



SOFTWARE TOOL ARTICLE

REVISED MendelianRandomization v0.5.0: updates to an R package for performing Mendelian randomization analyses using summarized data [version 2; peer review: 1 approved, 2 approved with reservations]

Jim R. Broadbent¹, Christopher N. Foley^{id}², Andrew J. Grant^{id}², Amy M. Mason^{id}¹, James R. Staley¹, Stephen Burgess^{id}^{1,2}

¹Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, CB2 0SR, UK

²MRC Biostatistics Unit, University of Cambridge, Cambridge, CB1 8RN, UK

V2 First published: 22 Oct 2020, 5:252
<https://doi.org/10.12688/wellcomeopenres.16374.1>
 Latest published: 10 Dec 2020, 5:252
<https://doi.org/10.12688/wellcomeopenres.16374.2>

Abstract

The MendelianRandomization package is a software package written for the R software environment that implements methods for Mendelian randomization based on summarized data. In this manuscript, we describe functions that have been added to the package or updated in recent years. These features can be divided into four categories: robust methods for Mendelian randomization, methods for multivariable Mendelian randomization, functions for data visualization, and the ability to load data into the package seamlessly from the PhenoScanner web-resource. We provide examples of the graphical output produced by the data visualization commands, as well as syntax for obtaining suitable data and performing a Mendelian randomization analysis in a single line of code.

Keywords

Mendelian randomization, instrumental variable, summarized data, genetic epidemiology, post-GWAS analysis, causal inference, genetic associations.

Open Peer Review

Reviewer Status ? ? ✓

	Invited Reviewers		
	1	2	3
version 2 (revision) 10 Dec 2020			✓ report
version 1 22 Oct 2020	? report	? report	? report

1. **Karl Smith Byrne**, The International Agency for Research on Cancer, Lyon, France
2. **Wes Spiller**^{id}, University of Bristol, Bristol, UK
3. **Marilyn C. Cornelis**, Northwestern University, Chicago, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Stephen Burgess (sb452@medschl.cam.ac.uk)

Author roles: **Broadbent JR:** Software, Writing – Review & Editing; **Foley CN:** Software, Writing – Review & Editing; **Grant AJ:** Software, Writing – Review & Editing; **Mason AM:** Software, Writing – Review & Editing; **Staley JR:** Software, Writing – Review & Editing; **Burgess S:** Software, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work is supported by the Wellcome Trust through a Sir Henry Dale Fellowship jointly funded by the Royal Society to SB [204623]. This work was also supported by funding from the National Institute for Health Research (Cambridge Biomedical Research Centre at the Cambridge University Hospitals National Health Service Foundation Trust). The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research or the Department of Health and Social Care.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Broadbent JR *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Broadbent JR, Foley CN, Grant AJ *et al.* **MendelianRandomization v0.5.0: updates to an R package for performing Mendelian randomization analyses using summarized data [version 2; peer review: 1 approved, 2 approved with reservations]** Wellcome Open Research 2020, 5:252 <https://doi.org/10.12688/wellcomeopenres.16374.2>

First published: 22 Oct 2020, 5:252 <https://doi.org/10.12688/wellcomeopenres.16374.1>

REVISED Amendments from Version 1

We have edited the manuscript text to address the reviewers' comments. Specifically, we have added more information on how to access information on the various estimation commands: the options for inputs and the outputs reported by the various functions. We have added more detail on correlated variants and on multivariable Mendelian randomization. We have provided more information on how the inverse-variance weighted method is implemented in the univariable and multivariable settings. We have also added a new table that lists the various estimation methods, the corresponding commands, some brief strengths and weaknesses, and provided references.

Any further responses from the reviewers can be found at the end of the article

Introduction

Mendelian randomization is an epidemiological technique that uses genetic variants to link risk factors to outcomes^{1,2}. The MendelianRandomization package is a software package written for the R software environment³ that implements methods for Mendelian randomization based on summarized data⁴. Summarized data are genetic associations with risk factors and outcomes taken from regression analyses that have been performed for each genetic variant in turn⁵. Such data (beta-coefficients and standard errors) are generated in a genome-wide association study, and have been publicly reported for hundreds of thousands of variants by many large studies and consortia⁶. While the basic functionality and initial features of the package have been discussed previously⁷, several functions have been added to the package in recent years. These features can be divided into four categories: robust methods for Mendelian randomization, methods for multivariable Mendelian randomization, functions for data visualization, and the ability to load data into the package seamlessly from the PhenoScanner web-resource. We discuss each of these categories in turn, describing the various options available to investigators. A list of functions in the package is provided as [Table 1](#). We do not discuss in detail the properties of the various methods

Table 1. Functions available in the MendelianRandomization package.

Functions are divided into five categories: data entry functions, univariable estimation methods, multivariable estimation methods, data visualization functions, and functions that load data from PhenoScanner.

Function	Description	Status
mr_input	Data entry for univariable analysis	
mr_mvinput	Data entry for multivariable analysis	*
mr_ivw	Inverse-variance weighted (IVW) method	†
mr_median	Median method	
mr_egger	MR-Egger method	
mr_maxlik	Maximum likelihood method	
mr_mbe	Mode-based estimation method	*
mr_hetpen	Heterogeneity penalized method	*
mr_conmix	Contamination mixture method	*
mr_lasso	Lasso method	*
mr_allmethods	Runs several methods	
mr_mvivw	Multivariable IVW method	*
mr_mvmedian	Multivariable median-based method	*
mr_mvegger	Multivariable MR-Egger method	*
mr_mvlasso	Multivariable lasso method	*
mr_plot	Scatter plot	†
mr_forest	Forest plot	*
mr_funnel	Funnel plot	*
mr_loo	Leave-one-out plot	*
extract.pheno.csv	Data entry from PhenoScanner .csv file (legacy)	
pheno_input	Data entry from web-based PhenoScanner	*

or the reasons for choosing between the various options presented; we would encourage users to read the relevant references for the methods or the recently-published guidelines paper on performing Mendelian randomization investigations⁸. We also encourage users to consult the package documentation, which describes all the options available for each method in greater detail. The aim of this paper is to provide a broad overview of the package.

Methods

Implementation

The initial release of the MendelianRandomization package included four functions for the estimation of causal effects based on summarized genetic data in a univariable (that is, one risk factor) Mendelian randomization framework. These were `mr_ivw` (inverse-variance weighted method, IVW)⁹, `mr_egger` (MR-Egger method)¹⁰, `mr_median` (simple and weighted median methods)¹¹, and `mr_maxlik` (maximum likelihood method)⁹. Each of these estimation functions takes an *MRInput* object as input, created using the `mr_input` command. The syntax is:

```
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))
```

where `ldlc` and `ldlcse` are genetic associations with low-density lipoprotein (LDL) cholesterol and their standard errors for 28 genetic variants as previously reported by Waterworth *et al.*¹², and `chdlodds` and `chdloddsse` are genetic associations with coronary heart disease risk for the same variants. These data variables are provided with the package. Syntax for the default operation of the `mr_egger` and `mr_median` commands (and all the other univariable estimation commands discussed in this paper) is identical, although user-options and the output from each method is different.

Some methods rely on all variants being uncorrelated; others allow correlated variants using the `correl` option. Using correlated variants requires the specification of the correlation matrix between genetic variants, on the assumption that the correlations between the genetic variants are the same as the correlations between the genetic association estimates⁴. Correlations are typically estimated from reference data, such as those from European-descent participants of the 1000 Genomes Project that can be obtained using the `ld_matrix` command in the TwoSampleMR package¹³. Care should be taken that entries in the correlation matrix are harmonized to the same effect and reference alleles as the genetic associations¹⁴; if the correlation matrix was calculated with the effect and reference alleles reversed, then the positive and negative signs should be flipped for the relevant column and row of the matrix (the diagonal terms should remain as +1). Exemplar data on genetic associations with calcium and fasting glucose for correlated variants are provided in the package. The IVW method can be applied to these data using the syntax:

```
mr_ivw(mr_input(calcium, calciumse, fastgluc, fastglucose, corr=calc.rho))
```

where `calc.rho` is the correlation matrix.

All methods allow confidence intervals to be calculated using a t-distribution rather than a normal distribution (`distribution = "t-dist"`) or based on a different significance level (`alpha = 0.05` corresponds to a 95% confidence interval). Other options are specific to particular methods; a list of input options for each method can be found in the package documentation under the subheading “Arguments”; for the `mr_ivw` method, this is accessed by the command `?mr_ivw`.

Each method provides output in a slightly different format. Generally, the estimate from the method is in the *Estimate* slot, its standard error is in the *StdError* slot, and the lower and upper limits of the confidence interval for the estimate are in the *CILower* and *CIUpper* slots. For the `mr_ivw` command, these can be accessed via:

```
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))$Estimate
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))$StdError
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))$CILower
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))$CIUpper
```

A list of output slots for each method can be found in the package documentation under the subheading “Value”; for the `mr_ivw` method, this is accessed by the command `?mr_ivw`.

Operation

The R software environment runs on a wide variety of UNIX platforms, Windows, and MacOS, and requires minimal computer resources (256 kilobytes of RAM is recommended). The package requires R version 3.0.1 or higher.

Use cases

Robust methods for Mendelian randomization

A brief description of each method is given in [Table 2](#). These methods were discussed in greater detail and compared in a review of robust methods for Mendelian randomization¹⁵.

The IVW method is implemented by weighted linear regression of the genetic associations with the outcome on the genetic associations with the risk factor⁴. There are two options in the `mr_ivw` method that represent different robust methods. The `robust` option performs the IVW method method using robust regression (referred to as MR-Robust)¹⁶. The `penalized` option performs the IVW method with penalized weights¹⁶. The syntax is:

```
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse), robust=TRUE)
mr_ivw(mr_input(ldlc, ldlcse, chdlodds, chdloddsse), penalized=TRUE)
```

Other methods implemented in the package are the mode-based method (`mr_mbe`)¹⁷, the heterogeneity penalized method (`mr_hetpen`)¹⁸, the lasso method (`mr_lasso`)¹⁶, and the contamination mixture method (`mr_conmix`)¹⁹. As for the `mr_ivw` command, the syntax is:

```
mr_mbe(mr_input(ldlc, ldlcse, chdlodds, chdloddsse))
```

and similarly for the other methods.

The `mr_mbe` method has options `weighting = "weighted"` or `weighting = "simple"`, corresponding to weighted and unweighted versions of the method. It also has options `stderr = "simple"` or `stderr = "delta"` corresponding to first- and second-order standard errors.

Table 2. Comparison of univariable methods implemented in the MendelianRandomization package.

A more detailed comparison of robust methods for Mendelian randomization can be found in a recent review¹⁵. Abbreviation: InSIDE = instrument strength independent of direct effect.

Method	Function name	Strengths and weaknesses	Reference
Inverse variance weighted	<code>mr_ivw</code>	Most efficient (greatest statistical power), biased if average pleiotropic effect differs from zero	9
MR-Egger	<code>mr_egger</code>	Robust to pleiotropy under InSIDE assumption, sensitive to outliers, sensitive to violations of InSIDE assumption, InSIDE assumption often not plausible, often less efficient	10
Median	<code>mr_median</code>	Robust to outliers, sensitive to addition/removal of genetic variants	11
Maximum likelihood	<code>mr_maxlik</code>	Similar to IVW method, accounts for uncertainty in genetic associations with risk factor	9
MR-Robust	<code>mr_ivw(..., robust=TRUE)</code>	Downweights outliers, efficient with valid IVs, high false positive rate with several invalid IVs	16
Penalized weights	<code>mr_ivw(..., penalized=TRUE)</code>	Downweights outliers, efficient with valid IVs, high false positive rate with several invalid IVs	16
Mode-based estimation	<code>mr_mbe</code>	Robust to outliers, sensitive to bandwidth parameter and addition/removal of genetic variants, often less efficient	17
Heterogeneity-penalized	<code>mr_hetpen</code>	Robust to outliers, can only be implemented for a small number of variants due to computational efficiency	18
MR-Lasso	<code>mr_lasso</code>	Removes outliers, efficient with valid IVs, high false positive rate with several invalid IVs	16
Contamination mixture	<code>mr_conmix</code>	Robust to outliers, sensitive to variance parameter and addition/removal of genetic variants	19

The `mr_hetpen` method has options `prior` to set the prior probability of a genetic variant being a valid instrument (default is 0.5), and `CIMin`, `CIMax`, and `CISep` to allow feasible and efficient calculation of confidence intervals.

The `mr_conmix` method has options `psi` to set the value of the standard deviation of the distribution of invalid estimands (that is, how variable are the quantities targeted by genetic variants that are invalid instrumental variables), and `CIMin`, `CIMax`, and `CISep` as above.

The `mr_lasso` method has the option `lambda` to set the tuning parameter in the penalized (lasso) regression model.

Methods for multivariable Mendelian randomization

Multivariable Mendelian randomization is an extension of the standard Mendelian randomization paradigm to include multiple risk factors in a single analysis model^{20,21}. Typically, it is employed when several risk factors share genetic predictors, and so it is not possible to find genetic variants that are specific predictors of a particular risk factor. In multivariable Mendelian randomization, it is assumed that the genetic variants are specifically associated with any of a set of risk factors, such that any causal pathway from the genetic variants to the outcome passes via one or other of the risk factors. To perform multivariable Mendelian randomization with summarized data, genetic associations are required for each variant with all of the risk factors.

Methods for multivariable Mendelian randomization take an *MRMVInput* object as an input, created using the `mr_mvinput` command. Four functions are included for the estimation of causal effects based on summarized genetic data in a multivariable Mendelian randomization framework. These are `mr_mvivw` (multivariable IVW method)²², `mr_mvegger` (multivariable MR-Egger method)²³, `mr_mvmedian` (multivariable median-based method)²⁴, and `mr_mvlasso` (multivariable lasso method). The syntax is:

```
mr_mvivw(mr_mvinput(bx = cbind(ldlc, hdlc, trig),
  bxse = cbind(ldlcse, hdlcse, trigse),
  by = chdlodds, byse = chdloddsse))
```

where `hdlc` and `hdlcse` are genetic associations with high-density lipoprotein (HDL) cholesterol and their standard errors, and `trig` and `trigse` are genetic associations with triglycerides and their standard errors for the same 28 variants. Again, these data variables are provided with the package. Syntax for the `mr_mvegger`, `mr_mvmedian`, and `mr_mvlasso` commands is identical.

The multivariable IVW method is implemented similarly to the univariable IVW method, except using multivariable regression of the genetic associations with the outcome on the genetic associations with the risk factors. As in the univariable case, the `mr_mvivw` command can be implemented using robust regression with the `robust = TRUE` option²⁴. The `mr_mvivw` and `mr_mvegger` methods have a `correl` option to allow for correlated variants. The `mr_mvlasso` method has the `lambda` option to set the penalization parameter as in the univariable case. All methods have `distribution` and `alpha` options as discussed above.

Functions for data visualization

The initial release of the MendelianRandomization package included two options for data visualization, both implemented using the `mr_plot` function. Application of the `mr_plot` function to an *MRInput* object gave an interactive scatter plot of the genetic associations together with a line representing the IVW estimate. Genetic associations with the risk factor are plotted on the horizontal axis, and genetic associations with the outcome on the vertical axis. Application of the `mr_plot` function to an *MRAII* object plotted a similar (although non-interactive) scatter plot of the genetic associations together with lines representing the estimates from various methods. An *MRAII* object is created by the `mr_allmethods` function, which returns estimates from various estimation methods.

We have added functionality so that the `mr_plot` function can now be applied to an *MRMVInput* object. In this case, we still plot the estimated genetic associations with the outcome on the vertical axis. On the horizontal axis, we plot predicted genetic associations with the outcome. These are fitted values from the multivariable IVW method, which regresses the genetic associations with the outcome on the genetic associations with the risk factors. Horizontal error bars represent confidence intervals for these fitted values. These reflect uncertainty in

the multivariable IVW estimates, but not in the genetic associations with the risk factors, which are assumed to be known without error. A diagonal line is plotted with gradient 1 to help the detection of outliers, which may be pleiotropic variants. The syntax is:

```
mr_plot(mr_mvinput(bx = cbind(ldlc, hdlc, trig), # Figure 1
            bxse = cbind(ldlcse, hdlcse, trigse),
            by = chdlodds, byse = chdloddsse))
```

In the example of [Figure 1](#), we additionally set the option `interactive = FALSE` to produce a non-interactive version of this plot.

In updating the package, we have added several additional functions for data visualization. The default implementation of the `mr_forest` function plots the variant-specific estimates in a forest plot, with the pooled estimate from the IVW method at the bottom ([Figure 2a](#)). The variant-specific estimates are the ratio estimates from each genetic variant in turn. This plot allows the user to investigate heterogeneity in the variant-specific estimates, which indicates potential pleiotropy in the analysis²⁵. Heterogeneity can also be expressed numerically by Cochran's Q statistic (for the IVW method) or Rucker's Q statistic (for the MR-Egger method), which are reported as the "heterogeneity test statistic" by the relevant estimation functions. The `mr_forest` function can also be used to plot estimates from different methods, either in addition to the variant-specific estimates or without them ([Figure 2b](#)):

```
mr_forest(mr_input(ldlc, ldlcse, chdlodds, chdloddsse)) # Figure 2A
mr_forest(mr_input(ldlc, ldlcse, chdlodds, chdloddsse), # Figure 2B
            snp_estimates=FALSE,
            methods = c("ivw", "median", "wmedian", "egger", "maxlik", "mbe", "conmix"))
```

(For presentation purposes, in this and subsequent figures we provide plots for the first 9 variants in the package only.) The `mr_funnel` function is similar, except that the variant-specific estimates are plotted against their precision (that is, the reciprocal of their standard error). This plot also enables the user to investigate heterogeneity in the variant-specific estimates ([Figure 3](#)):

```
mr_funnel(mr_input(ldlc, ldlcse, chdlodds, chdloddsse)) # Figure 3
```

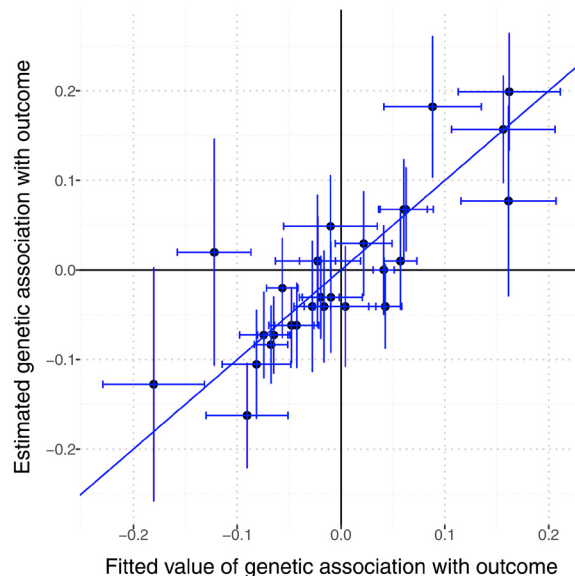


Figure 1. Scatter plot created by `mr_plot` command applied to a `MRMVIInput` object. Estimated genetic associations with the outcome (vertical axis) are plotted against predicted associations with the outcome from the multivariable inverse-variance weighted method (horizontal axis). Error bars are 95% confidence intervals, and the diagonal line has gradient 1.

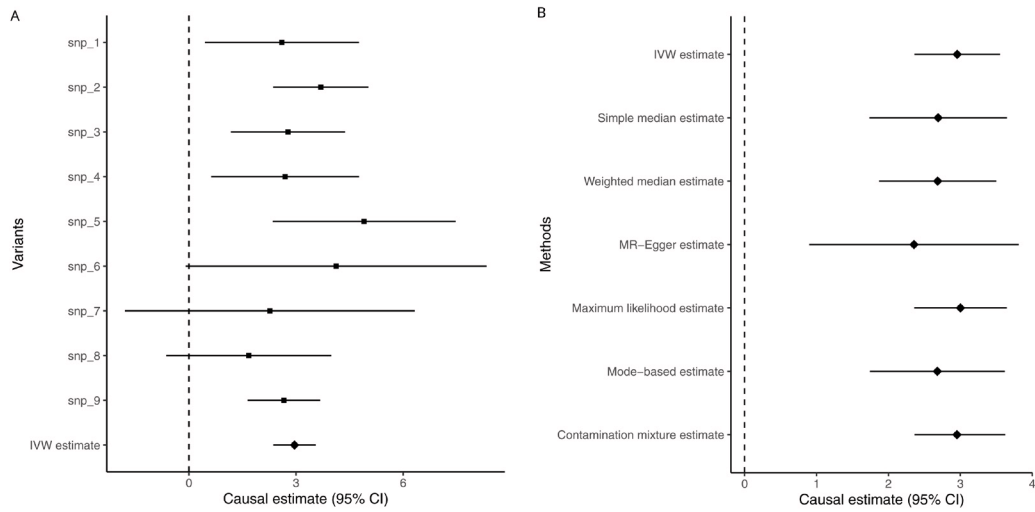


Figure 2. Forest plots created by `mr_forest` command. Left panel: comparison of variant-specific estimates plus inverse-variance weighted (IVW) estimate (default options). Right panel: comparison of estimates from different methods with variant-specific estimates switched off. Points represent estimates and horizontal error bars are 95% confidence intervals (CI).

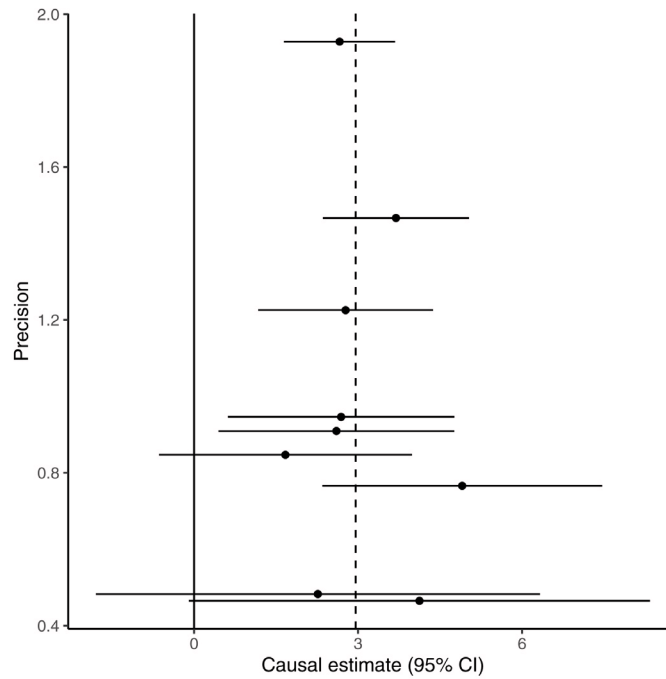


Figure 3. Funnel plot created by `mr_funnel` command. Points represent variant-specific estimates and horizontal error bars are 95% confidence intervals (CI).

The `mr_loo` function allows the user to investigate sensitivity of the IVW estimate to individual data points. This is implemented by calculating the IVW estimate omitting each variant from the analysis in turn (loo stands for ‘leave one out’). The IVW estimate based on all the variants is also plotted for reference (Figure 4):

```
mr_loo(mr_input(ldlc, ldlcse, chdlodds, chdloddsse)) # Figure 4
```

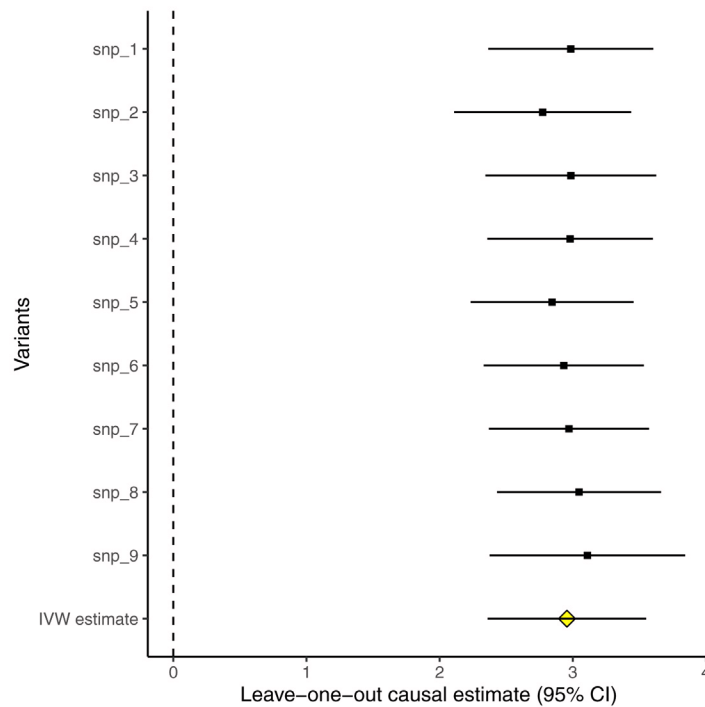



Figure 4. Leave-one-out plot created by `mr_loo` command. Points represent estimates from the inverse-variance weighted (IVW) method, omitting the variant indicated. Horizontal error bars are 95% confidence intervals (CI).

Output from each of these commands is a *ggplot* object, and so basic graphical parameters can be changed using functions from the *ggplot2* package²⁶. For example, the horizontal axis can be set to run from -5 to $+5$ using the following code:

```
library(ggplot2)
forest = mr_forest(mr_input(ldlc, ldlcse, chldodds, chldoddsse))
forest2 = forest + coord_cartesian(xlim=c(-5,5))
forest2
```

Loading data from PhenoScanner

The initial release of the MendelianRandomization package included a function called `extract.pheno.csv`. This function took a `.csv` file previously downloaded by the user from the PhenoScanner webtool (<http://www.phenoscanter.medschl.cam.ac.uk/>) and converted the file into an *MRInput* object, extracting the relevant genetic associations with the risk factor and outcome. PhenoScanner^{27,28} is a database of genetic associations that contains over 65 billion associations for over 150 million unique genetic variants, including genetic associations reported by major consortia, as well as those for the UK Biobank study reported by Ben Neale's team (<http://www.nealelab.is/uk-biobank>).

The `extract.pheno.csv` function is no longer maintained; however, it has been superseded by the `pheno_input` command, which calls PhenoScanner directly from R and creates an *MRInput* object. Using this command, the entire workflow of a Mendelian randomization analysis can be performed in a single line of code. For example:

```
mr_ivw(pheno_input(snps=c("rs12916", "rs2479409", "rs217434", "rs1367117",
                        "rs4299376", "rs629301", "rs4420638", "rs6511720"),
  exposure = "Low density lipoprotein", pmidE = "24097068", ancestryE = "European",
  outcome = "Coronary artery disease", pmidO = "26343387", ancestryO = "Mixed"))
```

This code first extracts data on eight genetic variants (their 'rsid' identifiers are listed above), and creates an *MRInput* object using the genetic associations with “low density lipoprotein” taken from the study with PubMed ID 24097068²⁹ in individuals of European descent as the summarized associations with the risk factor, and genetic associations with “coronary artery disease” taken from the study with PubMed ID 26343387³⁰ in a mixed ancestry sample as the summarized associations with the outcome. The triplet of trait name, PubMed ID, and ancestry is necessary to uniquely identify the correct dataset for genetic associations, as some publications report associations with multiple traits, or associations with the same trait in different ancestry groups. While the above code then implements the IVW method on this *MRInput* object, any other estimation or data visualization command that takes an *MRInput* object as input could be applied to the output of the `pheno_input` function.

Summary

In summary, the MendelianRandomization package has added a number of features since its initial release: to implement various robust estimation methods, to implement methods for multivariable Mendelian randomization, to enable a greater range of data visualization options, and to facilitate data entry. We conclude with the same warning that we stated at the end of the manuscript accompanying the initial package release⁷: while this software simplifies the operational aspects of a Mendelian randomization, the truly difficult parts of an analysis are choosing sensible risk factors and outcomes, selecting genetic variants that are plausible instrumental variables, performing a reasonable range of analyses, and interpreting the results with care and caution³¹. Software code for these aspects of an analysis cannot be written³².

Data availability

Underlying data

All data used in this article are distributed in the software package described, or can be freely downloaded using commands in the software package that are detailed in the text of the article.

Software availability

The MendelianRandomization package is available via the Comprehensive R Archive Network (CRAN)

The software package is available here: <https://cran.r-project.org/web/packages/MendelianRandomization/index.html>.

Source code is available from GitHub: <https://github.com/sb452/MendelianRandomization/tree/v0.5.0>

Archived source code at time of publication: <http://doi.org/10.5281/zenodo.4088672>³³

Software license: [GPL-2](#) | [GPL-3](#).

Acknowledgments

We thank users of the package for helpful feedback that has guided the development of the package.

References

- Smith GD, Ebrahim S: **'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?** *Int J Epidemiol.* 2003; **32**(1): 1–22. [PubMed Abstract](#) | [Publisher Full Text](#)
- Burgess S, Thompson SG: **Mendelian randomization: methods for using genetic variants in causal estimation.** Chapman & Hall, Boca Raton, FL, 2015. [Publisher Full Text](#)
- R Core Team: **R: A language and environment for statistical computing.** Version 4.0.2 (Taking Off Again). R Foundation for Statistical Computing, Vienna, Austria, 2020. [Reference Source](#)
- Burgess S, Dudbridge F, Thompson SG: **Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods.** *Stat Med.* 2016; **35**(11): 1880–1906. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bowden J, Del Greco FM, Minelli C, *et al.*: **A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization.** *Stat Med.* 2017; **36**(11): 1783–1802. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Burgess S, Scott RA, Timpson NJ, *et al.*: **Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors.** *Eur J Epidemiol.* 2015; **30**(7): 543–552. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Yavorska OO, Burgess S: **MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data.** *Int J Epidemiol.* 2017; **46**(6): 1734–1739. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Burgess S, Smith GD, Davies NM, *et al.*: **Guidelines for performing Mendelian randomization investigations [version 2; peer review: 2 approved].** *Wellcome Open Res.* 2020; **4**: 186. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

9. Burgess S, Butterworth A, Thompson SG: **Mendelian randomization analysis with multiple genetic variants using summarized data.** *Genet Epidemiol.* 2013; **37**(7): 658–665.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
10. Bowden J, Smith GD, Burgess S: **Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.** *Int J Epidemiol.* 2015; **44**(2): 512–525.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Bowden J, Smith GD, Haycock PC, *et al.*: **Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.** *Genet Epidemiol.* 2016; **40**(4): 304–314.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
12. Waterworth DM, Ricketts SL, Song K, *et al.*: **Genetic variants influencing circulating lipid levels and risk of coronary artery disease.** *Arterioscler Thromb Vasc Biol.* 2010; **30**(11): 2264–2276.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Hemani G, Zheng J, Elsworth B, *et al.*: **The MR-Base platform supports systematic causal inference across the human phenome.** *eLife.* 2018; **7**: e34408.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Hartwig FP, Davies NM, Hemani G, *et al.*: **Two-sample Mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique.** *Int J Epidemiol.* 2016; **45**(6): 1717–1726.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Slob EAW, Burgess S: **A comparison of robust Mendelian randomization methods using summary data.** *Genet Epidemiol.* 2020; **44**(4): 313–329.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Rees JMB, Wood AM, Dudbridge F, *et al.*: **Robust methods in Mendelian randomization via penalization of heterogeneous causal estimates.** *PLoS One.* 2019; **14**(9): e0222362.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. Hartwig FP, Smith GD, Bowden J: **Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption.** *Int J Epidemiol.* 2017; **46**(6): 1985–1998.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Burgess S, Zuber V, Gkatzionis A, *et al.*: **Modal-based estimation via heterogeneity-penalized weighting: model averaging for consistent and efficient estimation in Mendelian randomization when a plurality of candidate instruments are valid.** *Int J Epidemiol.* 2018; **47**(4): 1242–1254.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Burgess S, Foley CN, Allara E, *et al.*: **A robust and efficient method for Mendelian randomization with hundreds of genetic variants.** *Nat Commun.* 2020; **11**: 376.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Burgess S, Thompson SG: **Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects.** *Am J Epidemiol.* 2015; **181**(4): 251–260.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Sanderson E, Davey Smith G, Windmeijer F, *et al.*: **An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings.** *Int J Epidemiol.* 2019; **48**(3): 713–727.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Burgess S, Dudbridge F, Thompson SG: **Re: "Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects".** *Am J Epidemiol.* 2015; **181**(4): 290–291.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. Rees JMB, Wood AM, Burgess S: **Extending the MR-Egger method for multivariable Mendelian randomization to correct for both measured and unmeasured pleiotropy.** *Stat Med.* 2017; **36**(29): 4705–4718.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Grant AJ, Burgess S: **Pleiotropy robust methods for multivariable Mendelian randomization.** *arXiv.2008.11997.* 2020.
[Reference Source](#)
25. Bowden J, Hemani G, Davey Smith G: **Detecting individual and global horizontal pleiotropy in Mendelian randomization – a job for the humble heterogeneity statistic?** *Am J Epidemiol.* 2018; **187**(12): 2681–2685.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Wickham H: **ggplot2: Elegant Graphics for Data Analysis.** Springer-Verlag New York, 2016.
[Reference Source](#)
27. Staley JR, Blackshaw J, Kamat MA, *et al.*: **PhenoScanner: a database of human genotype-phenotype associations.** *Bioinformatics.* 2016; **32**(20): 3207–3209.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Kamat MA, Blackshaw JA, Young R, *et al.*: **PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations.** *Bioinformatics.* 2019; **35**(22): 4851–4853.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Willer CJ, Schmidt EM, Sengupta S, *et al.*: **Discovery and refinement of loci associated with lipid levels.** *Nat Genet.* 2013; **45**(11): 1274–1283.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Nikpay M, Goel A, Won HH, *et al.*: **A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease.** *Nat Genet.* 2015; **47**(10): 1121–1130.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Burgess S, Davey Smith G, Davies NM, *et al.*: **Guidelines for performing Mendelian randomization investigations [version 2; peer review: 2 approved].** *Wellcome Open Res.* 2020; **4**: 186.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Burgess S, Davey Smith G: **How humans can contribute to Mendelian randomization analyses.** *Int J Epidemiol.* 2019; **48**(3): 661–664.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. sb452: **sb452/mendelianrandomization version 0.5.0.** 2020.
<http://www.doi.org/10.5281/zenodo.4088672>

Open Peer Review

Current Peer Review Status: ? ? ✓

Version 2

Reviewer Report 16 December 2020

<https://doi.org/10.21956/wellcomeopenres.18125.r41795>

© 2020 Cornelis M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Marilyn C. Cornelis

Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

The authors have addressed my concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: genetic epidemiology, nutrition, cardiometabolic traits, aging traits, coffee, caffeine, metabolomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 November 2020

<https://doi.org/10.21956/wellcomeopenres.18010.r41072>

© 2020 Cornelis M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Marilyn C. Cornelis

Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

Broadbent *et al.* summarize updates made to their MendelianRandomization R package. The

package is widely used and thus a detailed description of updates is justified and warranted. I have only very minor suggestions.

1. The R package manual (updated September 30, 2020) details all functions described in this paper. The introduction or summary should briefly mention what the current paper adds above and beyond the R package manual.
2. The paper can use some minor re-organizing or heading revisions. Currently, "Introduction", "Methods", "Use cases", "Summary", "Data availability" and "Software availability" are the major sections. The "Use cases" is well written and describes features in four categories which the authors introduce in the beginning and constitutes the key/main section of the paper. The "Methods" section, however, seems out of place or mislabeled for this descriptive paper. Perhaps use "Software overview"?
3. Table 1. Please annotate functions that are new and not discussed in reference 5.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: genetic epidemiology, nutrition, cardiometabolic traits, aging traits, coffee, caffeine, metabolomics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 20 Nov 2020

Stephen Burgess, University of Cambridge, Cambridge, UK

Reviewer 3: Marilyn C. Cornelis

C0. Broadbent et al. summarize updates made to their MendelianRandomization R package. The package is widely used and thus a detailed description of updates is justified and warranted. I have only very minor suggestions.

- We thank the reviewer for their positive view of this submission.

C1. The R package manual (updated September 30, 2020) details all functions described in this paper. The introduction or summary should briefly mention what the current paper adds above and beyond the R package manual.

- The reviewer is correct that all of the material in the current paper is also available through the R package manual. However, there are reasons why we believe that a publication is worthwhile. First, this paper presents the contents of the package in a clearer and more holistic way than the manual. Secondly, it provides examples of usage and interpretation of the output. Thirdly, it creates a citation in the scientific literature so that we can track usage of the package.

- We have added text to the introduction to more clearly define the scope of this paper. We have also referenced the manual more clearly within the text, for example to obtain lists of the input options and output slots for each method.

C2. The paper can use some minor re-organizing or heading revisions. Currently, "Introduction", "Methods", "Use cases", "Summary", "Data availability" and "Software availability" are the major sections. The "Use cases" is well written and describes features in four categories which the authors introduce in the beginning and constitutes the key/main section of the paper. The "Methods" section, however, seems out of place or mislabeled for this descriptive paper. Perhaps use "Software overview"?

- Thank you for the suggestion. However, the format of the article is set by the journal to ensure homogeneity of papers across this article class. Hence we do not have complete freedom to determine the structure of the article. We have reviewed the guidelines and re-arranged where possible according to the journal guidelines, but the original structure largely remains.

C3. Table 1. Please annotate functions that are new and not discussed in reference 5.

- We have added annotation of those functions that are new and those that are updated since the publication of Yavorska and Burgess 2017 (now reference 7).

Competing Interests: No competing interests were disclosed.

Reviewer Report 05 November 2020

<https://doi.org/10.21956/wellcomeopenres.18010.r41070>

© 2020 Spiller W. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? Wes Spiller 

Department of Population Health Sciences, University of Bristol, Bristol, UK

This article demonstrates how a range of valuable Mendelian randomization approaches can be easily implemented using the MendelianRandomization R Package. It describes substantial additions to the existing software, which I believe would be of great value to researchers. This is especially true of integrating data from the PhenoScanner web-resource, which can be invaluable in identifying avenues for pleiotropic bias.

The material presented in the article is technically sound, though I believe it could be improved with minor revisions.

1. In the methods section, some of the wording is at times confusing, for example: "simple and weighted method methods)".
2. There is an understandably difficult balance between summarising methods previously described, and providing sufficient detail so as for researchers to understand how options relate to methods implemented. For example, while the `correl` option is referenced several times, there is no indication of appropriate usage (such as providing a suitable correlation matrix). Especially in the *Use cases* section, the details provided at times read like the help file for various functions when using R. This shifts from providing insufficient detail overall for using such methods (specifically novel options provided in the package) with confidence, to providing at times a puzzling focus on elements such as the standard error options for the modal estimator. If these options are important enough to be included in the manuscript, then this would warrant more description of their usage.
3. A brief overview of multivariable MR would potentially be helpful, as there is often confusion over whether only SNPs correlated with all exposures should be used, or a set of SNPs where each SNP is correlated with at least one exposure included in the model. I appreciate this has been explained in previous publications, but the clarification would be welcome.

Having performed an example analysis using this software as part of the review process, I commend the authors on providing helpful vignettes and documentation guiding the analysis process in the R package itself.

It is also very encouraging to see issues surrounding the appropriate selection of exposures and outcomes, as well as interpretation of results, afforded attention at times lacking in software showcasing novel statistical approaches.

With the above comments in mind, I believe the MendelianRandomization R package represents a valuable addition to a growing body of MR software ensuring emerging methods are easily implemented in applied research.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genetic Epidemiology, Applied Statistics, Mendelian Randomization, Transethnic MR

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 20 Nov 2020

Stephen Burgess, University of Cambridge, Cambridge, UK

Reviewer 2: Wes Spiller

B0. This article demonstrates how a range of valuable Mendelian randomization approaches can be easily implemented using the MendelianRandomization R Package. It describes substantial additions to the existing software, which I believe would be of great value to researchers. This is especially true of integrating data from the PhenoScanner web-resource, which can be invaluable in identifying avenues for pleiotropic bias.

- We thank the reviewer for their positive view of this submission.

The material presented in the article is technically sound, though I believe it could be improved with minor revisions.

B1. In the methods section, some of the wording is at times confusing, for example: "simple and weighted method methods)".

- Thank you for pointing out that error. The text should read "simple and weighted median methods".

B2. There is an understandably difficult balance between summarising methods previously described, and providing sufficient detail so as for researchers to understand how options relate to methods implemented. For example, while the correl option is referenced several times, there is no indication of appropriate usage (such as providing a suitable correlation matrix). Especially in the Use cases section, the details provided at times read like the help file for various functions when using R. This shifts from providing insufficient detail overall for using such methods

(specifically novel options provided in the package) with confidence, to providing at times a puzzling focus on elements such as the standard error options for the modal estimator. If these options are important enough to be included in the manuscript, then this would warrant more description of their usage.

- We appreciate the sentiment of the reviewer's comment. As per the response to A3, our preference is for this paper to be brief and to illustrate the options available rather than to wade into deep discussions about the properties of the various methods or the reasoning for choosing between the different options available. Hence, we are happy to provide the various options that can be chosen, but we would rather not discuss the merits and demerits of the options at length – these are already addressed elsewhere. This is now explicitly stated in the Introduction.

- If there are other specific examples where the reviewer feels that additional detail is required, then we would be happy to consider adding these. As for the specific examples that the reviewer calls out: 1) we are not aware of any novel options that are provided in the package – all the methods and options have been introduced elsewhere (see citations in the new Table 2); 2) we have added explanation and an example of the "correl" option; 3) the discussion of standard errors for the mode-based estimation method is one sentence. We do not think this is excessive. The original version of the mode-based estimation method provided by the authors gave 12 different estimates based on different options for the method (3 options for the phi parameter, weighted and unweighted, simple and second-order standard errors): we have condensed these down into the three options discussed in the submission.

B3. A brief overview of multivariable MR would potentially be helpful, as there is often confusion over whether only SNPs correlated with all exposures should be used, or a set of SNPs where each SNP is correlated with at least one exposure included in the model. I appreciate this has been explained in previous publications, but the clarification would be welcome.

- As per the response to A4, we have added a short description of multivariable MR to the manuscript. In particular, we clarify that genetic associations are required for all variants with each risk factor.

B4. Having performed an example analysis using this software as part of the review process, I commend the authors on providing helpful vignettes and documentation guiding the analysis process in the R package itself.

- We thank the reviewer for his positive feedback on the vignettes and documentation.

B5. It is also very encouraging to see issues surrounding the appropriate selection of exposures and outcomes, as well as interpretation of results, afforded attention at times lacking in software showcasing novel statistical approaches.

- We thank the reviewer for his positive feedback on our discussion of practical implementation of the methods and interpretation of the results.

B6. With the above comments in mind, I believe the MendelianRandomization R package represents a valuable addition to a growing body of MR software ensuring emerging methods are easily implemented in applied research.

- We thank the reviewer for his comments, which have strengthened and clarified this manuscript.

Competing Interests: No competing interests were disclosed.

Reviewer Report 02 November 2020

<https://doi.org/10.21956/wellcomeopenres.18010.r41160>

© 2020 Smith Byrne K. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Karl Smith Byrne

The International Agency for Research on Cancer, Lyon, France

This article describes an invaluable addition to the set of tools facilitating MR analyses, and provides a more up-to-date suite of methods.

I believe the paper, however, would benefit from some minor revisions:

1. In the introduction I advise the authors to amend the specification that MR is used on 'modifiable' risk factors and 'clinical' outcomes as MR can and often is conducted for risk factors that are not clearly modifiable, such as proteins, and similarly MR may be run on non-clinical outcomes.
2. Additionally, I would provide greater clarity for 'summarised data'; perhaps, summary level data for the association of SNPs with a trait, or facilitate a two-sample MR design.
3. For the description of methods, I would firstly comment that the degree of customisability of these methods (mr_mbe, for example), is rather impressive and to be lauded. However, I would perhaps add one or two sentences to each of the method paragraphs further clarifying the use of these options. For example, the description of the priors of SNPs with causal effects may be quite unclear for those coming new to these methods. While I acknowledge that ultimately the papers behind each method should be read by users of the software, one or two extra helper sentences for each method would be an easy way to improve the ease of use for your software.
4. Similarly, it would be useful to add some extra description to the MVMR introduction to explicitly specify what is going on; for example that the basic MVMR is just a regression with first order weights ($y \sim x_1 + x_2$) to make it less intimidating for users.
5. Is there a facility to report Cochran's or ruckers Q within the software? The authors discuss the value of visually inspecting plots for heterogeneity but it would wonderful if it were possible to return the statistics for this.
6. I also note that it is enjoyable that the plots can be directly edited as ggplot objects.

7. Finally, it would be excellent to have a brief paragraph explaining the format of the data frame of results being reported. It can be quite frustrating when using a package to have to manually explore the data object to try and sleuth out what exactly is present and what it means. I expect this is in the vignette in R but I recommend a summary of this here, perhaps, even only for the components reported in common for all MR methods.

In general, this is a welcome update that is described in a clear manner. With the minor updates I suggest above it will be a very useful tool for genetic epidemiology.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genetic and Molecular Cancer Epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 20 Nov 2020

Stephen Burgess, University of Cambridge, Cambridge, UK

Reviewer 1: Karl Smith Byrne

A0. This article describes an invaluable addition to the set of tools facilitating MR analyses, and provides a more up-to-date suite of methods.

- We thank the reviewer for their positive view of this submission.

I believe the paper, however, would benefit from some minor revisions:

A1. In the introduction I advise the authors to amend the specification that MR is used on

'modifiable' risk factors and 'clinical' outcomes as MR can and often is conducted for risk factors that are not clearly modifiable, such as proteins, and similarly MR may be run on non-clinical outcomes.

- We have modified the language as suggested.

A2. Additionally, I would provide greater clarity for 'summarised data'; perhaps, summary level data for the association of SNPs with a trait, or facilitate a two-sample MR design.

- We have added a brief description of what is meant by "summarized data".

A3. For the description of methods, I would firstly comment that the degree of customisability of these methods (mr_mbe, for example), is rather impressive and to be lauded. However, I would perhaps add one or two sentences to each of the method paragraphs further clarifying the use of these options. For example, the description of the priors of SNPs with causal effects may be quite unclear for those coming new to these methods. While I acknowledge that ultimately the papers behind each method should be read by users of the software, one or two extra helper sentences for each method would be an easy way to improve the ease of use for your software.

- We appreciate the sentiment of the reviewer. However, as the reviewer states, detailed descriptions of these methods are already available in the literature, and a paper reviewing the methods has already been written (Slob and Burgess, 2020). Our strong preference is that this paper remains short and focuses on the software package, rather than deviating too strongly into technical detail. To this end, we have included a brief description of the methods in a separate table (Table 2), so that this information is available to the interested reader, but it does not add to the length of the text. The aim of this paper is now more clearly stated in the introduction.

A4. Similarly, it would be useful to add some extra description to the MVMR introduction to explicitly specify what is going on; for example that the basic MVMR is just a regression with first order weights ($y \sim x_1 + x_2$) to make it less intimidating for users.

- We have added a brief description of multivariable Mendelian randomization (MVMR), and describe how the multivariable IVW method is implemented using multivariable regression.

A5. Is there a facility to report Cochran's or ruckers Q within the software? The authors discuss the value of visually inspecting plots for heterogeneity but it would wonderful if it were possible to return the statistics for this.

- The mr_ivw and mr_egger functions provide the "Heterogeneity test statistic" as part of the outcome. This is the same as Cochrane's Q statistic for the IVW method and Rucker's Q statistics for the MR-Egger method. This is clarified in the help files for the mr_ivw and mr_egger functions, and now in the manuscript.

A6. I also note that it is enjoyable that the plots can be directly edited as ggplot objects.

- That is wonderful!

A7. Finally, it would be excellent to have a brief paragraph explaining the format of the data frame of results being reported. It can be quite frustrating when using a package to have to manually explore the data object to try and sleuth out what exactly is present and what it means. I expect this is in the vignette in R but I recommend a summary of this here, perhaps, even only for the components reported in common for all MR methods.

- We understand the reviewer's request, and sympathize with this problem. As suggested, we have provided details of some key outputs that are common to all (or at least, most) methods. A full description of the output of each method is readily available in the package documentation.

- For example, typing `?mr_ivw` opens the description of the `mr_ivw` function. Under the subheading "Value", there is a clear list of the output for the IVW method. In this case, there are 15 slots: Model, Exposure, Outcome, Correlation, Robust, Penalized, Estimate, StdError, CILower, CIUpper, Alpha, Pvalue, SNPs, RSE, and Heter.Stat. We do not think it would be helpful to the reader to provide a detailed list of every slot for every method in this manuscript, but have added an explanation of how to access this information.

A8. In general, this is a welcome update that is described in a clear manner. With the minor updates I suggest above it will be a very useful tool for genetic epidemiology.

- We thank the reviewer for his comments, which have strengthened and clarified this manuscript.

Competing Interests: No competing interests were disclosed.