



REVIEW

**REVISED** Mendelian randomization for studying the effects of perturbing drug targets [version 2; peer review: 3 approved, 1 approved with reservations]

Dipender Gill <sup>1-5</sup>, Marios K. Georgakis<sup>6</sup>, Venexia M. Walker <sup>7-9</sup>, A. Floriaan Schmidt<sup>10,11</sup>, Apostolos Gkatzionis<sup>12</sup>, Daniel F. Freitag<sup>13</sup>, Chris Finan<sup>10,14,15</sup>, Aroon D. Hingorani<sup>10,14,15</sup>, Joanna M.M. Howson<sup>3</sup>, Stephen Burgess <sup>12,16</sup>, Daniel I. Swerdlow<sup>10</sup>, George Davey Smith <sup>7,8,17</sup>, Michael V. Holmes <sup>18</sup>, Martin Dichgans<sup>6,19,20</sup>, Robert A Scott<sup>21</sup>, Jie Zheng<sup>7,8</sup>, Bruce M. Psaty <sup>22,23</sup>, Neil M. Davies <sup>7,8,24</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

<sup>2</sup>Centre for Pharmacology and Therapeutics, Department of Medicine, Imperial College London, London, UK

<sup>3</sup>Novo Nordisk Research Centre, Oxford, UK

<sup>4</sup>Clinical Pharmacology and Therapeutics Section, Institute of Medical and Biomedical Education and Institute for Infection and Immunity, St George's, University of London, London, UK

<sup>5</sup>Clinical Pharmacology Group, Pharmacy and Medicines Directorate, St George's University Hospitals NHS Foundation Trust, London, UK

<sup>6</sup>Institute for Stroke and Dementia Research (ISD), University Hospital of Ludwig-Maximilians-University (LMU), Munich, Germany

<sup>7</sup>Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK

<sup>8</sup>Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

<sup>9</sup>Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

<sup>10</sup>Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, UK

<sup>11</sup>Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>12</sup>Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

<sup>13</sup>Bayer Pharmaceuticals, Open Innovation & Digital Technologies, Wuppertal, Germany

<sup>14</sup>UCL British Heart Foundation Research Accelerator, University College London, London, UK

<sup>15</sup>UCL Hospitals, NIHR Biomedical Research Centre, London, UK

<sup>16</sup>Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

<sup>17</sup>NIHR Bristol Biomedical Research Centre, University of Bristol, Bristol, UK

<sup>18</sup>Medical Research Council Population Health Research Unit, University of Oxford, Oxford, UK

<sup>19</sup>Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

<sup>20</sup>German Centre for Neurodegenerative Diseases (DZNE), Munich, Germany

<sup>21</sup>Human Genetics, GlaxoSmithKline, Stevenage, UK

<sup>22</sup>Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA, USA

<sup>23</sup>Kaiser Permanente Washington Health Research Institute, Seattle, WA, USA

<sup>24</sup>K.G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

**v2** First published: 28 Jan 2021, 6:16  
<https://doi.org/10.12688/wellcomeopenres.16544.1>

Latest published: 10 Feb 2021, 6:16  
<https://doi.org/10.12688/wellcomeopenres.16544.2>

Open Peer Review

Reviewer Status

### Abstract

Drugs whose targets have genetic evidence to support efficacy and safety are more likely to be approved after clinical development. In this paper, we provide an overview of how natural sequence variation in the genes that encode drug targets can be used in Mendelian randomization analyses to offer insight into mechanism-based efficacy and adverse effects. Large databases of summary level genetic association data are increasingly available and can be leveraged to identify and validate variants that serve as proxies for drug target perturbation. As with all empirical research, Mendelian randomization has limitations including genetic confounding, its consideration of lifelong effects, and issues related to heterogeneity across different tissues and populations. When appropriately applied, Mendelian randomization provides a useful empirical framework for using population level data to improve the success rates of the drug development pipeline.

### Keywords

Drugs, Genetics, Mendelian randomization

	Invited Reviewers			
	1	2	3	4
<b>version 2</b> (revision) 10 Feb 2021	 report	 report	 report	 report
<b>version 1</b> 28 Jan 2021				
<p>1. <b>Pedrum Mohammadi-Shemirani</b> , Hamilton Health Sciences, Hamilton, Canada Thrombosis and Atherosclerosis Research Institute, Hamilton, Canada McMaster University, Hamilton, Canada</p> <p><b>Michael Chong</b>, David Braley Cardiac, Vascular and Stroke Research Institute, Hamilton Health Sciences, Hamilton, Canada Thrombosis and Atherosclerosis Research Institute, Hamilton, Canada McMaster University, Hamilton, Canada McMaster University, Hamilton, Canada</p> <p>2. <b>Shiu Lun Ryan Au Yeung</b> , The University of Hong Kong, Hong Kong, Hong Kong</p> <p>3. <b>C. Mary Schooling</b> , City University of New York, New York, USA</p> <p>4. <b>Matthew Traylor</b> , Queen Mary University of London, London, UK</p> <p>Any reports and responses or comments on the article can be found at the end of the article.</p>				

**Corresponding author:** Dipender Gill ([dipender.gill@imperial.ac.uk](mailto:dipender.gill@imperial.ac.uk))

**Author roles:** **Gill D:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Georgakis MK:** Conceptualization, Funding Acquisition, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Walker VM:** Conceptualization, Funding Acquisition, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Schmidt AF:** Conceptualization, Funding Acquisition, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Gkatzionis A:** Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing; **Freitag DF:** Conceptualization, Writing – Review & Editing; **Finan C:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Hingorani AD:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Howson JMM:** Conceptualization, Writing – Review & Editing; **Burgess S:** Conceptualization, Funding Acquisition, Methodology, Writing – Review & Editing; **Swerdlow DI:** Conceptualization, Writing – Review & Editing; **Davey Smith G:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Holmes MV:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Dichgans M:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Scott RA:** Conceptualization, Writing – Review & Editing; **Zheng J:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Psaty BM:** Conceptualization, Funding Acquisition, Writing – Review & Editing; **Davies NM:** Conceptualization, Funding Acquisition, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** DG is employed part-time by Novo Nordisk. DFF is an employee of Bayer AG, Germany. JMMH is an employee of Novo Nordisk. DIS is an employee of Silence Therapeutics plc. MVH has collaborated with Boehringer Ingelheim in research, and in accordance with the policy of the Clinical Trial Service Unit and Epidemiological Studies Unit (University of Oxford), did not accept any personal payment. RAS is an employee and shareholder in GlaxoSmithKline. BMP serves on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. NMD reports funding from the Global Research Awards into Nicotine Dependence (GRAND) which is an independent grant awarding body funded by Pfizer. All other authors declare no conflicts of interest.

**Grant information:** DG is supported by the Wellcome Trust 4i Programme (203928/Z/16/Z) and British Heart Foundation Centre of Research Excellence (RE/18/4/34215) at Imperial College London, and by a National Institute for Health Research Clinical Lectureship at St. George's, University of London (CL-2020-16-001). MKG has received funding by the Onassis Foundation and the German Academic Exchange Service (DAAD). AG is funded by a Medical Research Council Methodology Research Panel Grant (RG88311). This project has received funding from the European Union's Horizon 2020 research and innovation programme (No 666881, to MD), SVDs@target (to MD) and No 667375, CoSTREAM (to MD); the DFG as part of the Munich Cluster for Systems Neurology (EXC 1010 SyNergy, to MD) and the CRC 1123 (B3, to MD); the Corona Foundation (to MD); the Fondation Leducq (Transatlantic Network of Excellence on the Pathogenesis of Small Vessel Disease of the Brain) (to MD). CF and ADH are supported by the UCL NIHR Biomedical Research Centre and UCL BHF Research Accelerator. ADH is an NIHR Senior Investigator. SB is supported by Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (204623/Z/16/Z). MVH works in a unit that receives funding from the UK Medical Research Council and is supported by a British Heart Foundation Intermediate Clinical Research Fellowship (FS/18/23/33512) and the National Institute for Health Research Oxford Biomedical Research Centre. JZ is funded by a Vice-Chancellor Fellowship from the University of Bristol. The Medical Research Council (MRC) and the University of Bristol support the MRC Integrative Epidemiology Unit (MC\_UU\_12013/1, MC\_UU\_12013/9, MC\_UU\_00011/1). The Norwegian Research Council support NMD grant number 295989. This work was supported by the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) National Heart, Lung, and Blood Institute (NHLBI) grant R01HL105756-09.

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

**Copyright:** © 2021 Gill D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Gill D, Georgakis MK, Walker VM *et al.* **Mendelian randomization for studying the effects of perturbing drug targets [version 2; peer review: 3 approved, 1 approved with reservations]** Wellcome Open Research 2021, 6:16 <https://doi.org/10.12688/wellcomeopenres.16544.2>

**First published:** 28 Jan 2021, 6:16 <https://doi.org/10.12688/wellcomeopenres.16544.1>

**REVISED Amendments from Version 1**

The following author was added: Robert A Scott. This author was omitted in the initial submission as an accident. All co-authors are aware (from previous drafts of the paper) that this author should be included. A conflict of interest statement was also added.

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

**Introduction**

The majority of small molecule and biologic drugs exert their effects by perturbing protein targets<sup>1</sup>. The identification of such targets is therefore central to drug discovery. Despite increasing investment in research and development within the pharmaceutical industry<sup>2</sup>, overall drug development failure rates remain high<sup>3-8</sup>, most notably for targets that represent novel mechanisms. Such failures result in increased costs and reduced availability of novel agents<sup>9</sup>.

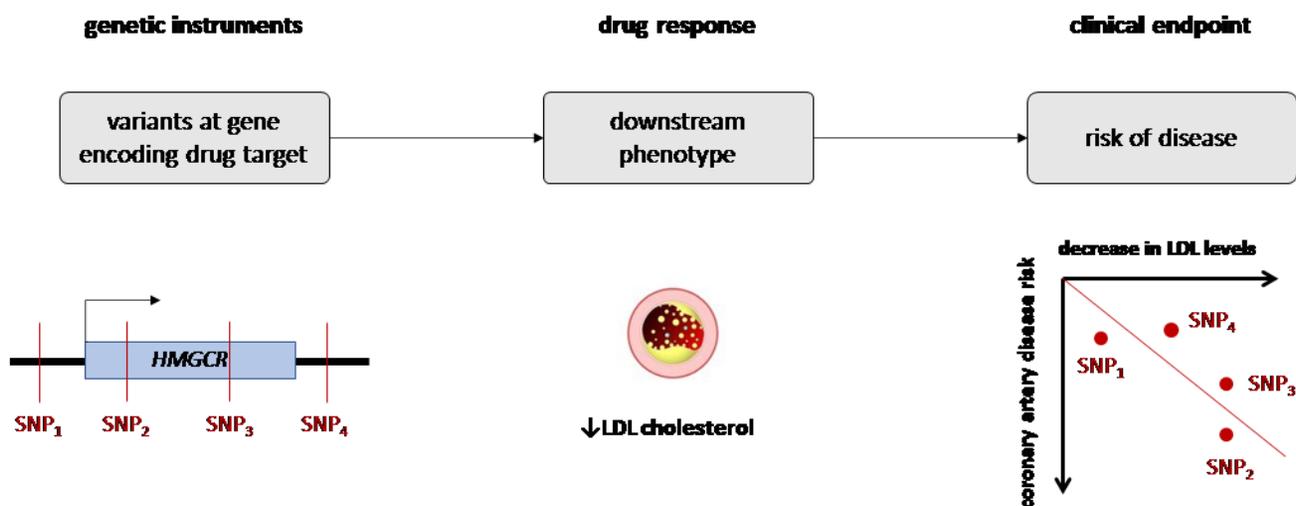
With the recent growth in genetic data<sup>10</sup>, there has been substantial progress in the identification of genes that are linked to human health and disease. Genetic data can potentially be used for identifying and prioritizing novel drug targets and indications<sup>2</sup>. For example, genome-wide association studies (GWAS) have corroborated approximately 70 of the 670 known effects of licensed drugs through associations at the loci of the genes coding for their corresponding target proteins<sup>11</sup>. Studies of drug development programs have also shown that targets with genomic support have a higher rate of success<sup>2,12-15</sup>.

**Mendelian randomization**

Through the random allocation of genetic variants at conception, genetic studies in human populations can imitate the design of randomized controlled trials (RCT)<sup>16,17</sup>. Such investigation uses genetic variants as instrumental variables for studying the effect of an exposure on an outcome, and has been referred to as Mendelian randomization (MR)<sup>18</sup>. Phenotypic observational studies are limited in their ability to draw causal inferences due to bias from confounding and reverse causation<sup>18</sup>. In contrast, MR uses the random allocation of genetic variants from parents to offspring during conception to guard against these biases.

MR requires the following instrumental variable assumptions: the genetic variant i) is associated with the exposure (relevance), ii) has no common cause with the outcome (independence), and iii) only affects the outcome via the exposure (exclusion restriction)<sup>19,20</sup>. The first of these is testable; the remaining assumptions are untestable but falsifiable. Assumption iii) the exclusion restriction, assumes that the genetic variant affects the outcome through the exposure and not any other horizontally pleiotropic pathways<sup>18,21</sup>. Further assumptions are also required to obtain valid point estimates, for instance, that the influence of the exposure on the outcome is the same for all individuals (effect homogeneity) or that the exposure is a monotonic (always increasing or always decreasing) function of the instrument for all individuals in the population (monotonicity)<sup>19</sup>. In addition, the interpretation of MR findings can have particular nuances, as previously described<sup>22</sup>.

Where the exposure under study is perturbation of a drug target, MR can be used to explore drug effects (Figure 1)<sup>23,24</sup>.



**Figure 1. Principles of Mendelian randomization studies (MR) studying drug effects.** MR makes use of genetic variants located within or close to a gene encoding a drug target (e.g. at *HMGCR* encoding the drug target of statins) that lead to downstream effects similar to the desired drug response (e.g. lowering of low-density lipoprotein [LDL] cholesterol) in order to explore effects on clinical outcomes (e.g. risk of coronary artery disease). SNP: single-nucleotide polymorphism.

For drug target MR specifically, genetic variants such as single-nucleotide polymorphisms (SNPs) related to the function or expression of the drug target protein can be used as instrumental variables to study the effect of perturbing that drug target<sup>25,26</sup>. These variants are typically in or near the gene that encodes the drug target (*cis*-variants). Such MR can be used in drug development to investigate the likely efficacy and safety of perturbing novel drug targets<sup>27,28</sup>, as well as explore the repurposing potential and adverse effects of existing drugs<sup>25</sup>.

The identification and validation of appropriate genetic variants as instrumental variables for an exposure is critical for the design and interpretation of all MR analyses<sup>29</sup>. While previous work has offered practical advice on selecting instruments for MR studies considering disease biomarkers<sup>24</sup>, the field is continuing to evolve rapidly<sup>30</sup>. The growth in genetic association study data that extends to tissue-specific gene expression<sup>31</sup>, circulating proteins<sup>32</sup>, metabolites<sup>33,34</sup> and cytokines<sup>35</sup>, has been coupled with increased efficiency of MR studies using automated software, databases, statistical packages and readily available code<sup>30,36–38</sup>. However, there is still no consensus on the strategy for identifying genetic instruments and exploring potential drug effects with MR. Here we discuss practical considerations while also offering illustrative examples for the most relevant points. We describe issues relating to selection of genetic variants as proxies for drug target perturbation, evaluation of the plausibility of genetic variants as proxies for drug target perturbation, generation and interpretation of MR estimates, and limitations of MR for investigating drug target perturbation. Finally, we offer a step-by-step framework for how to conduct a drug target MR study (Box 1).

#### Box 1. Step-by-step guide for conducting Mendelian randomization (MR) analyses of drug target perturbation

1. Determine the drug targets of interest
2. Identify the gene(s) encoding the relevant protein(s)
3. Choose data source for identifying instruments
4. Select genetic variants as instruments based on:
  - a. Strength of associations with downstream effects of drug target perturbation
  - b. Linkage disequilibrium structure
  - c. Distance from gene(s) encoding the drug target
5. Validate genetic variants for use as instruments by confirming that they recapitulate known on-target drug effects
6. Estimate effects of drug target perturbation on outcome(s) of interest using MR
  - a. Use appropriate method to account for linkage disequilibrium structure between variants
  - b. Scale estimates appropriately
  - c. Interpret MR as representing effects of lifelong drug target perturbation
7. Investigate potential adverse effects and repurposing opportunity using phenome-wide association study
8. Triangulate using other interventional, observational and experimental data

## Instrument selection

MR investigations of drug effects have mainly studied small molecule, peptide and biotherapeutic drugs<sup>39,40</sup>, where genetic instruments are selected as variants that mimic perturbation of their protein targets. Instrument selection can be considered in two parts: i) identifying the gene or group of genes corresponding to the drug target proteins and ii) selecting genetic variants to proxy perturbation of the drug targets. These steps are discussed in detail below, followed by consideration of drugs that have targets made up of multiple proteins.

### Identifying genes corresponding to drug target proteins

The key difference between conventional MR for an exposure and MR for the investigation of drug effects is that for the latter the instrument can be constructed in relation to the gene corresponding to the drug target, rather than genetic variants from across the genome (Table 1). The first step of this process is therefore to identify the drug target of interest and its corresponding gene. Resources such as DrugBank (which is freely available for non-commercial purposes) provide information about existing drugs, including their mechanism of action, targets and their corresponding gene(s), and indications<sup>41</sup>. Where the target of a drug is known, information regarding the corresponding gene can also be obtained from other databases such as Ensembl and UniProt<sup>42,43</sup>.

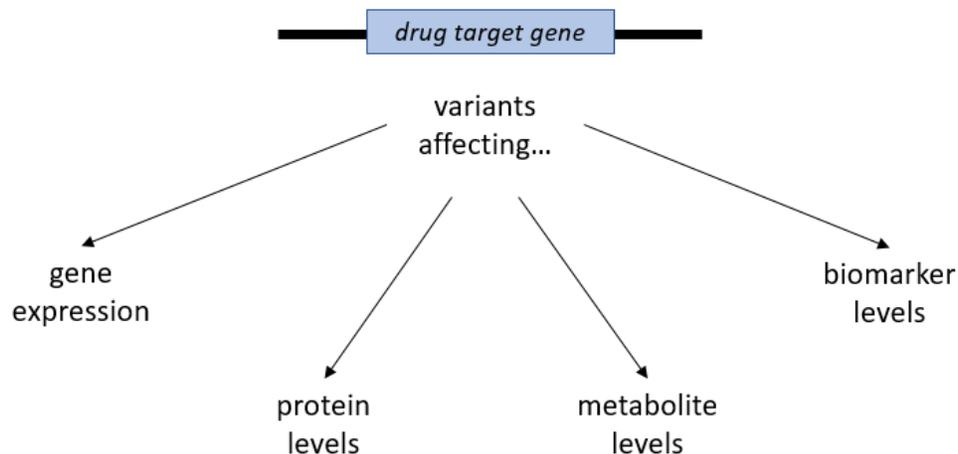
### Selecting genetic variants to proxy drug targets

Several factors need to be considered when selecting genetic variants to proxy the effects of drug target perturbation. If MR is being used to investigate effects of perturbing the target of a drug with an existing indication, then instruments can be selected based on their location at the corresponding gene and association with that indication. For instance, Gill *et al.* selected genetic variants to proxy antihypertensive drug class effects as those located at the gene corresponding to the drug target that also related to systolic blood pressure in a GWAS<sup>44</sup>. If the indication is not known, one possible approach is to use quantitative trait loci for expression of the gene encoding the drug target of interest (in relevant tissues or cell contexts) as instruments for drug target perturbation. An important limitation of gene expression is that variants affecting gene expression may not necessarily also affect protein expression, and *vice versa* (Figure 2)<sup>45</sup>. Furthermore, gene expression quantitative loci have been reported to account for little of the heritability of complex diseases<sup>46</sup>. Therefore, protein expression quantitative loci may make better instruments for proxying drug effects than gene expression data, if they are available in relevant tissues and contexts.

As mentioned above, selected instruments for drug target perturbation are often restricted to *cis*-acting genetic variants – those in or close to the gene of interest. In general, such variants are more likely to have effects specific to the protein of interest than genetic variants that are not located within the gene locus (*trans*-acting). This point is well highlighted by the example of C-reactive protein, for which MR analyses investigating effects of its circulating levels produce very different results depending on whether the instruments are selected from

**Table 1. Differences between conventional Mendelian randomization (MR) and MR specifically exploring drug target perturbation.**

	Conventional MR	MR investigating drug effects
<b>Aim of the analysis</b>	To investigate the effect of an exposure on an outcome	To investigate the effect of perturbing a drug target on an outcome
<b>Genomic location of instruments</b>	Genome-wide	Often restricted to the locus of the gene encoding the drug target under study
<b>Selection of genetic instruments</b>	Variants associated with the exposure under study	Variants associated with perturbation of the drug target under study
<b>Statistical analysis</b>	Typically uses uncorrelated variants; higher risk of pleiotropic effects on the outcome through pathways unrelated to the exposure	More frequent use of methods to account for correlation between instrument variants; lower risk of pleiotropic effects on the outcome through pathways unrelated to the drug target

**Figure 2. Potential strategies for selecting genetic variants as instruments for a protein drug target.** Variants within or close to the drug target gene might be selected on the basis of their associations with gene expression and levels of the target protein or known downstream functions such as effects on metabolite levels or biomarkers. Notably, variants influencing gene expression and protein levels do not always influence the function of the derived protein product and might not translate to downstream effects comparable to those achieved by the pharmacological modulation of the drug target.

the *CRP* locus or from throughout the genome<sup>24</sup>. Related considerations include how proximal *cis*-variants should be to the corresponding gene, and whether enhancer or promoter regions for the gene should be counted as *cis*-acting<sup>47</sup>. While some evidence currently supports that genetic variants affecting gene expression typically lie within 200kB of the gene locus<sup>48</sup>, there is no established consensus on the issue of proximity. The pertinent requirement is that the variant be related to the function or expression of the drug target. In the past, single-region MR analyses have sometimes used only the top variant in the region as a genetic instrument<sup>28,49,50</sup>. However, this approach can also be suboptimal if studying a region containing variants that have multiple conditionally independent associations with the exposure. For example, genetic association studies have suggested that the *SHBG* region encoding sex-hormone binding globulin (SHBG) may harbor up to nine variants independently associated with circulating SHBG concentration<sup>51</sup>, and

that using only the top variant may limit the statistical power of such MR analysis.

The degree to which variants at the same locus should be allowed to correlate with each other through linkage disequilibrium (LD) while still being modelled as independent also warrants attention. Unaccounted correlation between the variants used can result in underestimation of the standard error of MR estimates, yet there is no recommended LD threshold. To circumvent this issue, methods are available to adjust for LD between genetic variants used as instruments, which may help confirm the robustness of the findings and maximize statistical power<sup>52-54</sup>.

#### Investigating drugs with multiple targets

Many drugs do not have a target that is encoded by a single gene. For example, the calcium channel blocker class of

antihypertensive drugs have targets that are made up of proteins coded by several different genes<sup>44,55</sup>. At present there is no consensus on the best way to combine data from multiple genes corresponding to a single target into an instrument. Previous studies have selected genetic variants related to the individual genes and combined their data to investigate the effect of perturbing the drug target, while applying clumping to ensure independence as described above<sup>44,55</sup>.

### Instrument evaluation

Once the instrument has been selected, it can be evaluated to ascertain its validity for the analysis of interest. MR analyses exploring drug effects can be biased if the genetic variants incorporated as instruments have “horizontal” pleiotropic effects, where there are pathways from the variant to the specific outcome under consideration that do not pass through the exposure of interest<sup>56</sup>. In contrast, “vertical” pleiotropy lies on the causal pathway between the pharmacological mechanism and outcome<sup>56</sup>. Vertical pleiotropy does not bias MR estimates and is often of interest as it can provide insight into causal mechanisms and mediation. As with MR generally, one of the most useful approaches for evaluating instrument validity is to investigate its relation to a known effect of the exposure under consideration<sup>57</sup>. This approach is feasible for MR used to predict the effect of perturbing targets for which there are drugs with established indications and known associations with biomarkers<sup>58</sup>. For example, Walker *et al.* selected genetic variants to proxy antihypertensive drugs from gene expression data and validated these instruments through their associations with systolic blood pressure, prior to applying MR analyses investigating the outcome of interest, Alzheimer’s disease<sup>55</sup>. An instrument may also be examined in relation to potential confounders, in order to investigate violations of the independence and exclusion restriction assumptions necessary for MR<sup>56</sup>. Berry *et al.* illustrated such an approach during their evaluation of genetic proxies for vitamin D status<sup>59</sup>. In this study, the association of variants with social, dietary and lifestyle factors was investigated, to identify potential sources of confounding.

Complementary data may also be used for instrument evaluation. For example, MR studies designed to investigate the effect of genetically predicted variations in interleukin-6 (IL6) signaling would be expected to show that the selected instruments associate with molecules that are downstream of the pathway<sup>49</sup>. Genetic association estimates for the serum levels of several of these molecules are available, including IL6 and IL6 receptor (IL6R), C-reactive protein (CRP) and fibrinogen<sup>60</sup>. Hence, if the selected genetic instruments are valid proxies for IL6 signaling, they may be expected to show consistent effects across these molecules. An alternative example is provided by Wurtz *et al.* who demonstrated consistency between the metabolic changes associated with starting statins and metabolomic associations of the *HMGCR* variant rs12916 that was used to proxy statin effect<sup>61</sup>.

### Analysis

Given a set of genetic instruments, the statistical methods used for MR investigation of drug target perturbation are similar to those used for MR more generally<sup>62</sup>. Interpretability is often

facilitated by scaling of genetic associations to unit change in a trait related to drug target perturbation. For example, for analyses considering associations of variants in the *HMGCR* gene that are used to proxy statin drug effects, estimates may be scaled to change in low-density lipoprotein cholesterol levels<sup>50,63–65</sup>. As another example, for analyses investigating IL6R signaling using variants in the *IL6R* gene, effects may be scaled to downstream changes in CRP levels<sup>49,60</sup>. Care must be taken in the interpretation of such scaled estimates however, because although MR estimates may be directionally concordant to the effect of drug target perturbation on the biomarker, their magnitudes may not be comparable<sup>66</sup>.

Statistical approaches used to evaluate potential bias from horizontal pleiotropy in MR analyses can also be used in MR investigating drug target perturbation<sup>62</sup>. However, variants selected as instruments for drug target perturbation are often selected from within a specific locus rather than from throughout the genome, and may be limited in number. Statistical sensitivity analyses for investigating horizontal pleiotropy typically require large numbers of genetic variants, and so may not be suitable for many drug target MR analyses<sup>29</sup>. Assessment of heterogeneity between MR estimates produced by variants in a single locus is still possible however, and can be used to inform on potential bias related to horizontal pleiotropy<sup>53,67</sup>.

In an effort to better explore the target region and increase statistical power, genetic variants that have weaker associations with perturbation of the drug target may be considered as instruments<sup>27,68</sup>. Despite the potential benefits of this approach<sup>69</sup>, care must be taken to avoid weak instrument bias<sup>54,70</sup>. Under a two-sample design, weak instrument bias will attenuate MR estimates towards the null<sup>71</sup>.

MR can be used to assess a wide range of outcome traits and thus investigate potential effects of perturbing the drug target on these traits<sup>72</sup>. Such studies are often conducted as hypothesis-free, phenome-wide association analyses (PheWAS)<sup>73,74</sup>, and can be helpful for exploring potential adverse effects or identifying previously unknown re-purposing opportunities. For example, Schmidt *et al.* conducted a PheWAS of the *PCSK9* locus to assess potential adverse effects of PCSK9 inhibitor drugs<sup>75</sup>.

In addition to using MR, it is also possible to generate genetic evidence supporting a causal effect of drug target perturbation on an outcome by identifying proportionality of genetic associations with traits proxying drug target perturbation and the outcome, at the corresponding drug target gene locus. Such investigation is referred to as genetic colocalization, and can help distinguish causation from genetic confounding (such as may arise due to horizontal pleiotropy). Popular colocalization methods include coloc<sup>76</sup>, moloc<sup>77</sup>, eCAVIAR<sup>78</sup> and HEIDI<sup>79</sup>. However, a limitation of many colocalization approaches is that they assume there is only a single causal variant at the considered locus.

Triangulation, the practice of integrating evidence from several different methodological approaches and data sources that each

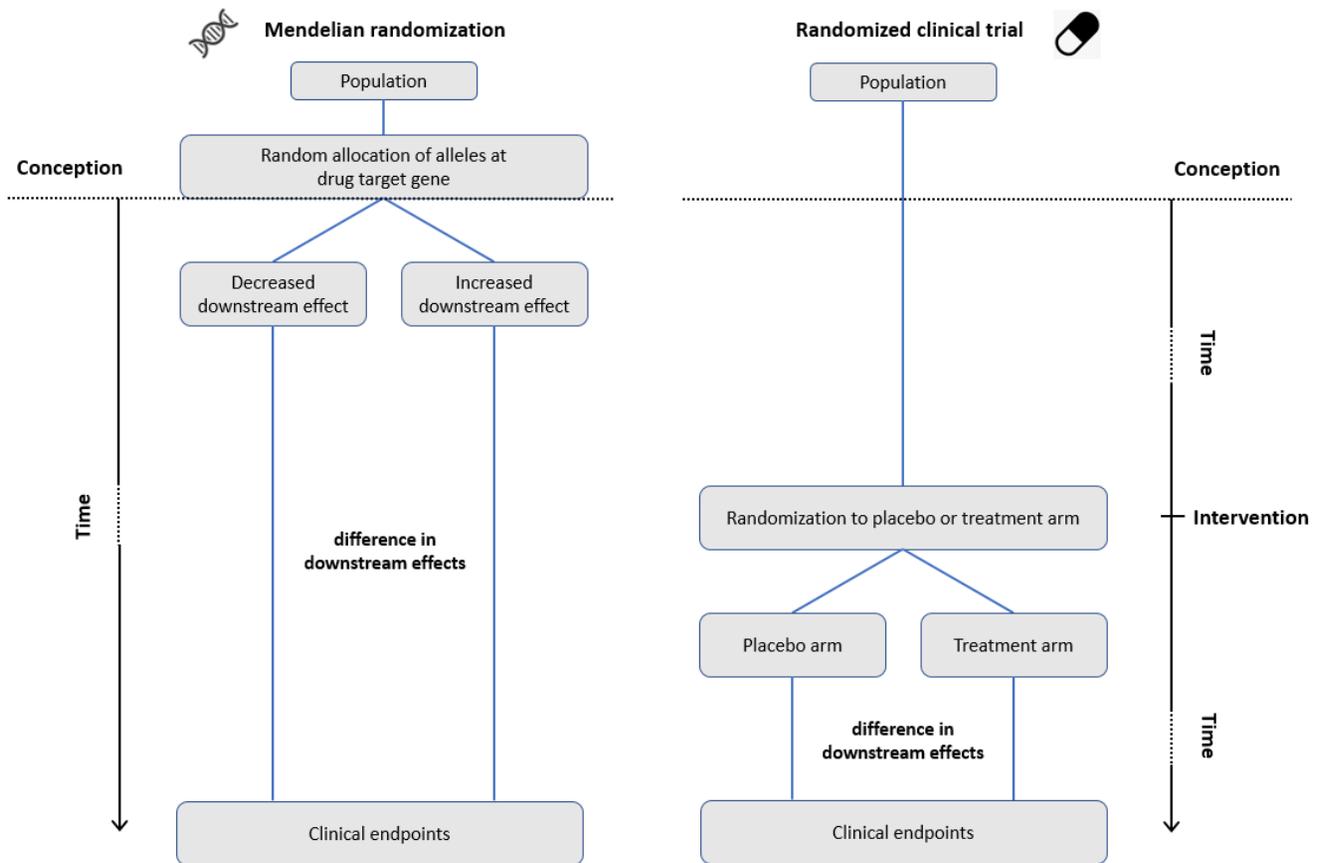
differ in their susceptibility to bias, is another important aspect of interpreting the analysis<sup>80</sup>. MR evidence should be considered alongside other study designs to increase confidence in findings<sup>58</sup>. For example, the European Atherosclerosis Society consensus statement on the role of low-density lipoproteins on atherosclerotic cardiovascular disease considers evidence from inherited disorders of lipid metabolism, prospective epidemiologic studies, MR investigations and RCTs<sup>81</sup>. Moreover, comparing instruments between different MR studies of the same exposure can provide additional evidence. For example, both Gill *et al.* and Walker *et al.* independently derived instruments for antihypertensive drug effects that perform comparably when tested against a common outcome<sup>44,55</sup>. Although different MR studies may use similar or overlapping data sources, different instrument selection approaches can make analyses vulnerable to distinct biases and so also have a role in triangulation of evidence.

### Limitations

As with all research methods, MR has limitations<sup>82</sup>. RCTs remain the best source of evidence evaluating drug efficacy and

guiding clinical practice<sup>83</sup>. While MR and RCTs have the same aim – reliable evidence of causation – they estimate different treatment parameters, which are not directly comparable. Genetic variants typically have smaller effects which accumulate across the entire life-course, whereas pharmacological agents are often prescribed later in life and typically have larger effects. Therefore, MR estimates reflect the lifelong effects of perturbing a drug target, which may not be equivalent to interventions given at a specific point in time and for a shorter time period (Figure 3). While these differences make it unlikely that MR estimates will accurately reflect the size of effect of a pharmacological intervention, they are still a useful indication of presence and direction of causal effects<sup>58</sup>.

A further limitation of MR for studying the effects of drug target perturbation is that it may not account for post-transcriptional and post-translational modification in the pathway from a gene to a biologically functional protein. Well-conducted MR analyses may be able to inform broadly on drug class effects, but not necessarily provide information on the effects of a specific pharmacological agent. For



**Figure 3. Comparison between Mendelian randomization (MR) study for drug effects and randomized clinical trial (RCT).** Similar to the randomization process of RCTs, the random allocation of alleles at a drug target gene in MR studies allows the distribution of individuals to groups that differ only regarding the downstream effects of the drug target and not other confounders. While the random allocation of alleles in MR studies happens at conception and leads to lifelong effects, the randomization in RCTs typically happens later in life and focuses on the effects of short-term interventions.

example, dihydropyridine and non-dihydropyridine subclasses of calcium-channel blocker antihypertensive drugs have distinct pharmacological effects. Genetic variants that affect blood pressure via calcium-channel blockade can estimate the effects of calcium-channel targeting drugs in general, but cannot differentiate the relative effects of dihydropyridine versus non-dihydropyridine subclasses<sup>44</sup>. Furthermore, MR in this context is applied to drug targets and not compounds – so it can be used to investigate the effects of perturbing a drug target, but is unlikely to be able to offer insight towards molecule specific effects<sup>28</sup>.

Drug effects also vary in different tissues and populations, and similarly MR estimates for the effects of perturbing drug targets may only be valid if genetic association data from the relevant tissues or populations are used. This limitation can have implications for both identifying instruments and using MR to study drug effects, as highlighted in an example that used gene expression data to identify instruments for antihypertensive drug classes in the investigation of repurposing potential for the prevention of Alzheimer's disease<sup>55</sup>. Here, it is not clear whether the same genetic variants related to gene expression in vascular, cardiac and brain tissue. Furthermore, to date, most genotyped samples have been sampled from European ancestry populations. While this approach minimizes the risk of population stratification and false-positive GWAS signals,

consideration of distinct ancestral groups is likely to offer novel insight. For example, genetic evidence on the effects of alcohol comes from variants in the *ALDH2* gene, which are common in Asian, but not European populations<sup>84</sup>.

## Conclusion

Over the last decade, MR has become a widely used epidemiological tool for estimating the causal effects of risk factors on clinical outcomes. On top of this well-studied application, there are now multiple examples highlighting its power for investigating drug effects. Despite its explicit assumptions, modern developments in statistical methodology and the widespread availability of multiple levels of omics data have provided the necessary resources to more reliably and efficiently use MR in order to study drug effects. As such, it has found a growing niche within the broader framework for exploring therapeutic targets, efficacy, adverse effects and repurposing potential. Given the high failure rates of clinical trials and that drug targets with genetic support are more likely to make it through the development pipeline<sup>13,15,23</sup>, MR can provide evidence for prioritizing agents to move forward in development.

## Data availability

No data is associated with this article.

## References

- Santos R, Ursu O, Gaulton A, et al.: **A comprehensive map of molecular drug targets.** *Nat Rev Drug Discov.* 2017; **16**(1): 19–34.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Plenge RM, Scolnick EM, Altshuler D: **Validating therapeutic targets through human genetics.** *Nat Rev Drug Discov.* 2013; **12**(8): 581–94.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Paul SM, Mytelka DS, Dunwiddie CT, et al.: **How to improve R&D productivity: the pharmaceutical industry's grand challenge.** *Nat Rev Drug Discov.* 2010; **9**(3): 203–14.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hay M, Thomas DW, Craighead JL, et al.: **Clinical development success rates for investigational drugs.** *Nat Biotechnol.* 2014; **32**(1): 40–51.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Munos B: **Lessons from 60 years of pharmaceutical innovation.** *Nat Rev Drug Discov.* 2009; **8**(12): 959–68.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pammolli F, Magazzini L, Riccaboni M: **The productivity crisis in pharmaceutical R&D.** *Nat Rev Drug Discov.* 2011; **10**(6): 428–38.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Scannell JW, Blanckley A, Boldon H, et al.: **Diagnosing the decline in pharmaceutical R&D efficiency.** *Nat Rev Drug Discov.* 2012; **11**(3): 191–200.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kola I, Landis J: **Can the pharmaceutical industry reduce attrition rates?** *Nat Rev Drug Discov.* 2004; **3**(8): 711–5.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wouters OJ, McKee M, Luyten J: **Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009–2018.** *JAMA.* 2020; **323**(9): 844–53.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Welter D, MacArthur J, Morales J, et al.: **The NHGRI GWAS Catalog, a curated resource of SNP-trait associations.** *Nucleic Acids Res.* 2014; **42**(Database issue): D1001–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Finan C, Gaulton A, Kruger FA, et al.: **The druggable genome and support for target identification and validation in drug development.** *Sci Transl Med.* 2017; **9**(383): eaag1166.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cook D, Brown D, Alexander R, et al.: **Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework.** *Nat Rev Drug Discov.* 2014; **13**(6): 419–31.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Nelson MR, Tipney H, Painter JL, et al.: **The support of human genetic evidence for approved drug indications.** *Nat Genet.* 2015; **47**(8): 856–60.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Sanseau P, Agarwal P, Barnes MR, et al.: **Use of genome-wide association studies for drug repositioning.** *Nat Biotechnol.* 2012; **30**(4): 317–20.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- King EA, Davis JW, Degner JF: **Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval.** *PLoS Genet.* 2019; **15**(12): e1008489.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hingorani A, Humphries S: **Nature's randomised trials.** *Lancet.* 2005; **366**(9501): 1906–8.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Thanassoulis G, O'Donnell CJ: **Mendelian randomization: nature's randomized trial in the post-genome era.** *JAMA.* 2009; **301**(22): 2386–8.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davey Smith G, 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, Small DS, Thompson SG: **A review of instrumental variable estimators for Mendelian randomization.** *Stat Methods Med Res.* 2017; **26**(5): 2333–55.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lawlor DA, Harbord RM, Sterne JAC, 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](#)
- 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](#)
- Holmes MV, Ala-Korpela M, Davey Smith G: **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](#)
23. Hingorani AD, Kuan V, Finan C, *et al.*: **Improving the odds of drug development success through human genomics: modelling study.** *Sci Rep.* 2019; **9**(1): 18911.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Swerdlow DI, Kuchenbaecker KB, Shah S, *et al.*: **Selecting instruments for Mendelian randomization in the wake of genome-wide association studies.** *Int J Epidemiol.* 2016; **45**(5): 1600–16.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. 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](#)
26. Schmidt AF, Finan C, Gordillo-Marañón M, *et al.*: **Genetic drug target validation using Mendelian randomization.** *bioRxiv.* 2019; 781039.  
[Publisher Full Text](#)
27. Ference BA, Ray KK, Catapano AL, *et al.*: **Mendelian Randomization Study of ACLY and Cardiovascular Disease.** *N Engl J Med.* 2019; **380**(11): 1033–42.  
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Sofat R, Hingorani AD, Smeeth L, *et al.*: **Separating the mechanism-based and off-target actions of cholesterol ester transfer protein inhibitors with CETP gene polymorphisms.** *Circulation.* 2010; **121**(1): 52–62.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Burgess S, Davey Smith G, Davies NM, *et al.*: **Guidelines for performing Mendelian randomization investigations [version 2; peer review: 2 approved].** *Wellcome Open Res.* 2019; **4**: 186.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Burgess S, Smith GD: **How humans can contribute to Mendelian randomization analyses.** *Int J Epidemiol.* 2019; **48**(3): 661–4.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. GTEx Consortium, Laboratory Data Analysis Coordinating Center Analysis Working Group, Statistical Methods Analysis Working Group, *et al.*: **Genetic effects on gene expression across human tissues.** *Nature.* 2017; **550**(7675): 204–13.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Sun BB, Maranville JC, Peters JE, *et al.*: **Genomic atlas of the human plasma proteome.** *Nature.* 2018; **558**(7708): 73–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Kettunen J, Demirkan A, Wurtz P, *et al.*: **Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA.** *Nat Commun.* 2016; **7**: 11122.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
34. Shin SY, Fauman EB, Petersen AK, *et al.*: **An atlas of genetic influences on human blood metabolites.** *Nat Genet.* 2014; **46**(6): 543–50.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Ahola-Olli AV, Wurtz P, Havulinna AS, *et al.*: **Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors.** *Am J Hum Genet.* 2017; **100**(1): 40–50.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. 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](#)
37. Yavorska OO, Burgess S: **MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data.** *Int J Epidemiol.* 2017; **46**(6): 1734–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
38. 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](#)
39. Johnson DE: **Biotherapeutics: Challenges and Opportunities for Predictive Toxicology of Monoclonal Antibodies.** *Int J Mol Sci.* 2018; **19**(11): 3685.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
40. Scott DE, Bayly AR, Abell C, *et al.*: **Small molecules, big targets: drug discovery faces the protein-protein interaction challenge.** *Nat Rev Drug Discov.* 2016; **15**(8): 533–50.  
[PubMed Abstract](#) | [Publisher Full Text](#)
41. Wishart DS, Feunang YD, Guo AC, *et al.*: **DrugBank 5.0: A major update to the DrugBank database for 2018.** *Nucleic Acids Res.* 2018; **46**(D1): D1074–D82.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
42. UniProt Consortium: **UniProt: a worldwide hub of protein knowledge.** *Nucleic Acids Res.* 2019; **47**(D1): D506–D15.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Hunt SE, McLaren W, Gil L, *et al.*: **Ensembl variation resources.** *Database (Oxford).* 2018; **2018**: bay119.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. Gill D, Georgakis MK, Koskeridis F, *et al.*: **Use of Genetic Variants Related to Antihypertensive Drugs to Inform on Efficacy and Side Effects.** *Circulation.* 2019; **140**(4): 270–9.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
45. Zheng J, Haberland V, Baird D, *et al.*: **Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases.** *bioRxiv.* 2019; 627398.  
[Publisher Full Text](#)
46. Yao DW, O'Connor LJ, Price AL, *et al.*: **Quantifying genetic effects on disease mediated by assayed gene expression levels.** *bioRxiv.* 2020; 730549.  
[Publisher Full Text](#)
47. Fishilevich S, Nudel R, Rappaport N, *et al.*: **GeneHancer: genome-wide integration of enhancers and target genes in GeneCards.** *Database (Oxford).* 2017; **2017**: bax028.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
48. Vösa U, Claringbould A, Westra HJ, *et al.*: **Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis.** *bioRxiv.* 2018; 447367.  
[Publisher Full Text](#)
49. Swerdlow DI, Holmes MV, Kuchenbaecker KB, *et al.*: **The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis.** *Lancet.* 2012; **379**(9822): 1214–24.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. 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](#)
51. Coviello AD, Haring R, Wellons M, *et al.*: **A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation.** *PLoS Genet.* 2012; **8**(7): e1002805.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
52. Burgess S, Zuber V, Valdes-Marquez E: **Mendelian randomization with fine-mapped genetic data: Choosing from large numbers of correlated instrumental variables.** *Genet Epidemiol.* 2017; **41**(8): 714–25.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. 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](#)
54. Patel A, Burgess S, Gill D, *et al.*: **Inference with many correlated weak instruments and summary statistics.** *arXiv: 2005.01765.* 2020.  
[Reference Source](#)
55. Walker VM, Kehoe PG, Martin RM, *et al.*: **Repurposing antihypertensive drugs for the prevention of Alzheimer's disease: a Mendelian randomization study.** *Int J Epidemiol.* 2019; **49**(4): 1132–1140.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
56. 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](#)
57. Scott RA, Freitag DF, Li L, *et al.*: **A genomic approach to therapeutic target validation identifies a glucose-lowering GLP1R variant protective for coronary heart disease.** *Sci Transl Med.* 2016; **8**(341): 341ra76.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
58. Gill D, Walker VM, Martin RM, *et al.*: **Comparison with randomized controlled trials as a strategy for evaluating instruments in Mendelian randomization.** *Int J Epidemiol.* 2020; **49**(4): 1404–1406.  
[PubMed Abstract](#) | [Publisher Full Text](#)
59. Berry DJ, Vimaleswaran KS, Whittaker JC, *et al.*: **Evaluation of genetic markers as instruments for Mendelian randomization studies on vitamin D.** *PLoS One.* 2012; **7**(5): e37465.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
60. Georgakis MK, Malik R, Gill DK, *et al.*: **Interleukin-6 signaling effects on ischemic stroke and other cardiovascular outcomes: a Mendelian Randomization study.** *medRxiv.* 2019; 19007682.  
[Publisher Full Text](#)
61. Wurtz P, Wang Q, Soininen P, *et al.*: **Metabolomic Profiling of Statin Use and Genetic Inhibition of HMG-CoA Reductase.** *J Am Coll Cardiol.* 2016; **67**(10): 1200–10.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
62. Slob EAW, Burgess S: **A Comparison Of Robust Mendelian Randomization Methods Using Summary Data.** *bioRxiv.* 2019; 577940.  
[Publisher Full Text](#)
63. Benn M, Nordestgaard BG, Frikke-Schmidt R, *et al.*: **Low LDL cholesterol, PCSK9 and HMGR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study.** *BMJ.* 2017; **357**: j1648.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
64. Ference BA, Majeed F, Penumetcha R, *et al.*: **Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGR, or both: a 2 x 2 factorial Mendelian randomization study.** *J Am Coll Cardiol.* 2015; **65**(15): 1552–61.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. Ference BA, Robinson JG, Brook RD, *et al.*: **Variation in PCSK9 and HMGR and Risk of Cardiovascular Disease and Diabetes.** *N Engl J Med.* 2016; **375**(22): 2144–53.  
[PubMed Abstract](#) | [Publisher Full Text](#)
66. Ference BA: **How to use Mendelian randomization to anticipate the results of randomized trials.** *Eur Heart J.* 2018; **39**(5): 360–2.  
[PubMed Abstract](#) | [Publisher Full Text](#)

67. Del Greco MF, Minelli C, Sheehan NA, *et al.*: **Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome.** *Stat Med.* 2015; **34**(21): 2926–40.  
[PubMed Abstract](#) | [Publisher Full Text](#)
68. Chong M, Sjaarda J, Pigeyre M, *et al.*: **Novel Drug Targets for Ischemic Stroke Identified Through Mendelian Randomization Analysis of the Blood Proteome.** *Circulation.* 2019; **140**(10): 819–30.  
[PubMed Abstract](#) | [Publisher Full Text](#)
69. Dudbridge F: **Power and predictive accuracy of polygenic risk scores.** *PLoS Genet.* 2013; **9**(3): e1003348.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
70. Burgess S, Thompson SG, Collaboration CCG: **Avoiding bias from weak instruments in Mendelian randomization studies.** *Int J Epidemiol.* 2011; **40**(3): 755–64.  
[PubMed Abstract](#) | [Publisher Full Text](#)
71. 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](#)
72. Allara E, Morani G, Carter P, *et al.*: **Genetic Determinants of Lipids and Cardiovascular Disease Outcomes: A Wide-Angled Mendelian Randomization Investigation.** *Circ Genom Precis Med.* 2019; **12**(12): e002711.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
73. Denny JC, Ritchie MD, Basford MA, *et al.*: **PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations.** *Bioinformatics.* 2010; **26**(9): 1205–10.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
74. Pendergrass SA, Brown-Gentry K, Dudek SM, *et al.*: **The use of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery.** *Genet Epidemiol.* 2011; **35**(5): 410–22.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
75. Schmidt AF, Holmes MV, Preiss D, *et al.*: **Phenome-wide association analysis of LDL-cholesterol lowering genetic variants in PCSK9.** *BMC Cardiovasc Disord.* 2019; **19**(1): 240.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
76. Giambartolomei C, Vukcevic D, Schadt EE, *et al.*: **Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics.** *PLoS Genet.* 2014; **10**(5): e1004383.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
77. Giambartolomei C, Liu JZ, Zhang W, *et al.*: **A Bayesian framework for multiple trait colocalization from summary association statistics.** *Bioinformatics.* 2018; **34**(15): 2538–45.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
78. Hormozdiari F, van de Bunt M, Segrè AV, *et al.*: **Colocalization of GWAS and eQTL Signals Detects Target Genes.** *Am J Hum Genet.* 2016; **99**(6): 1245–60.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
79. Zhu Z, Zhang F, Hu H, *et al.*: **Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets.** *Nat Genet.* 2016; **48**(5): 481–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
80. Lawlor DA, Tilling K, Smith GD: **Triangulation in aetiological epidemiology.** *Int J Epidemiol.* 2016; **45**(6): 1866–86.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
81. Borén J, Chapman MJ, Krauss RM, *et al.*: **Low-density lipoproteins cause atherosclerotic cardiovascular disease: pathophysiological, genetic, and therapeutic insights: a consensus statement from the European Atherosclerosis Society Consensus Panel.** *Eur Heart J.* 2020; **41**(24): 2313–2330.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
82. 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](#)
83. Bothwell LE, Greene JA, Podolsky SH, *et al.*: **Assessing the Gold Standard—Lessons from the History of RCTs.** *N Engl J Med.* 2016; **374**(22): 2175–81.  
[PubMed Abstract](#) | [Publisher Full Text](#)
84. Millwood IY, Walters RG, Mei XW, *et al.*: **Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China.** *Lancet.* 2019; **393**(10183): 1831–42.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Peer Review Status:    

---

## Version 2

Reviewer Report 01 March 2021

<https://doi.org/10.21956/wellcomeopenres.18307.r42576>

© 2021 Traylor 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.



**Matthew Traylor** 

Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK

This is a clear and timely review on the use of Mendelian Randomization for studying the effects of perturbing drug targets.

I have some minor comments for the authors' consideration:

In the discussion of 'selecting genetic variants to proxy drug targets', the authors discuss using instruments based on location at a corresponding gene. Although this approach is likely a valid one in many instances, there are exceptions where trans effects could confound associations. An example is the FTO association with BMI where variants close to the FTO gene have no impact on FTO and instead influence IRK3 / IRK5 expression via trans effects (<https://pubmed.ncbi.nlm.nih.gov/26287746/>).<sup>1</sup> I did therefore wonder if some discussion of the potential of this approach to be susceptible to failures of the relevance assumption of MR might be merited.

Similarly, I did wonder if more expanded discussion of the relative merits of the use of proteomics compared to transcriptomics, as discussed in recent work (<https://pubmed.ncbi.nlm.nih.gov/32591531/>)<sup>2</sup> could be included - as well as discussion of the importance of tissue specificity.

The section under analysis discussing colocalization approaches should be updated to reflect the latest publications, which offer more flexibility in terms of assumptions of numbers of causal variants (<https://www.biorxiv.org/content/10.1101/2021.02.23.432421v1>).<sup>3</sup>

### References

1. Claussnitzer M, Dankel SN, Kim KH, Quon G, et al.: FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N Engl J Med*. 2015; **373** (10): 895-907 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Schmidt A, Finan C, Gordillo-Marañón M, Asselbergs F, et al.: Genetic drug target validation

using Mendelian randomisation. *Nature Communications*. 2020; **11** (1). [Publisher Full Text](#)  
3. Wallace C: A more accurate method for colocalisation analysis allowing for multiple causal variants. *bioRxiv*. 2021. [Publisher Full Text](#)

**Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

**Are all factual statements correct and adequately supported by citations?**

Yes

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Statistical Genetics

**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.**

Reviewer Report 25 February 2021

<https://doi.org/10.21956/wellcomeopenres.18307.r42580>

© 2021 **Schooling C**. 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.



**C. Mary Schooling** 

School of Public Health and Health Policy, City University of New York, New York, NY, USA

Mendelian randomization (MR) is a fantastic method that has revolutionized our understanding of risk factors for many diseases, and whose further use in identifying drug targets is immensely valuable. This is a helpful guide to using Mendelian randomization (MR) to investigate drug targets but needs to address some additional relevant issues that could confuse researchers.

Key points that need to be made clearer:

1. MR does not quite imitate a randomized controlled trial (RCT). RCTs randomize to the drug or intervention after recruitment, meaning well-conducted RCTs are not open to confounding (bias typically from common causes of intervention/exposure and outcome) or selection bias (typically from selecting on exposure and outcome). MR studies usually use genetic studies recruited in mid-life well after the randomization of genetic material at

conception. As such, MR studies usually only concern survivors of their genetic make-up and also of the disease of interest, i.e., they may be selected on genetic variants and the disease of interest, meaning they are open to selection bias. In this situation estimates for the drug target on the disease of interest might be attenuated, reversed, or spurious. In addition, inadvertently selecting survivors of their genetic variants and competing risk of the disease of interest will similarly affect the estimate for the drug target on the disease of interest, even if nobody has died of the disease of interest before recruitment. Examples of this issue include the genetic variants corresponding to PCSK9 inhibitors appearing to have no effect on stroke<sup>1</sup> and the genetic variant corresponding to statins, i.e., rs12916, appearing to have no effect on stroke.<sup>2</sup> These findings conflict with RCTs, and are likely the result of selection bias, as described above.

2. Unlike RCTs, MR is best suited for investigations of the effects of drugs intended to reduce the risk of disease but is not necessarily suitable for studies of drug effects in patients. An MR study of a disease treatment drug in patients would be open to selection bias because patients are typically survivors of their genetic make-up and of the disease of interest. This point has been explained in relation to genetic studies,<sup>3</sup> and solutions are under-development for both genetic studies and MR. Researchers often want to use MR to assess effects of drugs in a patient sample just as they can in an RCT, so it needs to be explained when such an MR study would and would not be likely to give valid results.
3. Criteria for suitable samples for selecting genetic instruments for a drug need to be given. For example, should the sample be population-representative, in young people, from a trial of the relevant drug, from controls only of a case-control study?

There are also some other points that need clarification.

1. Paragraph 3, The major sources of bias in phenotypic observational studies are now understood to be confounding and selection bias, as explained here.<sup>4,5</sup> So, selection bias needs to be added to the second to last sentence of this paragraph. The last sentence of the same paragraph then needs to be amended slightly, because MR is open to selection bias, as explained above.
2. Paragraph 4, the elaboration of the exclusion restriction should explain what “horizontal pleiotropy” is because it has not been mentioned before. Please also make it clear that this assumption is violated if any open path from genetic variant to outcome exists that does not affect the exposure.<sup>6</sup> Instrument evaluation first paragraph, please similarly clarify the first sentence.
3. In the Limitations section, or elsewhere, please explain the differences between an RCT and an MR study in terms of selection bias and types of questions that can be answered by each design. Ensure Figure 3 reflects these differences.

Minor points:

1. Please avoid jargon, such as “clumping”.

## References

1. Hopewell J, Malik R, Valdés-Márquez E, Worrall B, et al.: Differential effects of PCSK9 variants on risk of coronary disease and ischaemic stroke. *European Heart Journal*. 2018; **39** (5): 354-359

[Publisher Full Text](#)

2. Schooling CM, Lopez PM, Yang Z, Zhao JV, et al.: Use of Multivariable Mendelian Randomization to Address Biases Due to Competing Risk Before Recruitment. *Front Genet.* 2020; **11**: 610852 [PubMed Abstract](#) | [Publisher Full Text](#)
3. Dudbridge F, Allen R, Sheehan N, Schmidt A, et al.: Adjustment for index event bias in genome-wide association studies of subsequent events. *Nature Communications.* 2019; **10** (1). [Publisher Full Text](#)
4. Hernán MA, Hernández-Díaz S, Robins JM: A structural approach to selection bias. *Epidemiology.* 2004; **15** (5): 615-25 [PubMed Abstract](#) | [Publisher Full Text](#)
5. Bareinboim E, Pearl J: Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences.* 2016; **113** (27): 7345-7352 [Publisher Full Text](#)
6. Pearl J: Causality: models, reasoning, and inference. *Cambridge University Press.* 2009.

**Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

**Are all factual statements correct and adequately supported by citations?**

Partly

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Partly

**Competing Interests:** No competing interests were disclosed.**Reviewer Expertise:** Epidemiological methods, drug effects, implications of the trade-off between reproduction and longevity for public health 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, however I have significant reservations, as outlined above.**

Reviewer Report 23 February 2021

<https://doi.org/10.21956/wellcomeopenres.18307.r42579>

© 2021 Au Yeung S. 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.

**Shiu Lun Ryan Au Yeung**

School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong,

Hong Kong

Review of the manuscript (Gill *et al.*, 2021)

This is a clearly written paper concerning the use of genetics to infer drug's effects on health outcomes. The paper summarises landmark studies/commentaries in this area of research and have covered the majority of the points. I have some comments for the authors' consideration.

### Major comments

- Although the use of genetics to infer effect of medication is promising, one of the challenges is the identification of a relevant variant to mimic medication use. Previous papers mainly use postulated phenotypes to identify the relevant causal variant (e.g. LDL cholesterol for HMGCR or systolic blood pressure for genetic targets of anti-hypertensives). However, the choice may not be that clear for other medications. For example, researchers may choose HbA1c over fasting glucose (PMID: 32748028)<sup>1</sup> to identify variants related to metformin because of sample size consideration and genotype coverage in the phenotype datasets (as per MAGIC database). Similarly, there may be pleiotropic effects of medication that are not mediated via the proposed phenotypes used to select the variants in Mendelian randomization and hence may not be able to capture the total effect of the medications. I think these are challenges that would be worth discussing to stimulate additional research. The use of proteomic data may help solve the abovementioned problem although proteomic datasets may not be as large and hence may be underpowered to identify relevant variants.
- Can the authors also comment on other approaches such as factorial MR design to investigate combined effects of medications (PMID: 27959767)?<sup>2</sup>
- Although Mendelian randomization studies are less vulnerable to confounding than observational studies, selection bias could still be an issue, in particular when this method is used to explore potential targets for diseases that have a late on-set year (e.g. Alzheimer's Disease).
- I think another challenge would be to investigate the effect of medication amongst patients (PMID: 28981501).<sup>3</sup>
- I am not sure if sensitivity analyses for correlated variants perform as well as uncorrelated variants. For example, a previous study commented that although MR-Egger can be adapted to correlated variants, the underlying assumption is unlikely met (PMID: 28527048).<sup>4</sup>

### Minor comments:

- Figure 3: I think it should be randomized controlled trial for RCT.
- For reference 26, I think this is now being published and can be updated (<https://www.nature.com/articles/s41467-020-16969-0>)

### References

1. Luo S, Schooling CM, Wong ICK, Au Yeung SL: Evaluating the impact of AMPK activation, a target of metformin, on risk of cardiovascular diseases and cancer in the UK Biobank: a Mendelian randomisation study. *Diabetologia*. **63** (11): 2349-2358 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Ference B, Robinson J, Brook R, Catapano A, et al.: Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *New England Journal of Medicine*. 2016; **375** (22): 2144-2153 [Publisher Full Text](#)
3. 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](#)
4. Burgess S, Thompson SG: Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol*. **32** (5): 377-389 [PubMed Abstract](#) | [Publisher Full Text](#)
5. Schmidt A, Finan C, Gordillo-Marañón M, Asselbergs F, et al.: Genetic drug target validation using Mendelian randomisation. *Nature Communications*. 2020; **11** (1). [Publisher Full Text](#)

**Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

**Are all factual statements correct and adequately supported by citations?**

Yes

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Yes

**Competing Interests:** I have previously co-authored papers with Stephen Burgess and Jie Zheng.

**Reviewer Expertise:** Cardiovascular diseases; Diabetes; 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.**

Reviewer Report 23 February 2021

<https://doi.org/10.21956/wellcomeopenres.18307.r42581>

© 2021 Chong M et al. 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.



**Pedrum Mohammadi-Shemirani** 

<sup>1</sup> Population Health Research Institute (PHRI), David Braley Cardiac, Vascular and Stroke Research Institute, Hamilton Health Sciences, Hamilton, ON, Canada

<sup>2</sup>

Thrombosis and Atherosclerosis Research Institute, Hamilton, ON, Canada

<sup>3</sup> Department of Pathology and Molecular Medicine, Michael G. DeGroot School of Medicine, McMaster University, Hamilton, ON, Canada

**Michael Chong**

<sup>1</sup> Population Health Research Institute (PHRI), David Braley Cardiac, Vascular and Stroke Research Institute, Hamilton Health Sciences, Hamilton, ON, Canada

<sup>2</sup> Thrombosis and Atherosclerosis Research Institute, Hamilton, ON, Canada

<sup>3</sup> Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON, Canada

<sup>4</sup> Department of Pathology and Molecular Medicine, Michael G. DeGroot School of Medicine, McMaster University, Hamilton, ON, Canada

Firstly, I thank the Wellcome Open Research editorial team for the opportunity to provide feedback on the review by Gill *et al.* (2021), entitled “Mendelian randomization for studying the effects of perturbing drug targets”. I have reviewed this work in detail in concert with Pedrum Mohammadi-Shemirani and the subsequent comments reflect our joint assessment. Gill *et al.* reviewed an increasingly popular application of Mendelian Randomization analyses to understand causal on-target effects of drug targets on disease risk. Current practices, considerations, and limitations of this approach were thoroughly summarized and thus we only provide minor comments. Firstly, while Figure 1 is a useful illustration, the last panel may be misinterpreted as decreasing LDL is associated with a greater risk of CAD. As such, Figure 1 should be amended for clarity. Secondly, as authors correctly acknowledge, restricting instruments to *cis* variation is the standard practice to mitigate the potential for genetic confounding, but it should also be noted that this still leaves open the possibility for genetic confounding, particularly in regions containing multiple protein-coding genes. Multivariable MR conditioning for the effects of the genetic instruments on the surrounding gene products may represent a way to assess whether findings may be confounded by proximal gene products when such data is available, an approach implemented by Porcu *et al.* (2019) (<https://www.nature.com/articles/s41467-019-10936-0>).<sup>1</sup> Similarly, in the section titled “Selecting genetic variants to proxy drug targets”, a qualifying statement may be warranted regarding the selection of genetic variants based on association with known indications. If pQTL data is available, this approach may better serve instrument validation while a selection of variants is supported by evidence of association with protein expression or function. If no such data is available, evidence of association with known indications can be used, but it warrants a cautious approach with careful consideration of the *cis* window and other protein-coding genes therein. Thirdly, authors should further comment on the careful use of trans-instruments to approximate drug targets as was executed by Georgakis *et al.* (2019) <https://pubmed.ncbi.nlm.nih.gov/30586705/>.<sup>2</sup> Fourth, we appreciate the authors’ decision to include example resources and databases to guide instrument selection (e.g., Ensembl, DrugBank, etc.), particularly for novice readers. Would the authors be able to provide similar suggested resources for other sections? For instance, databases to identify promoter and enhancer regions (e.g., GeneCards), or tools to help investigate violations of the independence and exclusion restriction assumptions (e.g., PhenoScanner). Lastly, we believe the following statement illustrates a very important point: “Moreover, comparing instruments between different MR studies of the same exposure can provide additional evidence. For example, both Gill *et al.* and Walker *et al.* independently derived instruments for antihypertensive drug effects that perform comparably when tested against a common outcome<sup>44,55</sup> Although different MR studies may use similar or overlapping data sources, different instrument selection approaches can make analyses vulnerable to distinct biases and so also

*have a role in triangulation of evidence.*" Perhaps this should be briefly commented on in the context of technological biases between different proteomics and metabolomic technologies (for example <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6277005/>)<sup>3</sup> thus further emphasizing the importance of ensuring the robustness of MR results using alternative instruments. As proteome-wide pQTL data becomes increasingly available to the public, such sensitivity analyses should become more commonplace. Overall, this review is exquisitely written, thoughtful, and appropriately referenced. Accordingly, we have no major concerns and only offer the aforementioned minor comments.

### References

1. eQTLGen Consortium, BIOS Consortium, Porcu E, Rüeger S, et al.: Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. *Nature Communications*. 2019; **10** (1). [Publisher Full Text](#)
2. Georgakis M, Gill D, Rannikmäe K, Traylor M, et al.: Genetically Determined Levels of Circulating Cytokines and Risk of Stroke. *Circulation*. 2019; **139** (2): 256-268 [Publisher Full Text](#)
3. Joshi A, Mayr M: In Aptamers They Trust. *Circulation*. 2018; **138** (22): 2482-2485 [Publisher Full Text](#)

**Is the topic of the review discussed comprehensively in the context of the current literature?**

Yes

**Are all factual statements correct and adequately supported by citations?**

Yes

**Is the review written in accessible language?**

Yes

**Are the conclusions drawn appropriate in the context of the current research literature?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Human genetics, genetic epidemiology

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

---