Systematic comparison of Mendelian randomization studies and randomized controlled trials using electronic databases

Maria K. Sobczyk[1]*, Jie Zheng[1,2,3], George Davey Smith[1], Tom R. Gaunt[1]

*To whom correspondence should be addressed.

1) MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK

2) Department of Endocrine and Metabolic Diseases, Shanghai Institute of Endocrine and Metabolic Diseases, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

3) 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, China

Corresponding author: Maria K. Sobczyk (maria.sobczyk@bristol.ac.uk)

Abstract

Mendelian Randomization (MR) uses genetic instrumental variables to make causal inferences. Whilst sometimes referred to as “nature’s randomized trial”, it has distinct assumptions that make comparisons between the results of MR studies with those of actual randomized controlled trials (RCTs) invaluable. To scope the potential for (semi-)-automated triangulation of MR and RCT evidence, 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. We found that only 11% 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 –25% 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 MeSH 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. 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.

Introduction

Randomized controlled trials (RCTs) are deemed the "gold standard" in evaluating the efficacy of interventions and guiding practice in clinical research, with well-established methodology[1]. 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, known as Mendelian Randomization (MR) – usually implemented as an instrumental variable (IV) analyses – has gained popularity[2,3]. This approach has been referred to as "nature's randomized trials"[4], and is based on the randomization from parents to offspring of genetic variants encapsulated in Mendel's laws of segregation and independent assortment[2,5]. At a population level the randomization is approximate, but still allows genetic variants that are robustly associated with the measured exposure to be used to estimate the unbiased causal effect of an exposure (generally acting across life) on health outcomes, as long as certain assumptions, discussed in detail elsewhere[2,3,6], are met.

Despite drawing on observational data, the MR approach broadly aligns with that of a randomized controlled trial, 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 randomization takes place at conception, the time lag to the start of outcome recoding is longer[7] compared to RCTs, where median duration of phase 3 trials is 40 months[8]. 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[9]. For instance, MR demonstrated the lack of effect of genetically-predicted concentrations of HDL-C on cardiovascular events[10–12] as well as selenium in prostate cancer prevention trials[13,14]. On the other hand, MR showed the beneficial effect of lifelong endogenous low LDL-C levels[15], HMG-CoA reductase inhibition (statin drug target) and PCSK9 inhibition on cardiovascular disease[15,16], 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 PCK9 inhibition[17].

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[18], for example, the effect on incident coronary heart disease risk of lowering LDL cholesterol by 1 mmol/l for 5 years. Whilst 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[19]. This is discussed further in the Supplementary Box.

Whilst 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[20], have short follow-up and typically take place after disease onset[21,22]. As in other study types, RCT results may be flawed due to poor design and execution, e.g. imperfect randomization, 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 prioritize intervention-condition pairs to assess in RCTs. Moreover, it has been proposed that MR also guide 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[1,23]. 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[24–26]. Finally, the extensive use of existing observational data for MR enables intervention targets to be evaluated in a wider range of sub-populations than is feasible for RCTs (improving generalizability), 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[27], EpigraphDB[28] 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”.

Results

ClinicalTrials.Gov data overview

In total, we found 379,094 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 2). In our analysis, we identified 166,954 RCT studies (44% of the total). To allow semi-automated comparison with MR studies, we focussed on the study subset which submitted their statistical analysis results to the database (referred to as the **main dataset,** Supplementary Dataset 1). However, we found that only 4% of studies – 13,807 met this criterion, along with including background information on the trial. To expand that number, we also considered an additional 23,080 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,** Supplementary Dataset 2).

The majority of RCTs in the main dataset followed parallel assignment of participants to treatment (Figure S1a), most were designed for treatment (n=10,812, Figure S1b), rather than prevention (n=1,422) and the vast majority of them had been completed (Figure S1c). More trials were observed to be in phase 3 than 4 (Figure S1d), most trials included both males and females (Figure S1e), and a great majority had 2 arms (Figure S1f). The median number of primary outcomes was 1 (Figure S2a), with a median of 5 secondary outcomes (Figure S2b). Over half of studies report at least 1 result with *p*-value less than 0.05 (Figure S2c). Comparison with features of all RCTs in the database showed that our selection was broadly representative (Supplementary 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 of annotation of RCTs with MeSH (Medical Subject Headings) in ClinicalTrials.Gov. The most common intervention was Drug (Table S1). Since we were only interested in the intervention types which can be instrumented by MR, we also focussed on the 4th and 7th most popular types of interventions: Behavioral 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 Behavioral 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 Behavioral 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 instrumental variables in MR can be used as proxies for pharmaceutical interventions in RCTs. Protein (pQTL) or expression (eQTL) quantitative trait loci (QTL), i.e. 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. (2020)[29]. We focussed 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[23].

We matched the drug target proteins in Zheng et al. (2020) with drug-gene associations sourced from EpigraphDB[28]. This allowed us to merge the Zheng et al. (2020) 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 low number of hits. The results, displayed in Table 2 show overlap of the RCT and MR datasets. We found 4 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[30]. The Zheng et al. (2020) 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 hyperlipidemia, hypercholesterolemia and dyslipidemias in general. Similarly, reduced expression/activity of apoB in MR and 6 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 dyslipidemias, 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)[31]. 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 leukemia, with the exception of ACHE (acetylocholinoesterase) whose inhibitors (galantamine, donepezil, rivastigmine) are used for treatment of cognitive decline in Alzheimer's disease[32].

We also compared the RCT dataset with Zheng et al. blood transcript expression (eQTL)-derived MR analysis (available in EpigraphDB: https://epigraphdb.org/xqtl). In total, we identified 15 drug target-disease matches in the eQTL dataset (Table S2), although unlike in the pQTL matches, the direction of effect in MR was incorrect in 8 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[33], respectively, and blood pressure-lowering action of ACE inhibitors[34].

PubMed-sourced MR and RCT studies

In addition to searching through the ClincalTrials.Gov database, we also queried PubMed for RCT and MR publications. In total, we found 3,538 MR studies published since mid-2000s and 63,187 RCTs published since 1970 until 2020 (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[35] provides a vast repository of semantic predications (subject-predicate-object triple, e.g. LDL-C causes ischemic heart disease). We linked the MR and RCT publications identified by our PubMed search to their corresponding SemMed triples in EpigraphDB using PubMed ID. Overall, only 12% and 25% 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 10,875 unique exposure-outcome pairs, discussed in detail in the Supplementary Note. However, only 125 of these were found to be shared across MR and RCT studies.

We then investigated the 125 matching subject-object pairs between MR and RCT studies (Table S3), as well as individual top counts among subjects (Table S4) and objects (Table S5). T2D, insulin and obesity were found among the top shared risk factors, along with lipids and vitamin D. Top outcomes included T2D, cardiovascular disease, asthma and Alzheimer's disease.

Case studies of matching MR and RCTs

Since our semi-automatic 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, glycemic, neuropsychiatric, musculoskeletal, autoimmune and cancer outcome phenotypes. In total, we surveyed 54 MR and 77 randomized controlled trial publications (RCTs and meta-analysis of RCTs, Supplementary Dataset 4, Figure 4) which were systematically compared across several criteria shown in sample Table S6, and encompass those in the popular PICO (Population, Intervention, Comparison, and Outcome) framework[36].

There, we compare an MR study and two RCT meta-analyses on the effect of vitamin D supplementation in multiple sclerosis (MS)[37–39]. Whilst the RCTs looked at potential therapeutic effect of vitamin D in diagnosed MS patients over 6 months-2 years: measured disability (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 5 small meta-analyzed 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.

1) *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)[40,41] 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[42], aerobic exercise[43,44], progressive resistance training[45] and maximal strength training[46] as interventions. While a MR study[41] and two trials[46,47] 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[40] reporting null effect and trials generally finding positive effects on HDL-C concentration and negative on LDL-C, total cholesterol and triglycerides blood levels[43–45,48].

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[49] 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 (≥700 IU daily) on bone fractures in the elderly seen in RCTs[50–52] may not have been accessible in the MR study.

2) 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[53,54], while 3 meta-analyses of RCTs found a consistent effect of a variety of exercise types on improving total and negative symptoms of schizophrenia[55–57] (Figure 4).

3) Outcome

The short duration of RCTs mean some outcomes (e.g. 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 cardiovascular disease (CVD) outcomes, we found matching conclusions with elevated SBP increasing the risk of CVD both in RCT[58–60] and MR[61–63] studies (Figure 4), with MR studies using both single disease outcomes and a composite outcome. However, MR studies analysing the impact of BMI on cardiovascular disease found reduced adiposity led to reduction in arterial hypertension, CVD and stroke[64,65], which contrasted with the results of one of the biggest RCTs to date. The Look AHEAD RCT in older type 2 diabetics 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 hospitalization for angina[66].

Secondly, RCT outcomes, are often on a quantitative scale measuring symptom strength according to established metrics, (e.g. depressive symptoms on Hamilton Depression Rating Scale[67]). However, the best disease 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[68,69] and RCTs[55,67,70–73], the differences in outcome phenotypes could potentially contribute to null MR results[74,75] and positive effect of vitamin D on attenuating eczema symptoms in RCTs[76–79].

4) 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[23]. For example, while in the MR study conducted in general population, there was no strong significant effect of exercise on glycemic markers: HbA1c, fasting glucose and HOMA-IR[80], a significant reduction was found in the meta-analysis of 32 RCTs involving patients with T2D[81].

On the other hand, 5 Mendelian randomization[82–86] studies along with 3 large RCTs[87–89] consistently provide evidence that weight loss is causally associated with reduced risk of T2D (Figure 4), despite MR including the general population and RCTs focussing 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[74,75], while RCTs were conducted separately in children in Mongolia[76] and Boston, USA[77] with winter atopic dermatitis and in adolescent and adult Iranians[78,79].

5) Comparator group

Firstly, 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[23]. Nevertheless, the two outcomes analysed here: hypertension for alcohol intake and lipids for smoking showed generally congruent results across study types (Supplementary Dataset 4).

Secondly, where it would be unethical to withhold already available efficacious treatments, trials will often include another active intervention in the comparator group[23], for example statins in the trials of effect of PCSK9 inhibitors[90,91] 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[16].

6) Duration of intervention

While the magnitude of effect seen in trials with long (> 3 years: weight loss to treat hypertension[92,93]/type 2 diabetes[87–89], blood pressure reduction to lower cardiovascular disease risk[58–60]) and short (< 6 months: alcohol intake reduction to lower blood pressure[94–97], exercise to benefit bone mineral density[46]/depression[55,67,71–73,98]) intervention may vary, we find both can result in directional effects consistent with MR results (Figure 4), although with exceptions[66,99].

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[100–103] and RCT[92,93] studies, high BMI on T2D risk in MR[82–86,104] and RCTs[87–89], and the null effect of vitamin D on various glycemic markers in diverse populations in MR[105], RCTs[106,107] and RCT systematic review[108].

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, 2 MR studies[109,110], a meta-analysis of 5 RCT studies[111] and two RCTs[112,113] indicate no effect of vitamin E on prostate cancer incidence with one outlier RCT[114] showing benefit of vitamin E supplementation in older smokers. Similar contrary findings were found for 1 RCT[115] showing beneficial effect of vitamin D on preventing depressive symptoms, as opposed to null effect in 4 MR studies[116–119] and 2 RCTs[120,121].

On the other hand, MR analyses can show spurious disagreement with the rest of the evidence base. For instance, 2 MR papers[122,123] and a meta-analysis of 16 RCT[123] studies reveal no significant effect of vitamin D on blood pressure in the general population, with the exception of one MR study[124] that indicated a blood pressure-lowering effect of higher vitamin D status. Similarly, a range of study types: one MR analysis[125], one RCT[126] and a meta-analysis of 27 prospective cohorts[127] (only some of them RCTs) confirm a negative impact of smoking on HDL-C levels, bar one MR study showing no significant effect[128]. A series of RCT meta-analyses[129–131] support an effect of coffee consumption (especially unfiltered) on unfavourable blood profile, although this is likely explained by diterpenes[132,133] rather than caffeine, as the latter shows evidence of cardioprotective effects[134]. However, only the recent biggest MR study[133] to date found a significant effect of coffee consumption on LDL-C and total-C levels, unlike two previous smaller MR analyses[135,136], 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 <1,000 studies, out of ~167,000 analysed RCTs were started before 2000). Nonetheless, we found that only 11% 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)[137], only 38% of eligible trials for 2008-2012 submitted their results at any time[138] which rose to 64% for 2018-2019[139]. Furthermore, 60% of studies for failed agents are reported not to be published in peer reviewed journals[140], 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 25% of MR and 12% of RCT studies associated with >=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 complimentary evidence on the effectiveness of interventions. Powerful examples include congruence of positive effect of high BMI on hypertension across MR[100–103] and RCT[92,93] studies, high BMI on T2D risk in MR[82–86,104] and RCTs[87–89], and the null effect of vitamin D on various glycemic markers in diverse populations in MR[105], RCTs[106,107] and RCT systematic review[108]. 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=1,002) and phenotypes (n=225, many non-diseases *per se*) in Zheng et al. (2020)[141], 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[140] and inclusion of non-disease phenotypes in MR analysis. Anecdotally, we found no MR evidence that decreased expression of PLA2G2A leads to reduced cardiovascular disease, which agrees with lack of efficacy of PLA2G2A inhibitor in clinical trials[142–144].

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[145], hidden pleiotropy in single instruments used in the MR analysis (directly observed for IL2RA)[146], presence of negative feedback loop involved in the drug mechanism[147], translation into protein isoforms with distinct biological effects[148] and differential cell-type specific drug effect[149].

The duration of intervention varies between RCTs and MR studies, with the former spanning no more than the duration of the trial, whilst 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)[7]. Moreover, intervention in RCTs with long follow-up is not necessarily similarly intensive throughout its duration, or may cease altogether after some time[67,92], i.e. 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[67] or weight loss programmes[87] 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[66]. That is why our analysis focussed on comparing direction of effect, while ignoring magnitude of effect[150]. However, in certain cases when enough reference data is available, it is feasible to compare MR and RCT effects on the same exposure difference scale[151].

Further impediments to direct comparison between MR and RCTs include differences in outcome definition (composite[58] versus single conditions[63]). Access to rare subpopulations with existing conditions, such as cancer patients receiving specific therapy[152] 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[153]. This is especially true of lifestyle interventions – such as different forms of psychological therapy, complex diet regimens[154] 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[155] which are often a (binary) single measurement, as opposed to multiple quantitative outcomes frequently measured in RCTs[156]. 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[157] as these studies are typically studies of convenience. This bias is less common in RCTs due to pre-registration of study design including power analysis[1], uncommon in MR[158]. Null effect in MR studies may be therefore spurious and not predictive of RCTs for that reason, as seen in two smaller MR studies[135,136] out of three[133] investigating the effect of coffee intake on blood lipids, contrasting with strong clinical trial[129–131] and biochemical evidence[132,133,159].

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[160,161]. 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 (Supplementary Box).

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[25]. 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[36,162,163].

Methods

ClinicalTrials.Gov *Data sources*

All ClinicalTrials.Gov study data is available for download as PostgreSQL database from The Database for Aggregate Analysis of ClinicalTrials.gov (AACT)[164] released by The Clinical Trials Transformation Initiative[165]. We downloaded its static release from 1st August 2021. 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 2 and provided in detail in Supplementary Note.

EpigraphDB *queries*

EpigraphDB[28] was used to collect information about confirmed drug-target associations which were initially sourced from the Open Targets Platform[166] and verified in DrugBank[167]. We also used EpigraphDB[28] to source SemMedDB[35] 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[168].

PubMed *Data harvesting*

We searched PubMed for all RCT and MR studies published before 2021 on 1st October 2021. 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 112,015 to 63,187. In order to retrieve potential MR studies, we used the keywords: "mendelian randomization" or "mendelian randomisation". We also considered using MeSH labels "Randomized Controlled Trial" and "Mendelian Randomization Analysis" but both in case of RCT and MR studies they returned unrealistically low number of hits (7,999 and 1,926, accordingly), a consequence of manual indexing.

Literature searches

We used Semantic Scholar and Google Scholar to survey MR and RCT literature indexed before 1st August 2021. 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 included original research MR, RCT studies as well as meta-analyses. As exposures, we chose common lifestyle risk factors for chronic disease[169] – body mass index (BMI), smoking, alcohol intake, blood pressure. Examples of dietary intervention (vitamin D/E, coffee) and behavioural intervention (exercise) were also selected. Outcomes included a number of cardiovascular disease and risk biomarkers, glycemic traits, neuropsychiatric disease, bone mineral density, atopic dermatitis and prostate cancer, amongst others (Supplementary Dataset 4). 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.

Acknowledgments

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.

Author contributions

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.

Competing interests

JZ, GDS and TRG receive funding from Biogen for unrelated research.

Data availability

Code used to carry out the analysis is available on GitHub: https://github.com/marynias/mr-rct. ClinicalTrials.Gov data was accessed via AACT: https://aact.ctti-clinicaltrials.org/snapshots and analysed data subset is available in Supplementary Datasets 1 & 2. pQTL and eQTL MR analysis results are available via EpigraphDB: https://epigraphdb.org/xqtl

References

1. Ference, B. A., Holmes, M. V & Smith, G. D. Using Mendelian Randomization to Improve the Design of Randomized Trials. *Cold Spring Harb. Perspect. Med.* **11**, (2021).

2. Davey Smith, G. & Ebrahim, S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int. J. Epidemiol.* **32**, 1–22 (2003).

3. Sanderson, E. *et al.* Mendelian randomization. *Nat. Rev. Methods Prim.* **2**, 6 (2022).

4. Hingorani, A. & Humphries, S. Nature's randomised trials. *Lancet* **366**, 1906–1908 (2005).

5. Davey Smith, G., Holmes, M. V, Davies, N. M. & Ebrahim, S. Mendel's laws, Mendelian randomization and causal inference in observational data: substantive and nomenclatural issues. *Eur. J. Epidemiol.* **35**, 99–111 (2020).

6. Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* **362**, (2018).

7. Swanson, S. A., Tiemeier, H., Ikram, M. A. & Hernán, M. A. Nature as a Trialist?: Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials. *Epidemiology* **28**, 653–659 (2017).

8. Martin, L., Hutchens, M. & Hawkins, C. Trial watch: Clinical trial cycle times continue to increase despite industry efforts. *Nat. Rev. Drug Discov.* **16**, 157 (2017).

9. Egger, M., Buitrago-Garcia, D. & Smith, G. D. Systematic Reviews of Epidemiological Studies of Etiology and Prevalence. in *Systematic Reviews in Health Care: Meta-analyses in Context* (eds. Egger, M., Higgins, J. P. T. & Smith, G. D.) Chapter 19 (Wiley-Blackwell, 2022).

10. Voight, B. F. *et al.* Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. *Lancet* **380**, 572–580 (2012).

11. Verdoia, M., Schaffer, A., Suryapranata, H. & De Luca, G. Effects of HDL-modifiers on cardiovascular outcomes: A meta-analysis of randomized trials. *Nutr. Metab. Cardiovasc. Dis.* **25**, 9–23 (2015).

12. Davey Smith, G. & Phillips, A. N. Correlation without a cause: an epidemiological odyssey. *Int. J. Epidemiol.* **49**, 4–14 (2020).

13. Yarmolinsky, J. *et al.* Circulating Selenium and Prostate Cancer Risk: A Mendelian Randomization Analysis. *JNCI J. Natl. Cancer Inst.* **110**, 1035–1038 (2018).

14. Lippman, S. M. *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* **301**, 39–51 (2009).

15. A., F. B. *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.* **60**, 2631–2639 (2012).

16. Ference, B. A. *et al.* Variation in PCSK9 and HMGCR and Risk of Cardiovascular Disease and Diabetes. *N. Engl. J. Med.* **375**, 2144–2153 (2016).

17. Hegele, R. A. Editorial comment: when Mendelian randomization goes astray. *Curr. Opin. Lipidol.* **32**, (2021).

18. Turner, R. M., Spiegelhalter, D. J., Smith, G. C. S. & Thompson, S. G. Bias modelling in evidence synthesis. *J. R. Stat. Soc. Ser. A (Statistics Soc.* **172**, 21–47 (2009).

19. Morris, T. T., Heron, J., Sanderson, E., Smith, G. D. & Tilling, K. Interpretation of mendelian randomization using one measure of an exposure that varies over time. *medRxiv* (2021) doi:10.1101/2021.11.18.21266515.

20. Victora, C. G., Habicht, J.-P. & Bryce, J. Evidence-Based Public Health: Moving Beyond Randomized Trials. *Am. J. Public Health* **94**, 400–405 (2004).

21. Chekroud, A. M. A Mendelian Randomization Approach for Assessing the Relationship Between Physical Activity and Depression. *JAMA Psychiatry* **76**, 361–362 (2019).

22. Harrison, S. *et al.* Long-term cost-effectiveness of interventions for obesity: A mendelian randomisation study. *PLOS Med.* **18**, e1003725 (2021).

23. Holmes, M. V, Richardson, T. G., Ference, B. A., Davies, N. M. & Davey Smith, G. Integrating genomics with biomarkers and therapeutic targets to invigorate cardiovascular drug development. *Nat. Rev. Cardiol.* **18**, 435–453 (2021).

24. Borén, J. *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.* **41**, 2313–2330 (2020).

25. Munafò, M. R., Higgins, J. P. T. & Smith, G. D. Triangulating Evidence through the Inclusion of Genetically Informed Designs. *Cold Spring Harb. Perspect. Med.* **11**, (2021).

26. Lawlor, D. A., Tilling, K. & Smith, G. D. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* **45**, 1866–1886 (2016).

27. Zarin, D. A., Tse, T., Williams, R. J., Califf, R. M. & Ide, N. C. The ClinicalTrials.gov Results Database — Update and Key Issues. *N. Engl. J. Med.* **364**, 852–860 (2011).

28. Liu, Y. *et al.* EpiGraphDB: a database and data mining platform for health data science. *Bioinformatics* 0–0 (2020).

29. Zheng, J. *et al.* Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* **52**, 1122–1131 (2020).

30. Cicero, A. F. G., Tartagni, E. & Ertek, S. Safety and tolerability of injectable lipid-lowering drugs: a review of available clinical data. *Expert Opin. Drug Saf.* **13**, 1023–1030 (2014).

31. Luo, J. *et al.* Structural Basis for the Dual Recognition of IL-12 and IL-23 by Ustekinumab. *J. Mol. Biol.* **402**, 797–812 (2010).

32. Hansen, R. A. *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* **3**, 211–225 (2008).

33. Ference, B. A. *et al.* Association of Genetic Variants Related to CETP Inhibitors and Statins With Lipoprotein Levels and Cardiovascular Risk. *JAMA* **318**, 947–956 (2017).

34. Piepho, R. W. Overview of the angiotensin-converting-enzyme inhibitors. *Am. J. Heal. Pharm.* **57**, S3–S7 (2000).

35. Kilicoglu, H., Shin, D., Fiszman, M., Rosemblat, G. & Rindflesch, T. C. SemMedDB: a PubMed-scale repository of biomedical semantic predications. *Bioinformatics* **28**, 3158–3160 (2012).

36. Marshall, I. J. *et al.* Trialstreamer: A living, automatically updated database of clinical trial reports. *J. Am. Med. Informatics Assoc.* **27**, 1903–1912 (2020).

37. Mokry, L. E. *et al.* Vitamin D and Risk of Multiple Sclerosis: A Mendelian Randomization Study. *PLOS Med.* **12**, e1001866 (2015).

38. James, E. *et al.* The effect of vitamin D-related interventions on multiple sclerosis relapses: a meta-analysis. *Mult. Scler. J.* **19**, 1571–1579 (2013).

39. Hempel, S. *et al.* A systematic review of the effects of modifiable risk factor interventions on the progression of multiple sclerosis. *Mult. Scler. J.* **23**, 513–524 (2017).

40. Zhuang, Z. *et al.* Association of physical activity, sedentary behaviours and sleep duration with cardiovascular diseases and lipid profiles: a Mendelian randomization analysis. *Lipids Health Dis.* **19**, 86 (2020).

41. Xu, F. *et al.* Estimates of the effects of physical activity on osteoporosis using multivariable Mendelian randomization analysis. *Osteoporos. Int.* **32**, 1359–1367 (2021).

42. Kelley, G. A., Kelley, K. S. & Tran, Z. V. Walking, lipids, and lipoproteins: a meta-analysis of randomized controlled trials. *Prev. Med. (Baltim).* **38**, 651–661 (2004).

43. Kelley, G. A. & Kelley, K. S. Aerobic exercise and lipids and lipoproteins in men: a meta-analysis of randomized controlled trials. *J. Men's Heal. Gend.* **3**, 61–70 (2006).

44. Kelley, G. A., Kelley, K. S. & Tran, Z. V. Aerobic Exercise and Lipids and Lipoproteins in Women: A Meta-Analysis of Randomized Controlled Trials. *J. Women's Heal.* **13**, 1148–1164 (2004).

45. Kelley, G. A. & Kelley, K. S. Impact of progressive resistance training on lipids and lipoproteins in adults: A meta-analysis of randomized controlled trials. *Prev. Med. (Baltim).* **48**, 9–19 (2009).

46. Mosti, M. P. *et al.* Maximal Strength Training Improves Bone Mineral Density and Neuromuscular Performance in Young Adult Women. *J. Strength Cond. Res.* **28**, (2014).

47. Allison, S. J., Folland, J. P., Rennie, W. J., Summers, G. D. & Brooke-Wavell, K. High impact exercise increased

femoral neck bone mineral density in older men: A randomised unilateral intervention. *Bone* **53**, 321–328 (2013).

48. Katzel, L. I. *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* **274**, 1915–1921 (1995).

49. Çolak, Y., Afzal, S. & Nordestgaard, B. G. 25-Hydroxyvitamin D and Risk of Osteoporotic Fractures: Mendelian Randomization Analysis in 2 Large Population-Based Cohorts. *Clin. Chem.* **66**, 676–685 (2020).

50. Bischoff-Ferrari, H. A. *et al.* A Pooled Analysis of Vitamin D Dose Requirements for Fracture Prevention. *N. Engl. J. Med.* **367**, 40–49 (2012).

51. Bischoff-Ferrari, H. A. *et al.* Prevention of Nonvertebral Fractures With Oral Vitamin D and Dose Dependency: A Meta-analysis of Randomized Controlled Trials. *Arch. Intern. Med.* **169**, 551–561 (2009).

52. Bischoff-Ferrari, H. A. *et al.* Fracture Prevention With Vitamin D SupplementationA Meta-analysis of Randomized Controlled Trials. *JAMA* **293**, 2257–2264 (2005).

53. Papiol, S. *et al.* Association Between Physical Activity and Schizophrenia: Results of a 2-Sample Mendelian Randomization Analysis. *JAMA Psychiatry* **78**, 441–444 (2021).

54. Sun, H. *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.* **263**, 598–604 (2020).

55. Rosenbaum, S., Tiedemann, A., Sherrington, C., Curtis, J. & Ward, P. B. Physical activity interventions for people with mental illness: A systematic review and meta-analysis. *The Journal of Clinical Psychiatry* vol. 75 964–974 (2014).

56. Firth, J., Cotter, J., Elliott, R., French, P. & Yung, A. R. A systematic review and meta-analysis of exercise interventions in schizophrenia patients. *Psychol. Med.* **45**, 1343–1361 (2015).

57. 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* **62**, 13–20 (2020).

58. Williamson, J. D. *et al.* Intensive vs Standard Blood Pressure Control and Cardiovascular Disease Outcomes in Adults Aged ≥75 Years: A Randomized Clinical Trial. *JAMA* **315**, 2673–2682 (2016).

59. 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.* **178**, 28–36 (2018).

60. Rahimi, K. *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* **398**, 1053–1064 (2021).

61. Gill, D. *et al.* Genetically Predicted Midlife Blood Pressure and Coronary Artery Disease Risk: Mendelian Randomization Analysis. *J. Am. Heart Assoc.* **9**, e016773 (2020).

62. Malik, R. *et al.* Relationship Between Blood Pressure and Incident Cardiovascular Disease: Linear and Nonlinear Mendelian Randomization Analyses. *Hypertension* **77**, 2004–2013 (2021).

63. Wan, E. Y. F. *et al.* Blood Pressure and Risk of Cardiovascular Disease in UK Biobank. *Hypertension* **77**, 367–375 (2021).

64. Riaz, H. *et al.* Association Between Obesity and Cardiovascular Outcomes: A Systematic Review and Meta-analysis of Mendelian Randomization Studies. *JAMA Netw. Open* **1**, e183788–e183788 (2018).

65. Larsson, S. C., Bäck, M., Rees, J. M. B., Mason, A. M. & Burgess, S. Body mass index and body composition in relation to 14 cardiovascular conditions in UK Biobank: a Mendelian randomization study. *Eur. Heart J.* **41**, 221–226 (2020).

66. Cardiovascular Effects of Intensive Lifestyle Intervention in Type 2 Diabetes. *N. Engl. J. Med.* **369**, 145–154 (2013).

67. Mather, A. S. *et al.* Effects of exercise on depressive symptoms in older adults with poorly responsive depressive disorder: Randomised controlled trial. *Br. J. Psychiatry* **180**, 411–415 (2002).

68. Choi, K. W. *et al.* Assessment of Bidirectional Relationships Between Physical Activity and Depression Among Adults: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry* **76**, 399–408 (2019).

69. Choi, K. W. *et al.* An Exposure-Wide and Mendelian Randomization Approach to Identifying Modifiable Factors for the Prevention of Depression. *Am. J. Psychiatry* **177**, 944–954 (2020).

70. Veale, D. *et al.* Aerobic exercise in the adjunctive treatment of depression: a randomized controlled trial. *J. R. Soc. Med.* **85**, 541–544 (1992).

71. Schuch, F. B. *et al.* Exercise as a treatment for depression: A meta-analysis adjusting for publication bias. *J. Psychiatr. Res.* **77**, 42–51 (2016).

72. Kvam, S., Kleppe, C. L., Nordhus, I. H. & Hovland, A. Exercise as a treatment for depression: A meta-analysis. *J. Affect. Disord.* **202**, 67–86 (2016).

73. Gordon, B. R. *et al.* Association of Efficacy of Resistance Exercise Training With Depressive Symptoms: Meta-analysis and Meta-regression Analysis of Randomized Clinical Trials. *JAMA Psychiatry* **75**, 566–576 (2018).

74. Manousaki, D. *et al.* Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: A Mendelian randomization study. *PLOS Med.* **14**, 1–16 (2017).

75. Drodge, D. R., Budu-Aggrey, A. & Paternoster, L. Causal Analysis Shows Evidence of Atopic Dermatitis Leading to an Increase in Vitamin D Levels. *J. Invest. Dermatol.* **141**, 1339–1341 (2021).

76. Camargo, C. A. *et al.* Randomized trial of vitamin D supplementation for winter-related atopic dermatitis in children. *J. Allergy Clin. Immunol.* **134**, 831-835.e1 (2014).

77. Sidbury, R., Sullivan, A. F., Thadhani, R. I. & Camargo, C. A. Randomized controlled trial of vitamin D supplementation for winter-related atopic dermatitis in Boston: a pilot study. *Br. J. Dermatol.* **159**, 245—247 (2008).

78. Amestejani, M. *et al.* Vitamin D supplementation in the treatment of atopic dermatitis: a clinical trial study. *J. Drugs Dermatology* **11**, (2012).

79. Javanbakht, M. H. *et al.* Randomized controlled trial using vitamins E and D supplementation in atopic dermatitis. *J. Dermatolog. Treat.* **22**, 144–150 (2011).

80. Meisinger, C., Linseisen, J., Leitzmann, M., Baurecht, H. & Baumeister, S. E. Association of physical activity and sedentary behavior with type 2 diabetes and glycemic traits: a two-sample Mendelian randomization study. *BMJ Open Diabetes Res. Care* **8**, (2020).

81. Shah, S. Z. 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.* **12**, 707–732 (2021).

82. Corbin, L. J. *et al.* BMI as a Modifiable Risk Factor for Type 2 Diabetes: Refining and Understanding Causal Estimates Using Mendelian Randomization. *Diabetes* **65**, 3002–3007 (2016).

83. Dale, C. E. *et al.* Causal Associations of Adiposity and Body Fat Distribution With Coronary Heart Disease, Stroke Subtypes, and Type 2 Diabetes Mellitus. *Circulation* **135**, 2373–2388 (2017).

84. Cheng, L. *et al.* Exposing the causal effect of body mass index on the risk of type 2 diabetes mellitus: a mendelian randomization study. *Front. Genet.* **10**, 94 (2019).

85. Wainberg, M. *et al.* Homogeneity in the association of body mass index with type 2 diabetes across the UK Biobank: A Mendelian randomization study. *PLOS Med.* **16**, e1002982 (2019).

86. 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.* **11**, 603 (2020).

87. Pan, X.-R. *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* **20**, 537–544 (1997).

88. Tuomilehto, J. *et al.* Prevention of Type 2 Diabetes Mellitus by Changes in Lifestyle among Subjects with Impaired Glucose Tolerance. *N. Engl. J. Med.* **344**, 1343–1350 (2001).

89. Reduction in the Incidence of Type 2 Diabetes with Lifestyle Intervention or Metformin. *N. Engl. J. Med.* **346**, 393–403 (2002).

90. Sabatine, M. S. *et al.* Efficacy and Safety of Evolocumab in Reducing Lipids and Cardiovascular Events. *N. Engl. J. Med.* **372**, 1500–1509 (2015).

91. Sabatine, M. S. *et al.* Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N. Engl. J. Med.* **376**, 1713–1722 (2017).

92. He, J., Whelton, P. K., Appel, L. J., Charleston, J. & Klag, M. J. Long-Term Effects of Weight Loss and Dietary Sodium Reduction on Incidence of Hypertension. *Hypertension* **35**, 544–549 (2000).

93. Long-Term Weight Loss and Changes in Blood Pressure: Results of the Trials of Hypertension Prevention, Phase II. *Ann. Intern. Med.* **134**, 1–11 (2001).

94. Puddey, I. B., Beilin, L. J., Vandongen, R., Rouse, I. L. & Rogers, P. Evidence for a direct effect of alcohol consumption on blood pressure in normotensive men. A randomized controlled trial. *Hypertension* **7**, 707–713 (1985).

95. Puddey, I. B., Beilin, L. J. & Vandongen, R. REGULAR ALCOHOL USE RAISES BLOOD PRESSURE IN TREATED HYPERTENSIVE SUBJECTS: A Randomised Controlled Trial. *Lancet* **329**, 647–651 (1987).

96. Xin, X. *et al.* Effects of Alcohol Reduction on Blood Pressure. *Hypertension* **38**, 1112–1117 (2001).

97. Puddey, I. B., Parker, M., Beilin, L. J., Vandongen, R. & Masarei, J. R. Effects of alcohol and caloric restrictions on blood pressure and serum lipids in overweight men. *Hypertension* **20**, 533–541 (1992).

98. Helgadóttir, B., Hallgren, M., Ekblom, Ö. & Forsell, Y. Training fast or slow? Exercise for depression: A randomized controlled trial. *Prev. Med. (Baltim).* **91**, 123–131 (2016).

99. Ma, C. *et al.* Effects of weight loss interventions for adults who are obese on mortality, cardiovascular disease, and cancer: systematic review and meta-analysis. *BMJ* **359**, (2017).

100. Timpson, N. J. *et al.* Does Greater Adiposity Increase Blood Pressure and Hypertension Risk? *Hypertension* **54**, 84–90 (2009).

101. Lyall, D. M. *et al.* Association of Body Mass Index With Cardiometabolic Disease in the UK Biobank: A Mendelian Randomization Study. *JAMA Cardiol.* **2**, 882–889 (2017).

102. Lee, M.-R., Lim, Y.-H. & Hong, Y.-C. Causal association of body mass index with hypertension using a Mendelian randomization design. *Medicine (Baltimore).* **97**, e11252–e11252 (2018).

103. van Oort, S., Beulens, J. W. J., van Ballegooijen, A. J., Grobbee, D. E. & Larsson, S. C. Association of Cardiovascular Risk Factors and Lifestyle Behaviors With Hypertension. *Hypertension* **76**, 1971–1979 (2020).

104. Rivera, C. M. & Ren, B. Mapping human epigenomes. *Cell* **155**, 39–55 (2013).

105. Wang, N. *et al.* Vitamin D, prediabetes and type 2 diabetes: bidirectional Mendelian randomization analysis. *Eur. J. Nutr.* **59**, 1379–1388 (2020).

106. Salehpour, A. *et al.* Does vitamin D3 supplementation improve glucose homeostasis in overweight or obese women? A double-blind, randomized, placebo-controlled clinical trial. *Diabet. Med.* **30**, 1477–1481 (2013).

107. Sollid, S. T. *et al.* No Effect of High-Dose Vitamin D Supplementation on Glycemic Status or Cardiovascular Risk Factors in Subjects With Prediabetes. *Diabetes Care* **37**, 2123–2131 (2014).

108. Zuk, A., Fitzpatrick, T. & Rosella, L. C. 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* **11**, e0154215 (2016).

109. Wright, M. E. *et al.* Association of Variants in Two Vitamin E Transport Genes with Circulating Vitamin E Concentrations and Prostate Cancer Risk. *Cancer Res.* **69**, 1429–1438 (2009).

110. Cheng, W.-W., Wang, Z.-K., Shangguan, H.-F., Zhu, Q. & Zhang, H.-Y. Are vitamins relevant to cancer risks? A Mendelian randomization investigation. *Nutrition* **78**, 110870 (2020).

111. Jiang, L. *et al.* Efficacy of Antioxidant Vitamins and Selenium Supplement in Prostate Cancer Prevention: A Meta-Analysis of Randomized Controlled Trials. *Nutr. Cancer* **62**, 719–727 (2010).

112. Gaziano, J. M. *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* **301**, 52–62 (2009).

113. Wang, L. *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.* **100**, 915–923 (2014).

114. Heinonen, O. P. *et al.* Prostate Cancer and Supplementation With α-Tocopherol and β-Carotene: Incidence and Mortality in a Controlled Trial. *JNCI J. Natl. Cancer Inst.* **90**, 440–446 (1998).

115. Vieth, R., Kimball, S., Hu, A. & Walfish, P. G. Randomized comparison of the effects of the vitamin D3adequate intake versus 100 mcg (4000 IU) per day on biochemical responses and the wellbeing of patients. *Nutr. J.* **3**, 8 (2004).

116. Michaëlsson, K., Melhus, H. & Larsson, S. C. Serum 25-Hydroxyvitamin D Concentrations and Major Depression: A Mendelian Randomization Study. *Nutrients* **10**, (2018).

117. Libuda, L. *et al.* Vitamin D and the Risk of Depression: A Causal Relationship? Findings from a Mendelian Randomization Study. *Nutrients* **11**, (2019).

118. Mulugeta, A., Lumsden, A. & Hyppönen, E. Relationship between Serum 25(OH)D and Depression: Causal Evidence from a Bi-Directional Mendelian Randomization Study. *Nutrients* **13**, (2021).

119. Milaneschi, Y. *et al.* A role for vitamin D and omega-3 fatty acids in major depression? An exploration using genomics. *Transl. Psychiatry* **9**, 219 (2019).

120. Bertone-Johnson, E. R. *et al.* Vitamin D Supplementation and Depression in the Women's Health Initiative Calcium and Vitamin D Trial. *Am. J. Epidemiol.* **176**, 1–13 (2012).

121. Kjærgaard, M. *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* **201**, 360–368 (2012).

122. Skaaby, T. *et al.* Vitamin D Status, Filaggrin Genotype, and Cardiovascular Risk Factors: A Mendelian Randomization Approach. *PLoS One* **8**, e57647 (2013).

123. Kunutsor, S. K., Burgess, S., Munroe, P. B. & Khan, H. Vitamin D and high blood pressure: causal association or epiphenomenon? *Eur. J. Epidemiol.* **29**, 1–14 (2014).

124. Vimaleswaran, K. S. *et al.* Association of vitamin D status with arterial blood pressure and hypertension risk: a mendelian randomisation study. *Lancet Diabetes Endocrinol.* **2**, 719–729 (2014).

125. Rosoff, D. B., Davey Smith, G., Mehta, N., Clarke, T.-K. & Lohoff, F. W. Evaluating the relationship between alcohol consumption, tobacco use, and cardiovascular disease: A multivariable Mendelian randomization study. *PLOS Med.* **17**, e1003410 (2020).

126. Gepner, A. D. *et al.* Effects of smoking and smoking cessation on lipids and lipoproteins: Outcomes from a randomized clinical trial. *Am. Heart J.* **161**, 145–151 (2011).

127. Maeda, K., Noguchi, Y. & Fukui, T. The effects of cessation from cigarette smoking on the lipid and lipoprotein profiles: a meta-analysis. *Prev. Med. (Baltim).* **37**, 283–290 (2003).

128. Åsvold, B. O. *et al.* Causal associations of tobacco smoking with cardiovascular risk factors: a Mendelian randomization analysis of the HUNT Study in Norway. *Int. J. Epidemiol.* **43**, 1458–1470 (2014).

129. Jee, S. H. *et al.* Coffee Consumption and Serum Lipids: A Meta-Analysis of Randomized Controlled Clinical Trials. *Am. J. Epidemiol.* **153**, 353–362 (2001).

130. Cai, L., Ma, D., Zhang, Y., Liu, Z. & Wang, P. The effect of coffee consumption on serum lipids: a meta-analysis of randomized controlled trials. *Eur. J. Clin. Nutr.* **66**, 872–877 (2012).

131. Du, Y., Lv, Y., Zha, W., Hong, X. & Luo, Q. Effect of coffee consumption on dyslipidemia: A meta-analysis of randomized controlled trials. *Nutr. Metab. Cardiovasc. Dis.* **30**, 2159–2170 (2020).

132. Urgert, R. & Katan, M. B. THE CHOLESTEROL-RAISING FACTOR FROM COFFEE BEANS. *Annu. Rev. Nutr.* **17**, 305–324 (1997).

133. Zhou, A. & Hyppönen, E. Habitual coffee intake and plasma lipid profile: Evidence from UK Biobank. *Clin. Nutr.* **40**, 4404–4413 (2021).

134. Lebeau, P. F. *et al.* Caffeine blocks SREBP2-induced hepatic PCSK9 expression to enhance LDLR-mediated cholesterol clearance. *Nat. Commun.* **13**, 770 (2022).

135. Kwok, M. K., Leung, G. M. & Schooling, C. M. Habitual coffee consumption and risk of type 2 diabetes, ischemic heart disease, depression and Alzheimer's disease: a Mendelian randomization study. *Sci. Rep.* **6**, 36500 (2016).

136. Nordestgaard, A. T., Thomsen, M. & Nordestgaard, B. G. Coffee intake and risk of obesity, metabolic syndrome and type 2 diabetes: a Mendelian randomization study. *Int. J. Epidemiol.* **44**, 551–565 (2015).

137. Zarin, D. A., Tse, T., Williams, R. J. & Carr, S. Trial Reporting in ClinicalTrials.gov — The Final Rule. *N. Engl. J. Med.*

375, 1998–2004 (2016).

138. Anderson, M. L. *et al.* Compliance with Results Reporting at ClinicalTrials.gov. *N. Engl. J. Med.* **372**, 1031–1039 (2015).

139. DeVito, N. J., Bacon, S. & Goldacre, B. Compliance with legal requirement to report clinical trial results on ClinicalTrials.gov: a cohort study. *Lancet* **395**, 361–369 (2020).

140. Hwang, T. J. *et al.* Failure of Investigational Drugs in Late-Stage Clinical Development and Publication of Trial Results. *JAMA Intern. Med.* **176**, 1826–1833 (2016).

141. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).

142. Holmes, M. V *et al.* Secretory Phospholipase A_2-IIA and Cardiovascular Disease. *J. Am. Coll. Cardiol.* **62**, 1966–1976 (2013).

143. Rosenson, R. S. *et al.* Effects of varespladib methyl on biomarkers and major cardiovascular events in acute coronary syndrome patients. *J. Am. Coll. Cardiol.* **56**, 1079–1088 (2010).

144. Nicholls, S. J. *et al.* Varespladib and cardiovascular events in patients with an acute coronary syndrome: the VISTA-16 randomized clinical trial. *JAMA* **311**, 252–262 (2014).

145. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).

146. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat. Genet.* **50**, 693–698 (2018).

147. Foulds, C. E. Disrupting a negative feedback loop drives endocrine therapy-resistant breast cancer. *Proc. Natl. Acad. Sci.* **115**, 8236–8238 (2018).

148. Cupido, A. J. *et al.* Dissecting the IL-6 pathway in cardiometabolic disease: A Mendelian randomization study on both IL6 and IL6R. *Br. J. Clin. Pharmacol.* **n/a**,.

149. Gallo, M. A. & Kaufman, D. Antagonistic and agonistic effects of tamoxifen: significance in human cancer. *Semin. Oncol.* **24**, S1—71—S1—80 (1997).

150. Ference, B. A. How to use Mendelian randomization to anticipate the results of randomized trials. *Eur. Heart J.* **39**, 360–362 (2017).

151. Holmes, M. V & Smith, G. D. Revealing the effect of CETP inhibition in cardiovascular disease. *Nat. Rev. Cardiol.* **14**, 635–636 (2017).

152. Kearns, A. E. *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* **18**, 321–328 (2010).

153. Li, J. *et al.* The Mediterranean diet, plasma metabolome, and cardiovascular disease risk. *Eur. Heart J.* **41**, 2645–2656 (2020).

154. Verfürden, M. L., Gilbert, R., Lucas, A., Jerrim, J. & Fewtrell, M. Effect of nutritionally modified infant formula on academic performance: linkage of seven dormant randomised controlled trials to national education data. *BMJ* **375**, e065805 (2021).

155. Paternoster, L., Tilling, K. & Davey Smith, G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges. *PLOS Genet.* **13**, e1006944 (2017).

156. Stubbs, B. *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 . *Eur. Psychiatry* **54**, 124–144 (2018).

157. Schooling, C. M., Freeman, G. & Cowling, B. J. Mendelian Randomization and Estimation of Treatment Efficacy for Chronic Diseases. *Am. J. Epidemiol.* **177**, 1128–1133 (2013).

158. Lor, G. C. Y. *et al.* Reporting and guidelines for mendelian randomization analysis: A systematic review of oncological studies. *Cancer Epidemiol.* **62**, 101577 (2019).

159. Poole, R. *et al.* Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes.

BMJ **359**, (2017).

160. Altman, D. G. & Bland, J. M. Statistics notes: Absence of evidence is not evidence of absence. *BMJ* **311**, 485 (1995).

161. Sterne, J. A. C. & Smith, G. D. Sifting the evidence—what's wrong with significance tests? *Phys. Ther.* **81**, 1464–1469 (2001).

162. Nye, B. E. *et al.* Understanding Clinical Trial Reports: Extracting Medical Entities and Their Relations. *AMIA Jt. Summits Transl. Sci. proceedings. AMIA Jt. Summits Transl. Sci.* **2021**, 485–494 (2021).

163. Whitton, J. & Hunter, A. Automated tabulation of clinical trial results: A joint entity and relation extraction approach with transformer-based language representations. (2021).

164. Tasneem, A. *et al.* The Database for Aggregate Analysis of ClinicalTrials.gov (AACT) and Subsequent Regrouping by Clinical Specialty. *PLoS One* **7**, e33677 (2012).

165. Tenaerts, P., Madre, L., Archdeacon, P. & Califf, R. M. The Clinical Trials Transformation Initiative: innovation through collaboration. *Nat. Rev. Drug Discov.* **13**, 797–798 (2014).

166. Ochoa, D. *et al.* Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Res.* **49**, D1302–D1310 (2020).

167. Wishart, D. S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2017).

168. Elsworth, B. & Gaunt, T. R. MELODI Presto: a fast and agile tool to explore semantic triples derived from biomedical literature. *Bioinformatics* **37**, 583–585 (2020).

169. Fine, L. J., Philogene, G. S., Gramling, R., Coups, E. J. & Sinha, S. Prevalence of multiple chronic disease risk factors: 2001 National Health Interview Survey. *Am. J. Prev. Med.* **27**, 18–24 (2004).

Table 1. Completeness of MeSH term annotation amongst 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	10,187	1,915	18.8%	936	9.2%
Literature	Drug	9,554	1,526	16.0%	914	9.6%
Main	Behavioral	1,013	872	86.1%	203	20.0%
Literature	Behavioral	2,427	2,248	92.6%	496	20.4%
Main	Dietary Supplement	215	102	47.4%	30	14.0%
Literature	Dietary Supplement	662	449	67.8%	118	17.8%

Table 2. Drug target-disease matches supported by evidence from MR (blood pQTL instruments, Zheng et al. 2020) 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	Hyperlipidemia, Dyslipidemia, Hypercholesterolemia, Mixed Dyslipidemia	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,	Hyperlipidemia, Dyslipidemia, Hypercholesterolemia, Mixed Dyslipidemia	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

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 analyzed exposures intervention intensity	different types of exercise 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 assesment scale, such as Hamilton Depression Rating Scale
	similarity between analyzed outcomes	different measures of adiposity
Source population	demographics health status	young adults or elderly 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

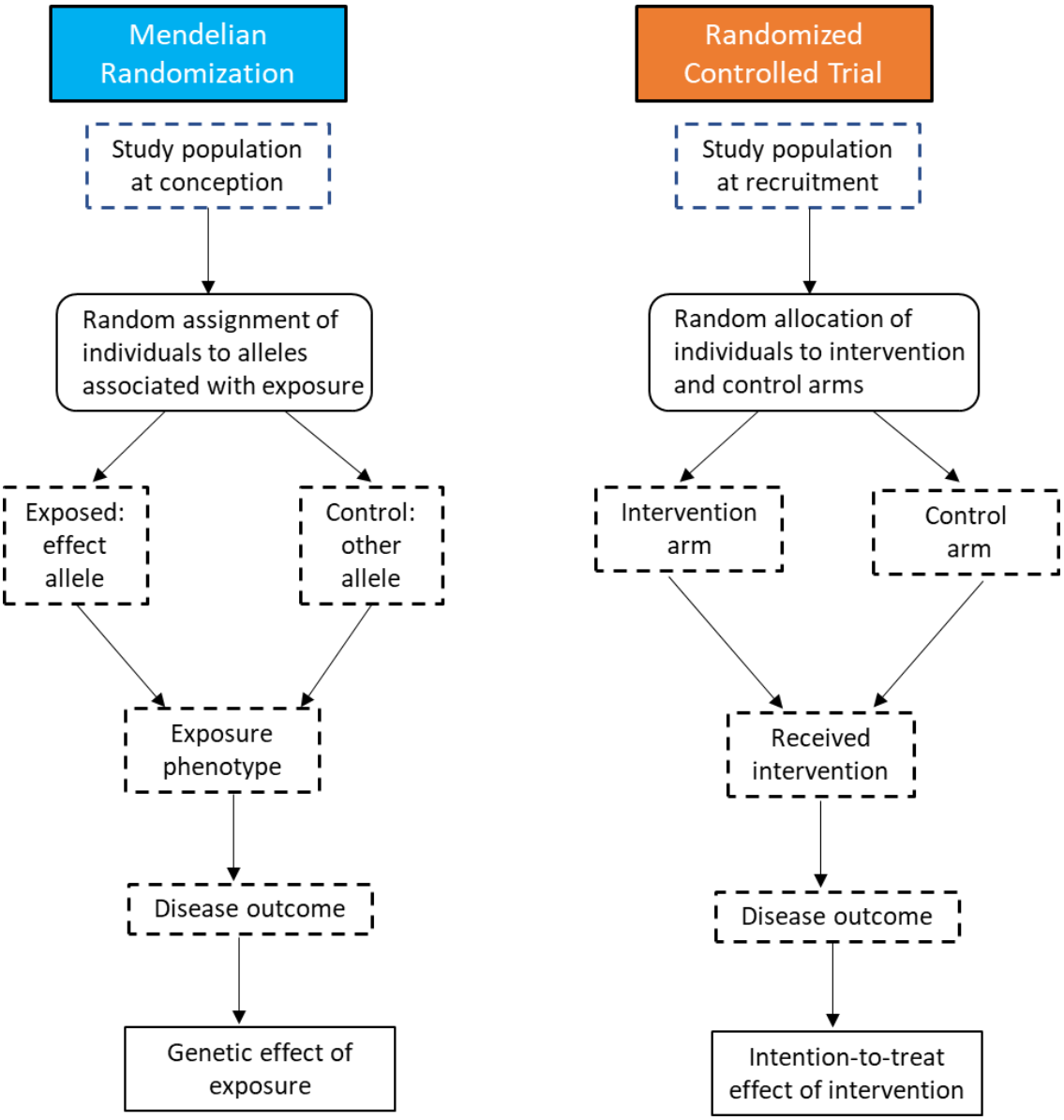

Figure 1. **Comparison of Mendelian Randomization and Randomized Controlled Trial design.** After: Nitsch *et al.* (2006), Ebrahim & Smith (2008), Ference (2018).

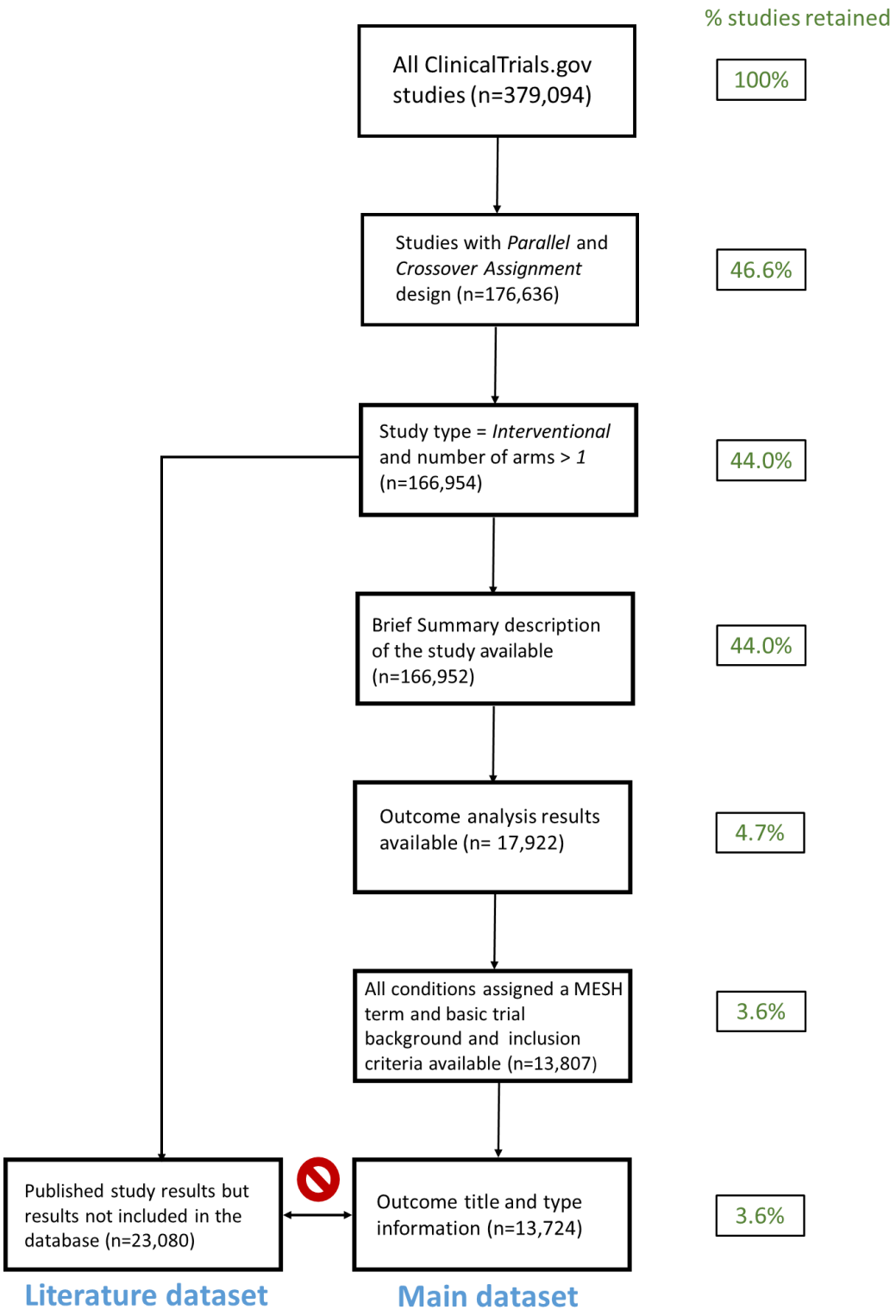

Figure 2. Filtering steps applied to ClinicalTrials.Gov database. Filtering was designed to identify RCT trials whose final results statistics were uploaded to the database (*Main dataset*). In addition, other RCTs which published their findings in scientific journals were identified (*Literature dataset*).

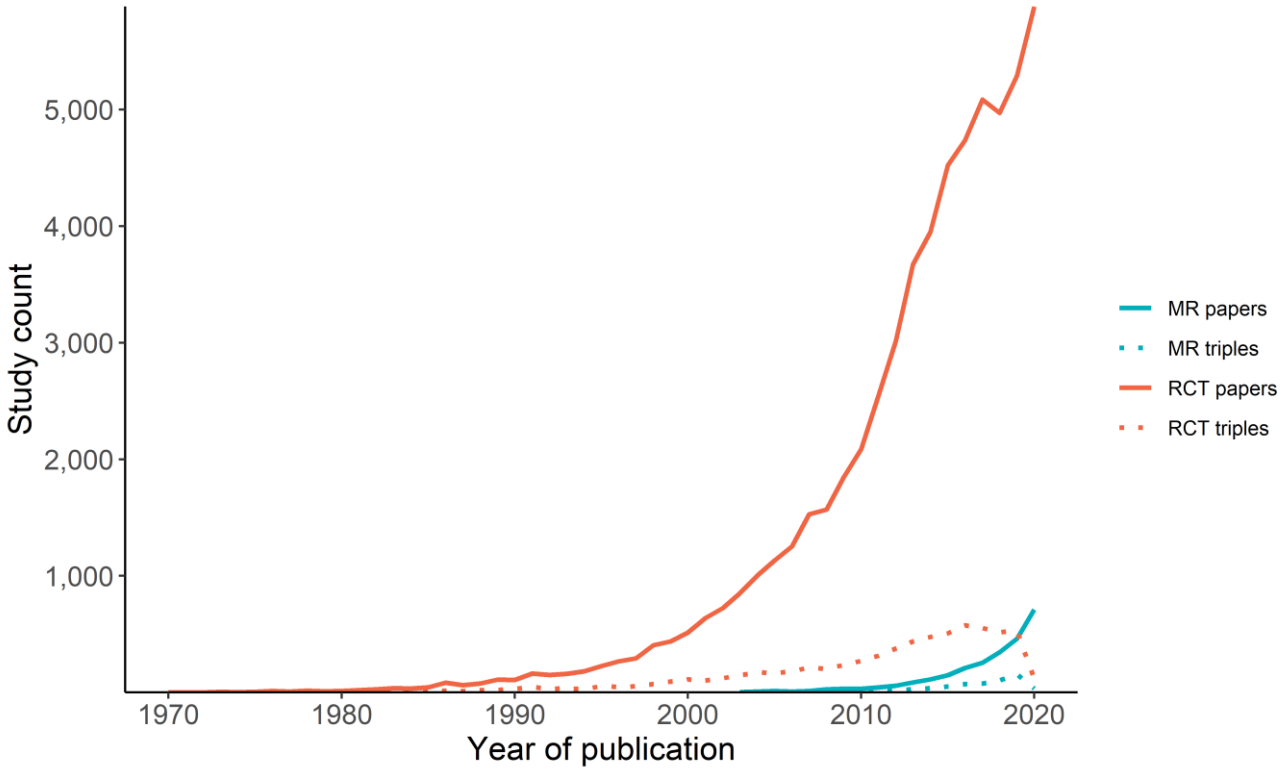

Figure 3. Popularity of MR and RCT studies over time. We compare counts of MR and RCT papers indexed by PubMed (solid lines) with number of semantic triples derived from them using SemMedDb (dashed lines).

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

- all positive effect size
- majority positive effect size
- all null
- majority null
- all negative effect size
- 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%).

Supplementary Files

Supplementary Box

Supplementary Note

Supplementary Figures

Figure S1. Overview of RCT general features in the *Main dataset* by their frequency: a) intervention model, b) primary purpose, c) trial status, d) trial phase, e) gender, f) number of arms.

Figure S2. Distribution of: a) primary outcome number among RCT studies, b) secondary outcome number among RCT studies, c) P-value among study results.

Figure S3. Overview of SemMedDb semantic triples derived from MR (blue, left-hand panels) and RCT (orange, right-hand panels) studies: a) & b) top 15 subjects, c) & d) top 15 predicates, d) & e) top 15 objects.

Supplementary Tables

Table S1. Comparison of the most common intervention types in the *Main* and *Literature* datasets.

Table S2. Drug target-disease matches supported by evidence from MR (blood eQTL instruments, https://epigraphdb.org/xqtl) and RCT studies (*Main dataset* from ClinicalTrials.Gov).

Table S3. Matching subject-object pairs between MR and RCT studies in semantic triples identified by SemMedDb.

Table S4. Matching subjects between MR and RCT studies in semantic triples identified by SemMedDb.

Table S5. Matching objects between MR and RCT studies in semantic triples identified by SemMedDb.

Table S6. Sample comparison of MR and RCT studies with matching exposures and outcomes across a range of criteria.

Supplementary Datasets

Supplementary Dataset 1. Select ClinicalTrials.Gov data fields for the RCT studies identified as the *main dataset*.

Supplementary Dataset 2. Select ClinicalTrials.Gov data fields for the RCT studies identified as the *literature dataset*.

Supplementary Dataset 3. Comparison of frequency of RCT general features between the *Main dataset* and background of all RCTs available in ClinicalTrials.Gov.

Supplementary Dataset 4. Case series of MR and RCT studies with matching exposures (interventions) and outcomes (conditions).