_id, dbp_id), pop = "EUR")
stroke_sbp_dbp_dat <- extract_outcome_data(outcomes = stroke_id, snps = sbp_dbp_dat$SNP)
sbpdbp_stroke_full_data <- mv_harmonise_data(exposure_dat = sbp_dbp_dat,
                                             outcome_dat = stroke_dbp_dat)
```

Or load the data from the course website:

```
sbpdbp_stroke_full_data <- readRDS("sbpdbp_stroke_full_data.RDS")
```

```
## mv-ivw
mvivw_res <- mv_multiple(sbpdbp_stroke_full_data)
mvivw_res$result
```

```
##   id.exposure                                exposure       id.outcome
## 1    ieu-b-38  systolic blood pressure || id:ieu-b-38 ebi-a-GCST006906
## 2    ieu-b-39 diastolic blood pressure || id:ieu-b-39 ebi-a-GCST006906
##                          outcome nsnp           b          se         pval
## 1 Stroke || id:ebi-a-GCST006906  165 0.031832618 0.007603289 2.830512e-05
## 2 Stroke || id:ebi-a-GCST006906  287 0.006373569 0.011029750 5.633637e-01
```

These results show no significant effect of diastolic blood pressure! From these results, we would estimate a causal odds ratio of 1.37 (95% CI: 1.27 to 1.48) per 10 unit increase in SBP and an estimated causal odds ratio of 1.03 (95% CI: 0.97 to 1.09) per 5 unit increase in DBP.

When using MVMR, it is a good idea to assess instrument strength. We can do this using the `strength_mvmr` function from the MVMR package (Sanderson, Spiller, and Bowden 2021).

```
mvmr_input <- with(sbpdbp_stroke_full_data,
                   format_mvmr(BXGs = exposure_beta,
                               BYG = outcome_beta,
                               seBXGs = exposure_se,
                               seBYG = outcome_se,
                               RSID = rownames(exposure_beta)))
strength_mvmr(mvmr_input)
```

```
## Warning in strength_mvmr(mvmr_input): Covariance between effect of genetic
## variants on each exposure not specified. Fixing covariance at 0.
```

```
##
## Conditional F-statistics for instrument strength
##
##             exposure1 exposure2
## F-statistic  5.889941   6.15238
```

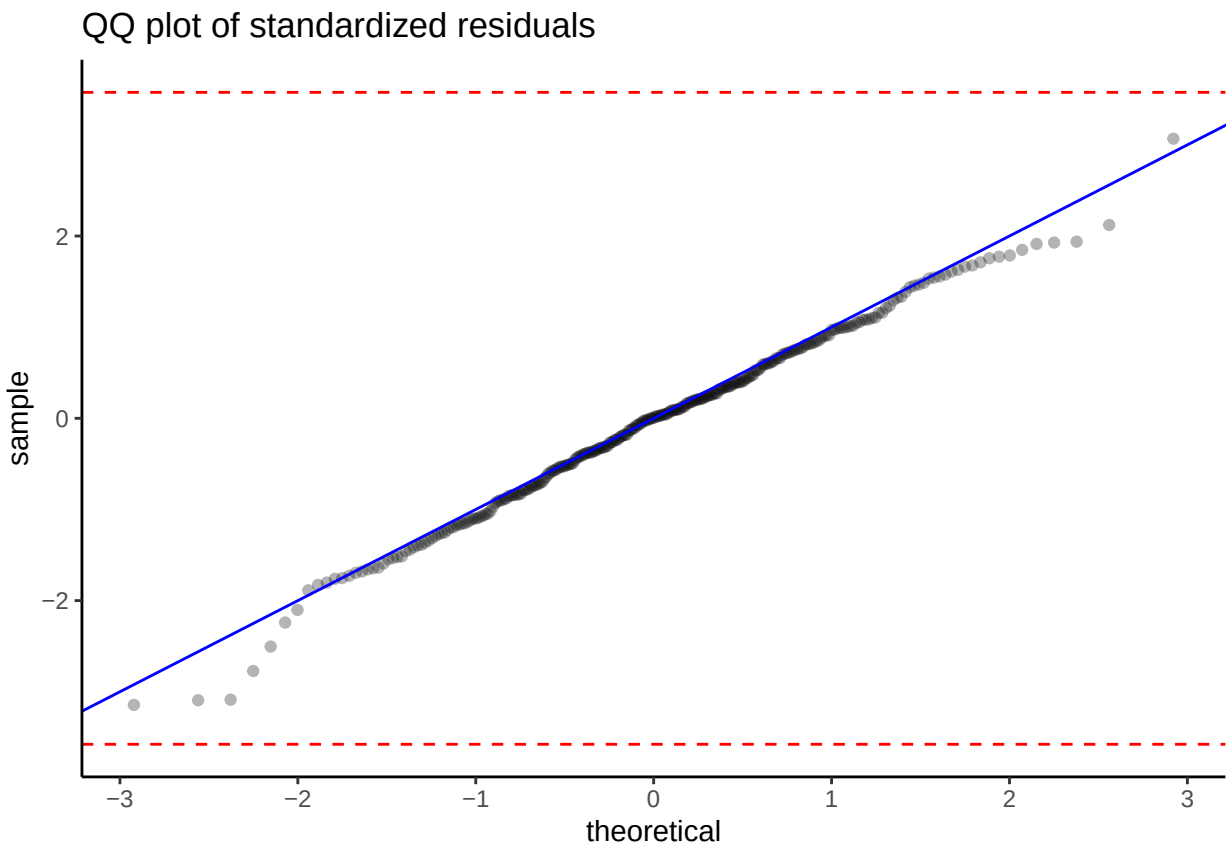
```
##             exposure1 exposure2
## F-statistic  5.889941   6.15238
```

Both of these values are under 10, so we might be concerned about weak instrument bias.

GRAPPLE can also perform MVMR and should be more robust to weak instruments than MV-IVW. It can also account for correlation due to sample overlap. Below, we create a residual correlation matrix from the estimates I obtained using cross-trait LD-score regression.

```
cor_mat <- matrix(c(1, 0.154, 0,
                    0.154, 1, 0,
                    0, 0, 1), nrow = 3, byrow = T)
grapple_res_sbpdbp <- sbpdbp_stroke_full_data %>%
               with(., data.frame(gamma_out1 = outcome_beta,
                                  se_out1 = outcome_se,
                                  gamma_exp1 = exposure_beta[,1],
                                  gamma_exp2 = exposure_beta[,2],
                                  se_exp1 = exposure_se[,1],
                                  se_exp2 = exposure_se[,2])) %>%
           grappleRobustEst(data = .,
                            loss.function = "tukey",
                            cor.mat = cor_mat,
                            plot.it = TRUE)
```

```
## [1] "Test of normality for the standardized residuals:"
## [1] "Anderson-Darling test: p-value = 0.634"
## [1] "Shapiro-Wilk test: p-value = 0.285"
## [1] "0 outliers detected!"
```




```
grapple_res_df <- data.frame(trait = c("sbp", "dbp"),
                             beta_hat = grapple_res_sbpdbp$beta.hat,
                             se = sqrt(diag(grapple_res_sbpdbp$beta.var)),
                             p = grapple_res_sbpdbp$beta.p.value)
grapple_res_df
```

```
##   trait   beta_hat          se          p
## 1   sbp 0.02750405 0.009035781 0.002335292
## 2   dbp 0.01183996 0.013075999 0.365215173
```

The GRAPPLE estimate for SBP is just a little smaller than the MV-IVW estimate.

Finally, we can include BMI to account for covariate adjustment in the blood pressure GWAS (Gilbody et al. 2022).

```
bmi_id <- "ukb-b-19953"
sbp_dbp_bmi_dat <- mv_extract_exposures(id_exposure = c(sbp_id, dbp_id, bmi_id), pop = "EUR")
stroke_sbp_dbp_bmi_dat <- extract_outcome_data(outcomes = stroke_id, snps = sbp_dbp_bmi_dat$SNP)
sbpdbpbmi_stroke_full_data <- mv_harmonise_data(exposure_dat = sbp_dbp_bmi_dat,
                                                outcome_dat = stroke_dbp_dat)
```

```
sbpdbpbmi_stroke_full_data <- readRDS("sbpdbpbmi_stroke_full_data.RDS")
mvivw_res <- mv_multiple(sbpdbpbmi_stroke_full_data)
mvivw_res$result
```

```
##   id.exposure                                exposure       id.outcome
## 1    ieu-b-38  systolic blood pressure || id:ieu-b-38 ebi-a-GCST006906
## 2    ieu-b-39 diastolic blood pressure || id:ieu-b-39 ebi-a-GCST006906
## 3 ukb-b-19953 Body mass index (BMI) || id:ukb-b-19953 ebi-a-GCST006906
##                         outcome nsnp          b          se        pval
## 1 Stroke || id:ebi-a-GCST006906  138 0.02305780 0.008556934 0.007046617
## 2 Stroke || id:ebi-a-GCST006906  220 0.02178947 0.012714534 0.086575629
## 3 Stroke || id:ebi-a-GCST006906    9 0.45958652 0.143464379 0.001357734
```

```
mvmr_input <- with(sbpdbpbmi_stroke_full_data,
                   format_mvmr(BXGs = exposure_beta,
                               BYG = outcome_beta,
                               seBXGs = exposure_se,
                               seBYG = outcome_se,
                               RSID = rownames(exposure_beta)))
strength_mvmr(mvmr_input)
```

```
## Warning in strength_mvmr(mvmr_input): Covariance between effect of genetic
## variants on each exposure not specified. Fixing covariance at 0.
```

```
##
## Conditional F-statistics for instrument strength
##
##             exposure1 exposure2 exposure3
## F-statistic  5.628177  5.846352  7.865408
```

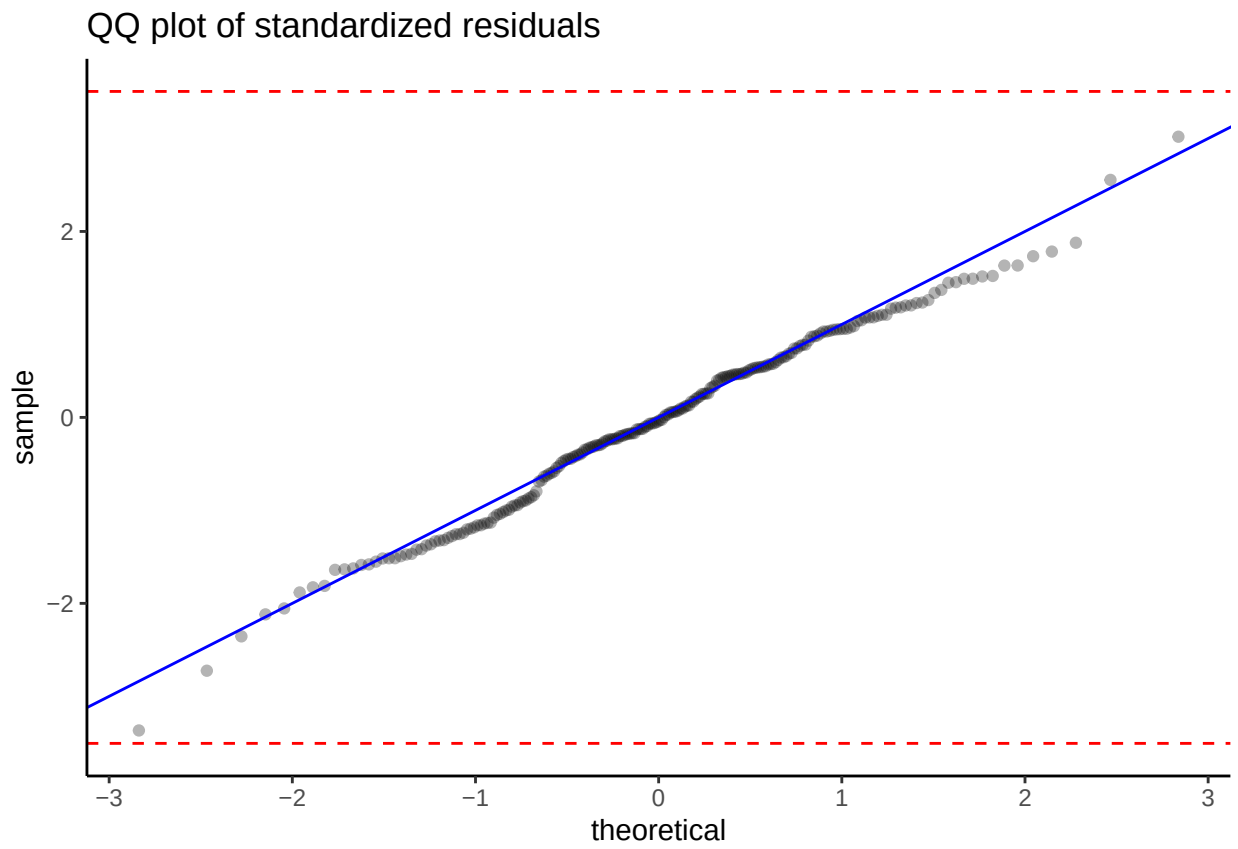
```
##             exposure1 exposure2 exposure3
## F-statistic  5.628177  5.846352  7.865408
```

The conditional F statistics for SBP and DBP are just slightly lower when accounting for BMI, so adding this trait should not increase weak instrument bias very much.

```
cor_mat <- matrix(c(1, 0.154, 0, 0,
                    0.154, 1, 0, 0,
                    0, 0, 1, 0,
                    0, 0, 0, 1), nrow = 4, byrow = T)
grapple_res_sbpdbpbmi <- sbpdbpbmi_stroke_full_data %>%
               with(., data.frame(gamma_out1 = outcome_beta,
                                  se_out1 = outcome_se,
                                  gamma_exp1 = exposure_beta[,1],
                                  gamma_exp2 = exposure_beta[,2],
                                  gamma_exp3 = exposure_beta[,3],
                                  se_exp1 = exposure_se[,1],
                                  se_exp2 = exposure_se[,2],
                                  se_exp3 = exposure_se[,3])) %>%
          grappleRobustEst(data = .,
                           loss.function = "tukey",
                           cor.mat = cor_mat,
                           plot.it = TRUE)
```

```
## Warning in grappleRobustEst(data = ., loss.function = "tukey", cor.mat = cor_mat, : Did not converge
##               Consider to increase niter or decrease tol.
```

```
## [1] "Test of normality for the standardized residuals:"
## [1] "Anderson-Darling test: p-value = 0.0759"
## [1] "Shapiro-Wilk test: p-value = 0.21"
## [1] "0 outliers detected!"
```




```
grapple_res_df <- data.frame(trait = c("sbp", "dbp", "bmi"),
                             beta_hat = grapple_res_sbpdbpbmi$beta.hat,
                             se = sqrt(diag(grapple_res_sbpdbpbmi$beta.var)),
                             p = grapple_res_sbpdbpbmi$beta.p.value)
grapple_res_df
```

```
##   trait   beta_hat         se           p
## 1   sbp 0.02242803 0.01036970 0.030553237
## 2   dbp 0.02109991 0.01542156 0.171246882
## 3   bmi 0.55286725 0.16949430 0.001106825
```

After adjusting for BMI, the estimate for SBP is somewhat attenuated, which could be a result of resolving bias due covariate adjustment in the GWAS, or a result of slightly increased weak instrument bias.

References

Evangelou, Evangelos, Helen R. Warren, David Mosen-Ansorena, Borbala Mifsud, Raha Pazoki, He Gao, Georgios Ntritsos, et al. 2018. "Genetic Analysis of over 1 Million People Identifies 535 New Loci Associated with Blood Pressure Traits." *Nature Genetics* 50 (12): 1755–55. https://doi.org/10.1038/s41588-018-0297-3.

Georgakis, Marios K., and Dipender Gill. 2021. "Mendelian Randomization Studies in Stroke: Exploration of Risk Factors and Drug Targets With Human Genetic Data." *Stroke* 52 (9): 2992–3003. https://doi.org/10.1161/STROKEAHA.120.032617.

Georgakis, Marios K., Dipender Gill, Alastair J. S. Webb, Evangelos Evangelou, Paul Elliott, Cathie L. M. Sudlow, Abbas Dehghan, Rainer Malik, Ioanna Tzoulaki, and Martin Dichgans. 2020. "Genetically Determined Blood Pressure, Antihypertensive Drug Classes, and Risk of Stroke Subtypes." *Neurology* 95 (4): e353–61. https://doi.org/10.1212/WNL.0000000000009814.

Gilbody, Joe, Maria Carolina Borges, George Davey Smith, and Eleanor Sanderson. 2022. "Multivariable MR Can Mitigate Bias in Two-Sample MR Using Covariable-Adjusted Summary Associations." Preprint. Epidemiology. https://doi.org/10.1101/2022.07.19.22277803.

Sanderson, Eleanor, Wes Spiller, and Jack Bowden. 2021. "Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomization." *Statistics in Medicine* n/a (n/a). https://doi.org/10.1002/sim.9133.

Surendran, Praveen, Elena V. Feofanova, Najim Lahrouchi, Ioanna Ntalla, Savita Karthikeyan, James Cook, Lingyan Chen, et al. 2020. "Discovery of Rare Variants Associated with Blood Pressure Regulation Through Meta-Analysis of 1.3 Million Individuals." *Nature Genetics* 52 (12): 1314–32. https://doi.org/10.1038/s41588-020-00713-x.