Multiple Sclerosis and Related Disorders 24 (2018) 72–78

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Review article

An introduction to Mendelian randomization with applications in neurology

Phillip H. Allman, Inmaculada B. Aban, Hemant K. Tiwari, Gary R. Cutter*

University of Alabama at Birmingham, Department of Biostatistics. Ryals Public Health Bldg. RPHB 327, AL, USA

ARTICLE INFO

Keywords:
Genetics
Mendelian randomization
Neurology
Multiple sclerosis
Alzheimer
Vitamin D

ABSTRACT

Mendelian randomization studies have become increasingly common due to the maturation of genome-wide association studies and its potential to ascertain causal relationships. With the increasing use of this method comes the need for medical practitioners and clinicians to develop an understanding of its rationale, limitations, and interpretation. Mendelian randomization attempts to ascertain a causal relationship between some risk factor of interest and some outcome or disease of interest. It exploits Mendel's law on the random assortment of genetic variants. This random assortment of genetic variants mimics the main principle of randomization used in clinical trials; with the genetic variant replacing the randomly allocated treatment. In this paper we provide a readable introduction to the rationale behind Mendelian randomization and its limitations. We also discuss and interpret several examples of Mendelian randomization analyses which pertain to neurological diseases.

1. Introduction

Mendelian randomization represents a novel epidemiological study design that incorporates genetic information into traditional epidemiological methods. Studies based on Mendelian randomization, and its applications to Neurology, have become increasingly common as genetic knowledge of health and disease expands with data from genome-wide association studies and whole genome sequencing. In Fig. 1 we can see the number of publications pertaining to Mendelian randomization in Neurology has risen substantially since 2011. Mendelian randomization provides an approach to addressing questions of causality without many of the typical biases that impact the validity of traditional or common epidemiological approaches. While Mendelian randomization can provide important suggestive evidence for causal relations between a risk factor and an outcome or disease, they are not true experiments, and are dependent on several strong assumptions. Evidence from randomized control trials, when possible, should continue to guide clinical decisions. However, Mendelian randomization studies are increasingly being used to identify potential targets for new drugs prior to embarking on costly randomized controlled trials. This paper will discuss the rationale and limitations of Mendelian randomization as a study design. We complete this paper with several examples of Mendelian randomization applied to Neurology and practical conclusions. But before we begin, it is first necessary to discuss causality, randomization, and confounding.

2. Causality

A variable (e.g., a risk factor, phenotype, or exposure) is said to be causal to some outcome if the variable directly leads to the occurrence of the outcome, and we can show that it does. The outcome here could be any number of things; such as a biological trait like blood pressure, an adverse event such as mortality, or a disease such as multiple sclerosis (MS).To show that a variable is causal for an outcome, we must satisfy several criteria, as outlined in the seminal paper by Hill (1965). These criteria include statistical information (i.e., a correlation), a solid biological argument, a proper sequence in time where the causal variable precedes the outcome, a similar result obtained from several investigations of the relationship, and we must be sure that the variable-outcome association is not distorted by any external factors (Hill, 1965). The previous point is crucial as it implies that data alone will typically not be enough to determine causality. That is, it says that finding a strong statistical correlation between a risk factor and an outcome is not sufficient to deduce a causal relationship. We need more than data if we wish to show causality. Furthermore, many things can simultaneously be causal for some outcome or adverse event; finding that one risk factor is causal for some outcome may not be the end of the story.

3. Randomization and clinical trials

A randomized clinical trial (RCT) is widely accepted to be the experimental design that gives us the best chance at ascertaining a causal

* Corresponding author.
E-mail addresses: allman@uab.edu (P.H. Allman), cutterg@uab.edu (G.R. Cutter).

https://doi.org/10.1016/j.msard.2018.06.017
Received 21 March 2018; Accepted 19 June 2018

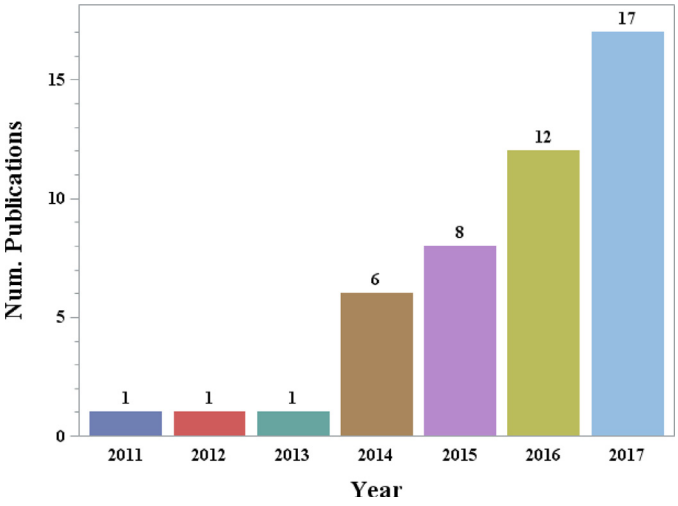

Fig. 1. Search results of publications on Mendelian randomization and neurology since 2011 available on Pubmed. (https://www.ncbi.nlm.nih.gov/pubmed/?term=Mendelian+Randomization+and+Neurology).

relationship between some risk factor and an outcome. They are considered the gold-standard of evidence, in fact Class I evidence, when designed and performed appropriately. The random allocation of a treatment or drug to study participants is the key characteristic of a RCT. This random allocation says that the treatment a patient receives is not influenced by some external factor, such as the patient's or clinician's preconceived opinions or socioeconomic factors. This randomization of patients to different treatments is the primary protector of our results from being affected by external factors; which may mask, skew, or distort our measurement of the true effect of the treatment. That is, randomization protects our results from being biased. If the treatment allocation in a RCT is truly random, then no external factors have influenced the assignment of patients to a particular treatment. In this case, we can be confident that our measured association between risk factor and outcome is reasonable and unbiased. It also strengthens the argument that, if we find a statistically significant association, then it is indeed a causal one. When a treatment allocation in a RCT is not random, for example when the patient's or clinician's beliefs or backgrounds played a role in treatment allocation, then it said to be biased. There may be some systematic differences between the different treatment groups. These systematic differences between the treatment groups may lead to a biased measurement of the risk factor-outcome association. Our measurement of the risk factor-outcome association has been distorted, and we are not sure if this measurement is indeed representative of the true treatment effect. In this case, the observed changes in the outcome may or may not be due solely to the treatment; they could also be due to the outside factor, or some combination of the treatment and the outside factor. Clearly, the randomness of treatment allocation is crucial in obtaining valid and unbiased results. Other design aspects of RCTs, such as blinding, help protect from the influence of outside factors.

It is crucial to recognize the sequence of events in a randomized study: 1) the treatment is assigned randomly and not influenced by some external factor; 2) the treatment then produces changes in the risk factor; and 3) the changes in the risk factor then affect the outcome after some time (if indeed there exists an association between the risk factor and outcome). Thus there are (at least) four variables at play in a randomized experiment: 1) the external factor, 2) the treatment assigned, 3) the risk factor of interest, and 4) the outcome or disease of interest. From these four variables, we choose any two and study the association between them, leading to (at least) six associations that we could investigate. For our purposes, we focus on the following four associations and their behavior:

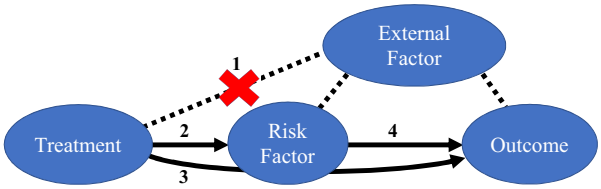

Fig. 2. Graphical representation of the associations between an external factor, treatment, risk factor, and outcome or disease of interest in a properly randomized experiment.

1. The association between the external factors and the treatment allocation
2. The association between treatment and the risk factor
3. The association between treatment assigned and the outcome
4. The association between risk factor and outcome.

These associations behave in a very particular way in a properly randomized experiment. We can see them graphically in Fig. 2, with numbers 1–4 corresponding to the list of associations above. The first association should be a null one. If we truly randomized patients to treatment groups, then external factors do not affect the treatment allocation (represented by the red X over arrow 1 in Fig. 2). The second association, between the treatment and the risk factor, should be a measurable and direct one. The treatment should induce changes in the risk factor; and the treatment is the one causing the changes in the risk factor, not the other way around. The third association, between the treatment and the outcome, should act only through the risk factor. The treatment affects the outcome only by the induced changes in the risk factor. The treatment should have no direct effect on the outcome. The fourth association, between the risk factor and the outcome, should be a direct one or a null one. If our study hypothesis is true, then changes in the risk factor should induce changes in the outcome. If our hypothesis is not true, then changes in the risk factor should not induce changes in the outcome. As we shall see, the associations in a proper Mendelian randomization analysis have these same behaviors.

4. Confounding and observational studies

The distortion of the relationship between risk factor and outcome by an external factor is most often referred to as confounding; and the variables which cause this distortion are called confounders. As described above, confounding can be mitigated rather well in a RCT through the use of randomization. However, other study designs may not be able to protect from confounding. Observational studies have long been performed when a RCT is not feasible or ethical. They are extremely useful for gathering large amount of data at low cost, and for gaining insight into relationships that are not well understood. However, there is a price to pay in the form of confounded results. When an association is found between some risk factor and some outcome or disease in observational data, there is no guarantee that it is a causal one. The relationship may have been distorted by some confounder, or reverse causality may have occurred (i.e., the outcome or disease preceded and caused changes in the risk factor). Indeed, results from observational data have often been overturned once the relationship was investigated in subsequent RCTs (Lawlor et al., 2004; Greenwald, 2003). But if we cannot perform a RCT, and the results from observational studies are unreliable and confounded, what are we to do? In specific situations, a study using Mendelian randomization may provide an answer. The relatively recent application of instrumental variable (IV) methods to the field of genetics, under the name of Mendelian randomization, may address the issue of confounded relationships in observational data when supplemented with genetic data (Almeida et al., 2014; Mokry et al., 2015; Pichler et al., 2013; Rhead et al., 2016; Smith and Ebrahim, 2003; Smith and Ebrahim, 2004;

Lawlor et al., 2008; Zhan et al., 2015).

5. Random assortment of genetic alleles

We may draw an analogy between randomly allocating some treatment in a RCT, and the random assortment of genetic alleles during reproduction (meiosis). Gregor Mendel was the first to theorize this random assortment of genes from parents to offspring in plants (Mendel, 1865). This has become known as Mendel's second law, and applies to humans as well. Mendel's law provides the basis for our analogy between random assignment of treatments within a RCT, and the random allocation of genetic alleles to offspring. In particular, his second law says that inherited genes are independent of external factors (i.e., confounders) such as patient's choices. To complete the analogy, we must be certain that the associations between the gene, risk factor, and outcome behave in the same way as they do in a RCT; with the gene replacing the assigned treatment. We specify that the gene of interest directly affects only the risk factor of interest, and that the gene affects the outcome only through the induced changes in the risk factor. When we find such a gene, we have found a serendipitous randomization, one which has been implemented by nature. A randomly assigned treatment has essentially been replaced with a randomly inherited gene; and, as we shall see, the behavior of the relationships between gene, risk factor, and outcome are the core assumptions of IV analyses and Mendelian randomization.

6. Precursors to Mendelian randomization

Katan was one of the first to realize that gene variants could be exploited for their random allocation to mitigate confounding (Katan, 1986). In the early 1980s there was considerable debate over the hypothesis that low serum cholesterol might directly increase cancer risk. However, there were many competing hypothesis: 1) reverse causation could be at play, latent tumors preceded and lead to low levels of cholesterol with later diagnoses of cancer; 2) smoking was associated with lower cholesterol which in turn increased one's cancer risk; 3) diet led to low cholesterol and higher incidence of cancer; etc. Essentially, researchers were not sure of the direction of the relationship between the risk factor and disease outcome and the role of confounders. Katan knew that it had been observed that individuals with abetalipoproteinaemia (or Bassen-Kornzweig syndrome) have negligible levels of serum cholesterol, and that these individuals did not seem predisposed to cancer. The apolipoprotein E (*ApoE*) gene was known to affect serum cholesterol, specifically the *ApoE2* variant was associated with lower levels of cholesterol. Katan's idea was that many people will carry the *ApoE2* variant and naturally will have lower levels of cholesterol from birth. Since the genes are randomly assigned at meiosis, these *ApoE2* carriers will not systematically differ from the carriers of other alleles. Thus there should be no confounding between low cholesterol induced by the genetics and the outcome of cancer. If low levels of cholesterol are causative of cancer, then cancer patients should have more *ApoE2* alleles compared to controls. Otherwise, the distribution of *ApoE2* alleles should be similar in both cases and controls. He essentially argued that the *ApoE* gene variant could serve as a substitute for random treatment allocation in a RCT because it directly impacts cholesterol, and that its effect on cancer was only due to the induced changes in cholesterol. Katan's reasoning corresponds to IV methods, and similar applications of his idea to genetics later became known as Mendelian randomization.

7. Mendelian randomization: instrumental variables with genetic variants

Instrumental variables (IVs) have long been applied to econometrics, and more recently to epidemiology and medicine (Goldberger, 1972; Greenland, 2000; Newhouse and McClellan, 1998; Robins, 1989; Sheiner and Rubin, 1995; Wright, 1928). To find an IV is to find a variable which plays a similar role to the random treatment allocation in a RCT. Finding an IV is not easy; strict and specific requirements are needed for an IV to be a valid one. These requirements pertain to the relationships between the IV, the risk factor of interest, and the outcome of interest. They essentially state that these relationships behave as they would in a RCT. Specifically, for an IV to be a valid one, we require three core conditions to be met (Moffitt, 1996; Angrist et al., 1996):

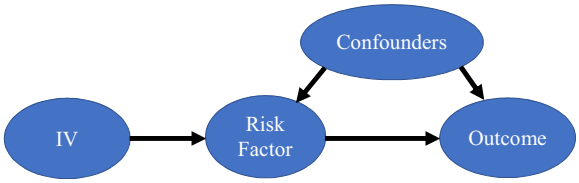

Fig. 3. Graphical representation of the associations required between an instrumental variable, risk factor, and outcome.

1. The IV must be independent of all confounders.
2. The IV must be directly associated with the risk factor of interest.
3. The IV must only affect the outcome of interest via the induced changes in the risk factor.

Condition 3 warrants a little more explanation. It says that the IV has no direct effect on the outcome. In other words, it states that the IV is independent of the outcome given the risk factor and confounders. The confounders in these conditions may be any number of variables; including socio-economic factors, medical history, demographics, etc. They could even be unmeasured or unimagined entirely. Fig. 3 graphically displays a scenario in which the IV meets all three assumptions. The lack of an arrow connecting the IV and the confounders represents condition 1 above. The arrow directed from the IV to the risk factor represents condition 2 above. The fact that the only path from the IV to the outcome passes through the risk factor represents condition 3 above. Violations of these three conditions lead to doubt in the validity of our chosen IV, as well as the results from the analysis. In Mendelian randomization studies, these violations are often due to linkage disequilibrium, pleiotropy, population stratification, canalization, and gene-environment interactions; described below. If these three conditions are met, then we may test whether or not there is any causal association between the risk factor and outcome. Two additional conditions are required if we wish to get an estimate of the magnitude of this causal relationship. These conditions are often difficult to verify with statistics alone. Specifically, core condition 3 above is sometimes impossible to verify statistically. Instead, a combination of statistics and knowledge of the biological mechanisms is used to verify these conditions. Any analysis which uses Mendelian randomization should provide some evidence, both statistical and biological, that these conditions are met.

Mendelian randomization is simply an application of IV methods in which the IV is specified to be a genetic variant or genetic score. The genetic IV could also be some combination of several genetic variants with similar characteristics and biological pathways, so long as the IV assumptions are still met by this aggregated genetic variable. Mendelian randomization is most useful when the association between risk factor and outcome is considerably confounded, or when measurement error is substantial. To perform a reliable Mendelian randomization analysis, we need a well-understood biological framework for the relationship between the genetic variant and the risk factor of interest. When the biological framework for a genetic variant is not well-understood, then we are unable to determine if the three core IV assumptions for our gene of choice are met. Mendelian randomization cannot serve as a replacement for RCTs; violations of the IV assumptions may be difficult to detect, and a valid genetic IV may not always

be available. Even when the IV assumptions are met, the results of a Mendelian randomization analysis may not agree with results from RCTs. In particular, the patient is affected by their genotype since birth; whereas the interventions performed in RCTs are typically over much smaller time scales. The pathway in which the genetic IV affects the risk factor may also differ from the pathway in which a RCTs intervention affects the risk factor. Large samples are required in a Mendelian randomization analysis; often, thousands of patients are needed to get reasonably accurate estimates of effect size and variance. Despite these limitations, Mendelian randomization may serve as a better way to ascertain causality, and better precursors to future RCTs, than observational studies when a valid genetic IV is available.

8. Evaluating the causal effect in Mendelian randomization

If we are only interested in whether or not there is a nonzero causal effect, then we must first check that the three core IV conditions above are met. Only then may test whether or not any causal effect exists. This is done with an appropriate statistical hypothesis test, chosen depending on the nature of the gene and outcome. For example, if both the gene and outcome are categorical variables and we have an adequate sample size, then a Chi-square test may be performed. If we find a statistically significant association between the genetic variant and outcome (and the three core assumptions are met), then it suggests a causal effect of the risk factor on the outcome. If, in addition, we wish to *estimate* the causal effect, then we must meet two additional assumptions: linearity and no interactions in the relationship between risk factor and outcome. By assuming linearity, we mean that the risk factor (X) and outcome (Y) are linearly related. By assuming no interactions, we require that the risk factor-outcome association is the same across various subgroups. With all five assumption met, the causal effect may be estimated as the ratio of 1) the gene-outcome association and 2) the gene-risk factor association. Thus, when all five assumptions hold, we may evaluate

$$\text{Causal Association of risk factor and disease} = \frac{\text{Association of gene and disease}}{\text{Association of gene and risk factor}} \quad (1)$$

to obtain an estimate of the causal effect of the risk factor on the outcome.

9. Limitations to Mendelian randomization and assumption violations

Limitations to Mendelian randomization consist of some of the usual limitations in observational studies, and violations of the IV assumptions. These limitations have been discussed at length elsewhere (Smith and Ebrahim, 2003; Lawlor et al., 2008; Thanassoulis and O'Donnell, 2018). We present some of these issues below.

9.1. Pleiotropy, linkage disequilibrium, and population stratification

Pleiotropy, linkage disequilibrium, and population stratification are fundamentally different phenomena. Despite their differences, their effect on the IV assumptions is the same. All three phenomena induce violation of the third IV assumption and confound or bias the risk factor-disease relationship. They are all characterized by the genetic IV having a relationship with the outcome which passes through some other variable that is not the risk factor of interest. For our purposes we refer to the variable that opens an alternative path from the genetic IV to the outcome as a *distorting variable*, because it distorts the risk factor-outcome association. The behavior of this distorting variable is indeed analogous to a confounding variable. Fig. 4 displays a scenario in which pleiotropy, linkage disequilibrium, or population stratification is present. The difference between these issues lies in the type of distorting variable. With pleiotropy, the distorting variable is an alternative risk factor which is not the one of interest. With linkage disequilibrium (LD), the distorting variable is another correlated genetic variant which is nearby our genetic variant. With population stratification (PS), the distorting variable is some external factor; usually being a demographic variable such as race, ethnicity, sex, or age in the example below (Cardon and Palmer, 2003). For a more detailed discussion on these issues, we refer the reader to Smith and Ebrahim (2003) and Thanassoulis and O'Donnell (2018).

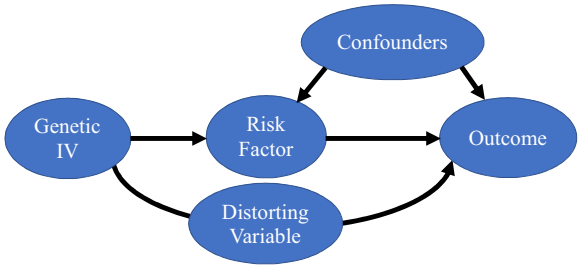

Fig. 4. Diagram of a Mendelian randomization in an alternative pathway exists between the genetic IV and the outcome or disease of interest, causing confounding of the risk factor-outcome association, and violating a core IV assumption.

9.2. Canalization

Canalization (a.k.a., genetic buffering or developmental compensation) essentially states that an organism may compensate for the effects of having a particular genotype (Waddington, 1942). This may occur over time or during the early developmental stages. In a Mendelian randomization setting, the association between the genotype and an outcome may be weakened and biased due to canalization. This issue is typically not present in randomized clinical studies. The 'randomization' via the gene occurs at birth and prior to development, while randomization done in an RCT is typically performed on a fully developed adult. Thus the body may be able to dilute the effects of a particular genotype, thanks to a compensatory response from the development of other systems and the extended time for adaptation. This phenomenon can be seen in knockout studies of mice (Garry et al., 1998). The behavior of the compensatory response is often not well-understood; thus the magnitude of this issue in Mendelian randomization studies is not well-understood either.

9.3. Unreliable estimates of associations

Recall that in a Mendelian randomization analysis, we must evaluate the association between the outcome and the genotype, and additionally the association between the risk factor and genotype if we want to estimate the magnitude of the causal effect. These association estimates may be considered unreliable if they lack adequate precision (i.e., confidence intervals that are too wide), or are inaccurate (i.e., biased). If these association estimates are unreliable, then the results of our Mendelian randomization analysis are unreliable as well. They may not be reliable due to the same issues present in all genetic association studies, and often present themselves as replication issues. These issues include lack of power, chance, publication bias, genotyping errors, and measurement error in the risk factor. This issue may be addressed by obtaining association estimates from multiple independent samples and populations.

9.4. Lack of a suitable genetic IV

We may not be able to determine a gene variant that clearly affects our risk factor of interest, and sometimes one may not exist at all. A genetic IV is considered weak if it has a weak association with the risk

factor of interest (i.e., the variance in the risk factor explained by the gene variant is a small portion of its total variance). Weak genetic IVs may result in poor precision and accuracy of the causal effect estimate. Even if we do find a suitable gene, it may not have enough common variants. An example of this occurs in the vitamin C and coronary heart disease (CHD) association (Smith and Ebrahim, 2003; Erichsen et al., 2001). This issue may be resolved over time, as genomics research uncovers the pathways between genetic variants and risk factors.

10. Example 1

A recent paper by Rhead et al. used Mendelian randomization to investigate the effect of low serum 25(OH)D levels on risk of multiple sclerosis (MS) (Rhead et al., 2016). The authors perform this analysis using two independent populations, adding strength to their results if they indeed are similar in each population. One of these study samples included 1056 patients with MS and 9015 controls from the Northern California region using data from the Kaiser Permanente Northern California (KPNC). The other study sample consisted of 6335 MS patients and 5762 controls from Sweden using data from the Epidemiological Investigation of Multiple Sclerosis (EIMS) and the Genes Environment in Multiple Sclerosis (GEMS) study. Both samples consist entirely of non-Hispanic whites, and thus their results may not generalize to other ethnicities.

The authors created a genetic IV comprising of three published genetic variants from Genome-Wide Association Studies (GWASs): rs2282679-A in the *GC* gene, rs2060793-A upstream of the *CYP2R1* gene, and rs3829251-G in the *NADSYN1* gene. The authors create a single genetic IV by combining these three SNPs into a weighted genetic risk score (wGRS); using a weighting average of their associations with vitamin D. Thus, their Mendelian randomization has the single aggregated genetic IV, vitamin D as the risk factor of interest, and developing MS as the outcome of interest.

These genetic variants are purported to drive the vitamin D levels which in turn predispose the individual to MS. There appears to be a well-understood role of these genes in affecting vitamin D, and the authors provide some evidence that vitamin D may regulate genes that play a role in the development of MS. Thus the biological framework seems well established, again adding validity to their results. The authors also do a reasonable job of investigating potential violations of the three core IV assumptions above. The do however discuss a potential issue that pertains to the variant rs3829251-G; in which this allele may play a role in cholesterol, this in turn may affect MS risk and severity (i.e., pleiotropy). When the IV assumptions for one of the genetic variants may be violated, a common approach is to perform a sensitivity analysis; in which the questionable genetic variant is removed and results re-calculated. The validity of their results would be strengthened by doing so. Another potential issue is that the variance in Vitamin D levels explained by their aggregated genetic variants is relatively low at approximately 5%; despite the variants being in or close to the genes involved in vitamin D synthesis (Jiang et al., 2018, Wang et al., 2010, Ahn et al., 2013, Hiraki et al., 2013). This suggests that the IV may be a weak one; and may violate the second IV assumption. In addition, their method to create the wGRS may be improved. Recall that the authors create a wGRS using the associations observed from GWAS. These association measures (i.e., β coefficients) depend on the measurement units and techniques; which may be differ between the study samples. This in turn can induce measurement error when results are combined across study samples. Standardizing the weights using the standard error could be the better alternative. One final critique may be made to the presentation of their results. Recall that to estimate the causal effect, we divide the gene-outcome association by the gene-risk factor association, thus a clear presentation of results should include both association estimates for all genetic IVs. The authors do not present the gene-outcome associations for each of their genetic IVs, making it difficult to verify their results without referring to other articles.

Despite the limitations above, we proceed to interpret their aggregated results pertaining to the causal effect of vitamin D on MS risk. The authors estimate a 21% (95% CI: 1%–36%) reduction in MS risk caused by higher vitamin D in the KPNC study. They obtain a similar result in the EIMS/GEMS sample, finding a 14% (95% CI: 2%–24%) causal reduction in MS risk by higher vitamin D. Thus, based on these results, we conclude that lower vitamin D causes an increase in MS risk. But what do we do with this conclusion? Recall that if a person has a particular genetic variant, then they experience its effects over their entire lifetime. Meanwhile, if we were to introduce an intervention which increased the patient's vitamin D levels, then its effects would only be experienced while taking the treatment (perhaps after a washout period as well). Even if we assume that low vitamin D causes an increase in MS risk, then it is not clear how long we would have to increase their vitamin D levels to observe a beneficial effect. Additionally, this study finds evidence for a causal association between lower vitamin D and *onset* of MS. It does *not* provide evidence that increasing vitamin D in a patient who *already has* MS will be beneficial. Despite the interpretive issues, this article presents moderate evidence of a causal effect of vitamin D on MS risk. Perhaps future clinical trials will provide a more decisive answer.

11. Example 2

Mokry et al. recently investigated the causal role of vitamin D in Alzheimer disease (AD) risk from three different studies (Mokry et al., 2016). Observational studies find evidence to suggest that decreased vitamin D levels are associated with increased AD risk (Littlejohns et al., 2014; Afzal et al., 2014). However, there is potential that confounding or reversal causality have distorted these results; and there is a lack of data on this association from RCTs. Thus this situation lends itself well to Mendelian randomization. Mokry et al. begin by searching for genetic variants associated with vitamin D using data from the SUNLIGHT Consortium (Wang et al., 2010). From this they identify four SNPs associated with vitamin D and use them as genetic IVs. These genetic IVs include rs2282679 in *GC* (IV1), rs12785878 near *DHCR7* (IV2), rs10741657 near *CYP2R1* (IV3), and rs6013897 in *CYP24A1* (IV4). We define each of these genetic IVs as IV1-IV4 for purposes of this example. They investigate the risk factor-outcome association using each of these SNPs separately. With each of these separate estimates, they then average them (using a fixed-effects meta-analysis model) to obtain a pooled estimate of the risk factor-outcome association. Thus they set up a Mendelian randomization analysis for each of the four SNPs. The framework for their Mendelian randomization analyses has one of these four SNPs as the IV, vitamin D as the risk factor of interest, and AD as the outcome of interest.

The vitamin D pathway involving these genes and SNPs seems well-understood, adding strength to their analysis. They obtain estimates of the association of their selected genetic IVs with vitamin D levels from the Canadian Multicentre Osteoporosis Study sample. These association estimates control for sex, age, BMI, and seasonal variations. Estimates of the association between vitamin D and AD risk were obtained from a separate sample, the International Genomics of Alzheimer's Project. Obtaining these association estimates from separate samples in indeed a valid approach (Burgess et al., 2013).

The authors discuss possible violations of the three IV assumptions for these genetic variants. They find that for IV2, the third core instrumental variable assumption may be violated, due to potential pathway between IV2 and AD which passes through a demographic variable (i.e., population stratification). They also find evidence that the third core instrumental variable assumption could be violated for IV1, due to a potential pathway between IV1 and AD which passes through a vitamin D binding protein. Despite these potential violations of the Mendelian randomization assumptions, the authors proceed with the analysis using all four genetic IVs. Thus interpretations of results which use IV1 and IV2 should be made cautiously. Also, recall that we

must assume linearity and no interactions to obtain a *point estimate* of the causal effect of the risk factor on the outcome in Mendelian randomization. The authors indeed present point estimates of this causal effect, however, no discussion is given on the possibility that linearity and no interactions has been violated, it is merely assumed. Thus, again, we proceed with caution in interpreting the causal effect estimates.

When all four IVs are used separately to assess the association between vitamin D and AD risk, none are found to show a statistically significant association at the 0.05 level. For IV1, they find that a 1 standard deviation (SD) reduction in vitamin D is associated with a 45% (95% CI: −2%–114%) causal increase in the risk for AD. This result is not statistically significant at the 0.05 level because the 95% confidence interval (CI) includes the null point of 0%. Although this result is somewhat close to achieving statistical significant, and may be considered promising, the potential violation of the third instrumental variable assumption creates some additional doubt. Similar results are obtained for IV2-IV4. Only when all these estimates are aggregated to produce a summary measure do they obtain a statistically significant result for the causal effect of vitamin D on AD risk (25% risk increase; 95% CI: 3%–51%). It is hard to say if this is a real effect, if the one of the four genetic IV's has dominated the others when averaged together, or if assumption violations have played a substantial role. The authors present this aggregated estimate as their primary result, and conclude “Our results demonstrated that a decreased 25OHD level was associated with risk of AD, in which a 1-SD decrease in the natural log-transformed 25OHD resulted in a 25% increase in the risk of AD ($p = 0.021$)” (Mokry et al., 2016). This interpretation could be considered a bit overreaching considering the issues above. However, the authors do give an informative sensitivity analysis and appropriately state that their results may have been affected by violations of the IV assumptions.

12. Example 3

Not any characteristic may be used as an IV in a Mendelian randomization analysis. While the occurrence of a characteristic may seem random as allele frequency, keep in mind the importance of the assumptions. Consider the John Cunningham (JC) virus. One may consider the possibility that JC viral status (JC+, JC-) could be used as an IV in a Mendelian randomization based on the random assortment of JC viral statuses in the population. Suppose we have tested individuals and have JC+ versus JC-, and attempt to compare extended dosing of natalizumab (NTZ) every 8 weeks (EID) to the standard dosing of every 4 weeks (SD). Ryerson et al. conducted such an observational study and compare EID to SD in patients with MS (Ryerson et al., 2016). They concluded that “Dosing intervals up to 8 weeks 5 days did not diminish effectiveness of NTZ therapy. Further monitoring is ongoing to evaluate if the risk of PML is reduced in patients on EID”. Suppose we perform a Mendelian randomization using this study; where we let the IV be JC viral status; the annualized relapse (AR) serves as the outcome or disease of interest; and NTZ dosing schedule (EID, SD) takes the place of the risk factor.

Just because the occurrence of the JC virus may effectively be thought to randomize the population, it suffers from several limitations of our assumptions. The first assumption is that the IV is independent of confounders. The prevalence of positivity to the JC virus increases with age because of life long exposure. In addition, relapse rates decline with age thus the confounding does indeed impact both the IV and the outcome variable. The IV is associated with the treatment since it is much more likely to find an EID person to be positive for the JC virus and one might question the biological plausibility that the virus directly impacts the relapse rate. Thus, not every potential variable can be used as an IV in an analysis based on the concepts of instrumental variables.

13. Conclusions

A properly implemented Mendelian randomization is subject to less confounding than conventional observational data analyses. Thus Mendelian randomization may provide more reliable estimates of the causal effect of a risk factor on an outcome of interest when compared to those obtained from observational studies; so long as the three core IV assumptions are met. These three core assumptions in a Mendelian randomization (and all IV analyses) pertain to the behavior of the relationships between the chosen genetic IV, the risk factor of interest, and the outcome or disease of interest. When the three core IV assumptions are met, these relationships behave in a similar way as those seen in RCTs. This method may also be useful in situations where a RCT is unethical or infeasible, so long as we can find a suitable genetic IV. However, this method does not serve as a substitute for a well-designed RCT. A number of limitations may be present in Mendelian randomization analyses, and awareness of these is pivotal to interpreting their results. These limitations include confounding of the causal risk factor-outcome association by an unaddressed distorting variable, canalization, unreliable associations estimates between the genetic IV and risk factor or between the risk factor and outcome, the potential that a suitable genetic IV is not available, and the need for large sample sizes in Mendelian randomization. When interpreting the results of a Mendelian randomization analysis one must consider several things. These include assessing if the three IV assumptions are met, whether or not the association between risk factor and outcome is linear with no interactions, the potential that the limitations above have not been addressed or accounted for, if the biological framework is well-understood or not, and that results from a Mendelian randomization may not coincide with results from a RCT even when both are properly implemented.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

Afzal, S., Bojesen, S.E., Nordestgaard, B.G., 2014. Reduced 25-hydroxyvitamin D and risk of Alzheimer's disease and vascular dementia. Alzheimers Dement 10, 296–302.
Ahn, J., Yu, K., Stolzenberg-Solomon, R., et al., 2013. Genome-wide association study of circulating vitamin D levels. Hum. Mol. Genet. 19, 2739–2745.
Almeida, O.P., Hankey, G.J., Yeap, B.B., et al., 2014. Alcohol consumption and cognitive impairment in older men: a Mendelian randomization study. Neurology 82, 1038–1044.
Angrist, J.D., Imbens, G.W., Rubin, R.B., 1996. Identification of causal effects using instrumental variables (with discussion). J. Am. Stat. Assoc. 91, 444–472.
Burgess, S., Butterworth, A.S., Thompson, S.G., 2013. Mendelian randomization analysis with multiple genetic variants using summarize data. Genet. Epidemiol. 37, 658–665.
Cardon, L.R., Palmer, L.J., 2003. Population stratification and spurious allelic association. Lancet 361, 598–604.
Erichsen, H.C., Eck, P., Levine, M., Chanock, S., 2001. Characterization of the genomic structure of the human vitamin C transporter *SVCT1* (*SLC23A2*). J. Nutr. 131, 2623–2627.
Garry, D.J., Ordway, G.A., Lorenz, J.N., et al., 1998. Mice without myoglobin. Nature 395, 905–908.
Goldberger, A.S., 1972. Structural equations methods in the social sciences. Econometrica 40, 979–1001.
Greenland, S., 2000. An introduction to instrumental variables for epidemiologists. Int. J. Epidemiol. 29, 722–729.
Greenwald, P., 2003. β-carotene and lung cancer: a lesson for future chemoprevention investigations. J. Natl. Cancer Inst. 95, e1.
Hill, A.B., 1965. The environment and disease: association or causation? Proc. R. Soc. Med. 58, 295–300.
Hiraki, L.T., Major, J.M., Chen, C., et al., 2013. Exploring the genetic architecture of circulating 25-hydroxyvitamin D. Genet. Epidemiol. 37, 92–98.
Jiang, X., O'Reilly, P.F., Aschard, H., et al., 2018. Genome-wide association study in

79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat. Commun. 9, 260. http://dx.doi.org/10.1038/s41467-017-02662-2.
Katan, M.B., 1986. Apolipoprotein E isoforms, serum cholesterol, and cancer. Lancet 327, 507–508.
Lawlor, D.A., Harbord, R., Sterne, J., et al., 2008. Mendelian randomisation: using genes as instruments for making causal inferences in epidemiology. Stat. Med. 27, 1133–1163.
Lawlor, D.A., Smith, G.D., Kundu, D., et al., 2004. Those confounded vitamins: what can we learn from the differences between observational versus randomised trial evidence? Lancet 363, 1724–1727.
Littlejohns, T.J., Henley, W.E., Lang, I.A., et al., 2014. Vitamin D and the risk of dementia and Alzheimer disease. Neurology 83, 920–928.
Mendel, G., 1865. Experiments in plant hybridization. In: Meetings of the Brünn Natural History Society, 8 February, pp. 3–47.
Moffitt, R.A., 1996. Identification of causal effects using instrumental variables: comment. J. Am. Stat. Assoc. 91, 462–465.
Mokry, L.E., Ross, S., Ahmad, O.S., et al., 2015. Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. PLoS Med. 12, e1001866.
Mokry, L.E., Ross, S., Morris, J.A., et al., 2016. Genetically decreased vitamin D and risk of Alzheimer disease. Neurology 87, 2567–2574.
Newhouse, J.P., McClellan, M., 1998. Econometrics in outcomes research: the use of instrumental variables. Annu. Rev. Public Health 19, 17–34.
Pichler, I., Del Greco, M.F., Gögele, M., et al., 2013. Serum iron levels and the risk of Parkinsons disease: a Mendelian randomization study. PLoS Med. 10, e1001462.
Rhead, B., Bäärnhielm, M., Gianfrancesco, M., et al., 2016. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. Neurol. Genet. 2, e97.
Robins, J.M., 1989. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest, L., Freeman, H., Mulley, A. (Eds.), Health Service Research Methodology: A Focus on AIDS. Washington: US Public Health Service, pp. 113–159.
Ryerson, L.Z., Frohman, T.C., Foley, J., et al., 2016. Extended interval dosing of natalizumab in multiple sclerosis. J. Neurol. Neurosurg. Psychiatry 87, 885–889.
Sheiner, L.B., Rubin, D.B., 1995. Intention-to-treat analysis and the goals of clinical trials. Clin. Pharmacol. Ther. 57, 6–15.
Smith, G.D., Ebrahim, S., 2003. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int. J. Epidemiol. 32, 1–22.
Smith, G.D., Ebrahim, S., 2004. Mendelian randomization: prospects, potentials, and limitations. Int. J. Epidemiol. 33, 30–42.
Thanassoulis, G., O'Donnell, C.J., 2018. Mendelian randomization. In: UpToDate Raby, B.A., Tirnauer, J.S. (Eds.), Topic 13514 Version 12.0, . Available: https://www.uptodate.com/contents/Mendelian-randomization [Accessed 02.20].
Waddington, C.H., 1942. Canalization of development and the inheritance of acquired characteristics. Nature 150, 563–565.
Wang, T.J., Zhang, F., Richards, J.B., et al., 2010. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet 376, 180–188.
Wright S. In: Appendix, Wright P.G. (ed). 1928, The Tariff on Animal and Vegetable Oils Macmillan, New York.
Zhan, Y., Song, C., Karlsson, R., et al., 2015. Telomere length shortening and Alzheimer disease—a Mendelian randomization study. JAMA Neurol. 72, 1202–1203.