MR example for Supplementary Material

# Example R and Stata code to perform a multivariable Mendelian randomization (MVMR) analysis

## R code

* Read in the data

suppressPackageStartupMessages({  
 library(tidyverse)  
 library(haven)  
})  
dat <- read\_dta("https://raw.github.com/remlapmot/mrrobust/master/dodata.dta")  
dat <- dat %>% filter(ldlcp2 < 1e-8)

### Example code using the MendelianRandomization package

* Install the package and load it into memory

# install.packages("MendelianRandomization") # uncomment on first run  
library(MendelianRandomization)

* Convert our data frame to the required class

datfmt <- mr\_mvinput(  
 bx = as.matrix(cbind(dat$ldlcbeta, dat$hdlcbeta, dat$tgbeta)),  
 bxse = as.matrix(cbind(dat$ldlcse, dat$hdlcse, dat$tgse)),  
 by = dat$chdbeta,  
 byse = dat$chdse,  
 exposure = "exposure",  
 outcome = "outcome",  
 snps = "snp",  
 effect\_allele = dat$a1,  
 other\_allele = dat$a2,  
 eaf = NA  
)

* Fit an MVMR/MVIVW model

mvivwfit <- mr\_mvivw(datfmt)  
mvivwfit

##   
## Multivariable inverse-variance weighted method  
## (variants uncorrelated, random-effect model)  
##   
## Number of Variants : 73   
##   
## ------------------------------------------------------------------  
## Exposure Estimate Std Error 95% CI p-value  
## exposure\_1 0.429 0.061 0.309, 0.548 0.000  
## exposure\_2 -0.194 0.131 -0.451, 0.062 0.138  
## exposure\_3 0.226 0.123 -0.016, 0.468 0.067  
## ------------------------------------------------------------------  
## Residual standard error = 1.490   
## Heterogeneity test statistic = 155.3766 on 70 degrees of freedom, (p-value = 0.0000)

* Fit an MVMR-Egger model

mvmreggerfit <- mr\_mvegger(datfmt)  
mvmreggerfit

##   
## Multivariable MR-Egger method  
## (variants uncorrelated, random-effect model)  
##   
## Orientated to exposure : 1   
## Number of Variants : 73   
## ------------------------------------------------------------------  
## Exposure Estimate Std Error 95% CI p-value  
## exposure\_1 0.567 0.100 0.371, 0.764 0.000  
## exposure\_2 -0.136 0.133 -0.398, 0.125 0.306  
## exposure\_3 0.274 0.125 0.030, 0.518 0.028  
## (intercept) -0.009 0.005 -0.020, 0.001 0.084  
## ------------------------------------------------------------------  
## Residual standard error = 1.469   
## Heterogeneity test statistic = 148.9290 on 69 degrees of freedom, (p-value = 0.0000)

### Example code using the MVMR package

* Install the package and load it into memory

# remotes::install\_github("wspiller/mvmr") # uncomment on first run  
library(MVMR)

* Create a data object of the required structure

r\_input <- format\_mvmr(  
 BXGs = dat[,c("ldlcbeta","hdlcbeta","tgbeta")],  
 BYG = dat$chdbeta,  
 seBXGs = dat[,c("ldlcse","hdlcse","tgse")],  
 seBYG = dat$chdse,  
 RSID = dat$rsid  
)

* Fit an MVMR model

mvmrfit <- ivw\_mvmr(r\_input)

##   
## Multivariable MR  
##   
## Estimate Std. Error t value Pr(>|t|)  
## exposure1 0.4286200 0.0609661 7.030464 1.099077e-09  
## exposure2 -0.1941989 0.1308289 -1.484372 1.421994e-01  
## exposure3 0.2260456 0.1232828 1.833554 7.097168e-02  
##   
## Residual standard error: 1.49 on 70 degrees of freedom

* Heterogeneity statistic

strength\_mvmr(r\_input)

## Warning in strength\_mvmr(r\_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 126.7447 35.29937 39.32731

## exposure1 exposure2 exposure3  
## F-statistic 126.7447 35.29937 39.32731

## Stata code

* Load the Statamarkdown package to enable Stata code chunks in an R Markdown file

# remotes::install\_github("hemken/statamarkdown") # uncomment on first run  
library(Statamarkdown)

## Stata found at C:/Program Files (x86)/Stata15/StataSE-64.exe

## The 'stata' engine is ready to use.

* Read in the data and create an indicator variable to select observations with *p*-value between the genotype and LDL-C < 10-8

use https://raw.github.com/remlapmot/mrrobust/master/dodata, clear  
gen byte sel1 = (ldlcp2 < 1e-8)

### Example code using the mrrobust package

* Install the mrrobust package using the github package

// output suppressed  
net install github, from("https://haghish.github.io/github/") replace  
gitget mrrobust

* Fit and MVMR model with phenotypes LDL-c and HDL-c (Burgess, Dudbridge, and Thompson 2015).

mvmr chdbeta ldlcbeta hdlcbeta [aw=1/(chdse^2)] if sel1==1

Number of genotypes = 73  
 Number of phenotypes = 2  
 Standard errors: Random effect  
 Residual standard error = 1.514  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .4670719 .0581901 8.03 0.000 .3530214 .5811224  
 hdlcbeta | -.2930048 .1211822 -2.42 0.016 -.5305175 -.0554921  
------------------------------------------------------------------------------

* Additionally include a third phenotype – triglycerides.

mvmr chdbeta ldlcbeta hdlcbeta tgbeta [aw=1/(chdse^2)] if sel1==1

Number of genotypes = 73  
 Number of phenotypes = 3  
 Standard errors: Random effect  
 Residual standard error = 1.490  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .42862 .0609661 7.03 0.000 .3091286 .5481113  
 hdlcbeta | -.1941989 .1308289 -1.48 0.138 -.4506189 .0622211  
 tgbeta | .2260456 .1232828 1.83 0.067 -.0155842 .4676755  
------------------------------------------------------------------------------

* Report the QA statistic for instrument validity and the conditional F-statistics for instrument strength for each phenotype (Sanderson et al. 2019; Sanderson, Spiller, and Bowden 2020).

mvmr chdbeta ldlcbeta hdlcbeta tgbeta [aw=1/(chdse^2)] if sel1==1, ///  
 gxse(ldlcse hdlcse tgse)

> gxse(ldlcse hdlcse tgse)  
  
 Number of genotypes = 73  
 Number of phenotypes = 3  
 Standard errors: Random effect  
 Residual standard error = 1.490  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .42862 .0609661 7.03 0.000 .3091286 .5481113  
 hdlcbeta | -.1941989 .1308289 -1.48 0.138 -.4506189 .0622211  
 tgbeta | .2260456 .1232828 1.83 0.067 -.0155842 .4676755  
------------------------------------------------------------------------------  
 Q\_A statistic for instrument validity; chi2(70) = 152.88 (p = 0.0000)  
 Conditional F-statistics for instrument strength:  
 F\_x1 = 130.31 (ldlcbeta)  
 F\_x2 = 36.29 (hdlcbeta)  
 F\_x3 = 40.44 (tgbeta)

* Fit an MVMR-Egger regression (Rees, Wood, and Burgess 2017), orienting the model with respect to the first phenotype in the main *varlist*.

mrmvegger chdbeta ldlcbeta hdlcbeta tgbeta [aw=1/(chdse^2)] if sel1==1

MVMR-Egger model oriented wrt: ldlcbeta  
 Number of genotypes = 73  
 Number of phenotypes = 3  
 Residual standard error = 1.469  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .5672993 .1002611 5.66 0.000 .370791 .7638075  
 hdlcbeta | -.1364113 .1332727 -1.02 0.306 -.3976209 .1247983  
 tgbeta | .2739803 .1246927 2.20 0.028 .0295871 .5183735  
 \_cons | -.0093655 .0054187 -1.73 0.084 -.019986 .001255  
------------------------------------------------------------------------------

We can also orient the model with respect to HDL-C.

mrmvegger chdbeta ldlcbeta hdlcbeta tgbeta [aw=1/(chdse^2)] if sel1==1, ///  
 orient(2)

> orient(2)  
  
 MVMR-Egger model oriented wrt: hdlcbeta  
 Number of genotypes = 73  
 Number of phenotypes = 3  
 Residual standard error = 1.501  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .4286398 .0614056 6.98 0.000 .308287 .5489926  
 hdlcbeta | -.1989637 .1541909 -1.29 0.197 -.5011723 .1032449  
 tgbeta | .2256794 .1243221 1.82 0.069 -.0179875 .4693463  
 \_cons | .0002155 .0036218 0.06 0.953 -.006883 .0073141  
------------------------------------------------------------------------------

Or we can orient the model with respect to triglycerides.

mrmvegger chdbeta ldlcbeta hdlcbeta tgbeta [aw=1/(chdse^2)] if sel1==1, ///  
 orient(3)

> orient(3)  
  
 MVMR-Egger model oriented wrt: tgbeta  
 Number of genotypes = 73  
 Number of phenotypes = 3  
 Residual standard error = 1.499  
------------------------------------------------------------------------------  
 | Coef. Std. Err. z P>|z| [95% Conf. Interval]  
-------------+----------------------------------------------------------------  
chdbeta |  
 ldlcbeta | .4203073 .0660026 6.37 0.000 .2909447 .54967  
 hdlcbeta | -.1903089 .1321536 -1.44 0.150 -.4493252 .0687075  
 tgbeta | .2065651 .1365427 1.51 0.130 -.0610537 .474184  
 \_cons | .0013499 .003951 0.34 0.733 -.0063939 .0090936  
------------------------------------------------------------------------------

## R session information for reproducibility

if (!requireNamespace("sessioninfo", quietly = TRUE)) install.packages("sessioninfo")  
sessioninfo::session\_info()

- Session info -----------------------------------------------------------------------------------  
 setting value   
 version R version 4.0.2 (2020-06-22)  
 os Windows 10 x64   
 system x86\_64, mingw32   
 ui RTerm   
 language (EN)   
 collate English\_United Kingdom.1252   
 ctype English\_United Kingdom.1252   
 tz Europe/London   
 date 2020-10-06   
  
- Packages ---------------------------------------------------------------------------------------  
 package \* version date lib source   
 arrangements 1.1.9 2020-09-13 [1] CRAN (R 4.0.2)   
 assertthat 0.2.1 2019-03-21 [1] CRAN (R 4.0.0)   
 backports 1.1.10 2020-09-15 [1] CRAN (R 4.0.2)   
 blob 1.2.1 2020-01-20 [1] CRAN (R 4.0.2)   
 broom 0.7.1 2020-10-02 [1] CRAN (R 4.0.2)   
 cellranger 1.1.0 2016-07-27 [1] CRAN (R 4.0.0)   
 cli 2.0.2 2020-02-28 [1] CRAN (R 4.0.0)   
 codetools 0.2-16 2018-12-24 [2] CRAN (R 4.0.2)   
 colorspace 1.4-1 2019-03-18 [1] CRAN (R 4.0.0)   
 conquer 1.0.2 2020-08-27 [1] CRAN (R 4.0.2)   
 crayon 1.3.4 2017-09-16 [1] CRAN (R 4.0.0)   
 curl 4.3 2019-12-02 [1] CRAN (R 4.0.0)   
 data.table 1.13.0 2020-07-24 [1] CRAN (R 4.0.2)   
 DBI 1.1.0 2019-12-15 [1] CRAN (R 4.0.0)   
 dbplyr 1.4.4 2020-05-27 [1] CRAN (R 4.0.2)   
 DEoptimR 1.0-8 2016-11-19 [1] CRAN (R 4.0.0)   
 digest 0.6.25 2020-02-23 [1] CRAN (R 4.0.0)   
 dplyr \* 1.0.2 2020-08-18 [1] CRAN (R 4.0.2)   
 ellipsis 0.3.1 2020-05-15 [1] CRAN (R 4.0.2)   
 evaluate 0.14 2019-05-28 [1] CRAN (R 4.0.0)   
 fansi 0.4.1 2020-01-08 [1] CRAN (R 4.0.0)   
 forcats \* 0.5.0 2020-03-01 [1] CRAN (R 4.0.0)   
 foreach 1.5.0 2020-03-30 [1] CRAN (R 4.0.2)   
 fs 1.5.0 2020-07-31 [1] CRAN (R 4.0.2)   
 generics 0.0.2 2018-11-29 [1] CRAN (R 4.0.0)   
 ggplot2 \* 3.3.2 2020-06-19 [1] CRAN (R 4.0.2)   
 glmnet 4.0-2 2020-06-16 [1] CRAN (R 4.0.2)   
 glue 1.4.2 2020-08-27 [1] CRAN (R 4.0.2)   
 gmp 0.6-1 2020-10-01 [1] CRAN (R 4.0.2)   
 gtable 0.3.0 2019-03-25 [1] CRAN (R 4.0.0)   
 haven \* 2.3.1 2020-06-01 [1] CRAN (R 4.0.2)   
 hms 0.5.3 2020-01-08 [1] CRAN (R 4.0.0)   
 htmltools 0.5.0 2020-06-16 [1] CRAN (R 4.0.2)   
 htmlwidgets 1.5.2 2020-10-03 [1] CRAN (R 4.0.2)   
 httr 1.4.2 2020-07-20 [1] CRAN (R 4.0.2)   
 iterators 1.0.12 2019-07-26 [1] CRAN (R 4.0.0)   
 iterpc 0.4.2 2020-01-10 [1] CRAN (R 4.0.0)   
 jsonlite 1.7.1 2020-09-07 [1] CRAN (R 4.0.2)   
 knitr 1.30 2020-09-22 [1] CRAN (R 4.0.2)   
 lattice 0.20-41 2020-04-02 [2] CRAN (R 4.0.2)   
 lazyeval 0.2.2 2019-03-15 [1] CRAN (R 4.0.0)   
 lifecycle 0.2.0 2020-03-06 [1] CRAN (R 4.0.0)   
 lubridate 1.7.9 2020-06-08 [1] CRAN (R 4.0.2)   
 magrittr 1.5 2014-11-22 [1] CRAN (R 4.0.0)   
 Matrix 1.2-18 2019-11-27 [2] CRAN (R 4.0.2)   
 MatrixModels 0.4-1 2015-08-22 [1] CRAN (R 4.0.0)   
 matrixStats 0.57.0 2020-09-25 [1] CRAN (R 4.0.2)   
 MendelianRandomization \* 0.5.0 2020-09-30 [1] CRAN (R 4.0.2)   
 modelr 0.1.8 2020-05-19 [1] CRAN (R 4.0.2)   
 munsell 0.5.0 2018-06-12 [1] CRAN (R 4.0.0)   
 MVMR \* 0.2 2020-09-29 [1] Github (wspiller/mvmr@dde107a)   
 pillar 1.4.6 2020-07-10 [1] CRAN (R 4.0.2)   
 pkgconfig 2.0.3 2019-09-22 [1] CRAN (R 4.0.0)   
 plotly 4.9.2.1 2020-04-04 [1] CRAN (R 4.0.2)   
 purrr \* 0.3.4 2020-04-17 [1] CRAN (R 4.0.2)   
 quantreg 5.73 2020-10-02 [1] CRAN (R 4.0.2)   
 R6 2.4.1 2019-11-12 [1] CRAN (R 4.0.0)   
 Rcpp 1.0.5 2020-07-06 [1] CRAN (R 4.0.2)   
 readr \* 1.4.0 2020-10-05 [1] CRAN (R 4.0.2)   
 readxl 1.3.1 2019-03-13 [1] CRAN (R 4.0.0)   
 reprex 0.3.0 2019-05-16 [1] CRAN (R 4.0.0)   
 rjson 0.2.20 2018-06-08 [1] CRAN (R 4.0.0)   
 rlang 0.4.7 2020-07-09 [1] CRAN (R 4.0.2)   
 rmarkdown 2.4 2020-09-30 [1] CRAN (R 4.0.2)   
 robustbase 0.93-6 2020-03-23 [1] CRAN (R 4.0.2)   
 rstudioapi 0.11 2020-02-07 [1] CRAN (R 4.0.0)   
 rvest 0.3.6 2020-07-25 [1] CRAN (R 4.0.2)   
 scales 1.1.1 2020-05-11 [1] CRAN (R 4.0.2)   
 sessioninfo 1.1.1 2018-11-05 [1] CRAN (R 4.0.0)   
 shape 1.4.5 2020-09-13 [1] CRAN (R 4.0.2)   
 SparseM 1.78 2019-12-13 [1] CRAN (R 4.0.0)   
 Statamarkdown \* 0.5.3 2020-09-29 [1] Github (hemken/statamarkdown@89ff92f)  
 stringi 1.5.3 2020-09-09 [1] CRAN (R 4.0.2)   
 stringr \* 1.4.0 2019-02-10 [1] CRAN (R 4.0.0)   
 survival 3.2-7 2020-09-28 [2] CRAN (R 4.0.2)   
 tibble \* 3.0.3 2020-07-10 [1] CRAN (R 4.0.2)   
 tidyr \* 1.1.2 2020-08-27 [1] CRAN (R 4.0.2)   
 tidyselect 1.1.0 2020-05-11 [1] CRAN (R 4.0.2)   
 tidyverse \* 1.3.0 2019-11-21 [1] CRAN (R 4.0.0)   
 vctrs 0.3.4 2020-08-29 [1] CRAN (R 4.0.2)   
 viridisLite 0.3.0 2018-02-01 [1] CRAN (R 4.0.0)   
 withr 2.3.0 2020-09-22 [1] CRAN (R 4.0.2)   
 xfun 0.18 2020-09-29 [1] CRAN (R 4.0.2)   
 xml2 1.3.2 2020-04-23 [1] CRAN (R 4.0.2)   
 yaml 2.2.1 2020-02-01 [1] CRAN (R 4.0.0)   
  
[1] C:/Users/eptmp/OneDrive - University of Bristol/Documents/R/win-library/4.0  
[2] C:/Program Files/R/R-4.0.2/library

## Stata session information for reproducibility

di c(version)  
github list mrrobust

15.1  
  
 --------------------------------------------------------------------------  
 Date Name Version user/repository Latest release  
 --------------------------------------------------------------------------  
 6 Oct 2020 mrrobust v0.1.0 remlapmot/mrrobust v0.1.0  
 --------------------------------------------------------------------------

## References

Burgess, S, F Dudbridge, and SG Thompson. 2015. “Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects.” *American Journal of Epidemiology* 181 (4): 251–60. <https://doi.org/10.1093/aje/kwu283>.

Rees, J, A Wood, and S Burgess. 2017. “Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy.” *Statistics in Medicine* 36 (29): 4705–18. <https://doi.org/10.1002/sim.7492>.

Sanderson, E, G Davey Smith, F Windmeijer, and J Bowden. 2019. “An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings.” *International Journal of Epidemiology* 48 (3): 713–27. <https://doi.org/10.1093/ije/dyy262>.

Sanderson, E, W Spiller, and J Bowden. 2020. “Testing and Correcting for Weak and Pleiotropic Instruments in Two-Sample Multivariable Mendelian Randomisation.” *bioRxiv*. <https://doi.org/10.1101/2020.04.02.021980>.