



Sobczyk-Barad, M. K., Zheng, J., Davey Smith, G., & Gaunt, T. R. (2023). Systematic comparison of Mendelian randomisation studies and randomised controlled trials using electronic databases. *BMJ Open*, 13(9), [e072087]. <https://doi.org/10.1136/bmjopen-2023-072087>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1136/bmjopen-2023-072087](https://doi.org/10.1136/bmjopen-2023-072087)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via BMJ at <http://dx.doi.org/10.1136/bmjopen-2023-072087> . Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# BMJ Open Systematic comparison of Mendelian randomisation studies and randomised controlled trials using electronic databases

Maria K Sobczyk ,<sup>1</sup> Jie Zheng ,<sup>1,2,3</sup> George Davey Smith ,<sup>1</sup>  
Tom R Gaunt <sup>1</sup>

**To cite:** Sobczyk MK, Zheng J, Davey Smith G, *et al.* Systematic comparison of Mendelian randomisation studies and randomised controlled trials using electronic databases. *BMJ Open* 2023;**13**:e072087. doi:10.1136/bmjopen-2023-072087

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2023-072087>).

Received 23 January 2023  
Accepted 21 August 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Maria K Sobczyk;  
[maria.sobczyk@bristol.ac.uk](mailto:maria.sobczyk@bristol.ac.uk)

## ABSTRACT

**Objective** To scope the potential for (semi)-automated triangulation of Mendelian randomisation (MR) and randomised controlled trials (RCTs) evidence since the two methods have distinct assumptions that make comparisons between their results invaluable.

**Methods** We mined ClinicalTrials.Gov, PubMed and EpigraphDB databases and carried out a series of 26 manual literature comparisons among 54 MR and 77 RCT publications.

**Results** We found that only 13% of completed RCTs identified in ClinicalTrials.Gov submitted their results to the database. Similarly low coverage was revealed for Semantic Medline (SemMedDB) semantic triples derived from MR and RCT publications—36% and 12%, respectively. Among intervention types that can be mimicked by MR, only trials of pharmaceutical interventions could be automatically matched to MR results due to insufficient annotation with Medical Subject Headings ontology. A manual survey of the literature highlighted the potential for triangulation across a number of exposure/outcome pairs if these challenges can be addressed.

**Conclusions** We conclude that careful triangulation of MR with RCT evidence should involve consideration of similarity of phenotypes across study designs, intervention intensity and duration, study population demography and health status, comparator group, intervention goal and quality of evidence.

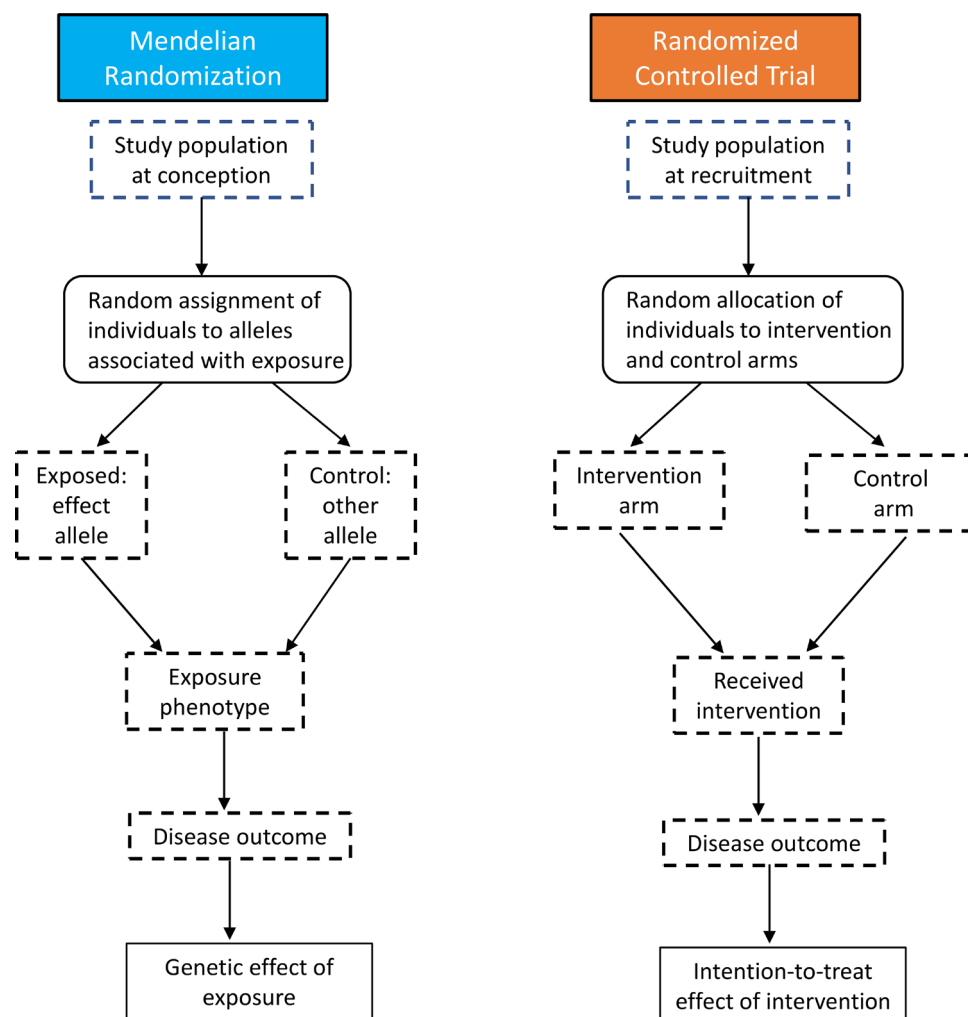
## BACKGROUND

Randomised controlled trials (RCTs) are deemed the ‘gold standard’ in evaluating the efficacy of interventions and guiding practice in clinical research, with well-established methodology.<sup>1</sup> In RCTs, a selection of individuals intended to represent the target population is randomly assigned to a treatment or control group, allowing estimation of the intervention’s effectiveness in the absence of confounding variables and reverse causality that are present in observational studies. In the past two decades, an approach to causal inference using natural genetic variation,

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ Mendelian randomisation (MR) has become a popular method in causal inference in genetic epidemiology, and while often used as proxy for to randomised clinical trials (RCTs), little is known about scope for automatic comparison between MR and RCT results.
- ⇒ Previous research has established conceptual similarities and differences between MR and RCT methodology, however, without focus on applied cases.
- ⇒ The study found that a low percentage of completed RCTs were submitted to ClinicalTrials.gov and that a similarly low coverage was found for MR and RCT publications in Semantic Medline. Only trials of pharmaceutical interventions could be automatically matched to MR results due to insufficient annotation with Medical Subject Headings ontology among other interventions.
- ⇒ Following manual extraction of MR and RCT literature, we assessed result concordance across the two methods and discussed multiple possible reasons for discrepancies.
- ⇒ Sparsity of data in electronic databases hinders the ability to automatically compare results of MR and RCT studies. In the absence of retrospective manual extraction of MR and RCT results from publications, more research effort needs to be spent developing machine-learning approaches to aid systematic comparisons. Our study helps identify study design features which need to be captured by such methods.

known as Mendelian randomisation (MR)—usually implemented as an instrumental variable (IV) analyses—has gained popularity.<sup>2,3</sup> This approach has been referred to as ‘nature’s randomised trials’<sup>4</sup> and is based on the randomisation from parents to offspring of genetic variants encapsulated in Mendel’s laws of segregation and independent assortment.<sup>2,5</sup> At a population level, the randomisation is approximate, but still allows genetic variants that are robustly associated with the measured exposure to be used to estimate



**Figure 1** Comparison of Mendelian randomisation and randomised controlled trial design. After: Nitsch *et al*,<sup>171</sup> Ebrahim and Smith<sup>172</sup> and Ference.<sup>157</sup>

the unbiased causal effect of an exposure (generally acting across life) on health outcomes, as long as certain assumptions, discussed in detail elsewhere,<sup>2,3,6</sup> are met.

Despite drawing on observational data, the MR approach broadly aligns with that of an RCT, where the goal is to estimate the causal effect of an intervention on the given endpoint based on groups (arms) which do not differ with respect to confounding variables (figure 1). However, since in MR randomisation takes place at conception, the time lag to the start of outcome recoding is longer<sup>7</sup> compared with RCTs, where median duration of phase 3 trials is 40 months.<sup>8</sup> Similarly to RCTs, most MR analyses should be free of confounding and reverse causation bias due to variants being allocated randomly before birth and outcome condition onset.

Previous research has shown examples of evidence triangulation where MR results predicted the overall RCT results based on totally orthogonal data with an unrelated set of systematic errors and biases.<sup>9</sup> For instance, MR demonstrated the lack of effect of genetically predicted concentrations of High-density lipoprotein cholesterol (HDL-C) on cardiovascular events<sup>10–12</sup> as well as selenium in prostate cancer prevention trials.<sup>13,14</sup> On the other

hand, MR showed the beneficial effect of lifelong endogenous low-density lipoprotein cholesterol (LDL-C) levels,<sup>15</sup> HMG-CoA reductase inhibition (statin drug target) and PCSK9 inhibition on cardiovascular disease (CVD),<sup>15,16</sup> while predicting also the increased risk of type 2 diabetes (T2D) as a side effect of statin usage. However, three independent MR studies were at odds with later RCTs by predicting increased risk of T2D also as a side effect of PCSK9 inhibition.<sup>17</sup>

There are several possible explanations for apparent or real discordance in the results of RCT and MR studies. These range from different durations, magnitude and time-varying nature of the exposure, origin of the study populations, and natural genetic variation imperfectly mimicking the molecular action of the drug, some of which we explore. The direct comparison of MR and RCT findings is facilitated by the use of a precisely defined estimand,<sup>18</sup> for example, the effect on incident coronary heart disease risk of lowering LDL cholesterol by 1 mmol/L for 5 years. While RCTs will estimate something close to this, and be scalable to it, with MR studies the exposure difference associated with the genetic instruments will often exist from birth (or before) and

may change in magnitude over time.<sup>19</sup> This is discussed further in online supplemental file 1.

While RCTs can provide the highest-quality evidence, they may have limitations. They are often expensive to carry out, can be of small size and lack external validity,<sup>20</sup> have short follow-up and typically take place after disease onset.<sup>21–22</sup> As in other study types, RCT results may be flawed due to poor design and execution, for example, imperfect randomisation, unblinding and differential loss to follow-up between study arms.

Unlike RCTs, MR studies are inexpensive and quick to perform when suitable genetic instruments are available. Therefore, they can potentially prioritise intervention–condition pairs to assess in RCTs. Moreover, it has been proposed that MR also guides the design of RCTs, improving eligibility criteria to prioritise groups most likely to benefit, suggesting diseases for composite endpoint construction and alerting to potential side effects.<sup>123</sup> Since MR analyses suffer a different set of biases than RCTs, MR evidence can be used to complement RCTs and other study designs in the triangulation framework to guide therapeutic development and clinical practice.<sup>24–26</sup> Finally, the extensive use of existing observational data for MR enables intervention targets to be evaluated in a wider range of subpopulations than is feasible for RCTs (improving generalisability), and allows comparisons to be made that might be unethical in experimental studies, for example, when there is strong evidence in favour of a particular treatment.

The goal of this research is to survey the extent of concordance between MR and RCT studies to date and identify possible factors for disparities in the direction of effect, which limit the ability to extrapolate from MR results to RCTs and increase the complexity of the triangulation process. In this study, we aimed to carry out a systematic analysis of MR and RCT results using automated mining of data in the public domain, including the ClinicalTrials.Gov,<sup>27</sup> EpigraphDB<sup>28</sup> and PubMed databases. We evaluate the comprehensiveness and scope of the data available and potential for comparative analyses between MR and RCTs. We then go on to develop a series of case studies looking in detail at MR and RCT comparisons across 26 exposure–outcome pairs. Throughout, we use the term ‘intervention’ as synonymous with ‘exposure’ and ‘condition’ as synonymous with ‘outcome’.

## METHODS

### ClinicalTrials.Gov data sources

All ClinicalTrials.Gov study data are available for download as PostgreSQL database from the database for Aggregate Analysis of ClinicalTrials.gov (AACT)<sup>29</sup> released by the Clinical Trials Transformation Initiative.<sup>30</sup> We downloaded its static release from 1 July 2023. Processing of the database files was carried out using custom Python and R scripts.

### ClinicalTrials.Gov data filtering

We filtered the ClinicalTrials.Gov studies using a number of criteria to identify RCTs with submitted results allowing direct comparison with MR studies. These are depicted in figure 2A and provided in detail in online supplemental note.

### EpigraphDB queries

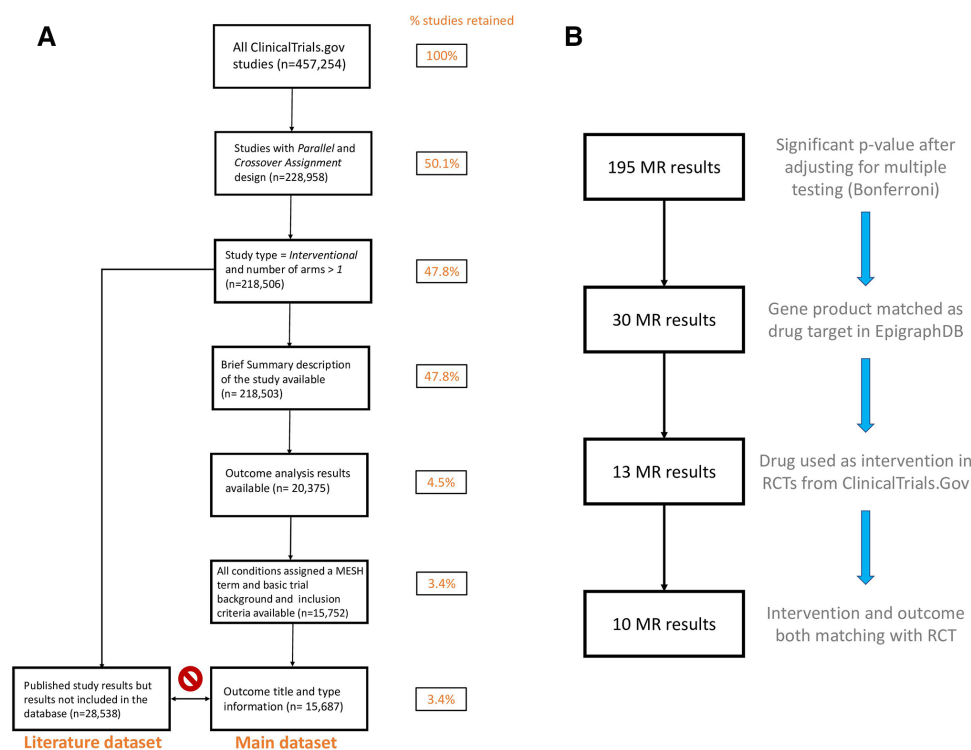
EpigraphDB<sup>28</sup> was used to collect information about confirmed drug–target associations which were initially sourced from the Open Targets Platform<sup>31</sup> and verified in DrugBank.<sup>32</sup> EpigraphDB was then used to retrieve expression quantitative trait loci (eQTL) and protein quantitative trait loci (pQTL) MR results previously described in Zheng *et al.*<sup>33</sup> We also used EpigraphDB<sup>28</sup> to source SemMedDB<sup>34</sup> V.1.8 semantic triples associated with select MR and RCT publications identified by PubMed. SemMedDB triples in EpigraphDB are prefiltered for annotation of epidemiological studies as described previously.<sup>35</sup>

### PubMed data harvesting

We searched PubMed for all RCT and MR studies published before 2023 on 1 July 2023. For RCTs, we searched titles and abstracts for keywords: “randomized controlled trial” or “RCT” and we used PubMed’s in-built *Randomized Controlled Trial* label filter to obtain more specific hits, reducing the number of hits from 129 077 to 74 559. In order to retrieve potential MR studies, we used the keywords: “mendelian randomization” or “mendelian randomisation”. We also considered using a Medical Subject Headings (MeSH) label “Mendelian Randomization Analysis” but it returned an unrealistically low number of hits (3174), a consequence of manual indexing.

### Literature searches

We used Semantic Scholar and Google Scholar to survey MR and RCT literature indexed before 1 July 2023. We queried the databases with the following search terms: “[exposure] [condition] Mendelian Randomization” and “[exposure] [condition] Randomized Controlled Trial”. The articles were initially screened by title and abstract. We considered original research MR, RCT studies as well as meta-analyses. We included 26 intervention–outcome pairs to represent a wide array of behavioural and nutritional interventions with a diverse set of common disease (cardiometabolic, neuropsychiatric, cancer, dermatological) and disease biomarker outcomes, based on our expert knowledge of the field. Our chosen exposures correspond to the top four modifiable risk factors accounting for 39% of deaths in the USA<sup>36</sup>: high alcohol intake, high body mass index (BMI), lack of exercise and smoking. In addition, since potentially preventative effects of nutritional factors are controversial and notoriously difficult to evaluate using non-randomised study designs,<sup>37</sup> we also included vitamins D and E as well as coffee as an intervention. We acknowledge that this



**Figure 2** (A) Filtering steps applied to ClinicalTrials.Gov database. Filtering was designed to identify RCTs whose final results statistics were uploaded to the database (main dataset). In addition, other RCTs which published their findings in scientific journals were identified (literature dataset). (B) Filtering steps applied to EpigraphDB database. Filtering was designed to identify protein QTL MR studies with intervention and exposure matching those of RCT published on ClinicalTrials.Gov. MR, Mendelian randomisation; QTL, quantitative trait loci; RCT, randomised controlled trial.

choice is somewhat subjective but we believe it to be illustrative of the current MR and RCT literature.

MR and RCT studies were compared across: population characteristics (sex, ethnicity, age, health status), comparator group, goal of intervention (prevention or treatment/slowing progression), direction of effect, length of follow-up, main test statistic in the study and its impact as judged by citation number.

### Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

## RESULTS

### ClinicalTrials.Gov data overview

In total, we found 457 254 individual studies were registered with a unique ClinicalTrials.gov identifier. We filtered them using a number of steps to identify RCTs and facilitate comparison with MR (figure 2A). In our analysis, we identified 218 506 RCT studies (48% of the total). To allow semiautomated comparison with MR studies, we focused on the study subset which submitted their statistical analysis results to the database (referred to as the *main dataset*, online supplemental dataset 1). However, we found that only 3.4% of studies—15 752 met this criterion, along with including background information on the trial. To expand that number, we also considered an

additional 28 538 RCT studies which did not publish their results in ClinicalTrials.gov but instead linked to a peer-reviewed publication (referred to as the *literature dataset*, online supplemental dataset 2).

The majority of RCTs in the main dataset followed parallel assignment of participants to treatment (online supplemental figure S1a), most were designed for treatment (n=12 336, online supplemental figure S1b), rather than prevention (n=1613) and the vast majority of them had been completed (online supplemental figure S1c). More trials were observed to be in phase 3 than 4 (online supplemental figure S1d), most trials included both males and females (online supplemental figure S1e) and a great majority had two arms (online supplemental figure S1f). The median number of primary outcomes was 1 (online supplemental figure S2a), with a median of 6 secondary outcomes (online supplemental figure S2b). Over half of studies report at least one result with p value less than 0.05 (online supplemental figure S2c). Comparison with features of all RCTs in the database showed that our selection was broadly representative (online supplemental dataset 3), although our dataset was enriched for completed and late-phase trials.

### Suitability of MeSH annotation

In order to attempt automated matching of RCTs and MRs involving similar interventions and outcomes for RCTs and MR, we needed to first establish the quality

**Table 1** Completeness of MeSH term annotation among the chosen intervention types in the *main* (RCT results available in ClinicalTrials.Gov) and *literature* (RCT results unavailable in ClinicalTrials.Gov but study linked to a publication with results) datasets

| Dataset    | Intervention type  | Total  | Intervention MeSH missing | % missing | Condition MeSH missing | % missing |
|------------|--------------------|--------|---------------------------|-----------|------------------------|-----------|
| Main       | Drug               | 11 537 | 2212                      | 19.2      | 992                    | 8.6       |
| Literature | Drug               | 10 927 | 1773                      | 16.2      | 1031                   | 9.4       |
| Main       | Behavioural        | 1239   | 1075                      | 86.8      | 242                    | 19.5      |
| Literature | Behavioural        | 3017   | 2815                      | 93.3      | 620                    | 20.5      |
| Main       | Dietary supplement | 242    | 116                       | 47.9      | 31                     | 12.8      |
| Literature | Dietary supplement | 830    | 581                       | 70.0      | 164                    | 19.8      |

MeSH, Medical Subject Headings; RCT, randomised controlled trial.

of annotation of RCTs with MeSH in ClinicalTrials.Gov. The most common intervention was drug (online supplemental table S1). Since we were only interested in the intervention types which can be instrumented by MR, we also focused on the fifth and seventh most popular types of interventions: behavioural and dietary supplement. We found that MeSH intervention annotations were missing for only 19% and 16% of drug interventions in the main and literature datasets, accordingly (table 1). However, the overwhelming majority of RCTs in the behavioural and dietary supplement category did not contain a MeSH intervention term. Due to well-standardised disease taxonomy, a much lower level of missing data was found for MeSH condition terms. This allowed us to proceed with automated analysis of drug RCT data; however, for behavioural and dietary supplement we were only able to do a manual screening for RCTs with corresponding MR studies.

### Pharmaceutical interventions in RCTs and MR

Genetic IVs in MR can be used as proxies for pharmaceutical interventions in RCTs. pQTL or eQTL, that is, variants associated with expression of protein drug targets are used to directly proxy the action of a drug. Here, we use the biggest MR dataset for drug target protein–disease associations, examined in whole blood, from Zheng *et al.*<sup>33</sup> We focused on *cis*-acting instruments as a more specific marker for drug efficacy as *trans*-instruments are more likely to be pleiotropic, potentially leading to spurious results.<sup>23</sup>

We matched the drug target proteins in Zheng *et al.*<sup>33</sup> with drug–gene associations sourced from EpigraphDB<sup>28</sup> (figure 2B). This allowed us to merge the Zheng *et al.*<sup>33</sup> dataset with the main and literature RCT dataset via the drug listed in EpigraphDB and MeSH drug intervention term, accordingly. For the outcome, we were then able to match RCTs and MR manually due to the reasonably

**Table 2** Drug target–disease matches supported by evidence from MR (blood pQTL instruments<sup>33</sup>) and RCT studies (*main dataset* from ClinicalTrials.Gov)

| MR exposure | RCT drug intervention  | MR outcomes  | RCT conditions   | Matching trials | Concordant direction of effect? | xQTL |
|-------------|------------------------|--|--|-----------------|---------------------------------|------|
| PCSK9       | Evolocumab, alirocumab | Non-cancer illness code self-reported: high cholesterol    id:UKB-a:108  | Hyperlipidaemia, dyslipidaemia, hypercholesterolaemia, mixed dyslipidaemia           | 25              | Yes                             | pQTL |
| APOB        | Mipomersen             | LDL cholesterol    id:300, HDL cholesterol    id:299, triglycerides    id:302, non-cancer illness code self-reported: high cholesterol    id:UKB-a:108, total cholesterol    id:301, | Hyperlipidaemia, dyslipidaemia, hypercholesterolaemia, mixed dyslipidaemia           | 6               | Yes                             | pQTL |
| IL12B       | Ustekinumab            | Non-cancer illness code self-reported: psoriasis    id:UKB-a:100; ulcerative colitis    id:970; Crohn's disease    id:12; inflammatory bowel disease    id:294                       | Psoriasis, psoriatic arthritis, Crohn's disease, colitis, inflammatory bowel disease | 21              | Yes                             | pQTL |

MR, Mendelian randomisation; pQTL, protein quantitative trait loci; QTL, quantitative trait loci; RCT, randomised controlled trial.

low number of hits. The results displayed in [table 2](#) show overlap of the RCT and MR datasets. We found four drugs: evolocumab/alirocumab, ustekinumab and mipomersen that share support from both MR and RCT studies. Evolocumab/alirocumab and mipomersen inhibit key players (PCSK9 and apoB) in lipid transport helping to lower plasma LDL-C levels.<sup>38</sup> The Zheng *et al*<sup>33</sup> MR study showed a negative effect of reduced PCSK9 levels on high cholesterol in the UK Biobank, while in 25 RCT studies drug-induced abrogation of PCSK9 activity led to positive outcomes in the treatment of hyperlipidaemia, hypercholesterolaemia and dyslipidaemias in general. Similarly, reduced expression/activity of apoB in MR and six RCT studies resulted in genetically predicted lower levels of LDL cholesterol and total cholesterol in the UK Biobank as well as improved outcomes in the treatment of dyslipidaemias, respectively. The third example of a good match between RCT and MR studies concerns inhibition of the p40 subunit of interleukin 12 and 23 (IL12B).<sup>39</sup> Both MR and 21 RCTs show benefit of inhibition of p40 on immune-mediated disease: psoriasis and inflammatory bowel disease.

In general, pQTL MR-based prediction of drug target–condition pairs offered good recall when compared with the pairs in the Open Targets Platform for the proteins with MR evidence. The only drug target indications missing included conditions not analysed in the MR study, such as CD33 protein being the drug target for treatment of leukaemia, with the exception of acetylcholinesterase whose inhibitors (galantamine, donepezil, rivastigmine) are used for treatment of cognitive decline in Alzheimer's disease.<sup>40</sup>

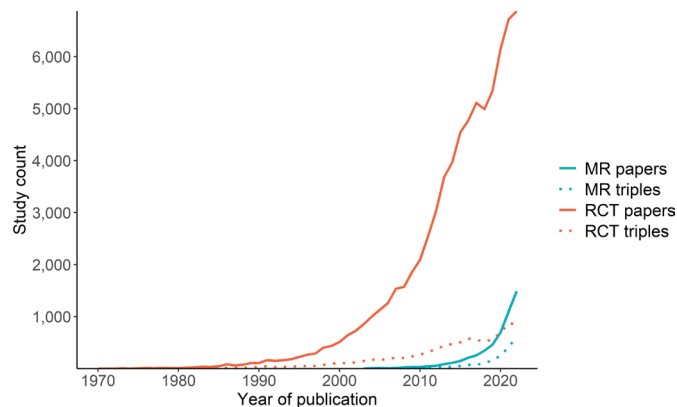
We also compared the RCT dataset with Zheng *et al* blood transcript expression (eQTL)-derived MR analysis (available in EpigraphDB: <https://epigraphdb.org/xqtl>, online supplemental figure S3). In total, we identified 15 drug target–disease matches in the eQTL dataset (online supplemental table S2), although unlike in the pQTL matches, the direction of effect in MR was incorrect in eight cases. Nevertheless, the eQTL MR results agreed with some well-known drug effects: HDL-C and LDL-C lowering action of CETP and HMGCR inhibitors,<sup>41</sup> respectively, and blood pressure-lowering action of ACE inhibitors.<sup>42</sup>

### PubMed-sourced MR and RCT studies

In addition to searching through the ClinicalTrials.Gov database, we also queried PubMed for RCT and MR publications. In total, we found 5135 MR studies published since mid-2000s and 73 306 RCTs published since 1970 until 2022 ([figure 3](#)).

### Semantic analysis with SemMedDB

We subsequently wanted to establish the thematic overlap between MR and RCT studies using an alternative method involving semantic analysis. SemMedDB<sup>34</sup> provides a vast repository of semantic predications (subject–predicate–object triple, for example, LDL-C causes ischaemic



**Figure 3** Popularity of MR and RCT studies over time. We compare counts of MR and RCT papers indexed by PubMed (solid lines) with number of semantic triples derived from them using SemMedDB (dashed lines). MR, Mendelian randomisation; RCT, randomised controlled trial.

heart disease). We linked the MR and RCT publications identified by our PubMed search to their corresponding SemMed triples in EpigraphDB using PubMed ID (online supplemental dataset 4). Overall, only 12% and 36% of RCT and MR papers, respectively, had a semantic triple associated with them ([figure 3](#)). When ignoring the predicate, and focusing only on the subject and object, we found a total of 15 113 unique exposure–outcome pairs (online supplemental figure S4), discussed in detail in the online supplemental note. However, only 221 of these were found to be shared across MR and RCT studies.

We then investigated the 221 matching subject–object pairs between MR and RCT studies (online supplemental table S3), as well as individual top counts among subjects (online supplemental table S4) and objects (online supplemental table S5). T2D, insulin and obesity were found among the top shared risk factors, along with lipids and vitamin D. Top outcomes included T2D, CVD, COVID-19 and Alzheimer's disease.

### Case studies of matching MR and RCTs

Since our semiautomatic mining of MR and RCT literature brought limited results for behavioural and nutritional interventions, we selected 26 intervention–outcome case studies by manual mining of the literature representing common lifestyle risk factors, dietary and behavioural exposures, paired with common cardiovascular, glycaemic, neuropsychiatric, musculoskeletal, autoimmune and cancer outcome phenotypes. In total, we surveyed 54 MR and 77 RCT publications (RCTs and meta-analysis of RCTs, online supplemental dataset 5, [figure 4](#)) which were systematically compared across several criteria shown in sample online supplemental table S6, and encompass those in the popular PICO (Population, Intervention, Comparison and Outcome) framework.<sup>43</sup>

There, we compare an MR study and two RCT meta-analyses on the effect of vitamin D supplementation in multiple sclerosis (MS).<sup>44–46</sup> While the RCTs looked at

| exposure        | outcome                | MR | RCT |
|-----------------|------------------------|----|-----|
| alcohol         | hypertension           | 2  | 4   |
| blood pressure  | cardiovascular disease | 3  | 3   |
| body mass index | cardiovascular disease | 2  | 2   |
| body mass index | hypertension           | 4  | 3   |
| body mass index | type 2 diabetes        | 5  | 3   |
| coffee          | glycemic biomarkers    | 1  | 2   |
| coffee          | lipids                 | 3  | 3   |
| diary intake    | blood pressure         | 1  | 1   |
| exercise        | bone mineral density   | 1  | 2   |
| exercise        | depression             | 2  | 7   |
| exercise        | glycemic biomarkers    | 1  | 2   |
| exercise        | lipids                 | 1  | 6   |
| exercise        | pain                   | 1  | 1   |
| exercise        | schizophrenia          | 2  | 3   |
| smoking         | lipids                 | 2  | 2   |
| vitamin D       | atopic dermatitis      | 2  | 4   |
| vitamin D       | blood pressure         | 3  | 2   |
| vitamin D       | body mass index        | 2  | 2   |
| vitamin D       | bone fractures         | 1  | 5   |
| vitamin D       | bone mineral density   | 3  | 2   |
| vitamin D       | cardiovascular disease | 1  | 2   |
| vitamin D       | depression             | 4  | 3   |
| vitamin D       | glycemic biomarkers    | 1  | 3   |
| vitamin D       | lipids                 | 3  | 4   |
| vitamin D       | multiple sclerosis     | 1  | 2   |
| vitamin E       | prostate cancer        | 2  | 4   |

| Legend  |                               |
|---|-------------------------------|
| <span style="display:inline-block; width:10px; height:10px; background-color:red;"></span>        | all positive effect size      |
| <span style="display:inline-block; width:10px; height:10px; background-color:lightcoral;"></span> | majority positive effect size |
| <span style="display:inline-block; width:10px; height:10px; background-color:gray;"></span>       | all null                      |
| <span style="display:inline-block; width:10px; height:10px; background-color:lightgray;"></span>  | majority null                 |
| <span style="display:inline-block; width:10px; height:10px; background-color:blue;"></span>       | all negative effect size      |
| <span style="display:inline-block; width:10px; height:10px; background-color:lightblue;"></span>  | majority negative effect size |

**Figure 4** Summary of case series of MR and RCT studies with matching exposures (interventions) and outcomes (conditions). The values correspond to the number of analysed studies in a given category, while the cell background colour indicates summary direction of effect on the outcome when exposure is increased—we report direction of effect found either in all analysed studies or their majority (>50%). MR, Mendelian randomisation; RCT, randomised controlled trial.



**Table 3** Overview of discussed criteria for assessment of alignment of MR and RCT study features

| Match criterion                    | Issues to consider                             | Example  |
|------------------------------------|--|--|
| Exposure (intervention)            | Similarity between analysed exposures          | Different types of exercise  |
|                                    | Intervention intensity                         | Vitamin D dosage   |
| Intervention goal                  | Prevention or treatment                        | Schizophrenia onset or treatment   |
| Outcome (condition)                | Single or composite outcome                    | Single or composite cardiovascular outcomes  |
|                                    | Binary versus categorical outcome              | Depression or rating on depression assessment scale, such as Hamilton Depression Rating Scale      |
|                                    | Similarity between analysed outcomes           | Different measures of adiposity  |
| Source population                  | Demographics                                   | Young adults or elderly  |
|                                    | Health status                                  | Diabetic or healthy  |
| Comparator group                   | Exposure-naïve or previously exposed           | Ex-smoker or never-smoker  |
|                                    | Active intervention or placebo                 | Statin as comparator or placebo  |
| Duration of intervention/follow-up | Length of intensive intervention and follow-up | Short duration of intervention (<6 months) or long duration and follow-up (>3 years) in RCT and MR |
|                                    | Not uniform intervention intensity or duration | Weekly counselling during the initial phase of the trial or throughout                             |

MR, Mendelian randomisation; RCT, randomised controlled trial.

potential therapeutic effect of vitamin D in patients with diagnosed MS over 6 months–2 years: measured disability (Expanded disability status scale (EDSS) Score) and recorded relapses as outcomes, the MR study took place in the general population and measured the causal effect of genetically predicted lifetime circulating vitamin D concentrations on prevention of MS. The conclusions of MR and RCT studies did not align well, with MR analyses providing evidence for reduced risk of MS conferred by higher vitamin D levels, but no significant therapeutic effect of vitamin D in existing MS was found in the five small meta-analysed trials. Differences which may impact on the ability for MR to complement RCT studies are summarised in [table 3](#) and discussed below based on this series of case studies.

#### Exposure/intervention

We found that overlapping MR and RCT interventions are often not perfectly identical which may impact on the estimated direction of effect. For example, MR exercise exposures are based on genetic variants associated with self-reported physical activity (moderate-to-vigorous and vigorous)<sup>47 48</sup> in studies assessing the effect on both lipids and bone mineral density (BMD). However, the corresponding RCTs used particular types of exercise, such as walking,<sup>49</sup> aerobic exercise,<sup>50 51</sup> progressive resistance training<sup>52</sup> and maximal strength training<sup>53</sup> as interventions. While an MR study<sup>48</sup> and two trials<sup>53 54</sup> showed concordant ([figure 4](#)), positive effect of exercise on BMD, we found that the effect of exercise on lipids did not match between MR and RCTs, with MR study<sup>47</sup> reporting null effect and trials generally finding positive effects

on HDL-C concentration and negative on LDL-C, total cholesterol and triglycerides blood levels.<sup>50–52 55</sup>

Furthermore, intensity of intervention can affect the comparative value of MR and RCT study conclusions. The MR study of vitamin D levels on bone fractures<sup>56</sup> was only able to assess linear effects of the normal range of circulating 25-hydroxyvitamin D concentrations. Consequently, the positive effect of high-concentration vitamin D ( $\geq 700$  IU daily) on bone fractures in the elderly seen in RCTs<sup>57–59</sup> may not have been accessible in the MR study.

#### Intervention goal

The intervention goal between MR and RCT studies can match (both prevention or treatment) or be misaligned which can potentially impact the ultimate conclusions of the study. We found the latter to be the case for the effect of exercise on schizophrenia. Two MR studies found a null preventative effect of exercise on schizophrenia,<sup>60 61</sup> while three meta-analyses of RCTs found a consistent effect of a variety of exercise types on improving total and negative symptoms of schizophrenia<sup>62–64</sup> ([figure 4](#)).

#### Outcome

The short duration of RCTs mean some outcomes (eg, myocardial infarction) do not accumulate enough events to detect a significant effect, therefore composite measures grouping related diseases are often used. When comparing the effect of systolic blood pressure (SBP) on CVD outcomes, we found matching conclusions with elevated SBP increasing the risk of CVD both in RCT<sup>65–67</sup> and MR<sup>68–70</sup> studies ([figure 4](#)), with MR studies using both single disease outcomes and a composite outcome.

However, MR studies analysing the impact of BMI on CVD found reduced adiposity led to reduction in arterial hypertension, CVD and stroke,<sup>71 72</sup> which contrasted with the results of one of the biggest RCTs to date. The Look AHEAD RCT in older patients with T2D found no preventative effect of weight loss on a composite outcome relating to mortality from cardiovascular causes, non-fatal myocardial infarction, non-fatal stroke or hospitalisation for angina.<sup>73</sup>

Second, RCT outcomes are often on a quantitative scale measuring symptom strength according to established metrics, (eg, depressive symptoms on Hamilton Depression Rating Scale<sup>74</sup>). However, the best disease Genome-wide association study (GWAS) used to identify MR instruments often represent binary disease outcomes, which could potentially lead to differential conclusions due to reduced power to detect subtler therapeutic effects. While exercise is causally associated with reduced depression and depressive symptoms both in MR<sup>75 76</sup> and RCTs,<sup>62 74 77–80</sup> the differences in outcome phenotypes could potentially contribute to null MR results<sup>81 82</sup> and positive effect of vitamin D on attenuating eczema symptoms in RCTs.<sup>83–86</sup>

#### Source population

MR studies are likely to draw from a wider demographic than RCTs due to use of biobanks and GWAS consortia, while RCTs focus on high risk groups.<sup>23</sup> For example, while in the MR study conducted in general population, there was no strong significant effect of exercise on glycaemic markers: Hemoglobin A1c (HbA1c), fasting glucose and Homeostatic model assessment for insulin resistance (HOMA-IR),<sup>87</sup> a significant reduction was found in the meta-analysis of 32 RCTs involving patients with T2D.<sup>88</sup>

On the other hand, five MR<sup>89–93</sup> studies along with three large RCTs<sup>94–96</sup> consistently provide evidence that weight loss is causally associated with reduced risk of T2D (figure 4), despite MR including the general population and RCTs focusing on at-risk individuals with impaired glucose tolerance.

As another example of possible demographics-driven differences in trial and MR results, MR studies on the relationship between vitamin D levels and atopic dermatitis were conducted in the general population,<sup>81 82</sup> while RCTs were conducted separately in children in Mongolia<sup>83</sup> and Boston, USA<sup>84</sup> with winter atopic dermatitis and in adolescent and adult Iranians.<sup>85 86</sup>

#### Comparator group

First, due to ethical considerations, trials of harmful behaviours such as alcohol drinking and smoking focus on cessation or reduction in existing users, and do not include never smokers or never drinkers as controls, unlike MR studies, which can potentially lead to differences in effect.<sup>23</sup> Nevertheless, the two outcomes analysed here: hypertension for alcohol intake and lipids for smoking showed generally congruent results across study types (online supplemental dataset 5).

Second, where it would be unethical to withhold already available efficacious treatments, trials will often include another active intervention in the comparator group,<sup>23</sup> for example, statins in the trials of effect of PCSK9 inhibitors<sup>97 98</sup> on LDL-C and cardiovascular events. Such RCT design can be mimicked by factorial MR estimating the interaction of multiple exposures, as shown in matching results of the equivalent MR study.<sup>16</sup>

#### Duration of intervention

While the magnitude of effect seen in trials with long (>3 years: weight loss to treat hypertension<sup>99 100</sup>/T2D,<sup>94–96</sup> blood pressure reduction to lower CVD risk<sup>65–67</sup>) and short (<6 months: alcohol intake reduction to lower blood pressure,<sup>101–104</sup> exercise to benefit BMD<sup>53</sup>/depression<sup>62 74 78–80 105</sup>) intervention may vary, we find both can result in directional effects consistent with MR results (figure 4), although with exceptions.<sup>73 106</sup>

#### Triangulation of MR and RCT results

Combining RCT and MR results can offer complimentary evidence on the effectiveness of interventions. Powerful examples include congruence of positive effect of high BMI on hypertension across MR<sup>107–110</sup> and RCT<sup>99 100</sup> studies, high BMI on T2D risk in MR<sup>89–93 111</sup> and RCTs<sup>94–96</sup> and the null effect of vitamin D on various glycaemic markers in diverse populations in MR,<sup>112</sup> RCTs<sup>113 114</sup> and RCT systematic review.<sup>115</sup>

We also found cases, where the majority of studies pointed to one direction of effect, with one MR or RCT identified as an outlier study. In these cases, having a wide array of MR and RCT studies (ideally meta-analysed) can be helpful in establishing the likely true causal direction of effect. For instance, two MR studies,<sup>116 117</sup> a meta-analysis of five RCT studies<sup>118</sup> and two RCTs<sup>119 120</sup> indicate no effect of vitamin E on prostate cancer incidence with one outlier RCT<sup>121</sup> showing benefit of vitamin E supplementation in older smokers. Similar contrary findings were found for one RCT<sup>122</sup> showing beneficial effect of vitamin D on preventing depressive symptoms, as opposed to null effect in four MR studies<sup>123–126</sup> and two RCTs.<sup>127 128</sup>

On the other hand, MR analyses can show spurious disagreement with the rest of the evidence base. For instance, 2 MR papers<sup>129 130</sup> and a meta-analysis of 16 RCT<sup>130</sup> studies reveal no significant effect of vitamin D on blood pressure in the general population, with the exception of one MR study<sup>131</sup> that indicated a blood pressure-lowering effect of higher vitamin D status. Similarly, a range of study types: 1 MR analysis,<sup>132</sup> 1 RCT<sup>133</sup> and a meta-analysis of 27 prospective cohorts<sup>134</sup> (only some of them RCTs) confirm a negative impact of smoking on HDL-C levels, bar one MR study showing no significant effect.<sup>135</sup> A series of RCT meta-analyses<sup>136–138</sup> support an effect of coffee consumption (especially unfiltered) on unfavourable blood profile, although this is likely explained by diterpenes<sup>139 140</sup> rather than caffeine, as the latter shows evidence of cardioprotective effects.<sup>141</sup> However, only the recent biggest MR study<sup>140</sup> to date



found a significant effect of coffee consumption on LDL-C and total-C levels, unlike two previous smaller MR analyses,<sup>142,143</sup> which found a non-significant directionally consistent relationship.

## DISCUSSION

Our study highlighted that sparsity of data in the electronic databases seriously hampers the ability to automatically parse and compare results of MR and RCT studies. Released for the first time in 2000, ClinicalTrials.gov is the most comprehensive resource for modern RCT (only <1000 studies, out of ~2 18 000 analysed RCTs were started before 2000). Nonetheless, we found that only 13% of all completed RCTs submitted their results to ClinicalTrials.gov, with median trial start date in 2012. Despite 2007 legislation requiring submission of RCT results to ClinicalTrials.gov within 1 year of completion (with exceptions),<sup>144</sup> only 38% of eligible trials for 2008–2012 submitted their results at any time<sup>145</sup> which rose to 64% for 2018–2019.<sup>146</sup> Furthermore, 60% of studies for failed agents are reported not to be published in peer-reviewed journals,<sup>147</sup> and in the work presented here we found MeSH annotations were missing from the majority of complex, behavioural and dietary interventions. These factors significantly hamper efforts to systematically triangulate RCT evidence with other studies.

Next, semantic triples describing conclusions of MR and RCT studies automatically extracted from literature abstracts using rule-based methods also had low coverage, with only 36% of MR and 12% of RCT studies associated with  $\geq 1$  triples. Consequently, we instead decided to focus on a detailed qualitative investigation of a series of case studies to identify the issues associated with triangulating MR and RCT studies

Combining RCT and MR results can offer complementary evidence on the effectiveness of interventions. Powerful examples include congruence of positive effect of high BMI on hypertension across MR<sup>107–110</sup> and RCT<sup>99,100</sup> studies, high BMI on T2D risk in MR<sup>89–93,111</sup> and RCTs<sup>94–96</sup> and the null effect of vitamin D on various glycaemic markers in diverse populations in MR,<sup>112</sup> RCTs<sup>113,114</sup> and RCT systematic review.<sup>115</sup> We also found cases, where the majority of studies pointed to one direction of effect, with one MR or RCT identified as an outlier study. In these cases, having a wide array of MR and RCT studies (ideally meta-analysed) can be helpful in establishing the likely true causal direction of effect.

Our analysis of genetically predicted effects of perturbation of drug target protein expression on a number of conditions with trials submitted to ClinicalTrials.gov revealed good concordance with established therapeutics for pQTLs. However, due to the limited number of proteins ( $n=1002$ ) and phenotypes ( $n=225$ , many non-diseases per se) in Zheng *et al.*,<sup>33,148</sup> the comparison is necessarily very preliminary. We identify only true positive cases, as false positives and true negatives are difficult to evaluate due to sparsity of drug clinical trial results

in ClinicalTrials.gov/literature<sup>147</sup> and inclusion of non-disease phenotypes in MR analysis. Anecdotally, we found no MR evidence that decreased expression of PLA2G2A leads to reduced CVD, which agrees with lack of efficacy of PLA2G2A inhibitor in clinical trials.<sup>149–151</sup>

The mixed reliability of eQTL instruments in predicting direction of effect on the outcome could be due to a number of factors such as less than perfect correlation between mRNA and protein levels,<sup>152</sup> hidden pleiotropy in single instruments used in the MR analysis (directly observed for IL2RA),<sup>153</sup> presence of negative feedback loop involved in the drug mechanism,<sup>154</sup> translation into protein isoforms with distinct biological effects<sup>155</sup> and differential cell-type specific drug effect.<sup>156</sup>

The duration of intervention varies between RCTs and MR studies, with the former spanning no more than the duration of the trial, while the latter can represent durations as long as the entire lifetime (although many exposures, such as alcohol intake, will be over a shorter time period).<sup>7</sup> Moreover, intervention in RCTs with long follow-up is not necessarily similarly intensive throughout its duration, or may cease altogether after some time,<sup>74,99</sup> that is, duration of follow-up is longer than duration of intervention in order to allow accumulation of enough events and/or confirm durability of intervention effect. Examples include lifestyle interventions, such as exercise<sup>74</sup> or weight loss programmes<sup>94</sup> like the Look AHEAD trial, with median follow-up of 9.6 years, where group and individual counselling sessions took place weekly in the first 6 months and tapered off over time.<sup>73</sup> That is why our analysis focused on comparing direction of effect, while ignoring magnitude of effect.<sup>157</sup> However, in certain cases when enough reference data are available, it is feasible to compare MR and RCT effects on the same exposure difference scale.<sup>158</sup>

Further impediments to direct comparison between MR and RCTs include differences in outcome definition (composite<sup>65</sup> vs single conditions<sup>70</sup>). Access to rare subpopulations with existing conditions, such as cancer patients receiving specific therapy<sup>159</sup> which are routinely exclusively enrolled into RCTs, can be difficult in MR due to the size of GWAS biobanks relative to  $N$  required for good power.

There are also a number of interventions and outcomes with no single phenotype which could be instrumented with GWAS variants, making MR approaches difficult, although sometimes possible with innovative MR approaches.<sup>160</sup> This is especially true of lifestyle interventions—such as different forms of psychological therapy, complex diet regimens<sup>161</sup> and fasting. Absent or limited heritability of a number of interventions and conditions, such as rehabilitation and traumatic injury makes MR approaches inaccessible.

The majority of MR studies track the onset rather than progression of disease due to availability of GWAS phenotypes<sup>162</sup> which are often a (binary) single measurement, as opposed to multiple quantitative outcomes frequently measured in RCTs.<sup>163</sup> For that reason, triangulation of

MR of onset with RCTs whose intervention is targeting progression of disease, may or may not result in agreement, as seen in our comparison of the effect of exercise on schizophrenia onset/progression (discordant) or depression (concordant) and vitamin D effect on atopic dermatitis onset/progression (discordant).

Many MR studies may be underpowered due to large sample required in indirect estimation<sup>164</sup> as these studies are typically studies of convenience. This bias is less common in RCTs due to preregistration of study design including power analysis,<sup>1</sup> uncommon in MR.<sup>165</sup> Null effect in MR studies may be therefore spurious and not predictive of RCTs for that reason, as seen in two smaller MR studies<sup>142 143</sup> out of three<sup>140</sup> investigating the effect of coffee intake on blood lipids, contrasting with strong clinical trial<sup>136–138</sup> and biochemical evidence.<sup>139 140 166</sup>

Furthermore, the presented literature survey used a simple heuristic of reported statistically significant evidence ( $p$  value  $< 0.05$  after multiple testing correction) to compare conclusions across MR and RCT studies, which has well-known limitations.<sup>167 168</sup> Inclusion of the full-spectrum of scaled point estimates along with their confidence intervals will reveal a more detailed picture in triangulation of MR and RCT evidence (online supplemental box).

Overall, we find that due to difficulty in identifying sufficient number of MR–RCT pairs matched for the same exposure and outcome, we cannot derive a numerical model to quantify reliability and importance of features of MR analysis in predicting the outcome of a future RCT. However, we make several general observations regarding usefulness of triangulation<sup>26</sup> of RCT with MR to guide MR studies. If an RCT shows a causal relationship between an intervention and an outcome which corresponds to the one observed in MR, it can help validate the use of these genetic variants as instruments in future MR studies. Moreover, RCTs can inform MR analyses about the plausible effect sizes and so can be useful for power calculations in MR. RCTs can help identify important interactions and subgroup effects, which can further inform MR study design. For instance, if an RCT identifies that a treatment has a stronger effect in a particular subgroup of individuals (eg, women, children), they could be analysed separately using one-sample MR.

## CONCLUSIONS

Our research highlights the challenges and benefits of triangulation of MR with RCT evidence. Future efforts, outside of the scope of this work, will focus on fully quantitative approaches towards triangulation, involving magnitude of effect size and not just its presence and direction.<sup>25</sup> Developers of such methods will need to be mindful of discrepancies in research hypothesis, duration and intensity of exposure, outcome measures, intervention aim, underlying population characteristics, violations of test assumptions as well as statistical power of the analysis. Furthermore, automated triangulation based

on electronic databases requires intensive effort towards structured capture of both MR and RCT study results and associated meta-data, as well as annotation with shared ontologies, which is still challenging using current natural language processing methods, despite constant progress.<sup>43 169 170</sup>

## Author affiliations

<sup>1</sup>MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, UK

<sup>2</sup>Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China

<sup>3</sup>Shanghai National Clinical Research Center for Metabolic Diseases, Key Laboratory for Endocrine and Metabolic Diseases of the National Health Commission of the PR China, Shanghai Key Laboratory for Endocrine Tumor, State Key Laboratory of Medical Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China

**Contributors** TRG conceived and supervised the study as well as acquired funding. MKS curated the data and performed the main analyses. JZ inspired the study and supplied xQTL analyses. MKS, GDS, JZ and TRG wrote the manuscript. MKS is the guarantor for this work and accepts full responsibility for the work.

**Funding** We would like to acknowledge UK Medical Research Council [mc\_uu\_00011/4] which provided funding for our work carried out at the University of Bristol MRC Integrative Epidemiology Unit. Our funder played no role in the design of the study, analysis and interpretation of data and in writing of the manuscript. GDS works within the MRC Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council (MC\_UU\_00011/1).

**Competing interests** JZ, GDS and TG receive funding from Biogen for unrelated research.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. All data generated or analysed during this study are included in this published article and its supplementary information files. Code used to carry out the analysis is available on GitHub: <https://github.com/marynias/mr-rct>. ClinicalTrials.gov data were accessed via AACT: <https://aact.ctti-clinicaltrials.org/download> and analysed data subset is featured in Supplementary datasets 1 and 2. pQTL and eQTL MR analysis results are available via EpigraphDB: <https://epigraphdb.org/xqtl>. PubMed database can be accessed on <https://pubmed.ncbi.nlm.nih.gov/> and analysed data subset is featured in Supplementary dataset 4. SemMedDB can be accessed on [https://lhncbc.nlm.nih.gov/ii/tools/SemRep\\_SemMedDB\\_SKR/SemMedDB\\_download.html](https://lhncbc.nlm.nih.gov/ii/tools/SemRep_SemMedDB_SKR/SemMedDB_download.html) and analysed data subset is featured in Supplementary dataset 4. Case series of MR and RCT studies with matching exposures (interventions) and outcomes (conditions) is featured in Supplementary dataset 5. Supplementary datasets 1–5 are available for download on Zenodo (<https://doi.org/10.5281/zenodo.8104176>).

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

## ORCID iDs

Maria K Sobczyk <http://orcid.org/0000-0003-0000-4100>  
 Jie Zheng <http://orcid.org/0000-0002-6623-6839>  
 George Davey Smith <http://orcid.org/0000-0002-1407-8314>  
 Tom R Gaunt <http://orcid.org/0000-0003-0924-3247>

## REFERENCES

- Ference BA, Holmes MV, Smith GD. Using Mendelian randomization to improve the design of randomized trials. *Cold Spring Harb Perspect Med* 2021;11:a040980.
- Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease. *Int J Epidemiol* 2003;32:1–22.
- Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers* 2022;2:6.
- Hingorani A, Humphries S. Nature's randomised trials. *Lancet* 2005;366:1906–8.
- Davey Smith G, Holmes MV, Davies NM, et al. Mendel's laws, Mendelian randomization and causal inference in observational data: substantive and nomenclatural issues. *Eur J Epidemiol* 2020;35:99–111.
- Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601.
- Swanson SA, Tiemeier H, Ikram MA, et al. Nature as a trialist?: Deconstructing the analogy between Mendelian randomization and randomized trials. *Epidemiology* 2017;28:653–9.
- Martin L, Hutchens M, Hawkins C. Trial watch: clinical trial cycle times continue to increase despite industry efforts. *Nat Rev Drug Discov* 2017;16:157.
- Egger M, Buitrago-Garcia D, Smith GD. Chapter 19: systematic reviews of epidemiological studies of etiology and prevalence. In: Egger M, Higgins JPT, Smith GD, eds. *Systematic reviews in health care: meta-analyses in context*. Wiley-Blackwell, 2022.
- Voight BF, Peloso GM, Orho-Melander M, et al. Plasma HDL cholesterol and risk of myocardial infarction: a Mendelian randomisation study. *Lancet* 2012;380:572–80.
- Verdoia M, Schaffer A, Suryapranata H, et al. Effects of HDL-modifiers on cardiovascular outcomes: a meta-analysis of randomized trials. *Nutr Metab Cardiovasc Dis* 2015;25:9–23.
- Davey Smith G, Phillips AN. Correlation without a cause: an epidemiological odyssey. *Int J Epidemiol* 2020;49:4–14.
- Yarmolinsky J, Bonilla C, Haycock PC, et al. Circulating selenium and prostate cancer risk: a Mendelian randomization analysis. *J Natl Cancer Inst* 2018;110:1035–8.
- Lippman SM, Klein EA, Goodman PJ, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the selenium and vitamin E cancer prevention trial (SELECT). *JAMA* 2009;301:39–51.
- Ference BA, Yoo W, Alesh I, et al. Effect of long-term exposure to lower low-density lipoprotein cholesterol beginning early in life on the risk of coronary heart disease. *J Am Coll Cardiol* 2012;60:2631–9.
- Ference BA, Robinson JG, Brook RD, et al. Variation in Pcsk9 and HMGCR and risk of cardiovascular disease and diabetes. *N Engl J Med* 2016;375:2144–53.
- Hegele RA. Editorial comment: when Mendelian randomization goes astray. *Curr Opin Lipidol* 2021;32:79–80.
- Turner RM, Spiegelhalter DJ, Smith GCS, et al. Bias modelling in evidence synthesis. *J R Stat Soc Ser A Stat Soc* 2009;172:21–47.
- Morris TT, Heron J, Sanderson E, et al. Interpretation of Mendelian randomization using one measure of an exposure that varies over time. *Epidemiology* [Preprint] 2021.
- Victoria CG, Habicht J-P, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health* 2004;94:400–5.
- Chekroud AM. A Mendelian randomization approach for assessing the relationship between physical activity and depression. *JAMA Psychiatry* 2019;76:361–2.
- Harrison S, Dixon P, Jones HE, et al. Long-term cost-effectiveness of interventions for obesity: a Mendelian randomisation study. *PLOS Med* 2021;18:e1003725.
- Holmes MV, Richardson TG, Ference BA, et al. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat Rev Cardiol* 2021;18:435–53.
- 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;41:2313–30.
- Munafò MR, Higgins JPT, Smith GD. Triangulating evidence through the inclusion of genetically informed designs. *Cold Spring Harb Perspect Med* 2021;11:a040659.
- Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol* 2016;45:dyw314.
- Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov results database--update and key issues. *N Engl J Med* 2011;364:852–60.
- Liu Y, Elsworth B, Erola P, et al. EpiGraphDB: a database and data mining platform for health data science. *Bioinformatics* [Preprint].
- Tasneem A, Aberle L, Ananth H, et al. The database for aggregate analysis of ClinicalTrials.gov (AACT) and subsequent regrouping by clinical specialty. *PLoS One* 2012;7:e33677.
- Tenaerts P, Madre L, Archdeacon P, et al. The clinical trials transformation initiative: innovation through collaboration. *Nat Rev Drug Discov* 2014;13:797–8.
- Ochoa D, Hercules A, Carmona M, et al. Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Res* 2021;49:D1302–10.
- 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:D1074–82.
- Zheng J, Haberland V, Baird D, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet* 2020;52:1122–31.
- Kilicoglu H, Shin D, Fiszman M, et al. SemMedDB: a PubMed-scale repository of biomedical semantic predications. *Bioinformatics* 2012;28:3158–60.
- Elsworth B, Gaunt TR, Wren J. MELODI presto: a fast and agile tool to explore semantic triples derived from biomedical literature. *Bioinformatics* 2021;37:583–5.
- Mokdad AH, Marks JS, Stroup DF, et al. Actual causes of death in the United States, 2000. *JAMA* 2004;291:1238–45.
- Carnegie R, Zheng J, Sallis HM, et al. Mendelian randomisation for nutritional psychiatry. *Lancet Psychiatry* 2020;7:208–16.
- Cicero AFG, Tartagni E, Ertek S. Safety and tolerability of injectable lipid-lowering drugs: a review of available clinical data. *Expert Opin Drug Saf* 2014;13:1023–30.
- Luo J, Wu S-J, Lacy ER, et al. Structural basis for the dual recognition of IL-12 and IL-23 by Ustekinumab. *J Mol Biol* 2010;402:797–812.
- Hansen RA, Gartlehner G, Webb AP, et al. Efficacy and safety of donepezil, galantamine, and rivastigmine for the treatment of Alzheimer's disease: a systematic review and meta-analysis. *Clin Interv Aging* 2008;3:211–25.
- Ference BA, Kastelein JJP, Ginsberg HN, et al. Association of genetic variants related to CETP inhibitors and statins with lipoprotein levels and cardiovascular risk. *JAMA* 2017;318:947–56.
- Piepho RW. Overview of the angiotensin-converting-enzyme inhibitors. *Am J Health Syst Pharm* 2000;57:S3–7.
- Marshall IJ, Nye B, Kuiper J, et al. Trialstreamer: a living, automatically updated database of clinical trial reports. *J Am Med Assoc* 2020;27:1903–12.
- Mokry LE, Ross S, Ahmad OS, et al. Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. *PLoS Med* 2015;12:e1001866.
- James E, Dobson R, Kuhle J, et al. The effect of vitamin D-related interventions on multiple sclerosis relapses: a meta-analysis. *Mult Scler* 2013;19:1571–9.
- Hempel S, Graham GD, Fu N, et al. A systematic review of the effects of Modifiable risk factor interventions on the progression of multiple sclerosis. *Mult Scler* 2017;23:513–24.
- Zhuang Z, Gao M, Yang R, et al. Association of physical activity, sedentary behaviours and sleep duration with cardiovascular diseases and lipid profiles: a Mendelian randomization analysis. *Lipids Health Dis* 2020;19:86.
- Xu F, Zhang Q, Wang L-K, et al. Estimates of the effects of physical activity on osteoporosis using multivariable Mendelian randomization analysis. *Osteoporos Int* 2021;32:1359–67.
- Kelley GA, Kelley KS, Tran ZV. Walking, lipids, and lipoproteins: a meta-analysis of randomized controlled trials. *Prev Med* 2004;38:651–61.
- Kelley GA, Kelley KS. Aerobic exercise and lipids and lipoproteins in men: a meta-analysis of randomized controlled trials. *J Mens Health Gen* 2006;3:61–70.
- Kelley GA, Kelley KS, Tran ZV. Aerobic exercise and lipids and lipoproteins in women: a meta-analysis of randomized controlled trials. *J Womens Health (Larchmt)* 2004;13:1148–64.

- 52 Kelley GA, Kelley KS. Impact of progressive resistance training on lipids and lipoproteins in adults: a meta-analysis of randomized controlled trials. *Prev Med* 2009;48:9–19.
- 53 Mosti MP, Carlsen T, Aas E, et al. Maximal strength training improves bone mineral density and neuromuscular performance in young adult women. *J Strength Cond Res* 2014;28:2935–45.
- 54 Allison SJ, Folland JP, Rennie WJ, et al. High impact exercise increased femoral neck bone mineral density in older men: a randomised unilateral intervention. *Bone* 2013;53:321–8.
- 55 Katzel LI, Bleecker ER, Colman EG, et al. Effects of weight loss vs aerobic exercise training on risk factors for coronary disease in healthy, obese, middle-aged and older men. A randomized controlled trial. *JAMA* 1995;274:1915–21.
- 56 Çolak Y, Afzal S, Nordestgaard BG. 25-Hydroxyvitamin D and risk of osteoporotic fractures: Mendelian randomization analysis in 2 large population-based cohorts. *Clin Chem* 2020;66:676–85.
- 57 Bischoff-Ferrari HA, Willett WC, Orav EJ, et al. A pooled analysis of vitamin D dose requirements for fracture prevention. *N Engl J Med* 2012;367:40–9.
- 58 Bischoff-Ferrari HA, Willett WC, Wong JB, et al. Prevention of nonvertebral fractures with oral vitamin D and dose dependency: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2009;169:551–61.
- 59 Bischoff-Ferrari HA, Willett WC, Wong JB, et al. Fracture prevention with vitamin D supplementation: a meta-analysis of randomized controlled trials. *JAMA* 2005;293:2257–64.
- 60 Papiol S, Schmitt A, Maurus I, et al. Association between physical activity and schizophrenia: results of a 2-sample Mendelian randomization analysis. *JAMA Psychiatry* 2021;78:441–4.
- 61 Sun H, Gao X, Que X, et al. The causal relationships of device-measured physical activity with bipolar disorder and schizophrenia in adults: a 2-sample Mendelian randomization study. *J Affect Disord* 2020;263:598–604.
- 62 Rosenbaum S, Tiedemann A, Sherrington C, et al. Physical activity interventions for people with mental illness: a systematic review and meta-analysis. *J Clin Psychiatry* 2014;75:964–74.
- 63 Firth J, Cotter J, Elliott R, et al. A systematic review and meta-analysis of exercise interventions in schizophrenia patients. *Psychol Med* 2015;45:1343–61.
- 64 Sabe M, Kaiser S, Sentissi O. Physical exercise for negative symptoms of schizophrenia: systematic review of randomized controlled trials and meta-analysis. *Gen Hosp Psychiatry* 2020;62:13–20.
- 65 Williamson JD, Supiano MA, Applegate WB, et al. Intensive vs standard blood pressure control and cardiovascular disease outcomes in adults aged  $\geq 75$  years: A randomized clinical trial. *JAMA* 2016;315:2673–82.
- 66 Brunström M, Carlberg B. Association of blood pressure lowering with mortality and cardiovascular disease across blood pressure levels: a systematic review and meta-analysis. *JAMA Intern Med* 2018;178:28–36.
- 67 Rahimi K, Bidel Z, Nazarzadeh M, et al. Age-stratified and blood-pressure-stratified effects of blood-pressure-lowering pharmacotherapy for the prevention of cardiovascular disease and death: an individual participant-level data meta-analysis. *Lancet* 2021;398:1053–64.
- 68 Gill D, Georgakis MK, Zuber V, et al. Genetically predicted midlife blood pressure and coronary artery disease risk: Mendelian randomization analysis. *J Am Heart Assoc* 2020;9:e016773.
- 69 Malik R, Georgakis MK, Vujkovic M, et al. Relationship between blood pressure and incident cardiovascular disease: linear and nonlinear Mendelian randomization analyses. *Hypertension* 2021;77:2004–13.
- 70 Wan EYF, Fung WT, Schooling CM, et al. Blood pressure and risk of cardiovascular disease in UK Biobank. *Hypertension* 2021;77:367–75.
- 71 Riaz H, Khan MS, Siddiqi TJ, et al. Association between obesity and cardiovascular outcomes: a systematic review and meta-analysis of Mendelian randomization studies. *JAMA Netw Open* 2018;1:e183788.
- 72 Larsson SC, Bäck M, Rees JMB, et al. Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *Eur Heart J* 2020;41:221–6.
- 73 The Look AHEAD Research Group. Cardiovascular effects of intensive lifestyle intervention in type 2 diabetes. *N Engl J Med* 2013;369:145–54.
- 74 Mather AS, Rodriguez C, Guthrie MF, et al. Effects of exercise on depressive symptoms in older adults with poorly responsive depressive disorder: randomised controlled trial. *Br J Psychiatry* 2002;180:411–5.
- 75 Choi KW, Chen C-Y, Stein MB, et al. Assessment of bidirectional relationships between physical activity and depression among adults: a 2-sample Mendelian randomization study. *JAMA Psychiatry* 2019;76:399–408.
- 76 Choi KW, Stein MB, Nishimi KM, et al. An exposure-wide and Mendelian randomization approach to identifying modifiable factors for the prevention of depression. *Am J Psychiatry* 2020;177:944–54.
- 77 Veale D, Le Fevre K, Pantelis C, et al. Aerobic exercise in the adjunctive treatment of depression: a randomized controlled trial. *J R Soc Med* 1992;85:541–4.
- 78 Schuch FB, Vancampfort D, Richards J, et al. Exercise as a treatment for depression: a meta-analysis adjusting for publication bias. *J Psychiatr Res* 2016;77:42–51.
- 79 Kvam S, Kleppe CL, Nordhus IH, et al. Exercise as a treatment for depression: a meta-analysis. *J Affect Disord* 2016;202:67–86.
- 80 Gordon BR, McDowell CP, Hallgren M, et al. Association of efficacy of resistance exercise training with depressive symptoms: meta-analysis and meta-regression analysis of randomized clinical trials. *JAMA Psychiatry* 2018;75:566–76.
- 81 Manousaki D, Paternoster L, Standl M, et al. Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: a Mendelian randomization study. *PLoS Med* 2017;14:e1002294.
- 82 Drodge DR, Budu-Aggrey A, Paternoster L. Causal analysis shows evidence of atopic dermatitis leading to an increase in vitamin D levels. *J Invest Dermatol* 2021;141:1339–41.
- 83 Camargo CA, Ganmaa D, Sidbury R, et al. Randomized trial of vitamin D supplementation for winter-related Atopic dermatitis in children. *J Allergy Clin Immunol* 2014;134:831–5.
- 84 Sidbury R, Sullivan AF, Thadhani RI, et al. Randomized controlled trial of vitamin D supplementation for winter-related atopic dermatitis in Boston: a pilot study. *Br J Dermatol* 2008;159:245–7.
- 85 Amestegani M, Salehi BS, Vasigh M, et al. Vitamin D supplementation in the treatment of atopic dermatitis: a clinical trial study. *J Drugs Dermatology* 2012;11.
- 86 Javanbakht MH, Keshavarz SA, Djalali M, et al. Randomized controlled trial using vitamins E and D supplementation in atopic dermatitis. *J Dermatolog Treat* 2011;22:144–50.
- 87 Meisinger C, Linseisen J, Leitzmann M, et al. Association of physical activity and sedentary behavior with type 2 diabetes and Glycemic traits: a two-sample Mendelian randomization study. *BMJ Open Diab Res Care* 2020;8:e001896.
- 88 Shah SZA, Karam JA, Zeb A, et al. Movement is improvement: the therapeutic effects of exercise and general physical activity on Glycemic control in patients with type 2 diabetes mellitus: A systematic review and meta-analysis of randomized controlled trials. *Diabetes Ther* 2021;12:707–32.
- 89 Corbin LJ, Richmond RC, Wade KH, et al. BMI as a Modifiable risk factor for type 2 diabetes: refining and understanding causal estimates using Mendelian randomization. *Diabetes* 2016;65:3002–7.
- 90 Dale CE, Fatemifar G, Palmer TM, et al. Causal associations of adiposity and body fat distribution with coronary heart disease, stroke subtypes, and type 2 diabetes mellitus. *Circulation* 2017;135:2373–88.
- 91 Cheng L, Zhuang H, Ju H, et al. Exposing the causal effect of body mass index on the risk of type 2 diabetes mellitus: a Mendelian randomization study. *Front Genet* 2019;10:94.
- 92 Wainberg M, Mahajan A, Kundaje A, et al. Homogeneity in the association of body mass index with type 2 diabetes across the UK Biobank: a Mendelian randomization study. *PLOS Med* 2019;16:e1002982.
- 93 Xu H, Jin C, Guan Q. Causal effects of overall and abdominal obesity on insulin resistance and the risk of type 2 diabetes mellitus: A two-sample Mendelian randomization study. *Front Genet* 2020;11:603.
- 94 Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the DA Qing IGT and diabetes study. *Diabetes Care* 1997;20:537–44.
- 95 Tuomilehto J, Lindström J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001;344:1343–50.
- 96 Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393–403.
- 97 Sabatine MS, Giugliano RP, Wiviott SD, et al. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. *N Engl J Med* 2015;372:1500–9.
- 98 Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and clinical outcomes in patients with cardiovascular disease. *N Engl J Med* 2017;376:1713–22.

- 99 He J, Whelton PK, Appel LJ, *et al.* Long-term effects of weight loss and dietary sodium reduction on incidence of hypertension. *Hypertension* 2000;35:544–9.
- 100 Stevens VJ. Long-term weight loss and changes in blood pressure: results of the trials of hypertension prevention, phase II. *Ann Intern Med* 2001;134:1.
- 101 Puddey IB, Beilin LJ, Vandongen R, *et al.* Evidence for a direct effect of alcohol consumption on blood pressure in normotensive men. A randomized controlled trial. *Hypertension* 1985;7:707–13.
- 102 Puddey IB, Beilin LJ, Vandongen R. Regular alcohol use raises blood pressure in treated hypertensive subjects. A randomised controlled trial. *Lancet* 1987;1:647–51.
- 103 Xin X, He J, Frontini MG, *et al.* Effects of alcohol reduction on blood pressure. *Hypertension* 2001;38:1112–7.
- 104 Puddey IB, Parker M, Beilin LJ, *et al.* Effects of alcohol and caloric restrictions on blood pressure and serum lipids in overweight men. *Hypertension* 1992;20:533–41.
- 105 Helgadóttir B, Hallgren M, Ekblom Ö, *et al.* Training fast or slow? Exercise for depression: a randomized controlled trial. *Prev Med* 2016;91:123–31.
- 106 Ma C, Avenell A, Bolland M, *et al.* Effects of weight loss interventions for adults who are obese on mortality, cardiovascular disease, and cancer: systematic review and meta-analysis. *BMJ* 2017;359:j4849.
- 107 Timpson NJ, Harbord R, Davey Smith G, *et al.* Does greater Adiposity increase blood pressure and hypertension risk. *Hypertension* 2009;54:84–90.
- 108 Lyall DM, Celis-Morales C, Ward J, *et al.* Association of body mass index with Cardiometabolic disease in the UK Biobank: A Mendelian randomization study. *JAMA Cardiol* 2017;2:882–9.
- 109 Lee M-R, Lim Y-H, Hong Y-C. Causal Association of body mass index with hypertension using a Mendelian randomization design. *Medicine (Baltimore)* 2018;97:e11252.
- 110 van Oort S, Beulens JWW, van Ballegooijen AJ, *et al.* Association of cardiovascular risk factors and lifestyle behaviors with hypertension. *Hypertension* 2020;76:1971–9.
- 111 Rivera CM, Ren B. Mapping human epigenomes. *Cell* 2013;155:39–55.
- 112 Wang N, Wang C, Chen X, *et al.* Vitamin D, prediabetes and type 2 diabetes: bidirectional Mendelian randomization analysis. *Eur J Nutr* 2020;59:1379–88.
- 113 Salehpour A, Shidfar F, Hosseinpanah F, *et al.* Does vitamin D3 supplementation improve glucose homeostasis in overweight or obese women? A double-blind, randomized, placebo-controlled clinical trial. *Diabet Med* 2013;30:1477–81.
- 114 Sollid ST, Hutchinson MYS, Fuskevåg OM, *et al.* No effect of high-dose vitamin D supplementation on glycemic status or cardiovascular risk factors in subjects with prediabetes. *Diabetes Care* 2014;37:2123–31.
- 115 Zuk A, Fitzpatrick T, Rosella LC. Effect of vitamin D3 supplementation on inflammatory markers and glycemic measures among overweight or obese adults: a systematic review of randomized controlled trials. *PLoS One* 2016;11:e0154215.
- 116 Wright ME, Peters U, Gunter MJ, *et al.* Association of variants in two vitamin E transport genes with circulating vitamin E concentrations and prostate cancer risk. *Cancer Res* 2009;69:1429–38.
- 117 Cheng W-W, Wang Z-K, Shangquan H-F, *et al.* Are vitamins relevant to cancer risks? A Mendelian randomization investigation. *Nutrition* 2020;78:110870.
- 118 Jiang L, Yang K, Tian J, *et al.* Efficacy of antioxidant vitamins and selenium supplement in prostate cancer prevention: a meta-analysis of randomized controlled trials. *Nutr Cancer* 2010;62:719–27.
- 119 Gaziano JM, Glynn RJ, Christen WG, *et al.* Vitamins E and C in the prevention of prostate and total cancer in men: the physicians' health study II randomized controlled trial. *JAMA* 2009;301:52–62.
- 120 Wang L, Sesso HD, Glynn RJ, *et al.* Vitamin E and C supplementation and risk of cancer in men: posttrial follow-up in the physicians' health study II randomized trial. *Am J Clin Nutr* 2014;100:915–23.
- 121 Heinonen OP, Koss L, Albanes D, *et al.* Prostate cancer and supplementation with A-tocopherol and B-carotene: incidence and mortality in a controlled trial. *JNCJ* 1998;90:440–6.
- 122 Vieth R, Kimball S, Hu A, *et al.* Randomized comparison of the effects of the vitamin D3 Adequate intake versus 100 Mcg (4000 IU) per day on biochemical responses and the wellbeing of patients. *Nutr J* 2004;3:8.
- 123 Michaëlsson K, Melhus H, Larsson SC. Serum 25-hydroxyvitamin D concentrations and major depression: a Mendelian randomization study. *Nutrients* 2018;10:1987.
- 124 Libuda L, Laabs B-H, Ludwig C, *et al.* Vitamin D and the risk of depression: A causal relationship? findings from a Mendelian randomization study. *Nutrients* 2019;11:1085.
- 125 Mulugeta A, Lumsden A, Hyppönen E. Relationship between serum 25(OH)D and depression: causal evidence from a bi-directional Mendelian randomization study. *Nutrients* 2021;13:109.
- 126 Milaneschi Y, Peyrot WJ, Nivard MG, *et al.* A role for vitamin D and Omega-3 fatty acids in major depression? an exploration using Genomics. *Transl Psychiatry* 2019;9:219.
- 127 Bertone-Johnson ER, Powers SI, Spangler L, *et al.* Vitamin D supplementation and depression in the women's health initiative calcium and vitamin D trial. *Am J Epidemiol* 2012;176:1–13.
- 128 Kjærgaard M, Waterloo K, Wang CEA, *et al.* Effect of vitamin D supplement on depression scores in people with low levels of serum 25-Hydroxyvitamin D: nested case–control study and randomised clinical trial. *Br J Psychiatry* 2012;201:360–8.
- 129 Skaaby T, Husemoen LLN, Martinussen T, *et al.* Vitamin D status, filaggrin genotype, and cardiovascular risk factors: a Mendelian randomization approach. *PLoS ONE* 2013;8:e57647.
- 130 Kunutsor SK, Burgess S, Munroe PB, *et al.* Vitamin D and high blood pressure: causal association or epiphenomenon. *Eur J Epidemiol* 2014;29:1–14.
- 131 Vimalaswaran KS, Cavadino A, Berry DJ, *et al.* Association of vitamin D status with arterial blood pressure and hypertension risk: a Mendelian Randomisation study. *Lancet Diabetes Endocrinol* 2014;2:719–29.
- 132 Rosoff DB, Davey Smith G, Mehta N, *et al.* Evaluating the relationship between alcohol consumption, tobacco use, and cardiovascular disease: A multivariable Mendelian randomization study. *PLoS Med* 2020;17:e1003410.
- 133 Gepner AD, Piper ME, Johnson HM, *et al.* Effects of smoking and smoking cessation on lipids and lipoproteins: outcomes from a randomized clinical trial. *Am Heart J* 2011;161:145–51.
- 134 Maeda K, Noguchi Y, Fukui T. The effects of cessation from cigarette smoking on the lipid and lipoprotein profiles: a meta-analysis. *Prev Med* 2003;37:283–90.
- 135 Åsvold BO, Bjørngaard JH, Carlslake D, *et al.* Causal associations of tobacco smoking with cardiovascular risk factors: a Mendelian randomization analysis of the HUNT study in Norway. *Int J Epidemiol* 2014;43:1458–70.
- 136 Jee SH, He J, Appel LJ, *et al.* Coffee consumption and serum lipids: a meta-analysis of randomized controlled clinical trials. *Am J Epidemiol* 2001;153:353–62.
- 137 Cai L, Ma D, Zhang Y, *et al.* The effect of coffee consumption on serum lipids: a meta-analysis of randomized controlled trials. *Eur J Clin Nutr* 2012;66:872–7.
- 138 Du Y, Lv Y, Zha W, *et al.* Effect of coffee consumption on dyslipidemia: a meta-analysis of randomized controlled trials. *Nutr Metab Cardiovasc Dis* 2020;30:2159–70.
- 139 Urgert R, Katan MB. The cholesterol-raising factor from coffee beans. *Annu Rev Nutr* 1997;17:305–24.
- 140 Zhou A, Hyppönen E. Habitual coffee intake and plasma lipid profile: evidence from UK Biobank. *Clin Nutr* 2021;40:4404–13.
- 141 Lebeau PF, Byun JH, Platko K, *et al.* Caffeine blocks SREBP2-induced hepatic PCSK9 expression to enhance LDLR-mediated cholesterol clearance. *Nat Commun* 2022;13:770.
- 142 Kwok MK, Leung GM, Schooling CM. Habitual coffee consumption and risk of type 2 diabetes, ischemic heart disease, depression and Alzheimer's disease: a Mendelian randomization study. *Sci Rep* 2016;6:36500.
- 143 Nordestgaard AT, Thomsen M, Nordestgaard BG. Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study. *Int J Epidemiol* 2015;44:551–65.
- 144 Zarin DA, Tse T, Williams RJ, *et al.* Trial reporting in ClinicalTrials.gov – the final rule. *N Engl J Med* 2016;375:1998–2004.
- 145 Anderson ML, Chiswell K, Peterson ED, *et al.* Compliance with results reporting at ClinicalTrials.gov. *N Engl J Med* 2015;372:1031–9.
- 146 DeVito NJ, Bacon S, Goldacre B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* 2020;395:361–9.
- 147 Hwang TJ, Carpenter D, Lauffenburger JC, *et al.* Failure of investigational drugs in late-stage clinical development and publication of trial results. *JAMA Intern Med* 2016;176:1826–33.
- 148 Understanding Society Scientific Group, International Consortium for Blood Pressure, Blood Pressure-International Consortium of Exome Chip Studies. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet* 2019;51:51–62.
- 149 Holmes MV, Simon T, Exeter HJ. Secretory Phospholipase A2-IIA and cardiovascular disease. *J Am Coll Cardiol* 2013;62:1966–76.

- 150 Rosenson RS, Hislop C, Elliott M, *et al.* Effects of Varespladib methyl on biomarkers and major cardiovascular events in acute coronary syndrome patients. *J Am Coll Cardiol* 2010;56:1079–88.
- 151 Nicholls SJ, Kastelein JJP, Schwartz GG, *et al.* Varespladib and cardiovascular events in patients with an acute coronary syndrome: the VISTA-16 randomized clinical trial. *JAMA* 2014;311:252–62.
- 152 Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. *Cell* 2016;165:535–50.
- 153 Verbanck M, Chen C-Y, Neale B, *et al.* Detection of widespread horizontal Pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 2018;50:693–8.
- 154 Foulds CE. Disrupting a negative feedback loop drives endocrine therapy-resistant breast cancer. *Proc Natl Acad Sci U S A* 2018;115:8236–8.
- 155 Cupido AJ, Asselbergs FW, Natarajan P, *et al.* Dissecting the IL-6 pathway in cardiometabolic disease: a Mendelian randomization study on both IL6 and IL6R. *Br J Clin Pharmacol* 2022;88:2875–84.
- 156 Gallo MA, Kaufman D. Antagonistic and agonistic effects of tamoxifen: significance in human cancer. *Semin Oncol* 1997;24:S1–71.
- 157 Ference BA. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur Heart J* 2018;39:360–2.
- 158 Holmes MV, Smith GD. Revealing the effect of CETP inhibition in cardiovascular disease. *Nat Rev Cardiol* 2017;14:635–6.
- 159 Kearns AE, Northfelt DW, Dueck AC, *et al.* Osteoporosis prevention in prostate cancer patients receiving androgen ablation therapy: placebo-controlled double-blind study of estradiol and Risedronate: N01C8. *Support Care Cancer* 2010;18:321–8.
- 160 Li J, Guasch-Ferré M, Chung W, *et al.* The Mediterranean diet, plasma Metabolome, and cardiovascular disease risk. *Eur Heart J* 2020;41:2645–56.
- 161 Verfürden ML, Gilbert R, Lucas A, *et al.* Effect of Nutritionally modified infant formula on academic performance: linkage of seven dormant randomised controlled trials to national education data. *BMJ* 2021;375:e065805.
- 162 Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease Therapeutics: conceptual and methodological challenges. *PLoS Genet* 2017;13:e1006944.
- 163 Stubbs B, Vancampfort D, Hallgren M, *et al.* EPA guidance on physical activity as a treatment for severe mental illness: a meta-review of the evidence and position statement from the European psychiatric Association (EPA), supported by the International Organization of Physical Therapists in Mental Health (LOPTMH). *Eur Psychiatry* 2018;54:124–44.
- 164 Schooling CM, Freeman G, Cowling BJ. Mendelian randomization and estimation of treatment efficacy for chronic diseases. *Am J Epidemiol* 2013;177:1128–33.
- 165 Lor GCY, Risch HA, Fung WT, *et al.* Reporting and guidelines for Mendelian randomization analysis: a systematic review of oncological studies. *Cancer Epidemiol* 2019;62:101577.
- 166 Poole R, Kennedy OJ, Roderick P, *et al.* Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *BMJ* 2017;359:j5024.
- 167 Altman DG, Bland JM. Statistics notes: absence of evidence is not evidence of absence. *BMJ* 1995;311:485.
- 168 Sterne JAC, Smith GD. Sifting the evidence—what's wrong with significance tests *Phys Ther* 2001;81:1464–9.
- 169 Nye BE, DeYoung J, Lehman E, *et al.* Understanding clinical trial reports: extracting medical entities and their relations. *AMIA Jt Summits Transl Sci* 2021;2021:485–94. Available: <https://pubmed.ncbi.nlm.nih.gov/34457164>
- 170 Whitton J, Hunter A. Automated tabulation of clinical trial results: A joint entity and relation extraction approach with transformer-based language representations. [Preprint] 2021.
- 171 Nitsch D, Molokhia M, Smeeth L, *et al.* Limits to causal inference based on Mendelian randomization: a comparison with randomized controlled trials. *American Journal of Epidemiology* 2006;163:397–403.
- 172 Ebrahim S, Davey Smith G. Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? *Human Genetics* 2008;123:15–33.