

REVIEW

Open Access



Common pitfalls in drug target Mendelian randomization and how to avoid them

Dipender Gill^{1,2*}, Marie-Joe Dib³, H el ene T. Cronj e^{1,4}, Ville Karhunen^{1,4}, Benjamin Woolf^{4,5,6}, Eloi Gagnon⁷, Iyas Daghlas⁸, Michael Nyberg⁹, Donald Drakeman^{10,11} and Stephen Burgess^{1,4,12}

Abstract

Background Drug target Mendelian randomization describes the use of genetic variants as instrumental variables for studying the effects of pharmacological agents. The paradigm can be used to inform on all aspects of drug development and has become increasingly popular over the last decade, particularly given the time- and cost-efficiency with which it can be performed even before commencing clinical studies.

Main body In this review, we describe the recent emergence of drug target Mendelian randomization, its common pitfalls, how best to address them, as well as potential future directions. Throughout, we offer advice based on our experiences on how to approach these types of studies, which we hope will be useful for both practitioners and those translating the findings from such work.

Conclusions Drug target Mendelian randomization is nuanced and requires a combination of biological, statistical, genetic, epidemiological, clinical, and pharmaceutical expertise to be utilized to its full potential. Unfortunately, these skillsets are relatively infrequently combined in any given study.

Keywords Mendelian randomization, Drug target, Pharmacology, Drug development

*Correspondence:

Dipender Gill
dipender.gill@sequoiagenetics.com

¹ Sequoia Genetics, London, UK

² Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, 90 Wood Lane, London, W12 0BZ, UK

³ Cardiovascular Division, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA

⁴ Medical Research Council Biostatistics Unit, University of Cambridge, Cambridge, UK

⁵ School of Psychological Science, University of Bristol, Bristol, UK

⁶ Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK

⁷ Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie de Qu ebec, Laval University, Qu ebec, Canada

⁸ Department of Neurology, University of California San Francisco, San Francisco, CA, USA

⁹ Cardiovascular Biology, Global Drug Discovery, Novo Nordisk A/S, Maaloev, Denmark

¹⁰ University of Cambridge Centre for Health Leadership & Enterprise, Judge Business School, Trumpington Street, Cambridge, UK

¹¹ Advent Venture Partners, London, UK

¹² Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK



  The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Drug development has a notoriously high failure rate, with less than 10% of drugs entering clinical study eventually being approved for use in patients [1]. This low success rate may be at least partly attributed to most of the preclinical evidence for drug effects coming from animal studies and traditional epidemiological associations, which are respectively limited in their translatability to humans and their ability to draw causal inferences [2].

Mendelian randomization is a method that uses genetic variants as instrumental variables for studying the causal effect of an exposure on an outcome [3]. These variants should relate to the exposure of interest, and to the outcome under investigation only through the exposure, and not through alternative pleiotropic pathways [3, 4]. The Mendelian randomization paradigm helps overcome the limitations of traditional epidemiology because it uses human genetic data to infer the causal effects of drug target perturbation in humans [5, 6]. Given that most drug targets are proteins, and that proteins are coded for by genes [3], investigation of drug target effects is particularly amenable to study through genetics-based causal inference methods such as Mendelian randomization [6]. In fact, this approach is increasingly being used to study

the potential effect of pharmacological perturbation of drug targets on clinical outcomes for the purposes of informing drug development strategies, a practice known as ‘drug target Mendelian randomization’ (Fig. 1) [1, 6]. Interestingly, drug targets with human genetic evidence have been shown to be at least twice as likely to make it through clinical development [7, 8]. Given that the average new drug requires more than 10 years and 1 billion US dollars to obtain regulatory approval [9], insights from Mendelian randomization studies have the potential to tremendously improve the efficiency of generating effective treatments to prevent and treat disease.

Judicious use of human genetic data has the potential to inform on various critical aspects of drug development, including on-target efficacy, safety, repurposing, biomarker selection, effect heterogeneity, and interactions [10]. Given the potential impact on informing drug development efforts, and in conjunction with increasing data availability, it is not surprising that there has been an explosion in the number of drug target Mendelian randomization analyses performed over recent years, and that most large pharmaceutical companies now incorporate human genetic evidence in their drug development pipelines [11]. However, the increasing number of drug target

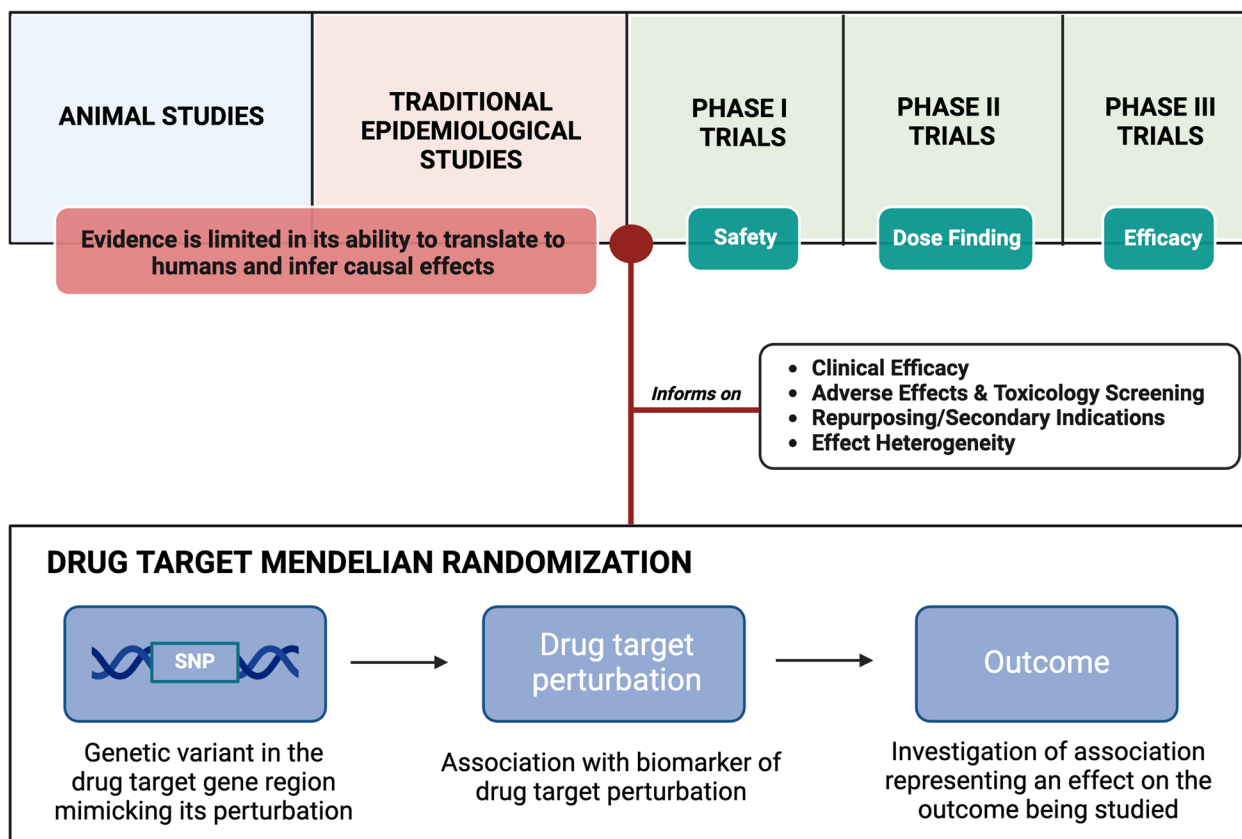


Fig. 1 Drug target Mendelian randomization within the context of drug development

Mendelian randomization analyses has, in some cases, not been accompanied by a corresponding improvement in quality, resulting in increasing uncertainty around their utility and translatability.

The purpose of this review is to share insights gained from over a decade of experience in drug target Mendelian randomization, through which we have actively contributed by publishing numerous methods-focused and applied research papers. The intended audience for this review includes experts in the field of Mendelian randomization who are interested in effectively applying methods for maximal clinical impact, and individuals in the drug development field who are in the pursuit of leveraging genetics-based insights. Given the relatively recent emergence of this paradigm, we are all on a steep learning curve, and it is therefore important for us to share our knowledge along the way. This review is not meant to be a prescriptive set of guidelines [12, 13], a repetition of the general principles of Mendelian randomization [14, 15] and drug target Mendelian randomization [1, 6, 16], or a statistical handbook on performing these analyses [17, 18], as these topics have already been extensively covered previously. The aim of this review is rather to share our learnings effectively, describing what we have found to be the common pitfalls in drug target Mendelian randomization, along with strategies to help avoid them. We acknowledge that we have previously published studies that directly contradict some of our current recommendations [19, 20]. However, we consider this to be a hallmark of science—a process that involves the intricate refining ideas and concepts through a continuous cycle of learning and improvement. We hope to utilize our insights in this review to advance the applications of drug target Mendelian randomization within the scientific community.

We divide the aforementioned pitfalls into sections related to (i) selecting exposures specific to the research question, (ii) identifying biologically plausible instruments for the exposure under study, (iii) utilizing appropriate outcome genetic association data, and (iv) defending against possible false negative and false positive results. We then offer some advice on general considerations when contextualizing a drug target Mendelian randomization study, before providing our conclusions on the current state of the field, its gaps, and future directions.

Selecting exposures specific to the research question

Misspecification of the exposure relevant to the research question is perhaps the most fundamental cause of a drug target Mendelian randomization analysis giving misleading results. A requisite for success in defining the exposure is an understanding of the mechanism of action for

the drug target under investigation. This can then enable identification of appropriate biomarkers for the exposure, to thus help identify suitable instruments.

Drug targets with multiple mechanisms

Drugs exerting their effects via multiple mechanisms are not necessarily amenable to drug target Mendelian randomization, or require thorough considerations of all the pathways they involve. To illustrate this, one Mendelian randomization study claimed to be investigating the effect of metformin, a commonly used anti-diabetic drug [21]. However, the actual analysis performed was concerned with the effects of genetically predicted growth differentiation factor 15 levels, whose altered circulating levels represent only one possible mechanism by which metformin exerts its effects [22, 23]. Consideration of the various targets through which metformin exerts its effects may have yielded contrasting results that are more reflective of its various mechanisms. Furthermore, in instances where the drug target is not known, as is the case for even some drugs routinely used in clinical practice [24], drug target Mendelian randomization will not be viable.

Drug use versus drug target effects as exposures

Another example of this is consideration of drug use (e.g. electronic health care drug prescription records or self-reported medication use), rather than the mechanisms through which the drug target elicits its clinical effects (e.g. reduction in systolic blood pressure through calcium channel blockade), as the exposure of interest. Genetic predictors of drug use are more likely to relate to the underlying condition that the drug is used to treat, socioeconomic factors related to medication access, or to generic determinants of medication adherence, rather than the effects of the drug target [25, 26]. In contrast, genetic variants predicting drug target effects would be expected to specifically mimic pharmacological perturbation of the protein target, rather than factors that reflect its use. This discrepancy has been studied in detail for caffeine, where genetic variants predicting the effects of plasma caffeine levels yield different Mendelian randomization estimates compared to those predicting its consumption [27]. These examples reinforce the importance of asking the right research question, and carefully selecting relevant exposure data.

Multi-protein complex drug targets

Situations where the drug target is made up of multiple proteins can become particularly challenging. For example, antihypertensives of the calcium channel blocker class target several protein subunits that collectively form the calcium channel. Mendelian randomization analyses

of this drug target have pooled variants in the genes coding for any of the protein subunits to generate instruments for drug target perturbation [28–30]. However, this does not account for unequal contributions of the different protein subunits to the target's overall effects, nor potential interactions between them.

Non-protein drug targets

There is even greater complexity when the therapeutic target is not a protein. For example, some Mendelian randomization investigations into the effect of increasing blood metabolite levels have considered genetic variants from the entire genome as instruments [31]. These will likely represent heterogeneous mechanisms, with potentially distinct effects. Interpreting which of these mechanisms or pathways are most representative of the therapeutic intervention being considered will require biological insight into how the different genetic variants may be exerting their effects, in relation to the intervention under study.

Drug targets with no valid genetic instruments

Not all drug targets, nor therapeutic interventions, may be amenable to study using Mendelian randomization. In our experience, genetic instruments are identifiable for the majority, but not all drug targets. This is broadly consistent with the estimated proportion of two thirds of approved drugs that have genetic support [32]. This is likely partially attributable to the lack of available genetic proxies for the remaining third of drug targets. Examples of drug classes that have recently been approved but have not been robustly instrumented in Mendelian randomization are sodium-glucose co-transporter 2 inhibitors and interleukin-17 inhibitors. Drug target Mendelian randomization can only be used to study host protein drug targets, and so is not applicable to the study of pharmacological agents directly targeting microorganisms. In this way, drug target Mendelian randomization successfully predicted immunomodulatory drug targets that would be efficacious in severe covid-19 [33], but not successful antiviral agents.

Long versus short-acting pharmacological perturbation

Genetic variants typically predict lifelong changes in drug targets, and so tend to resemble the effect of long-term pharmacological perturbations. This is an important consideration in scenarios where the effects of long-term pharmacological perturbation are different to those with short-acting mechanisms. An example of this is glucose-dependent insulinotropic polypeptide receptor (GIPR) signalling, where in some tissues the effects of long-term, in contrast to short-term, agonism are believed to mimic antagonistic effects via receptor desensitization [34].

Biological considerations when selecting instrument variants

Our approach for undertaking a drug target Mendelian randomization study almost always begins with a detailed interrogation of the known biology of the target being investigated. In this way, it is possible to optimize instrument selection by exploring the possible data sources that may be relevant to use. We recommend some considerations in this section.

Acknowledging the limitations of expression and protein quantitative trait loci

Instruments in Mendelian randomization analyses should associate with the exposure and only associate with the outcome through the exposure and not some alternative pleiotropic pathway [35, 36]. In drug target Mendelian randomization, the exposure is the perturbation of a drug target. There are few genetic association data sets available that directly measure drug target perturbation, so it is necessary to use surrogates or 'proxies' for this. It has become common practice to use genetic variants associated with gene expression or protein abundance as such proxies. These are called expression quantitative trait loci (eQTLs) for gene expression and protein quantitative trait loci (pQTLs) for protein expression. The advantages of using these traits to select instruments are that such data are widely available [37–41], and this approach requires little biological insight to be blindly applied at scale [42–44]. Superficially, this strategy also makes sense, because variants associating with altered levels of the drug target would be expected to serve as instruments for its perturbation. Unfortunately, the reality is quite different for several critical reasons.

Firstly, genetic variants predicting gene expression typically relate to transcription of a gene region and affect expression levels of several genes simultaneously [45], thereby introducing inherent pleiotropy that risks biasing Mendelian randomization analyses. This issue was likely why a drug target Mendelian randomization analysis of fibroblast growth factor 21 gave misleading results [46], with unintentional flipping of the direction of association [47]. Secondly, such genetic predictors of gene and protein expression can be tissue specific [48]. Indeed, this heterogeneity in the genetic predictors of gene and protein expression across different tissues has allowed for multivariable drug target Mendelian randomization analyses that investigate the specific tissue in which a target is likely exerting its clinical effects [49]. However, it also means that using genetic association estimates from mis-specified tissues will yield potentially misleading results [50]. This is challenging considering that tissues driving disease associations can be unknown, especially when dealing with circulating factors. Thirdly, genetic

predictors of gene and protein expression may vary in different physiological states [51], and data related to the specific context under investigation are scarcely available. Fourthly, there are issues around the specificity and sensitivity of assays used to measure gene expression and relative protein abundance which can limit the validity of any consequently identified genetic associations [52].

For these reasons, we recommend some degree of biological validation of instruments when they are selected using gene expression or proteomic data. A good example is the study of angiotensin converting enzyme (ACE) inhibition, whereby instruments were selected and validated by their associations with lower ACE gene expression and plasma ACE protein levels, as well as with lower blood pressure, a known clinical effect of ACE inhibition [53]. Similarly, in a study of phosphodiesterase type 5 inhibition, instruments were selected and validated using their association with lower gene expression, as well as lower blood pressure, and lower risk of erectile dysfunction and pulmonary hypertension, consistent with the known clinical effects of perturbing this drug target [54].

The use of clinical traits, biomarkers, and endophenotypes related to established or biologically expected drug target effects for identifying instruments offers other critical advantages over the use of gene expression or proteomic data. Specifically, there is greater potential for the link between a variant in the protein coding gene and the mechanism of action relevant to the clinical outcome being lost or confounded by pleiotropic associations when the biological distance from the variant to the trait used to select instruments is shorter (Table 1). This risk is reduced by identifying instrument variants using clinically relevant traits that are further along the putative causal pathway, and hence more proximal to the outcome

under investigation. As such, clinical traits related to the drug target should generally be preferred for selecting instrument variants over molecular traits. A challenge to this approach is that it requires biological understanding of the drug target and is therefore more difficult to scale in agnostic hypothesis-free analyses.

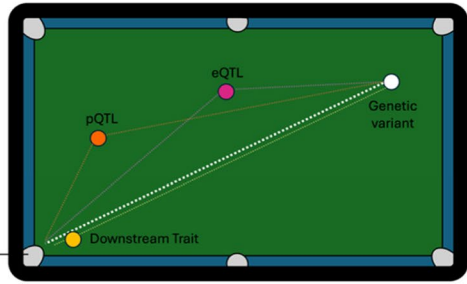
Considerations for protein function altering genetic variants

Other strategies for selecting instruments in drug target Mendelian randomization include variants that have effects on gene or protein function, rather than levels. Such instruments may be described as ‘loss-of-function’ or ‘gain-of-function’ variants. In practice, such terminology is potentially misleading, as protein function is scarcely describable on a binary or linear scale and is more commonly multi-dimensional. For example, for a protein that serves both as a cell surface receptor and circulating soluble receptor (e.g. interleukin-6 receptor), its reduced binding to the cell surface may increase some of its activities while reducing others. An example of a genetic instrument that exerts functional effects is a missense variant in the *GIPR* gene, which has been used to mimic long-term functional antagonism [55–57], although its biological effects are somewhat more complicated than this [34]. As with eQTLs and pQTLs, we similarly advocate validation of instrument selection using functionally annotated variants with biologically informed positive control associations. For the *GIPR* example above, the described missense variant used as an instrument for inhibition also strongly associates with lower body weight, consistent with what has been demonstrated in clinical trials [58]. Overall, it is important that as much evidence is obtained as possible to support the validity of the instruments employed, as without valid

Table 1 Billiard analogy

Billiard analogy: molecular vs clinical traits for instrument selection. There is likely greater potential for the link between a variant in the protein coding gene and the mechanism of action relevant to the clinical outcome being lost or confounded by pleiotropic associations when the biological distance from the variant to the trait used to select instrument variants is shorter. This is because many biological mechanisms can affect gene expression and protein levels. But, if the downstream trait is carefully chosen, only those causal pathways that act via the downstream trait are likely to affect the outcome. By selecting variants that associate with the downstream trait, we control a larger portion of the causal pathway from variant to outcome; the portion of the causal pathway that is unknown is shorter.

This is analogous to considerations in the game of billiards (or snooker or pool). The distance from the cue ball (genetic variant) to pocket (outcome) is fixed. It is typically easier to aim for a ball that is further away from the cue ball and hence closer to the pocket (downstream trait), rather than to aim for a ball that is close to the cue ball and further from the pocket (eQTL or pQTL). When selecting or validating a genetic variant as mimicking a druggable mechanism, it is typically more reliable to select a surrogate trait that is further downstream in the causal pathway.



instruments, the whole Mendelian randomization study will be misleading.

Incorporation of biological knowledge

In conclusion for this section, incorporating biological insights can be transformative for selecting appropriate and valid instruments. Even when instruments are derived from eQTL or pQTL data, pre-acquired biological insight would enable the selection of the appropriate tissue to consider [38]. Being aware of all options also allows for the most relevant data with greatest statistical power to be leveraged, along with as many traits as possible, to maximize interrogation of instrument validity. For example, drug target Mendelian randomization analyses of interleukin-6 receptor signalling leveraged genetic variants robustly associated with several biomarkers known from clinical practice to be affected by perturbation of this drug target [59]. While this instrument has also been validated through positive control associations with disease outcomes known to be affected by inhibition of this target [60], such a rigorous approach may not be feasible with novel drug targets.

Having identified genetic instruments to study pharmacological perturbation of the drug target under consideration, the next step is a thorough understanding of the outcome of interest. This will consequently allow for the relevant outcome genetic association data to be selected or compiled, along with the most appropriate statistical analysis plan. Every drug target is unique, and every research question is different, so such considerations cannot be easily automated to cater for all scenarios.

Outcome definition

As with all Mendelian randomization studies, it is important for the outcome under study to be suitably defined. In some scenarios, there may not be appropriate genetic association data available for the desired outcome, or at the desired level of phenotypic granularity. Therefore, it may be necessary to resort to the next best data source available. For example, heart failure is a heterogeneous disease with many distinct aetiological mechanisms, which vary in their treatment [61]. However, the most utilized genetic association data only consider heart failure as one entity [62]. The situation is similar for studies of chronic kidney disease [63]. Other common challenges in outcome definitions relate to misclassification, both in self-reported and clinician-ascribed diagnoses [64]. Similarly, studies of outcomes that are more common in older individuals may be vulnerable to survival bias [65]. Overall, high-quality curated data is essential for the success of drug target Mendelian randomization studies.

A fundamental distinction in outcome definitions separates risk of disease incidence and disease progression. For example, glucagon-like peptide 1 agonists are licensed for the treatment of type 2 diabetes mellitus; however, genetic association data for type 2 diabetes mellitus risk specifically relate to the risk of developing the disease [66]. In this instance, the mechanisms predisposing to type 2 diabetes overlap with those affecting its progression, so the distinction is not so critical, and genetic association data for type 2 diabetes risk generally correlate well with those for glycaemic control [67]. However, in scenarios where the genetic predictors of disease risk are quite different to those of disease progression, the implications can be more pronounced. For example, ischemic stroke typically arises from an obstruction of blood flow to the brain, most commonly due to a blood clot. In contrast, progression of disability after stroke is more related to neurological and inflammatory mechanisms affecting brain injury and recovery, as well as social factors related to engagement with rehabilitation. As such, genetic association data for stroke risk cannot be used to study drug targets for stroke recovery, and vice versa. Consistent with this, there is little correlation between the genetic predictors of ischemic stroke risk and recovery [68].

Statistical power considerations and interpretation of associations

Statistical power considerations

It is not uncommon for statistical power calculations to be expected when planning a drug target Mendelian randomization study. However, this would require an estimate of the proportion of variance of the drug target effect that is explained by the genetic instrument [12]. With perhaps the exception of homozygote ‘complete-loss-of-function’ mutation carriers, we are not aware of any consistent means of reliably estimating the proportion of variance in drug target effects explained by genetic variants. As such, it seems implausible to perform reliable statistical power calculations for most drug target Mendelian randomization scenarios. Further, even if such calculations could be performed, it is not clear how they would translate clinically, as the effect of any drug compound or therapeutic intervention in clinical practice would be a function of the pharmacological properties of that compound, including both its pharmacodynamic and pharmacokinetic characteristics.

Instead, to gauge the relative power of drug target Mendelian randomization analyses, it may be more appropriate to compare estimates against those obtained from similar analyses of related traits. For example, in

a drug target Mendelian randomization study into the potential repurposing opportunity and adverse effects of commonly prescribed antihypertensive drug classes, comparisons were made to the Mendelian randomization analyses of genetically predicted lower blood pressure by any mechanism after sampling the same number of variants [69]. An analogous approach was taken in a study comparing effects of the two C-type natriuretic peptide receptors [70].

When interpreting the results of a drug target Mendelian randomization analysis, as with all analyses, it is important to consider that the absence of any strong associations may also be attributed to low statistical power. This may yield false negative findings. For example, there is strong genetic evidence implicating both lipoprotein lipase (LPL) and angiotensin-like protein 4 (ANGPTL4) as therapeutic targets for lipid-lowering and cardiovascular disease reduction [71]. However, such evidence implicating angiotensin-like protein 3 (ANGPTL3) as a target for cardiovascular disease is more limited when considering common genetic variants [72]. Given that this evidence is stronger when considering rarer, functionally relevant variants in *ANGPTL3* that have a greater magnitude of effect, we speculate that this discrepancy may be attributable to low statistical power in some of the studies [73]. Consistent with this, *ANGPTL3* is expressed in the liver and then circulates through the plasma to its relevant sites of action, while *LPL* and *ANGPTL4* are expressed more widely [74]. Thus, the 'biological distance' between *ANGPTL3* and cardiovascular disease may be greater than for *LPL* and *ANGPTL4*, suggesting that statistical power may also be lower, even if the relative efficacy of these drug targets might be similar in clinical practice. This claim holds even if we do not measure *ANGPTL3* expression in the liver, as what is important is the distance between the mechanism of interest and the outcome, regardless of which biomarker is used to measure the mechanism.

Making sense of false positive findings

False positive findings are also rife in drug target Mendelian randomization. With the explosion in data availability and the increasing ease by which they may be conducted [75], an unprecedented number of analyses are being performed daily. The pressure to publish for individuals working in academia means that scientists will be eager to share any potentially noteworthy associations that they identify, including at the cost of bypassing appropriate levels of scrutiny, replication, or validation. Such studies may be 'false positives' because of publication bias, as well as selective sharing of significant findings without appropriate transparency or correction for multiple testing.

Other common causes for false positive findings in drug target Mendelian randomization relate to genetic confounding through variants in linkage disequilibrium. Such possibilities can be explored using statistical colocalization methods as follow-up analyses for drug target Mendelian randomization, as previously described [76]. For example, a previous study used colocalization to provide evidence that circulating GIP levels and cardiovascular disease risk are driven by distinct causal variants in the *GIPR* gene, and so any identified Mendelian randomization associations supporting an effect of GIP levels on cardiovascular disease risk at this locus are likely attributable to genetic confounding through this correlated variant [77].

Even for the most comprehensive drug target Mendelian randomization analysis, it is still critical to triangulate the evidence with all other available forms of evidence when drawing conclusions to inform drug development efforts. For example, drug target Mendelian randomization analyses have consistently generated evidence to support that cholesteryl ester transfer protein inhibition will be an efficacious strategy for reducing cardiovascular disease risk [78], yet until recently, clinical trial efforts in this area have been unsuccessful [79]. This illustrates that even if there is genetic evidence that a target is likely to be efficacious, generating a drug that achieves this result is still dependent on its pharmacological properties, the population that is treated, the timing of treatment, and the specific outcome that is studied.

Human genetic insights in the context of preclinical evidence

In conjunction with genetic evidence, a series of preclinical validation studies are needed to provide mechanistic insight into how a novel target is linked to disease. These studies are crucial for improving confidence in a target playing a causal role but are also key for highlighting how the target should be pharmaceutically modulated. As an example, a protein may have several functions that relate to different structural elements and thus it may be pivotal to understand which epitope needs to be modulated to achieve the desired alteration of protein function.

Preclinical experiments should preferably be performed in assays comprised of human cells displaying the relevant disease phenotype. Recapitulating such phenotypes has typically been difficult. However, with the emergence of single-cell transcriptomic and proteomic profiling that includes sufficient spatial resolution, it is possible to map out the molecular fingerprint of spatially distributed cellular populations in disease tissue [80]. Here, it is important to underscore that recapitulating such a phenotype *in vitro* is no simple feat. This is exemplified in the case for stem cell-derived cardiomyocytes,

where it has proven difficult to induce a phenotype that fully resembles that of an adult myocyte [81]. When such challenges are overcome, it is possible to explore the effects of perturbing a molecular target and thereby provide crucial insights into how this may hold a therapeutic potential in a given cell type and tissue relevant for human disease.

There may also be targets that cannot be validated in single- or even multicellular systems as the mechanism in scope can only be captured in a model system allowing for integrative responses between numerous tissues and organs. To avoid that a potential human relevant therapeutic effect is lost in translation, animal models can be used if the mechanism identified is also operating in the selected model that recapitulates key features of human disease. Such model understanding requires intricate characterization of the model and thus comes with a substantial investment in time and resources, which to some extent may explain why inadequate models are often chosen. Failure to choose an animal model system that sufficiently mimics relevant human pathophysiology and the molecular mechanism in scope is commonly seen and is a key cause of low human translatability of findings as evidenced by efforts made within the area of atherosclerosis research [82].

It is important to highlight that the information obtained from the above-described preclinical studies should be contextualized with other layers of human data. Hence, findings from preclinical experiments should be used to define and refine genetic analyses and vice versa. This iterative process will enhance the probability of success of the drug discovery program by providing the basis for asking the relevant research questions.

Future directions

With the increasing popularity of drug target Mendelian randomization, there is also paralleled potential to optimize the quality of studies performed and the resultant insights that can be gained from them. Paradoxically, the growth in the number of published analyses has apparently not yet translated to directly informing drug target discovery and development efforts [83]. However, we acknowledge that the actual impact of genetics in this regard may be masked by a need for pharmaceutical and biotechnology companies to protect commercially sensitive findings.

Although the approach is potentially impactful, it should always be interpreted within the context of its limitations. Mendelian randomization analyses typically consider the lifelong effects of small changes arising from genetic variation, which is rarely the same thing as a discrete clinical intervention observed at a specific timepoint in the life course [6]. This is on top of the other potential limitations encountered with

the method, including absence of reliable instruments, bias from genetic confounding through variants in linkage disequilibrium, and unavailability of relevant genetic association data. We emphasize that drug target Mendelian randomization should by no means be considered a substitute for comprehensive preclinical safety and toxicology assessment, or for thorough evaluation through clinical trials. Instead, it should be used in conjunction with other sources of evidence to help prioritize and inform clinical development efforts.

Historically, many drug target Mendelian randomization efforts have focused on cardiometabolic and anti-inflammatory targets. This is likely attributable to the amenability of these targets to such analyses, as well as their relevance to clinical translation efforts. In contrast, this has left gaps in other therapeutic areas, particularly related to psychiatric and cancer outcomes. For psychiatric outcomes, this may be explained by the complexity of the central nervous system and related challenges in identifying valid instruments and designing translatable drug target Mendelian randomization analyses. For cancer outcomes, most malignancies arise due to somatic mutations, whereas Mendelian randomization is typically concerned with the effects of germline genetic variation. Additionally, most genetic association data for cancer and psychiatric outcomes relate to disease incidence rather than progression, whereas most of their therapeutics are focused on disease treatment rather than prevention. For these reasons, further efforts are required to understand the degree to which pharmacological targets for prevention are also applicable to disease treatment, and vice versa.

With the continued access to larger, more diverse, and higher-quality genetic association and epidemiological data, in conjunction with advancing statistical methods and the emergence of artificial intelligence, there has never been a more exciting time for the field of drug target Mendelian randomization. While artificial intelligence is yet to provide much tangible impact in drug target Mendelian randomization, it has offered some promise for functional annotation of genetic variants that may serve as instruments [84], to help inform on their mechanism of effect and suitability for any given study. Given the dramatic growth in our understanding of applications of artificial intelligence, it seems inevitable that more uses will emerge over the next few years. However, we believe that automation of drug target Mendelian randomization is still a while away. The nuanced interplay of biology, statistics, genetics, epidemiology, clinical translation, and drug development essential for successful implementation of the paradigm likely means that specialist and experienced skillsets will continue to be required, and such a combination is rarely amenable to automation.

EXAMPLE CONSIDERATIONS IN DRUG TARGET MENDELIAN RANDOMIZATION

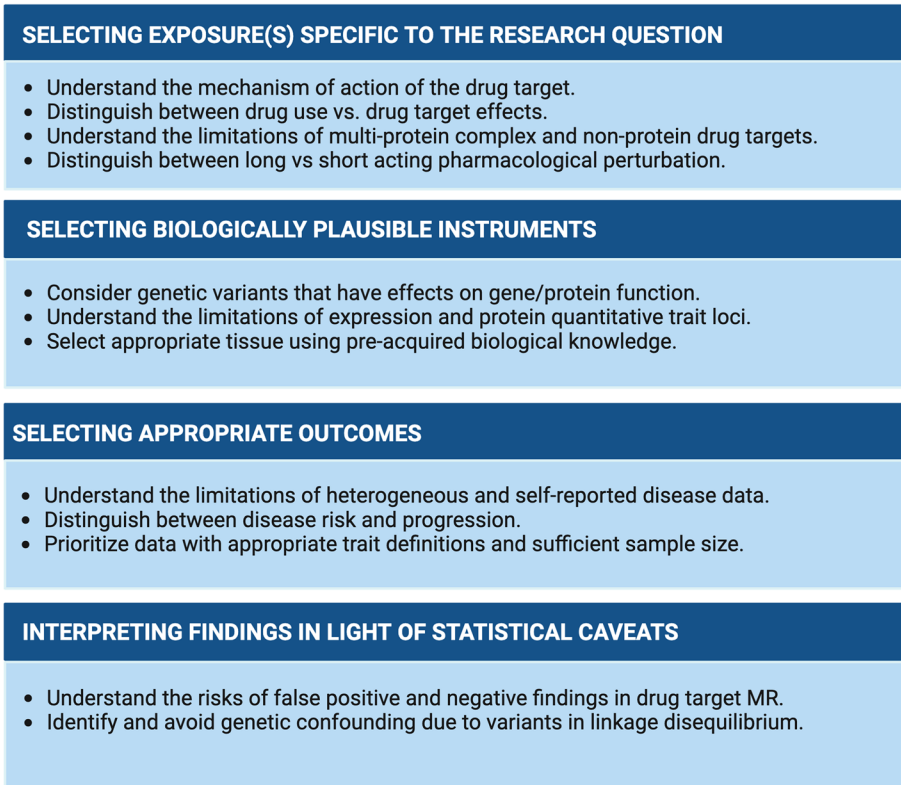


Fig. 2 Some key considerations in drug target Mendelian randomization

Conclusions

We believe that it is through ongoing study of a greater diversity of pharmacological targets, therapeutic areas, and data sources that the field of drug target Mendelian randomization will continue to develop. Continual sharing of insights and learnings through scientific publications and support to methodologists and practitioners will enable the field to move forward efficiently (Fig. 2), bringing more successful drugs to market. Close collaboration with those that directly invest in and undertake drug development is also essential, to ensure that the findings can have maximum impact. All of this will serve individuals who stand to benefit from more effective medicines.

Abbreviations

ACE	Angiotensin-converting enzyme
ANGPTL3	Angiopoietin-like protein 3
ANGPTL4	Angiopoietin-like protein 4
eQTL	Expression quantitative trait locus
GIPR	Glucose-dependent insulinotropic polypeptide receptor
LPL	Lipoprotein lipase
pQTL	Protein quantitative trait locus

Acknowledgements

We thank all those that have contributed to the field of drug target Mendelian randomization.

Authors' contributions

DG, MJD, HTC, MN and SB drafted the main manuscript. VK, BW, EG, ID and DD critically revised the manuscript for intellectual content. All authors read and approved the final manuscript.

Authors' Twitter handles

Dipender Gill: @dpsg108; Marie-Joe Dib: @mariejoedib; H el ene T. Cronj e: @htcronje; Ville Karhunen: @karhunen_v; Benjamin Woolf: @BarWoolf; Eloi Gagnon: @egagnon_phd; Iyas Daghlas: @iyas_daghlas; Stephen Burgess: @stevesphd.

Funding

EG is supported by a doctoral scholarship awarded by the Fonds du Recherche du Qu ebec Sant e (FRQS). HTC and VK are supported by the Wellcome Trust (225790/Z/22/Z) and the United Kingdom Research and Innovation Medical Research Council (MC_UU_00040/01). SB is supported by the Wellcome Trust (225790/Z/22/Z) and the United Kingdom Research and Innovation Medical Research Council (MC_UU_00002/7, MC_UU_00040/01). This research was supported by the National Institute for Health Research Cambridge Biomedical Research Centre (NIHR203312).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

No ethical approval was required for this work. No consent to participate was required for this work.

Consent for publication

No consent for publication was required for this work.

Competing interests

Sequoia Genetics is a private limited company that works with investors, pharma and biotech companies by performing research that leverages genetic data to help inform drug discovery and development. DG is the Founder and Director of Sequoia Genetics. HTC, VK and SB are employed by Sequoia Genetics. MN is employed by Novo Nordisk. DG and DD have financial interests in several biotechnology companies. The remaining coauthors have no conflicts of interest to disclose.

Received: 1 August 2024 Accepted: 10 October 2024

Published online: 15 October 2024

References

- Daghlis I, Gill D. Mendelian randomization as a tool to inform drug development using human genetics. *Camb Prism Precis Med*. 2023;1:e16.
- Dowden H, Munro J. Trends in clinical success rates and therapeutic focus. *Nat Rev Drug Discov*. 2019;18(7):495–6.
- Smith GD, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.
- Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*. 2014;23(R1):R89–98.
- Hingorani A, Humphries S. Nature's randomised trials. *Lancet*. 2005;366(9501):1906–8.
- Gill D, Georgakis MK, Walker VM, Schmidt AF, Gkatzionis A, Freitag DF, et al. Mendelian randomization for studying the effects of perturbing drug targets. *Wellcome Open Res*. 2021;6:16.
- 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.
- Nelson MR, Tipney H, Painter JL, Shen J, Nicoletti P, Shen Y, et al. The support of human genetic evidence for approved drug indications. *Nat Genet*. 2015;47(8):856–60.
- Hinkson IV, Madej B, Stahlberg EA. Accelerating Therapeutics for Opportunities in Medicine: A Paradigm Shift in Drug Discovery. *Front Pharmacol*. 2020;11:770.
- Heilbron K, Mozaffari SV, Vacic V, Yue P, Wang W, Shi J, et al. Advancing drug discovery using the power of the human genome. *J Pathol*. 2021;254(4):418–29.
- Zhang X, Yu W, Li Y, Wang A, Cao H, Fu Y. Drug development advances in human genetics-based targets. *MedComm* (2020). 2024;5(2):e481.
- Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res*. 2019;4:186.
- Davies NM, Holmes MV, Davey SG. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601.
- Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2019;48(3):713–27.
- Larsson SC, Butterworth AS, Burgess S. Mendelian randomization for cardiovascular diseases: principles and applications. *Eur Heart J*. 2023;44(47):4913–24.
- Burgess S, Mason AM, Grant AJ, Slob EAW, Gkatzionis A, Zuber V, et al. Using genetic association data to guide drug discovery and development: Review of methods and applications. *Am J Hum Genet*. 2023;110(2):195–214.
- Burgess S, Cronje HT. Incorporating biological and clinical insights into variant choice for Mendelian randomisation: examples and principles. *eGastroenterology*. 2024;2(1):e100042.
- Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey SG. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr*. 2016;103(4):965–78.
- Karhunen V, Larsson SC, Gill D. Genetically proxied growth-differentiation factor 15 levels and body mass index. *Br J Clin Pharmacol*. 2021;87(10):4036–9.
- Ference BA, Ray KK, Catapano AL, Ference TB, Burgess S, Neff DR, et al. Mendelian Randomization Study of ACLY and Cardiovascular Disease. *N Engl J Med*. 2019;380(11):1033–42.
- Zhou H, Shen J, Fang W, Liu J, Zhang Y, Huang Y, et al. Mendelian randomization study showed no causality between metformin use and lung cancer risk. *Int J Epidemiol*. 2020;49(4):1406–7.
- Yarmolinsky J, Bull CJ, Walker VM, Nounu A, Davey SG. Mendelian randomization applied to pharmaceutical use: the case of metformin and lung cancer. *Int J Epidemiol*. 2020;49(4):1410–1.
- Rena G, Hardie DG, Pearson ER. The mechanisms of action of metformin. *Diabetologia*. 2017;60(9):1577–85.
- Davis RL. Mechanism of action and target identification: a matter of timing in drug discovery. *iScience*. 2020;23(9):101487.
- Woolf B, Cronjé HT, Zagkos L, Larsson SC, Gill D, Burgess S. Comparison of caffeine consumption behavior with plasma caffeine levels as exposure measures in drug-target Mendelian randomization. *Am J Epidemiol*. 2024;kwae143.
- Yuan S, Larsson SC, Gill D, Burgess S. Concerns about instrumental variable selection for biological effect versus uptake of proton pump inhibitors in Mendelian randomisation analysis. *Gut*. 2024;gutjnl-2024-332280.
- Zagkos L, Cronje HT, Woolf B, de La Harpe R, Burgess S, Mantzoros CS, et al. Genetic investigation into the broad health implications of caffeine: evidence from phenome-wide, proteome-wide and metabolome-wide Mendelian randomization. *BMC Med*. 2024;22(1):81.
- Gill D, Georgakis MK, Koskeridis F, Jiang L, Feng Q, Wei WQ, et al. Use of genetic variants related to antihypertensive drugs to inform on efficacy and side effects. *Circulation*. 2019;140(4):270–9.
- Ye C, Wang T, Wang H, Lian G, Xie L. Causal relationship between genetic proxies for calcium channel blockers and the risk of depression: a drug-target Mendelian randomization study. *Front Psychiatry*. 2024;15:1377705.
- Fan B, Schooling CM, Zhao JV. Genetic proxies for calcium channel blockers and cancer: a Mendelian randomization study. *J Hum Hypertens*. 2023;37(11):1028–32.
- Sae-Jie W, Supasai S, Kivimaki M, Price JF, Wong A, Kumari M, et al. Triangulating evidence from observational and Mendelian randomization studies of ketone bodies for cognitive performance. *BMC Med*. 2023;21(1):340.
- Rusina PV, Falaguera MJ, Romero JMR, McDonagh EM, Dunham I, Ochoa D. Genetic support for FDA-approved drugs over the past decade. *Nat Rev Drug Discov*. 2023;22(11):864.
- Gaziano L, Giambartolomei C, Pereira AC, Gaulton A, Posner DC, Swanson SA, et al. Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. *Nat Med*. 2021;27(4):668–76.
- Gabe MBN, van der Velden WJC, Gadgaard S, Smit FX, Hartmann B, Brauner-Osborne H, et al. Enhanced agonist residence time, internalization rate and signalling of the GIP receptor variant [E354Q] facilitate receptor desensitization and long-term impairment of the GIP system. *Basic Clin Pharmacol Toxicol*. 2020;126 Suppl 6(Suppl 6):122–32.
- Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(6):1102.
- Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology*. 2006;17(3):260–7.
- Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank*. 2015;13(5):307–8.
- Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*. 2015;348(6235):648–60.
- Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580–5.
- Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrismisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53(12):1712–21.
- Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622(7982):329–38.

42. Liu Y, Wang Q, Zhao Y, Liu L, Hu J, Qiao Y, et al. Identification of novel drug targets for multiple sclerosis by integrating plasma genetics and proteomes. *Exp Gerontol*. 2024;194:112505.
43. Liu M, Chen M, Tan J, Chen A, Guo J. Plasma proteins and inflammatory dermatoses: proteome-wide Mendelian randomization and colocalization analyses. *Arch Dermatol Res*. 2024;316(7):443.
44. Fan Q, Wen S, Zhang Y, Feng X, Zheng W, Liang X, et al. Assessment of circulating proteins in thyroid cancer: proteome-wide Mendelian randomization and colocalization analysis. *iScience*. 2024;27(6):109961.
45. Tian J, Keller MP, Broman AT, Kendziorski C, Yandell BS, Attie AD, et al. The dissection of expression quantitative trait locus hotspots. *Genetics*. 2016;202(4):1563–74.
46. Larsson SC, Michaelsson K, Mola-Caminal M, Hoijer J, Mantzoros CS. Genome-wide association and Mendelian randomization study of fibroblast growth factor 21 reveals causal associations with hyperlipidemia and possibly NASH. *Metabolism*. 2022;137: 155329.
47. Larsson SC, Michaelsson K, Mola-Caminal M, Hoijer J, Mantzoros CS. Corrigendum to “Genome-wide association and Mendelian randomization study of fibroblast growth factor 21 reveals causal associations with hyperlipidemia and possibly NASH” [Metab Volume 137, December 2022, 155329]. *Metabolism*. 2023;143: 155555.
48. Ongen H, Brown AA, Delaneau O, Panousis NI, Nica AC, Consortium G. Estimating the causal tissues for complex traits and diseases. *Nat Genet*. 2017;49(12):1676–83.
49. Patel A, Gill D, Shungin D, Mantzoros CS, Knudsen LB, Bowden J, et al. Robust use of phenotypic heterogeneity at drug target genes for mechanistic insights: Application of cis-multivariable Mendelian randomization to GLP1R gene region. *Genet Epidemiol*. 2024;48(4):151–63.
50. Gill D, Burgess S. The evolution of mendelian randomization for investigating drug effects. *PLoS Med*. 2022;19(2):e1003898.
51. Rhodes B, Merriman ME, Harrison A, Nissen MJ, Smith M, Stamp L, et al. A genetic association study of serum acute-phase C-reactive protein levels in rheumatoid arthritis: implications for clinical interpretation. *PLoS Med*. 2010;7(9):e1000341.
52. Lemmelä S, Wigmore EM, Benner C, Havulinna AS, Ong RMY, Kempf T, et al. Integrated analyses of growth differentiation factor-15 concentration and cardiometabolic diseases in humans. *Elife*. 2022;11:e76272.
53. Gill D, Arvanitis M, Carter P, Hernandez Cordero AI, Jo B, Karhunen V, et al. ACE inhibition and cardiometabolic risk factors, lung ACE2 and TMPRSS2 gene expression, and plasma ACE2 levels: a Mendelian randomization study. *R Soc Open Sci*. 2020;7(11):200958.
54. Woolf B, Rajasundaram S, Cronje HT, Yarmolinsky J, Burgess S, Gill D. A drug target for erectile dysfunction to help improve fertility, sexual activity, and wellbeing: mendelian randomisation study. *BMJ*. 2023;383:e076197.
55. Karhunen V, Daghlas I, Zuber V, Vujkovic M, Olsen AK, Knudsen LB, et al. Leveraging human genetic data to investigate the cardiometabolic effects of glucose-dependent insulinotropic polypeptide signalling. *Diabetologia*. 2021;64(12):2773–8.
56. Torekov SS, Harslof T, Rejnmark L, Eiken P, Jensen JB, Herman AP, et al. A functional amino acid substitution in the glucose-dependent insulinotropic polypeptide receptor (GIPR) gene is associated with lower bone mineral density and increased fracture risk. *J Clin Endocrinol Metab*. 2014;99(4):E729–33.
57. Rogers M, Gill D, Ahlqvist E, Robinson T, Mariosa D, Johansson M, et al. Genetically proxied impaired GIPR signaling and risk of 6 cancers. *iScience*. 2023;26(6):106848.
58. Veniant MM, Lu SC, Atangan L, Komorowski R, Stanislaus S, Cheng Y, et al. A GIPR antagonist conjugated to GLP-1 analogues promotes weight loss with improved metabolic parameters in preclinical and phase 1 settings. *Nat Metab*. 2024;6(2):290–303.
59. Georgakis MK, Malik R, Gill D, Franceschini N, Sudlow CLM, Dichgans M, et al. Interleukin-6 Signaling Effects on Ischemic Stroke and Other Cardiovascular Outcomes: A Mendelian Randomization Study. *Circ Genom Precis Med*. 2020;13(3): e002872.
60. Cupido AJ, Asselbergs FW, Natarajan P, Group CIW, Ridker PM, Hovingh GK, et al. Dissecting the IL-6 pathway in cardiometabolic disease: a Mendelian randomization study on both IL6 and IL6R. *Br J Clin Pharmacol*. 2022;88(6):2875–84.
61. Ziaiean B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*. 2016;13(6):368–78.
62. Lumbers RT, Shah S, Lin H, Czuba T, Henry A, Swerdlow DJ, et al. The genomics of heart failure: design and rationale of the HERMES consortium. *ESC Heart Fail*. 2021;8(6):5531–41.
63. Schlosser P, Tin A, Matias-Garcia PR, Thio CHL, Joehanes R, Liu H, et al. Meta-analyses identify DNA methylation associated with kidney function and damage. *Nat Commun*. 2021;12(1):7174.
64. Bassett E, Broadbent J, Gill D, Burgess S, Mason AM. Inconsistency in UK biobank event definitions from different data sources and its impact on bias and generalizability: a case study of venous thromboembolism. *Am J Epidemiol*. 2024;193(5):787–97.
65. Cho Y, Rau A, Reiner A, Auer PL. Mendelian randomization analysis with survival outcomes. *Genet Epidemiol*. 2021;45(1):16–23.
66. Suzuki K, Hatzikotoulas K, Southam L, Taylor HJ, Yin X, Lorenz KM, et al. Genetic drivers of heterogeneity in type 2 diabetes pathophysiology. *Nature*. 2024;627(8003):347–57.
67. Chen J, Spracklen CN, Marenne G, Varshney A, Corbin LJ, Luan J, et al. The trans-ancestral genomic architecture of glycemic traits. *Nat Genet*. 2021;53(6):840–60.
68. Soderholm M, Pedersen A, Lorentzen E, Stanne TM, Bevan S, Olsson M, et al. Genome-wide association meta-analysis of functional outcome after ischemic stroke. *Neurology*. 2019;92(12):e1271–83.
69. Walker VM, Kehoe PG, Martin RM, Davies NM. Repurposing antihypertensive drugs for the prevention of Alzheimer’s disease: a Mendelian randomization study. *Int J Epidemiol*. 2020;49(4):1132–40.
70. Cronje HT, Karhunen V, Hovingh GK, Coppieters K, Lagerstedt JO, Nyberg M, et al. Genetic evidence implicating natriuretic peptide receptor-3 in cardiovascular disease risk: a Mendelian randomization study. *BMC Med*. 2023;21(1):158.
71. Myocardial Infarction G, Investigators CAEC, Stitzel NO, Stirrups KE, Masca NG, Erdmann J, et al. Coding variation in ANGPTL4, LPL, and SVEP1 and the risk of coronary disease. *N Engl J Med*. 2016;374(12):1134–44.
72. Gobeil E, Bourgault J, Mitchell PL, Houessou U, Gagnon E, Girard A, et al. Genetic inhibition of angiopoietin-like protein-3, lipids, and cardiometabolic risk. *Eur Heart J*. 2024;45(9):707–21.
73. Dewey FE, Gusarova V, Dunbar RL, O’Dushlaine C, Schurmann C, Gottesman O, et al. Genetic and pharmacologic inactivation of ANGPTL3 and cardiovascular disease. *N Engl J Med*. 2017;377(3):211–21.
74. Kersten S. Role and mechanism of the action of angiopoietin-like protein ANGPTL4 in plasma lipid metabolism. *J Lipid Res*. 2021;62: 100150.
75. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenotype. *Elife*. 2018;7:7.
76. Zuber V, Grinberg NF, Gill D, Manipur I, Slob EAW, Patel A, et al. Combining evidence from Mendelian randomization and colocalization: review and comparison of approaches. *Am J Hum Genet*. 2022;109(5):767–82.
77. Bowker N, Hansford R, Burgess S, Foley CN, Auyeung VPW, Erzurumluoglu AM, et al. Genetically predicted Glucose-Dependent Insulinotropic Polypeptide (GIP) levels and cardiovascular disease risk are driven by distinct causal variants in the GIPR region. *Diabetes*. 2021;70(11):2706–19.
78. Schmidt AF, Hunt NB, Gordillo-Maranon M, Charoen P, Drenos F, Kivimaki M, et al. Cholesteryl ester transfer protein (CETP) as a drug target for cardiovascular disease. *Nat Commun*. 2021;12(1):5640.
79. Group HTRC, Bowman L, Hopewell JC, Chen F, Wallendszus K, Stevens W, et al. Effects of anacetrapib in Patients with Atherosclerotic Vascular Disease. *N Engl J Med*. 2017;377(13):1217–27.
80. Piwecka M, Rajewsky N, Rybak-Wolf A. Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. *Nat Rev Neurol*. 2023;19(6):346–62.
81. Ottaviani D, Ter Huurne M, Elliott DA, Bellin M, Mummery CL. Maturing differentiated human pluripotent stem cells in vitro: methods and challenges. *Development*. 2023;150(11):dev201103.
82. May LT, Bartolo BA, Harrison DG, Guzik T, Drummond GR, Figtree GA, et al. Translating atherosclerosis research from bench to bedside: navigating the barriers for effective preclinical drug discovery. *Clin Sci (Lond)*. 2022;136(23):1731–58.

83. Trajanoska K, Bherer C, Taliun D, Zhou S, Richards JB, Mooser V. From target discovery to clinical drug development with human genetics. *Nature*. 2023;620(7975):737–45.
84. Kim GB, Kim JY, Lee JA, Norsigian CJ, Palsson BO, Lee SY. Functional annotation of enzyme-encoding genes using deep learning with transformer layers. *Nat Commun*. 2023;14(1):7370.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.