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

```
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

```
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

```
library(Statamarkdown)
```

- 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

```
// Note: output suppressed
net install mrrobust, from("https://raw.github.com/remlapmot/mrrobust/master/")
mrdeps
```

- 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)
```

```
> e 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)
```

```
> t(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)
```

```
> t(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

```
library(sessioninfo)
session_info()
```

```
- Session info ---------------------------------------------------------------------------------------
 setting  value
 version  R version 4.1.1 (2021-08-10)
 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     2021-09-20

- Packages -------------------------------------------------------------------------------------------
 package               * version date       lib source
 arrangements            1.1.9   2020-09-13 [1] CRAN (R 4.1.0)
 assertthat              0.2.1   2019-03-21 [1] CRAN (R 4.1.0)
 backports               1.2.1   2020-12-09 [1] CRAN (R 4.1.0)
 broom                   0.7.9   2021-07-27 [1] CRAN (R 4.1.0)
 cellranger              1.1.0   2016-07-27 [1] CRAN (R 4.1.0)
 cli                     3.0.1   2021-07-17 [1] CRAN (R 4.1.0)
 codetools               0.2-18  2020-11-04 [2] CRAN (R 4.1.1)
 colorspace              2.0-2   2021-06-24 [1] CRAN (R 4.1.0)
 conquer                 1.0.2   2020-08-27 [1] CRAN (R 4.1.0)
 crayon                  1.4.1   2021-02-08 [1] CRAN (R 4.1.0)
 curl                    4.3.2   2021-06-23 [1] CRAN (R 4.1.0)
 data.table              1.14.0  2021-02-21 [1] CRAN (R 4.1.0)
 DBI                     1.1.1   2021-01-15 [1] CRAN (R 4.1.0)
 dbplyr                  2.1.1   2021-04-06 [1] CRAN (R 4.1.0)
```

```
DEoptimR                1.0-9   2021-05-24 [1] CRAN (R 4.1.0)
digest                  0.6.27  2020-10-24 [1] CRAN (R 4.1.0)
dplyr                 * 1.0.7   2021-06-18 [1] CRAN (R 4.1.0)
ellipsis                0.3.2   2021-04-29 [1] CRAN (R 4.1.0)
evaluate                0.14    2019-05-28 [1] CRAN (R 4.1.0)
fansi                   0.5.0   2021-05-25 [1] CRAN (R 4.1.0)
fastmap                 1.1.0   2021-01-25 [1] CRAN (R 4.1.0)
forcats               * 0.5.1   2021-01-27 [1] CRAN (R 4.1.0)
foreach                 1.5.1   2020-10-15 [1] CRAN (R 4.1.0)
fs                      1.5.0   2020-07-31 [1] CRAN (R 4.1.0)
generics                0.1.0   2020-10-31 [1] CRAN (R 4.1.0)
ggplot2               * 3.3.5   2021-06-25 [1] CRAN (R 4.1.0)
glmnet                  4.1-2   2021-06-24 [1] CRAN (R 4.1.0)
glue                    1.4.2   2020-08-27 [1] CRAN (R 4.1.0)
gmp                     0.6-2   2021-01-07 [1] CRAN (R 4.1.0)
gtable                  0.3.0   2019-03-25 [1] CRAN (R 4.1.0)
haven                 * 2.4.3   2021-08-04 [1] CRAN (R 4.1.0)
hms                     1.1.0   2021-05-17 [1] CRAN (R 4.1.0)
htmltools               0.5.2   2021-08-25 [1] CRAN (R 4.1.1)
htmlwidgets             1.5.4   2021-09-08 [1] CRAN (R 4.1.1)
httr                    1.4.2   2020-07-20 [1] CRAN (R 4.1.0)
iterators               1.0.13  2020-10-15 [1] CRAN (R 4.1.0)
iterpc                  0.4.2   2020-01-10 [1] CRAN (R 4.1.0)
jsonlite                1.7.2   2020-12-09 [1] CRAN (R 4.1.0)
knitr                   1.34    2021-09-09 [1] CRAN (R 4.1.1)
lattice                 0.20-44 2021-05-02 [2] CRAN (R 4.1.1)
lazyeval                0.2.2   2019-03-15 [1] CRAN (R 4.1.0)
lifecycle               1.0.0   2021-02-15 [1] CRAN (R 4.1.0)
lubridate               1.7.10  2021-02-26 [1] CRAN (R 4.1.0)
magrittr                2.0.1   2020-11-17 [1] CRAN (R 4.1.0)
Matrix                  1.3-4   2021-06-01 [2] CRAN (R 4.1.1)
MatrixModels            0.5-0   2021-03-02 [1] CRAN (R 4.1.0)
matrixStats             0.61.0  2021-09-17 [1] CRAN (R 4.1.1)
MendelianRandomization * 0.5.1   2021-04-16 [1] CRAN (R 4.1.0)
modelr                  0.1.8   2020-05-19 [1] CRAN (R 4.1.0)
munsell                 0.5.0   2018-06-12 [1] CRAN (R 4.1.0)
MVMR                  * 0.3     2021-08-11 [1] Github (wspiller/mvmr@a6388a8)
pillar                  1.6.2   2021-07-29 [1] CRAN (R 4.1.0)
pkgconfig               2.0.3   2019-09-22 [1] CRAN (R 4.1.0)
plotly                  4.9.4.1 2021-06-18 [1] CRAN (R 4.1.0)
purrr                 * 0.3.4   2020-04-17 [1] CRAN (R 4.1.0)
quantreg                5.86    2021-06-06 [1] CRAN (R 4.1.0)
R6                      2.5.1   2021-08-19 [1] CRAN (R 4.1.1)
RadialMR                1.0     2021-07-12 [1] Github (WSpiller/RadialMR@d63d3fc)
Rcpp                    1.0.7   2021-07-07 [1] CRAN (R 4.1.0)
readr                 * 2.0.1   2021-08-10 [1] CRAN (R 4.1.1)
readxl                  1.3.1   2019-03-13 [1] CRAN (R 4.1.0)
remotes                 2.4.0   2021-06-02 [1] CRAN (R 4.1.0)
reprex                  2.0.1   2021-08-05 [1] CRAN (R 4.1.0)
rjson                   0.2.20  2018-06-08 [1] CRAN (R 4.1.0)
rlang                   0.4.11  2021-04-30 [1] CRAN (R 4.1.0)
rmarkdown               2.11    2021-09-14 [1] CRAN (R 4.1.1)
robustbase              0.93-8  2021-06-02 [1] CRAN (R 4.1.0)
rstudioapi              0.13    2020-11-12 [1] CRAN (R 4.1.0)
```

```
 rvest                   1.0.1   2021-07-26 [1] CRAN (R 4.1.0)
 scales                  1.1.1   2020-05-11 [1] CRAN (R 4.1.0)
 sessioninfo           * 1.1.1   2018-11-05 [1] CRAN (R 4.1.0)
 shape                   1.4.6   2021-05-19 [1] CRAN (R 4.1.0)
 SparseM                 1.81    2021-02-18 [1] CRAN (R 4.1.0)
 Statamarkdown         * 0.7.0   2021-09-15 [1] Github (Hemken/Statamarkdown@a68a8b9)
 stringi                 1.7.4   2021-08-25 [1] CRAN (R 4.1.1)
 stringr               * 1.4.0   2019-02-10 [1] CRAN (R 4.1.0)
 survival                3.2-13  2021-08-24 [2] CRAN (R 4.1.1)
 tibble                * 3.1.4   2021-08-25 [1] CRAN (R 4.1.1)
 tidyr                 * 1.1.3   2021-03-03 [1] CRAN (R 4.1.0)
 tidyselect              1.1.1   2021-04-30 [1] CRAN (R 4.1.0)
 tidyverse             * 1.3.1   2021-04-15 [1] CRAN (R 4.1.0)
 tzdb                    0.1.2   2021-07-20 [1] CRAN (R 4.1.0)
 utf8                    1.2.2   2021-07-24 [1] CRAN (R 4.1.0)
 vctrs                   0.3.8   2021-04-29 [1] CRAN (R 4.1.0)
 viridisLite             0.4.0   2021-04-13 [1] CRAN (R 4.1.0)
 withr                   2.4.2   2021-04-18 [1] CRAN (R 4.1.0)
 xfun                    0.26    2021-09-14 [1] CRAN (R 4.1.1)
 xml2                    1.3.2   2020-04-23 [1] CRAN (R 4.1.0)
 yaml                    2.2.1   2020-02-01 [1] CRAN (R 4.1.0)

[1] C:/Users/eptmp/Documents/R/win-library/4.1
[2] C:/Program Files/R/R-4.1.1/library
```

Stata session information for reproducibility

```
di c(version)
ado describe mrrobust
```

```
16.1


------------------------------------------------------------------------------
[89] package mrrobust from https://raw.github.com/remlapmot/mrrobust/master
------------------------------------------------------------------------------

TITLE
      'mrrobust': Stata package for two-sample Mendelian randomization analyses

DESCRIPTION/AUTHOR(S)
      Author: Tom Palmer
      Distribution-Date: 20210917

INSTALLATION FILES
      m\mrmedian.ado
      m\mrmedian.sthlp
      m\mrmedianobs.ado
      m\mrmedianobs_work.ado
      m\mrmedianobs.sthlp
      m\mregger.ado
      m\mregger.sthlp
```

```
      m\mrrobust.sthlp
      m\mreggerplot.ado
      m\mreggerplot.sthlp
      m\mrmodal.ado
      m\mrmodal.sthlp
      m\mrratio.ado
      m\mrratio.sthlp
      m\mrivests.ado
      m\mrivests.sthlp
      m\mrforest.ado
      m\mrforest.sthlp
      m\mreggersimex.ado
      m\mreggersimex.sthlp
      m\mreggersimexonce.ado
      m\mrmodalplot.ado
      m\mrmodalplot.sthlp
      m\mrfunnel.ado
      m\mrfunnel.sthlp
      m\mrdeps.ado
      m\mrdeps.sthlp
      m\mr.ado
      m\mr.sthlp
      m\mrmvivw.ado
      m\mrmvivw.sthlp
      m\mvivw.ado
      m\mvivw.sthlp
      m\mvmr.ado
      m\mvmr.sthlp
      m\mrmvegger.ado
      m\mrmvegger.sthlp
      m\mrleaveoneout.ado
      m\mrleaveoneout.sthlp
      m\mrrobust-author.ihlp

INSTALLED ON
      20 Sep 2021
--------------------------------------------------------------------------------
```

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: 251--260. 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: 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: 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.