SOFTWARE TOOL ARTICLE

REVISED Using the MR-Base platform to investigate risk factors and drug targets for thousands of phenotypes [version 2; peer review: 3 approved]

Venexia M Walker [1,2], Neil M Davies [1,2], Gibran Hemani [1,2], Jie Zheng[1,2], Philip C Haycock[1,2], Tom R Gaunt[1-3], George Davey Smith [1-3], Richard M Martin [1-3]

[1]Medical Research Council Integrative, Epidemiology Unit, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK
[2]Bristol Medical School: Population Health Sciences, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK
[3]National Institute for Health Research Bristol Biomedical Research Centre, University of Bristol, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK

v2 **First published:** 29 Jul 2019, **4**:113 (https://doi.org/10.12688/wellcomeopenres.15334.1)
Latest published: 07 Nov 2019, **4**:113 (https://doi.org/10.12688/wellcomeopenres.15334.2)

Abstract

Mendelian randomization (MR) estimates the causal effect of exposures on outcomes by exploiting genetic variation to address confounding and reverse causation. This method has a broad range of applications, including investigating risk factors and appraising potential targets for intervention. MR-Base has become established as a freely accessible, online platform, which combines a database of complete genome-wide association study results with an interface for performing Mendelian randomization and sensitivity analyses. This allows the user to explore millions of potentially causal associations. MR-Base is available as a web application or as an R package. The technical aspects of the tool have previously been documented in the literature. The present article is complementary to this as it focuses on the applied aspects. Specifically, we describe how MR-Base can be used in several ways, including to perform novel causal analyses, replicate results and enable transparency, amongst others. We also present three use cases, which demonstrate important applications of Mendelian randomization and highlight the benefits of using MR-Base for these types of analyses.

Keywords

Mendelian randomization, GWAS, causal inference, causality, sensitivity analysis, genetics

Open Peer Review

Reviewer Status ✓ ✓ ✓

	Invited Reviewers		
	1	**2**	**3**
REVISED **version 2** published 07 Nov 2019			
version 1 published 29 Jul 2019	✓ report	✓ report	✓ report

1 **Bernard M. Y. Cheung**, The University of Hong Kong, Hong Kong, China

2 **C. Mary Schooling**, City University of New York, New York, USA
The University of Hong Kong, Hong Kong SAR, China

3 **Apostolos Gkatzionis**, University of Cambridge, Cambridge, UK

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

Corresponding author: Venexia M Walker (venexia.walker@bristol.ac.uk)

Author roles: Walker VM: Formal Analysis, Writing – Original Draft Preparation, Writing – Review & Editing; **Davies NM**: Writing – Review & Editing; **Hemani G**: Software, Writing – Review & Editing; **Zheng J**: Software, Writing – Review & Editing; **Haycock PC**: Software, Writing – Review & Editing; **Gaunt TR**: Funding Acquisition, Software, Supervision, Writing – Review & Editing; **Davey Smith G**: Funding Acquisition, Writing – Review & Editing; **Martin RM**: Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests: Tom Gaunt, George Davey Smith and Gibran Hemani report grants from GlaxoSmithKline and Biogen to support development of the MR-Base platform; Neil Davies has received a grant for unrelated research from the Global Research Awards for Nicotine Dependence, which is an Independent Competitive Grants Program supported by Pfizer; all other authors report no other conflicts of interest.

Grant information: This work was supported by the Wellcome Trust [208806], a Sir Henry Dale Fellowship awarded to Gibran Hemani, and the Integrative Epidemiology Unit. The Integrative Epidemiology Unit is supported by the Medical Research Council and the University of Bristol [MC_UU_00011/1, MC_UU_00011/4 and MC_UU_00011/6]. Richard Martin is supported by a Cancer Research UK programme grant, the Integrative Cancer Epidemiology Programme [C18281/A19169].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Walker VM, Davies NM, Hemani G *et al.* **Using the MR-Base platform to investigate risk factors and drug targets for thousands of phenotypes [version 2; peer review: 3 approved]** Wellcome Open Research 2019, **4**:113 (https://doi.org/10.12688/wellcomeopenres.15334.2)

First published: 29 Jul 2019, **4**:113 (https://doi.org/10.12688/wellcomeopenres.15334.1)

REVISED **Amendments from Version 1**

The major differences between this version of the article and the previously published version are as follows:

- The distinction between Mendelian randomization and causal inference has been clarified at several points in the article, including the abstract and introduction

- The breadth of the traits available in MR-Base is now listed in the introduction

- A section dedicated to the limitations of Mendelian randomization has been added after the section 'Principles of Mendelian randomization'

- Definitions for collider bias and the instrument strength independent of direct effect assumption (INSIDE) have been included in Table 1

- Further details regarding the estimates and p-values returned by MR-Base have been added to Table 3

- Table 4, and Figure 2, Figure 3, Figure 4 and Figure 5 have been added to show exemplar results for use case 2

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

Introduction

Mendelian randomization (MR) is a causal inference method used to study the effects of risk factors and exposures on outcome traits by exploiting genetic variation to address confounding and reverse causation[1]. Two-sample Mendelian randomization is an extension of this method that allows the use of summary statistics from genome-wide association studies (GWASs) in place of individual-level genetic data. Mendelian randomization can be used across multiple health outcomes for several applications, as detailed in Box 1. However, the data required to perform the analysis and knowledge of the latest methods can be inaccessible. MR-Base[2] combines a database of summary statistics on traits and health outcomes from over 20,000 GWASs, with an interface for performing two-sample Mendelian randomization to simplify the implementation of this method. As of February 2019, the repository was populated by curated and harmonized datasets corresponding to over 250 billion single nucleotide polymorphisms (SNP)-trait associations[3]. Traits include anthropometric measures, risk factors, metabolites, metals, vitamins, circulating proteins, disease outcomes, and disease intermediates (such as LDL cholesterol and systolic blood pressure), as well as DNA methylation and gene expression phenotypes. MR-Base is available via a web interface or through the package 'TwoSampleMR' for R. Useful links, including these, can be found in Box 2.

Box 1. *Applications of MR-Base*

Subject to suitable data and appropriate methods being available, Mendelian randomization can be implemented across multiple health outcomes to:

- Identify novel (or confirm previously reported) risk and prognostic factors
- Evaluate potential interventions for follow-up in independent replication or experimental studies, based on robust causal analysis and data-integration across multiple study designs, without exposure of patients
- Predict unexpected effects (adverse and beneficial) of an intervention
- Provide causal estimates based on exploratory analyses from clinical trials
- Investigate potential biological mechanisms underpinning risk factor-disease associations

Box 2. *Useful links*

- MR-Base web application: http://www.mrbase.org/
- Exemplar code for the use cases: https://github.com/MRCIEU/mrbase_casestudies
- MR-Base PheWAS web application: http://phewas.mrbase.org/
- TwoSampleMR R package: https://github.com/MRCIEU/TwoSampleMR/
- MRInstruments R package: https://github.com/MRCIEU/MRInstruments/
- TwoSampleMR R package wiki: https://mrcieu.github.io/TwoSampleMR/
- Mendelian randomization primer: https://youtu.be/LoTgfGotaQ4
- Mendelian randomization podcast: https://soundcloud.com/bmjpodcasts/mendelian-randomisation-for-the-moderately-intelligent
- Mendelian randomization webinar: https://www.youtube.com/watch?v=pc3uQz06gO8&feature=youtu.be&app=desktop

The rationale for the development of MR-Base was to provide easy access to analysis-ready data and allow systematic application of Mendelian randomization methods. The tool was developed in the R statistical environment and has an application programming interface that controls user interaction with the underlying database, where curated GWAS data are stored and can be queried. Further technical details can be found in the existing MR-Base article[3]. The aim of this article is to describe how the MR-Base platform can be used in practice (for example, for triangulation and transparency) and demonstrate these uses, through examples, to new audiences. It is complimentary to the existing MR-Base article,

which focuses on describing and demonstrating MR-Base as a resource.

Principles of Mendelian randomization

Mendelian randomization is a method to assess the causal effect of an exposure on an outcome using an instrument, defined by one or more SNPs, as a proxy for the exposure. The SNPs are used as instrumental variables and must meet three conditions: (i) they must be associated with the exposure; (ii) they must only affect the outcome via the exposure; and (iii) there must be no factor that causes both the SNP and outcome. These conditions are known as the instrumental variable assumptions and are illustrated in Figure 1. SNPs are plausible instruments because they are determined at conception and generally cannot be subsequently affected by the environment[4]. If these assumptions hold, then Mendelian randomization effect estimates are unlikely to be due to confounding or reverse causation. However, Mendelian randomization is still subject to important limitations (see *Limitations of Mendelian randomization*).

Methodological advances mean that Mendelian randomization can be implemented using summary statistics from GWAS, without individual level data[5]. This requires SNP-exposure associations and SNP-outcome associations obtained from separate datasets and is known as two-sample Mendelian randomization[5]. This is in contrast to one-sample Mendelian randomization, where both the exposure and outcome are measured in all individuals from the same sample. As a result, two-sample Mendelian randomization can exploit much larger sample sizes and estimate effects with higher precision than is typically possible using any single sample. It also allows access to large case control studies of disease outcomes that may not have measured the exposure of interest. This can be particularly beneficial when considering expensive or difficult to measure phenotypes, such as DNA methylation, metabolomics, proteomics, and gene expression[6]. Note that the use of GWAS in this way for two-sample Mendelian randomization also requires us to make an additional assumption that the GWAS used are providing unbiased genetic estimates.

Detailed discussion of the theory and interpretation of Mendelian randomization results can be found elsewhere[1,6–11]. Key definitions used throughout the present discussion are given in Table 1.

Limitations of Mendelian randomization

Mendelian randomization is subject to several limitations. For example, effect sizes may not be indicative of the effects of a clinical intervention later in life. This can occur for several reasons, such as cumulative exposure, where Mendelian randomization estimates may reflect the effect of lifelong exposure, or time-dependent exposure, where intervention outside of a critical period does not have an effect despite the Mendelian randomization estimates suggesting an effect exists. Mendelian randomization estimates may also be affected by issues such as horizontal pleiotropy, whereby the SNPs chosen to proxy the exposure may affect the outcome by pathways other than the exposure of interest leading to biased results[1,6–12]. Furthermore, investigators are reliant on relevant data being made available and appropriate methods being developed. For instance, it is currently difficult to study prognostic factors due to the lack of GWAS conducted for disease progression outcomes and the susceptibility of current

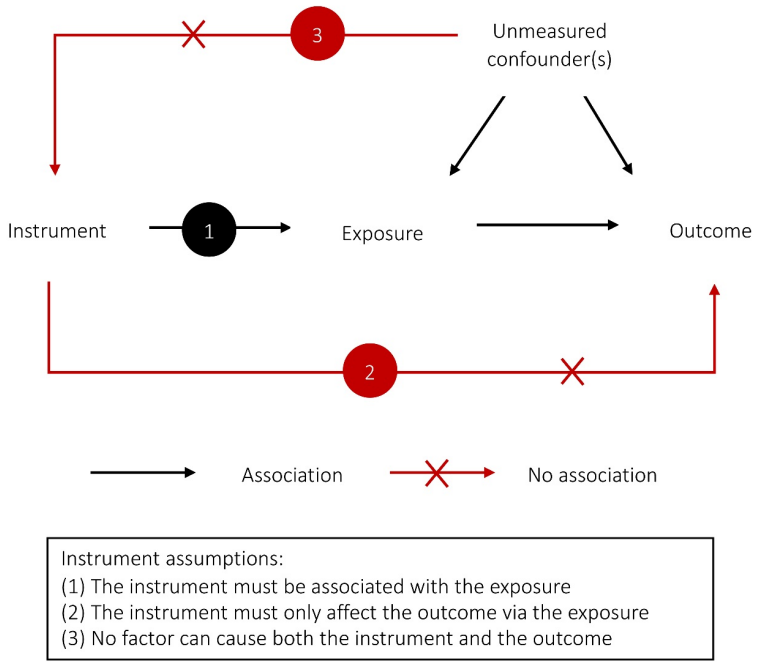

Figure 1. Overview of the instrument assumptions.

Table 1. Key definitions.

Term	Definition
Mendelian randomization	Mendelian randomization is a method to assess the causal effect of an exposure on an outcome using an instrument, defined by one or more single nucleotide polymorphisms, as a proxy for exposure.
Genome-wide association study (GWAS)	Genome-wide association studies identify the genetic variants that are associated with a given phenotype.
Single nucleotide polymorphism	A single nucleotide polymorphism is a difference in the DNA nucleotides between individuals.
Triangulation	"The practice of obtaining more reliable answers to research questions through integrating results from several different approaches, where each approach has different key sources of potential bias that are unrelated to each other[13]."
Pleiotropy	Pleiotropy is when genetic variants effect multiple phenotypes that appear to be unrelated.
Horizontal pleiotropy	Horizontal pleiotropy occurs when the outcome is affected by the instrument single nucleotide polymorphism(s) through a pathway that is independent of the exposure and invalidates the second Mendelian randomization assumption. This is opposed to vertical pleiotropy, which occurs when the instrument single nucleotide polymorphism(s) are associated with other phenotypes that occur between exposure and outcome or after the outcome of interest and does not invalidate the Mendelian randomization assumptions.
Collider bias	A form of bias introduced as a result of conditioning on a variable that is both a consequence of the exposure and the outcome and hence changing the relationship between them.
Genome-wide significance	A conventional threshold, defined as p-values less than 5e-8, that is commonly used to determine which genetic variants are 'hits' in a genome-wide association study.
Allele harmonization	Allele harmonization is the process of specifying the effect and other alleles in the same way in both the outcome and exposure data.
Clumping	Clumping is a method for identifying the independent signals among correlated SNPs.
Linkage disequilibrium	Two genetic variants are in linkage disequilibrium if their alleles are associated.
Heterogeneity	Heterogeneity is defined as the variation in the causal estimate across SNPs.
Palindromic single nucleotide polymorphism (SNP)	A SNP is described as palindromic if the pair of alleles on the forward-strand are the same as the pair of alleles on the reverse strand (i.e. G/C or A/T SNPs).
Minor allele frequency (MAF)	The MAF is a measure of how common the least common allele is for a given genetic variant.
Funnel plot	Funnel plots present the effect estimates against a measure of precision – in the case of MR-Base, the inverse standard error of the instrument – to allow visual assessment of heterogeneity.
No measurement error (NOME) assumption	The NOME assumption assumes that the variance of the instrument-exposure association is negligible and so can be ignored.
Quantitative trait locus (QTL)	A QTL is a DNA variant associated with the variation that is observed in a phenotype.
Zero modal pleiotropy assumption (ZEMPA)	An assumption that the mode of the bias terms for individual instruments is zero.
Instrument strength independent of direct effect assumption (INSIDE)	An assumption that there is zero correlation between the SNP-exposure associations (i.e. the instrument strength) and the SNP-outcome associations (i.e. the direct effect of the instruments on the outcome)[14]. Required for some MR methods, such as MR-Egger.

methods to collider bias[15]. Selection bias, which occurs when the individuals included within the analysis are not representative of the population that you are trying to study, can also be an issue[16]. Careful consideration of the GWAS used for two-sample Mendelian randomization is therefore encouraged to minimize the effect of this bias on estimates. An additional limitation, specific to two-sample Mendelian randomization, is the requirement for no sample overlap. While using different samples for the exposure and outcome data is an advantage of this method as it can increase power, it can be difficult to determine whether individual participants appear in both datasets. If they do, simulations have suggested there is a potential for "substantial bias and inflated Type 1 error rates[17]". While MR-Base includes multiple sensitivity analyses that aim to address some of these limitations, such as horizontal pleiotropy. Others, such as sample overlap, are not accounted for and therefore must be considered carefully by users of this tool. Further details and limitations of Mendelian randomization can be found in the literature[1,6–11].

Potential applications of MR-Base

MR-Base[2] is suitable for a broad range of applications and, consequently, is intended for use by a broad range of professionals. These include clinical and non-clinical researchers, public health specialists, policy makers and those in the pharmaceutical industry. Some of the key ways in which the

platform can be used are summarised below and specific use cases are discussed in the *Use cases* section.

- MR-Base can be used to rapidly implement two-sample Mendelian randomization to investigate potential risk and prognostic factors (use cases 1 and 2) and evaluate potential drug targets (use case 3). The GWAS database and online analytical platform provided by MR-Base allow two-sample Mendelian randomization to be implemented quickly and easily to test associations for a range of traits (behavioural, physiological, hormonal, epigenomic, metabolomic, microbiomic) in relation to outcomes. This can be used without the need to generate new data, for example, when exploring new research ideas. Note that these investigations are subject to the relevant data being made available and appropriate methods being developed (see *Limitations of Mendelian randomization*). However, new data are regularly added to the database and the platform is regularly updated to incorporate the latest methods, which should help to overcome these issues in the future.

- MR-Base can conduct sensitivity analyses (use case 2). As with all analytical methods, Mendelian randomization methods are based on assumptions which may not hold. MR-Base offers many Mendelian randomization methods for investigators to choose from and conducts several standard sensitivity analyses that allow relaxation of the assumptions and provide ways of assessing potential pleiotropy. Some of the more commonly used Mendelian randomization methods are selected by default in the platform, including the Wald ratio[18], MR-Egger[14] and the inverse variance weighted method[14]. Use of multiple methods is recommended as they differ in their strengths, limitations and efficiency. For example, MR-Egger has been developed to detect assumption violations such as invalid instruments due to pleiotropy but can lack precision. The best way to assess the reliability of the causal estimates obtained from Mendelian randomization is to triangulate across multiple Mendelian randomization methods and with findings from non-Mendelian randomization study designs[13]. This is demonstrated in use case 2. Summary tables for the methods and the output of Mendelian randomization analyses conducted using MR-Base (as of October 2018) are provided in Table 2 and Table 3, respectively. The platform continues to be under active development and, as highlighted before, new methods are added as they arise.

Table 2. Overview of MR methods available in MR-Base.

Method	Details	References
Wald ratio	The Wald ratio method is also known as the ratio of coefficients method. It divides the regression coefficient of the instrument on the outcome by the regression coefficient of the instrument on the exposure and can be used when only one instrument SNP is available.	18
Maximum likelihood	This method maximizes the likelihood of a model, which is based on the exposure-outcome relationship and the distribution of the estimates of the genetic association, to obtain a causal estimate.	19
MR Egger regression	MR Egger calculates Wald ratios for each of the instruments and combines the results using an adapted Egger regression. The causal effect is the Egger regression slope coefficient and the intercept is an estimate of the average pleiotropic effect across instruments. Bootstrapping can help to improve the reliability of standard error estimates for non-zero causal effects.	14
MR Egger (bootstrap)		
Simple median	These methods calculate Wald ratios for each of the instruments and select the median value (according to the specified method) as the causal estimate. They provide valid estimates when more than half of the SNPs satisfy the instrumental variable assumptions.	20,21
Weighted median		
Penalised weighted median		
Inverse variance weighted	This method calculates the Wald ratio for each of the instruments and combines the results using an inverse-variance weighted meta-analysis approach. The slope from this approach can be interpreted at the causal effect of the exposure on the outcome. The variance of the effect can be estimated using either a fixed or multiplicative random effects model. The latter is usually implemented unless there is under-dispersion in the effect estimates, in which case a fixed effects model is used.	19,22
Inverse variance weighted (multiplicative random effects)		
Inverse variance weighted (fixed effects)		
Simple mode	The mode-based methods use the causal effect estimates for individual SNPs to form clusters. The causal effect estimate is then taken as the causal effect estimate from the largest cluster of SNPs. The weighted mode methods use the same process but assign weights to each SNP. Mode-based methods require ZEMPA, which states that the mode of the bias terms for the individual instruments is zero.	23
Weighted mode		
Weighted mode (NOME)		
Simple mode (NOME)		

Table 3. Overview of the tables and graphs included in the MR-Base platform.

Tab	Details
MR results	A table with the causal estimates resulting from each MR method that was implemented. See Table 4 for an example based on use case 2. Estimates are presented in the units of the exposure SNP(s). Estimates are beta coefficients for the outcome and should be exponentiated if the unit of the outcome was a log odds ratio. P-values are calculated using a t-distribution.
Heterogeneity statistics	A table with statistics indicating the variation in the causal estimate across SNPs, i.e. heterogeneity. Lower heterogeneity indicates better reliability of results.
Causal direction test	The results of a test that uses variation explained in both the exposure and outcome to assess whether the direction of the results is likely to be correct. Note the test cannot determine whether a causal association exists.
Horizontal pleiotropy	The Egger regression intercept with its standard error and a p-value.
Single SNP analysis	A summary graph showing the individual effects of SNPs, calculated using the Wald ratio, along with the overall results to assess the consistency across SNPs. See Figure 2 for an example.
Method comparison plot	A graphical representation of the results given under the 'MR results' tab. This graph shows the effect of the SNP(s) on exposure against the effect of the SNP(s) on the outcome. The graph is structured so that the effect of the SNP(s) on the exposure is always positive and the effect of the SNP(s) on the outcome is directed accordingly. See Figure 3 for an example based on use case 2.
Leave-one-out analysis	A graph showing the results of MR analyses using the inverse variance weighted method when leaving one SNP out each time. This analysis can be used to assess whether the SNPs are consistent in terms of their effect on the overall outcome or whether the results are being driven by a single outlying SNP. See Figure 4 for an example based on use case 2.
Funnel plot	A graph to visually assess heterogeneity, particularly horizontal pleiotropy. Horizontal pleiotropy is likely if points are spread. Directional horizontal pleiotropy may be present if the graph is not symmetrical. See Figure 5 for an example based on use case 2.

- MR-Base can replicate results (use cases 2 and 3). MR-Base can be used to replicate the results of studies, regardless of whether they originally used Mendelian randomization or MR-Base, if relevant GWAS are available. This may be useful in several situations, including when appraising studies in the literature.
- MR-Base can be used to support triangulation of evidence (use case 2). Triangulation has been defined as "the practice of obtaining more reliable answers to research questions through integrating results from several different approaches, where each approach has different key sources of potential bias that are unrelated to each other[13]." Mendelian randomization, implemented in MR-Base, can be linked with other designs which are intended to reveal biases (for example, a negative control study) or exploit different confounder structures (for example, a cross-context comparison as a source of evidence in a triangulation framework).
- MR-Base can enable transparency (use case 1). MR-Base has been developed to encourage transparency by providing the analysis code needed to replicate the analysis in the output. Further to this, studies that use data from the platform can be directly replicated by others as they can access the same data that has been formatted in a consistent manner via the provided allele harmonization procedures.

Methods

Implementation

MR-Base[2] can perform two-sample Mendelian randomization and provide summary statistics from a range of GWAS for this purpose. As highlighted previously, it can be accessed through a web application or as an R package. Data can either be accessed through the platform or be uploaded by the user, both of which are demonstrated in the following *Use cases* section. Data harmonization between the SNP-exposure associations and the SNP-outcome associations and Mendelian randomization are then performed according to options specified by the user using buttons on the web application or commands for the R package.

Operation

The web application can be accessed from any platform that allows the use of a java-script compatible graphical web browser. The R package can be accessed from any platform where R version 3.5 or later can be installed. Step-by-step instructions for the web application and code for the R package are available for each of the use cases (see *Software availability*)[24]. A generalized workflow for using the MR-Base web interface is provided in Box 3.

Box 3. Generalized workflow for using the MR-Base web interface

Mendelian randomization analyses can be performed using the MR-Base web interface as detailed below:

1. Access the platform (http://www.mrbase.org/) and sign the data access agreement using a Google account.
2. Define the exposure according to one of the following options:
   a. By selecting the relevant GWAS from an existing source, such as the MR-Base GWAS catalog. This is demonstrated in use cases 1 and 2.
   b. By uploading an instrument file, specifying the delimiter for the file and filling in the form to map the column names to those supplied in the file. Columns not included in the file can be left blank in the mapping. This is demonstrated in use case 3.
3. Define the outcome by selecting the relevant GWAS from the MR-Base GWAS catalog.
4. Specify the analysis settings:
   - Set linkage disequilibrium (LD) clumping preference – by default this will be 'Do not check for LD between SNPs'.
   - Set linkage disequilibrium proxies preference – by default this will be to use proxies with a minimum linkage disequilibrium R squared value of 0.8 and allow palindromic SNPS with a minor allele frequency threshold up to 0.3.
   - Set allele harmonisation preference – by default this will be 'Attempt to align strands for palindromic SNPs'. If used, this setting will remove palindromic SNPs with minor allele frequencies close to 0.5 as the effect allele will be ambiguous.
   - Select the methods for analysis – by default this will be the Wald ratio, MR Egger, weighted median, inverse variance weighted and weighted mode methods.
5. Select the 'perform MR analysis' button and save the results, including the citations that are to be referenced in any published work arising from this analysis, on the following screen.

Note that the MR-Base web interface will provide the analysis code as an output if you wish to recreate your analysis in R. Also, note that there will be no graphical results produced for single SNP instruments as the sensitivity analyses, which are illustrated in the graphs, can only be conducted when there are multiple SNPs.

Use cases

We discuss three studies that demonstrate important applications of Mendelian randomization and highlight the benefits of using MR-Base for these types of analyses in the following sections.

Use case 1: subjective wellbeing and cardiometabolic health

Our first use case demonstrates the rapid implementation of Mendelian randomization using MR-Base to investigate a risk factor proposed in the observational literature and enable transparency of the Mendelian randomization study. The specific workflow for this case study is provided, alongside the necessary code, data and results, on GitHub (see *Software availability*)[24]. It is based on work by Wootton *et al.* that used MR-Base to investigate the association between subjective wellbeing and 11 measures of cardiometabolic health[25]. It has been reproduced based on information in the paper and, in particular, their code on GitHub. Studies with data on both subjective wellbeing and measures of cardiometabolic health are rare. Therefore, the authors chose to use two-sample Mendelian randomization so that they could use separate samples for their exposure (sample 1) and outcome (sample 2) phenotypes and use UK Biobank as a single-sample Mendelian randomization sensitivity analysis. In addition to this, the largest available GWAS of subjective wellbeing at the time of the study had identified just three SNPs to instrument this phenotype at the conventional genome-wide significance level and only one of these three SNPs replicated in an independent sample. Consequently, the authors used a lower p-value threshold of $P < 5\times10^{-5}$ to increase the number of SNPs in their instrument. This potentially increases their instrument strength but may also increase their susceptibility to weak instrument bias or pleiotropy.

It was straightforward to use two-sample Mendelian randomization with MR-Base in this study as it allowed the authors to consider multiple outcomes simultaneously and to use summary data that was already formatted for analysis. It also allowed specification of their preferred p-value threshold for the instrument. To make allowance for the non-independence of the selected SNPs, it was necessary to prune the SNPs to determine an independent set for the analysis. MR-Base allows users to select independent SNPs through a process known as 'clumping', which identifies independent signals by considering the linkage disequilibrium between SNPs. Linkage disequilibrium refers to the allelic association between groups of SNPs, which are typically located in a similar region of the genome. Failure to consider the linkage disequilibrium between SNPs can lead to overestimation of instrument strength and overly precise effect estimates. MR-Base overcomes this by picking the SNP from the group of SNPs in linkage disequilibrium that has the strongest evidence of association with the exposure for use in the Mendelian randomization analysis. In the Wootton *et al.* study, clumping reduces the subjective wellbeing instrument from 724 SNPs to 84 SNPs, highlighting the importance of this step in the analysis.

Use case 2: systolic blood pressure and coronary heart disease

The second use case demonstrates how MR-Base can be used to conduct sensitivity analyses and triangulate evidence. The specific workflow for this case study is provided, alongside the necessary code, data and results, on GitHub (see *Software availability*)[24]. Sample code for using the 'TwoSampleMR'

R package based on this use case is provided in Box 4 and exemplar output is provided in Table 4 and Figure 2–Figure 5. It is based on work by Ference *et al.* that examined the effect of systolic blood pressure on coronary heart disease[26]. Here, we recreate the Mendelian randomization component of the work. This can be triangulated with evidence from meta-analyses of prospective observational studies and randomized controlled trials. These meta-analyses found that lower systolic blood pressure reduced risk of coronary heart disease with odds ratios of 0.75 (95% CI: 0.71 to 0.78; p = 0.006) and 0.83 (95% CI: 0.76 to 0.90; p = 0.001), respectively. A full discussion of the triangulation element of this research is provided by Lawlor *et al.*[13]

Box 4. Sample code for using the 'TwoSampleMR' R package based on use case 2

```
# Load the TwoSampleMR package
library(TwoSampleMR)

# List the outcomes available in MR-Base
ao <- available_outcomes()

# Extract the instruments from the systolic blood pressure GWAS (ID: 'UKB-a:360')
exposure_dat <- extract_instruments(c('UKB-a:360'))

# Extract the outcome data from the coronary heart disease GWAS (ID: 7)
outcome_dat <- extract_outcome_data(exposure_dat$SNP, c('7'),
                                    proxies = 1, rsq = 0.8, align_alleles = 1,
                                    palindromes = 1, maf_threshold = 0.3)

# Harmonize the exposure and outcome data

dat <- harmonise_data(exposure_dat, outcome_dat, action = 2)

# Perform MR analysis

mr_results <- mr(dat)
```

Table 4. Exemplar MR results table based on use case 2.

id.exposure	id.outcome	outcome	exposure	method	nsnp	b	se	pval
UKB-a:360	7	Coronary heart disease || id:7	Systolic blood pressure automated reading || id:UKB-a:360	MR Egger	157	0.9711	0.2917	0.001091
UKB-a:360	7	Coronary heart disease || id:7	Systolic blood pressure automated reading || id:UKB-a:360	Weighted median	157	0.571	0.07664	9.226e-14
UKB-a:360	7	Coronary heart disease || id:7	Systolic blood pressure automated reading || id:UKB-a:360	Inverse variance weighted	157	0.5663	0.0905	3.924e-10
UKB-a:360	7	Coronary heart disease || id:7	Systolic blood pressure automated reading || id:UKB-a:360	Weighted mode	157	0.571	0.1744	0.00131

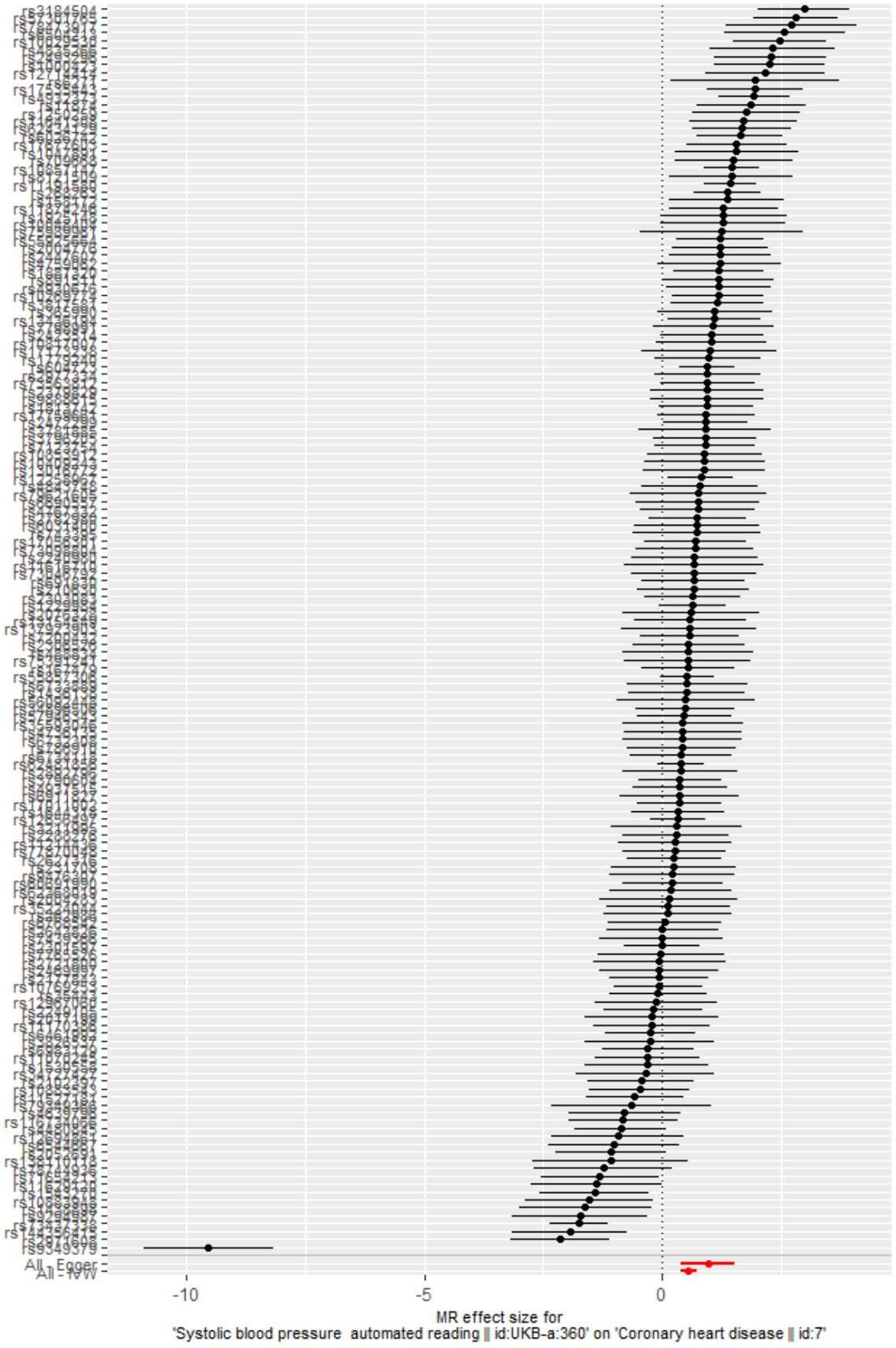

Figure 2. Exemplar single SNP analysis plot based on use case 2.

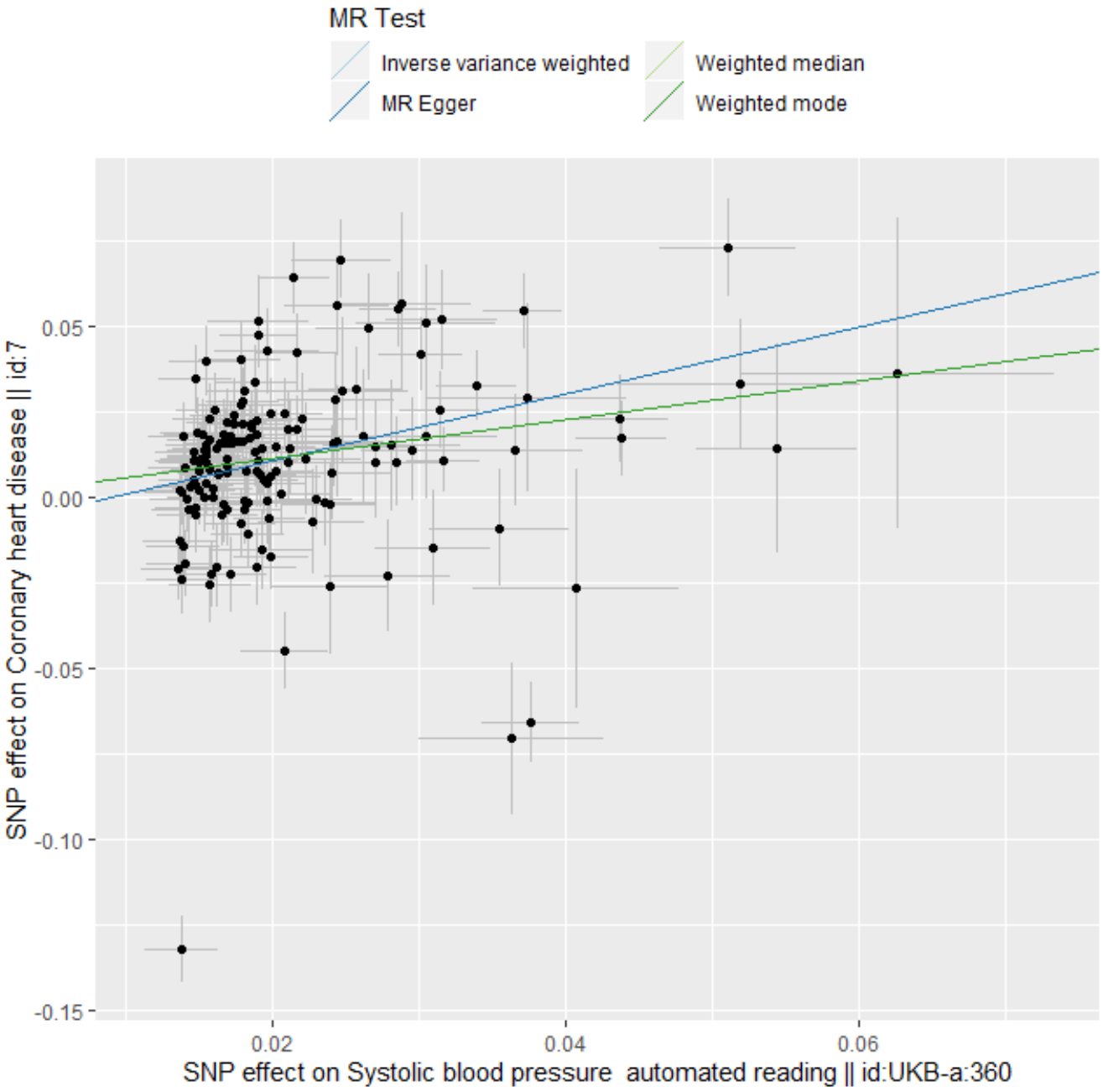

Figure 3. Exemplar method comparison plot based on use case 2.

In our results, as with the original paper, we were concerned about directional pleiotropy, which occurs when genetic variants affect the outcome independently of the exposure. This is because large GWAS, such as the GWAS of systolic blood pressure we use here, may identify SNPs of unknown function[27]. To assess the effect of this upon our results, we can look at the MR-Egger regression intercept provided by default on the MR-Base web application and calculable using the TwoSampleMR package for R. The intercept provides an estimate of the magnitude of horizontal pleiotropy and in our case is -0.0087 (SE: 0.0059; p = 0.147). This suggests limited evidence for directional pleiotropy among our results. We also used several Mendelian randomization methods for our analysis as a further sensitivity analysis and found consistent results for the effect of increased systolic blood pressure on coronary heart disease, regardless of the method used (IVW - OR: 1.76, 95% CI: 1.48 to 2.10, p = 3.92e-10; MR

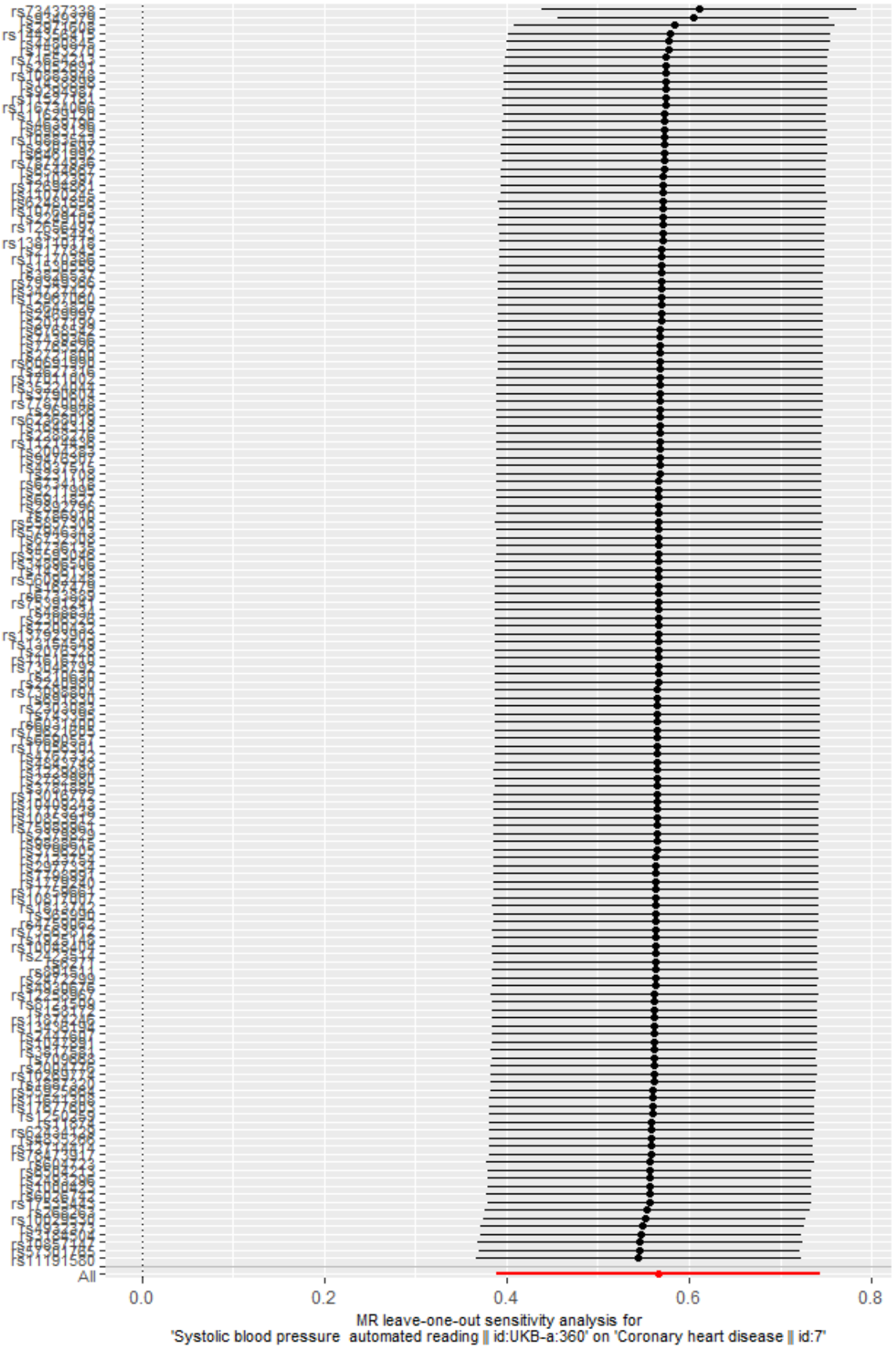

Figure 4. Exemplar leave-one-out analysis plot based on use case 2.

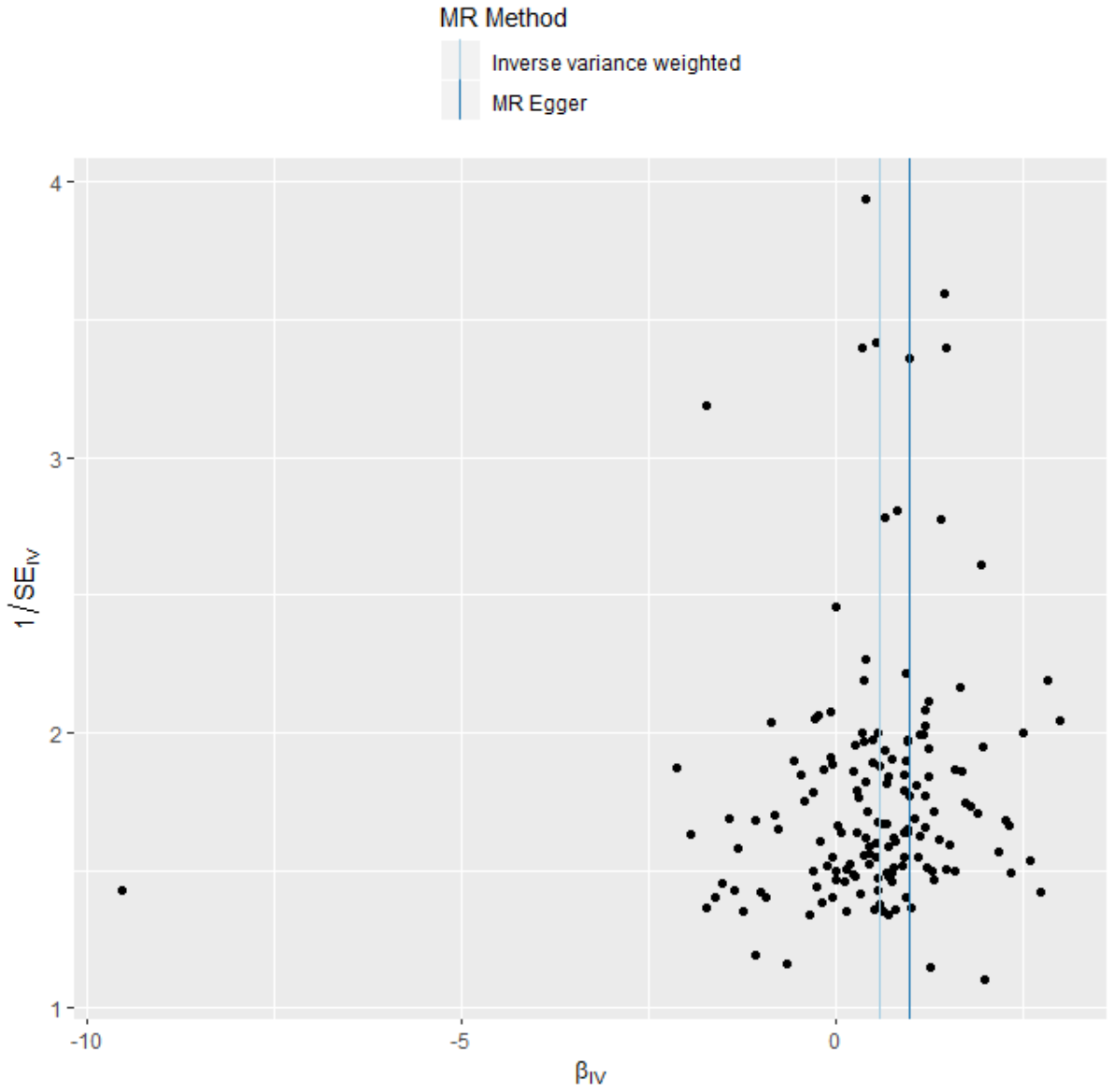

Figure 5. Exemplar funnel plot based on use case 2.

Egger - OR: 2.64, 95% CI: 1.49 to 4.68, p = 1.09e-03; Weighted median – OR: 1.77, 95% CI: 1.52 to 2.05, p = 3.92e-10; Weighted mode – OR: 1.77, 95% CI: 1.26 to 2.49, p = 1.31e-3). Note that the estimates returned by MR-Base are beta coefficients for the outcome. Binary outcomes are commonly reported as log odds ratios and so will need to be exponeniated in order to obtain an odds ratio, as was done here.

Use case 3: HMGCR and type 2 diabetes

Our final use case demonstrates how MR-Base can be used to replicate a study and appraise a potential pharmaceutical intervention. The specific workflow for this case study is provided, alongside the necessary code, data and results, on GitHub (see *Software availability*)[24]. It is based on research by Swerdlow *et al.* that investigated the effect of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), the target of statins, on risk of type 2 diabetes[28]. This study used a single SNP as an instrument: rs17238484. To demonstrate the features of MR-Base, we have uploaded the data necessary to define the instrument for this analysis, instead of using data already within the platform. These data were extracted from the 2013 GWAS by the Global Lipids Genetics Consortium[29]. Once uploaded, the column names must be mapped to those used by MR-Base before the analysis can be run. If you are using the R package, there are equivalent commands that perform data formatting (see this guide). Although units are not required for the analysis to run, it is important that the units of the instrument-exposure and instrument-outcome effects are known, as this determines the interpretation of the effect estimate obtained by Mendelian randomization.

If you use MR-Base, please cite the resource using reference[3]. We also ask that you cite and acknowledge the studies that contributed the data and methodology used in your analysis.

Conclusions

Mendelian randomization is a method for estimating causal effects of an exposure on an outcome that are unlikely to be due to confounding or reverse causation. The method has a broad range of applications, including the investigation of risk factors and the appraisal of potential targets for pharmaceutical intervention. MR-Base eases the implementation of this method by combining a database of GWAS results with an analysis interface to allow Mendelian randomization to be used in several ways, such as for transparency and replication. Consequently, novice users can now perform sophisticated, causal appraisals of exposures by implementing Mendelian randomization using this powerful but accessible tool.

Data availability

Underlying data

Source data for the use cases are available through the database integrated into the MR-Base platform. The database is large and contains data from over 20,000 GWAS; therefore, it is not possible to host this data on an external repository. The data can be accessed via the MR-Base web application or by using the TwoSampleMR package for R to interact with the application programming interface. Users are required to accept the data access agreement by logging in with a Google account before access to the data is granted.

Software availability

MR-Base software:

- Software available from: http://www.mrbase.org/
- Source code available from: https://github.com/MRCIEU/TwoSampleMR
- Archived source code at time of publication: https://doi.org/10.5281/zenodo.3298001[2]
- License: GPL-3.0

Use cases workflow and code:

- Source code available from: https://github.com/MRCIEU/mrbase_casestudies
- Archived source code at time of publication: http://doi.org/10.5281/zenodo.3239316[24]
- License: MIT

References

1. Davies NM, Holmes MV, Davey Smith G: **Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians.** *BMJ.* 2018; **362**: k601.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Hemani G; mightyphil2000, qingyuanzhao, *et al.*: **MRCIEU/TwoSampleMR: WellcomeOpen**. 2019.
http://www.doi.org/10.5281/zenodo.3298001
3. 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
4. Greenland S: **An introduction to instrumental variables for epidemiologists.** *Int J Epidemiol.* 2000; **29**(4): 722–9.
PubMed Abstract | Publisher Full Text
5. Lawlor DA: **Commentary: Two-sample Mendelian randomization: opportunities and challenges.** *Int J Epidemiol.* 2016; **45**(3): 908–15.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Holmes MV, Ala-Korpela M, Smith GD: **Mendelian randomization in cardiometabolic disease: challenges in evaluating causality.** *Nat Rev Cardiol.* 2017; **14**(10): 577–90.
PubMed Abstract | Publisher Full Text | Free Full Text
7. 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
8. Davey Smith G, Hemani G: **Mendelian randomization: genetic anchors for causal inference in epidemiological studies.** *Hum Mol Genet.* 2014; **23**(R1): R89–98.
PubMed Abstract | Publisher Full Text | Free Full Text
9. Haycock PC, Burgess S, Wade KH, *et al.*: **Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies.** *Am J Clin Nutr.* 2016; **103**(4): 965–78.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Zheng J, Baird D, Borges MC, *et al.*: **Recent Developments in Mendelian Randomization Studies.** *Curr Epidemiol Rep.* 2017; **4**(4): 330–45.
PubMed Abstract | Publisher Full Text | Free Full Text
11. Walker VM, Davey Smith G, Davies NM, *et al.*: **Mendelian randomization: a novel approach for the prediction of adverse drug events and drug repurposing opportunities.** *Int J Epidemiol.* 2017; **46**(6): 2078–89.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Pingault JB, O'Reilly PF, Schoeler T, *et al.*: **Using genetic data to strengthen causal inference in observational research.** *Nat Rev Genet.* 2018; **19**(9): 566–580.
PubMed Abstract | Publisher Full Text
13. Lawlor DA, Tilling K, Davey Smith G: **Triangulation in aetiological epidemiology.** *Int J Epidemiol.* 2016; **45**(6): 1866–86.
PubMed Abstract | Publisher Full Text | Free Full Text
14. Bowden J, Davey Smith G, Burgess S: **Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.** *Int J Epidemiol.* 2015; **44**(2): 512–25.
PubMed Abstract | Publisher Full Text | Free Full Text
15. Paternoster L, Tilling K, Davey Smith G: **Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges.** *PLoS Genet.* 2017; **13**(10): e1006944.
PubMed Abstract | Publisher Full Text | Free Full Text
16. Gkatzionis A, Burgess S: **Contextualizing selection bias in Mendelian randomization: how bad is it likely to be?** *Int J Epidemiol.* 2019; **48**(3): 691–701.
PubMed Abstract | Publisher Full Text | Free Full Text
17. Burgess S, Davies NM, Thompson SG: **Bias due to participant overlap in two-sample Mendelian randomization.** *Genet Epidemiol.* 2016; **40**(7): 597–608.
PubMed Abstract | Publisher Full Text | Free Full Text
18. Lawlor DA, Harbord RM, Sterne JA, *et al.*: **Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology.** *Stat Med.* 2008; **27**(8): 1133–63.
PubMed Abstract | Publisher Full Text
19. Burgess S, Butterworth A, Thompson SG: **Mendelian randomization analysis with multiple genetic variants using summarized data.** *Genet Epidemiol.* 2013; **37**(7): 658–65.
PubMed Abstract | Publisher Full Text | Free Full Text
20. Kang H, Zhang A, Cai TT, *et al.*: **Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization.** *J Am Stat Assoc.* 2016; **111**(513): 132–144.
Publisher Full Text

21. Bowden J, Davey Smith G, Haycock PC, *et al.*: **Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator.** *Genet Epidemiol.* 2016; **40**(4): 304–14.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Burgess S, Bowden J: **Integrating summarized data from multiple genetic variants in Mendelian randomization: bias and coverage properties of inverse-variance weighted methods.** *arXiv.* 2015; 04486.
Reference Source

23. Hartwig FP, Davey Smith G, 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

24. Walker V, Gaunt T: **MRCIEU/mrbase_casestudies: Initial release.** 2019.
http://www.doi.org/10.5281/zenodo.3239316

25. Wootton RE, Lawn RB, Millard LA, *et al.*: **Evaluation of the causal effects between subjective wellbeing and cardiometabolic health: mendelian randomisation study.** *BMJ.* 2018; **362**: k3788.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Ference BA, Julius S, Mahajan N, *et al.*: **Clinical effect of naturally random allocation to lower systolic blood pressure beginning before the development of hypertension.** *Hypertension.* 2014; **63**(6): 1182–8.
PubMed Abstract | Publisher Full Text

27. **Rapid GWAS of thousands of phenotypes for 337,000 samples in the UK Biobank.** [Internet]. Neale lab. [cited 2018 Aug 1].
Reference Source

28. Swerdlow DI, Preiss D, Kuchenbaecker KB, *et al.*: **HMG-coenzyme A reductase inhibition, type 2 diabetes, and bodyweight: evidence from genetic analysis and randomised trials.** *Lancet.* 2015; **385**(9965): 351–61.
PubMed Abstract | Publisher Full Text | Free Full Text

29. Global Lipids Genetics Consortium: **Discovery and refinement of loci associated with lipid levels.** *Nat Genet.* 2013; **45**(11): 1274–83.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔

Version 1

Reviewer Report 23 September 2019

https://doi.org/10.21956/wellcomeopenres.16742.r36074

✔ **Apostolos Gkatzionis**
MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

The paper by Walker *et al*. provides an introduction to the already popular MR-Base platform, a software tool developed by the authors for accessing GWAS results and conducting Mendelian randomization analyses in a fast and efficient way. Prior to reading the paper, I have not had the opportunity to use MR-Base. While reviewing, I used the web application to reproduce the authors' Use Cases 1 and 2 as a tutorial. In my experience, the software is user-friendly, straightforward in its implementation and produces a concise and informative report. Likewise, the paper serves well as an instructions manual that is easy for applied practitioners to follow. It is therefore a welcome addition to the literature.

My comments are minor and mainly concerned with the presentation of the paper:

- I agree with previous reviewers that the article would benefit from a short discussion of the challenges involved in performing Mendelian randomization. The MR-Base software efficiently automates the process of implementing an MR analysis, which is naturally very useful for practical purposes. However, this makes it easy to ignore some of the assumptions that underpin Mendelian randomization and limitations of the approach. Especially topics that are not automatically addressed by MR-Base, such as selection bias and the need for SNP-exposure and SNP-outcome associations to be obtained from similar populations, could be highlighted to alert applied practitioners using the software to potential limitations of their analyses.

- As suggested by the authors, MR-Base can greatly contribute to promoting transparency and reproducibility of research results in Mendelian randomization. Given that the software is already quite popular, it could be useful if the authors incorporated a searchable database (similar to the GWAS catalogue) of published Mendelian randomization analyses that were conducted using MR-Base. This would, for example, allow researchers who want to investigate a specific trait to easily access previous research findings concerning causal relationships between that trait and other common risk factors or outcomes.

- It is helpful that the paper is complemented with links to GitHub pages, containing instructions on how to implement the authors' analysis for each of the three Use Cases. Nevertheless, I think it could improve the presentation of the manuscript if the authors incorporated some of the results tables and associated graphs obtained from MR-Base into the paper for illustration purposes.

Overall, the MR-Base software and associated GWAS catalogue are valuable tools for applied researchers and can greatly simplify the process of conducting Mendelian randomization. I enjoyed learning about them for this review and look forward to using them in my 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?
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: Biostatistics, Statistical Genetics, Mendelian randomization.

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.

Author Response 04 Nov 2019

Venexia Walker, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK

Thank you for your comments on our article. We have responded to each of them below.

"I agree with previous reviewers that the article would benefit from a short discussion of the challenges involved in performing Mendelian randomization. The MR-Base software efficiently automates the process of implementing an MR analysis, which is naturally very useful for practical purposes. However, this makes it easy to ignore some of the assumptions that underpin Mendelian randomization and limitations of the approach. Especially topics that are not automatically addressed by MR-Base, such as selection bias and the need for SNP-exposure and SNP-outcome associations to be obtained from similar populations, could be highlighted to alert applied practitioners using the software to potential limitations of their analyses."

We have added a section titled 'Limitations of Mendelian randomization' to the article, which

includes discussion of selection bias and the requirement for non-overlapping samples in two-sample Mendelian randomization. It reads: “Mendelian randomization is subject to several limitations. For example, effect sizes may not be indicative of the effects of a clinical intervention later in life. This can occur for several reasons, such as cumulative exposure, where Mendelian randomization estimates may reflect the effect of lifelong exposure, or time-dependent exposure, where intervention outside of a critical period does not have an effect despite the Mendelian randomization estimates suggesting an effect exists. Mendelian randomization estimates may also be affected by issues such as horizontal pleiotropy, whereby the SNPs chosen to proxy the exposure may affect the outcome by pathways other than the exposure of interest leading to biased results [1, 5–11]. Furthermore, investigators are reliant on relevant data being made available and appropriate methods being developed. For instance, it is currently difficult to study prognostic factors due to the lack of GWAS conducted for disease progression outcomes and the susceptibility of current methods to collider bias [13]. A related issue is selection bias, which often occurs due to the presence of a collider. This form of bias is present when the individuals included within the analysis are not representative of the population that you are trying to study. [https://doi.org/10.1093/ije/dyy202] Careful consideration of the GWAS used for two-sample Mendelian randomization is therefore encouraged to minimize the effect of this bias on estimates. An additional limitation, specific to two-sample Mendelian randomization, is the requirement for no sample overlap. While using different samples for the exposure and outcome data is an advantage of this method as it can increase power, it can be difficult to determine whether individual participants appear in both datasets. If they do, simulations have suggested there is a potential for “substantial bias and inflated Type 1 error rates”. [https://doi.org/10.1002/gepi.21998] While MR-Base includes multiple sensitivity analyses that aim to address some of these limitations, such as horizontal pleiotropy. Others, such as sample overlap, are not accounted for and therefore must be considered carefully by users of this tool. Further details and limitations of Mendelian randomization can be found in the literature [1, 5–10].”

“As suggested by the authors, MR-Base can greatly contribute to promoting transparency and reproducibility of research results in Mendelian randomization. Given that the software is already quite popular, it could be useful if the authors incorporated a searchable database (similar to the GWAS catalogue) of published Mendelian randomization analyses that were conducted using MR-Base. This would, for example, allow researchers who want to investigate a specific trait to easily access previous research findings concerning causal relationships between that trait and other common risk factors or outcomes.”

We do not require users to directly notify us of published work that has used MR-Base however, we do ask that they cite the original MR-Base article in any published research. Reported research findings can therefore be found by looking through the citations for this article. There are currently no plans to include a searchable database of published Mendelian randomization analyses because of the large amount of manual curation of studies that would be needed to create such a database.

“It is helpful that the paper is complemented with links to GitHub pages, containing instructions on how to implement the authors’ analysis for each of the three Use Cases. Nevertheless, I think it could improve the presentation of the manuscript if the authors incorporated some of the results tables and associated graphs obtained from MR-Base into the paper for illustration purposes.”

We agree that exemplar tables and graphs that demonstrate the output from MR-Base would

improve the article. We have therefore added exemplar output for use case 2 – see Table 4 and Figures 2-5.

Competing Interests: No competing interests were disclosed.

Reviewer Report 14 August 2019

https://doi.org/10.21956/wellcomeopenres.16742.r36078

C. Mary Schooling
[1] Graduate School of Public Health and Health Policy, City University of New York, New York, NY, USA
[2] School of Public Health, The University of Hong Kong, Hong Kong SAR, China

This is an excellent instructional resource concerning use of Mendelian randomization (MR) and MR-Base. In that context it would be very helpful if the paper included material explaining the interpretation of MR-Base estimates in relation to causal inference, consolidated important limitations of MR-Base into one comprehensive section and reflected more on the interpretation of the MR estimate. Some of this information on limitations and interpretation of MR is already in the paper, but could be consolidated and made more comprehensive for the intended target audience, while the first item should be made clearer. Specifically,

1. The paper should make a clearer distinction between MR and causal inference. The paper explains that MR addresses confounding and reverse causation. However, valid causal inference requires estimates free from confounding and free from selection bias. So, the paper should draw the reader's attention to this difference. As such, it would be better to explain throughout that MR gives unconfounded estimates rather than to equate MR with strengthening causal inference. As a general "health warning" it should also be pointed out that all MR studies depend on the assumption that the underlying genome wide association studies give unbiased genetic estimates. That is not to detract from the strengths of MR-Base and MR but just to be clear for the reader.

2. The sections explaining the limitations of MR and MR-Base would be better consolidated under a heading "Limitations" for easy reference. The material in the second half of the paragraph on page 3 headed "Principles of Mendelian randomization" could be moved to this section. The sentences at the bottom of page 4 explaining that MR cannot be used to assess disease progression could be moved to this section. It might also be worth mentioning that MR may not give reliable estimates of the effects of risk factors in samples of patients.

3. More information on the interpretation of the MR estimate would also be helpful. Many MR estimates rely on the INSIDE (instrument strength independent of the direct effect) assumption which seems to imply that the MR estimate means that the exposure of interest or a precursor sharing genetic predictors with the exposure could be causal.

MR-Base is a wonderful tool that generally gives the same results as MR conducted by hand or using other packages. In that context, there are a few clarifications that would be helpful.

1. MR-Base users may not realize that the results are given in beta coefficients for the outcome which needs to be exponentiated if the outcome is reported as logodds. Maybe this point could be added to Table 3.

2. MR-Base has sometimes given different p-values for MR-Egger than the MendelianRandomization package, possibly because MR-Base is using a t-distribution and MendelianRandomization is using a normal distribution to obtain these p-values. Maybe this this difference is no longer relevant but if it is still occurring it might be worth addressing it in some way.

3. The clump_data function of MR-Base could make it clear whether genetic variants are excluded because they are correlated with the ones retained or because they are not in the 1000 genomes catalog.

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?
Partly

Competing Interests: One of my PhD students is currently at the University of Bristol working on a project supervised by Tom Gaunt.

Reviewer Expertise: Epidemiological Methods, Non-communicable diseases, Evolutionary Biology, Interventions.

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.

Author Response 04 Nov 2019

Venexia Walker, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK

Thank you for your comments regarding our article. Please see our response below.

“The paper should make a clearer distinction between MR and causal inference. The paper explains that MR addresses confounding and reverse causation. However, valid causal inference requires estimates free from confounding and free from selection bias. So, the paper should draw

the reader's attention to this difference. As such, it would be better to explain throughout that MR gives unconfounded estimates rather than to equate MR with strengthening causal inference."

We have revised the language in the article to better distinguish MR and causal inference. For example: the first sentence, which read "Mendelian randomization (MR) is a method used to study the effects of risk factors and exposures on outcome traits using genetic variation" now reads "Mendelian randomization (MR) is a causal inference method used to study the effects of risk factors and exposures on outcome traits by exploiting genetic variation to address confounding and reverse causation".

"As a general "health warning" it should also be pointed out that all MR studies depend on the assumption that the underlying genome wide association studies give unbiased genetic estimates. That is not to detract from the strengths of MR-Base and MR but just to be clear for the reader."

We have added this warning to the 'Principles of Mendelian randomization' section that introduces two-sample Mendelian randomization. It reads: "Note that the use of GWAS in this way for two-sample Mendelian randomization also requires us to make an additional assumption that the GWAS used are providing unbiased genetic estimates."

"The sections explaining the limitations of MR and MR-Base would be better consolidated under a heading "Limitations" for easy reference. The material in the second half of the paragraph on page 3 headed "Principles of Mendelian randomization" could be moved to this section. The sentences at the bottom of page 4 explaining that MR cannot be used to assess disease progression could be moved to this section. It might also be worth mentioning that MR may not give reliable estimates of the effects of risk factors in samples of patients."

We have added a section titled 'Limitations of Mendelian randomization' to the article that includes the sentences you highlighted from pages three and four. It reads: "Mendelian randomization is subject to several limitations. For example, effect sizes may not be indicative of the effects of a clinical intervention later in life. This can occur for several reasons, such as cumulative exposure, where Mendelian randomization estimates may reflect the effect of lifelong exposure, or time-dependent exposure, where intervention outside of a critical period does not have an effect despite the Mendelian randomization estimates suggesting an effect exists. Mendelian randomization estimates may also be affected by issues such as horizontal pleiotropy, whereby the SNPs chosen to proxy the exposure may affect the outcome by pathways other than the exposure of interest leading to biased results [1, 5–11]. Furthermore, investigators are reliant on relevant data being made available and appropriate methods being developed. For instance, it is currently difficult to study prognostic factors due to the lack of GWAS conducted for disease progression outcomes and the susceptibility of current methods to collider bias [13]. A related issue is selection bias, which often occurs due to the presence of a collider. This form of bias is present when the individuals included within the analysis are not representative of the population that you are trying to study. [https://doi.org/10.1093/ije/dyy202] Careful consideration of the GWAS used for two-sample Mendelian randomization is therefore encouraged to minimize the effect of this bias on estimates. An additional limitation, specific to two-sample Mendelian randomization, is the requirement for no sample overlap. While using different samples for the exposure and outcome data is an advantage of this method as it can increase power, it can be difficult to determine whether individual participants appear in both datasets. If they do, simulations have suggested there is a potential for "substantial bias and inflated Type 1 error rates". [https://doi.org/10.1002/gepi.21998] While MR-Base includes multiple sensitivity analyses that aim

to address some of these limitations, such as horizontal pleiotropy. Others, such as sample overlap, are not accounted for and therefore must be considered carefully by users of this tool. Further details and limitations of Mendelian randomization can be found in the literature [1, 5–10]."

"More information on the interpretation of the MR estimate would also be helpful. Many MR estimates rely on the INSIDE (instrument strength independent of the direct effect) assumption which seems to imply that the MR estimate means that the exposure of interest or a precursor sharing genetic predictors with the exposure could be causal."

We deemed detailed discussion of the theory and interpretation of Mendelian randomization results to be beyond the scope of this article, which deals with the practical considerations when conducting this type of analysis using MR-Base. We refer readers at the end of the section 'Principles of Mendelian randomization' to the literature where further information on theory and interpretation of results can be found. However, we agree that the INSIDE assumption is an important consideration when using some Mendelian randomization methods and should have been included with the other assumptions in 'Table 1. Key definitions.' Consequently, we have added the following definition to that table: "An assumption that there is zero correlation between the SNP-exposure associations (i.e. the instrument strength) and the SNP-outcome associations (i.e. the direct effect of the instruments on the outcome). [15] Required for some MR methods, such as MR-Egger."

"MR-Base users may not realize that the results are given in beta coefficients for the outcome which needs to be exponentiated if the outcome is reported as logodds. Maybe this point could be added to Table 3."

We have highlighted this point in Table 3 as suggested and reiterated it at the end of Use Case 2, which makes use of this transformation, with the following statement: "Note that the estimates returned by MR-Base are beta coefficients for the outcome. Binary outcomes are commonly reported as log odds ratios, as was the case here, and so will need to be exponeniated in order to obtain an odds ratio."

"MR-Base has sometimes given different p-values for MR-Egger than the MendelianRandomization package, possibly because MR-Base is using a t-distribution and MendelianRandomization is using a normal distribution to obtain these p-values. Maybe this difference is no longer relevant but if it is still occurring it might be worth addressing it in some way."

As you mention, MR-Base uses a t-distribution to calculate p-values and other packages may use alternative distributions for this calculation. We have stated that this distribution is used in Table 3, so users are informed. Generally, we would encourage users to avoid using p-values to indicate the relevance of their findings and, instead, focus on effect estimates and confidence intervals.

"The clump_data function of MR-Base could make it clear whether genetic variants are excluded because they are correlated with the ones retained or because they are not in the 1000 genomes catalog."

As of 19th September 2019, the 'clump_data' function states the reason for SNP removal. The TwoSampleMR package will need to be updated if the currently installed version was downloaded prior to this date.

Competing Interests: No competing interests were disclosed.

Reviewer Report 08 August 2019

https://doi.org/10.21956/wellcomeopenres.16742.r36141

Bernard M. Y. Cheung
Department of Medicine, The University of Hong Kong, Hong Kong, China

General comments:

The article by Walker *et al.* on using the MR-Base platform to investigate risk factors and drug targets for thousands of phenotypes is a welcome addition to the literature. Naturally, the authors wax lyrical about the MR-Base platform they created and developed. It is indeed an attractive platform in that it holds summary statistics on traits and health outcomes from over 20,000 GWAS. To search for relevant GWAS and retrieve the data from scratch would have taken a researcher weeks and months, whereas with this integrated platform, you can do this in a day once you have learnt how to use it. Another major advantage of this platform is that it is transparent and enables other researchers to confirm the results using the same database and the published analysis codes.

The article is concise and focuses on the information the end-user would want. Notably, three case studies, or 'use cases', are discussed to illustrate the utility and advantages of MR-Base. The explanation of two-sample Mendelian randomization tends to be brief, but this is reasonable, as the authors have previously published extensively on the methodology.

All in all, MR-Base is an exciting innovation that has already produced some important findings and the present article provides much practical information to enable interested researchers to explore and exploit this new research tool.

Specific comments:

- As 'thousands of phenotypes' is in the title, there should be some information on the quantity, breadth and depth of the phenotypic variables. Part of the information can be given as a table if appropriate.
- In lieu of Discussion, there is a paragraph of conclusions. Regardless of what the section is called, it would be nice if there were some discussions on the limitations of MR-Base.
- Would this database continue to be free or would it require subscription or payment in the future?
- In the Abstract, complementary was misspelt as complimentary.

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: Cardiovascular risk factors; meta-analysis; clinical pharmacology

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.

Author Response 04 Nov 2019

Venexia Walker, University of Bristol, Oakfield House, Oakfield Grove, Bristol, UK

Thank you for your feedback on our article. We have addressed your specific comments as follows.

“As ‘thousands of phenotypes’ is in the title, there should be some information on the quantity, breadth and depth of the phenotypic variables. Part of the information can be given as a table if appropriate.”

We have intentionally provided the scale, but not a specific quantity, of phenotypes as the database continues to grow and there is no clear way to define distinct traits. However, we have added the following text to the introduction that highlights the breadth and depth of the phenotypic variables: “Traits include anthropometric measures, risk factors, metabolites, metals, vitamins, circulating proteins, disease outcomes, and disease intermediates (such as LDL cholesterol and systolic blood pressure), as well as DNA methylation and gene expression phenotypes.”

“In lieu of Discussion, there is a paragraph of conclusions. Regardless of what the section is called, it would be nice if there were some discussions on the limitations of MR-Base.”

We have added a section titled ‘Limitations of Mendelian randomization’ to the article. It reads: “Mendelian randomization is subject to several limitations. For example, effect sizes may not be indicative of the effects of a clinical intervention later in life. This can occur for several reasons, such as cumulative exposure, where Mendelian randomization estimates may reflect the effect of lifelong exposure, or time-dependent exposure, where intervention outside of a critical period does not have an effect despite the Mendelian randomization estimates suggesting an effect exists. Mendelian randomization estimates may also be affected by issues such as horizontal pleiotropy,

whereby the SNPs chosen to proxy the exposure may affect the outcome by pathways other than the exposure of interest leading to biased results [1, 5–11]. Furthermore, investigators are reliant on relevant data being made available and appropriate methods being developed. For instance, it is currently difficult to study prognostic factors due to the lack of GWAS conducted for disease progression outcomes and the susceptibility of current methods to collider bias [13]. A related issue is selection bias, which often occurs due to the presence of a collider. This form of bias is present when the individuals included within the analysis are not representative of the population that you are trying to study. [https://doi.org/10.1093/ije/dyy202] Careful consideration of the GWAS used for two-sample Mendelian randomization is therefore encouraged to minimize the effect of this bias on estimates. An additional limitation, specific to two-sample Mendelian randomization, is the requirement for no sample overlap. While using different samples for the exposure and outcome data is an advantage of this method as it can increase power, it can be difficult to determine whether individual participants appear in both datasets. If they do, simulations have suggested there is a potential for "substantial bias and inflated Type 1 error rates". [https://doi.org/10.1002/gepi.21998] While MR-Base includes multiple sensitivity analyses that aim to address some of these limitations, such as horizontal pleiotropy. Others, such as sample overlap, are not accounted for and therefore must be considered carefully by users of this tool. Further details and limitations of Mendelian randomization can be found in the literature [1, 5–10]."

"Would this database continue to be free or would it require subscription or payment in the future?"

As stated in the disclaimers within the data access agreement, we do not guarantee that the Platform or any GWAS Data will always be available or interrupted. However, there are no current plans for the database to require a subscription or payment in the future.

"In the Abstract, complementary was misspelt as complimentary."

Thank you – we have corrected this misspelling.

Competing Interests: No competing interests were disclosed.