Title:

Mendelian randomization studies of periodontitis – understanding benefits and natural limitations in an applied context**.**

Authors:

Simon Haworth[1], Nicholas J. Timpson[2], Kimon Divaris[3]

1. Bristol Dental School, University of Bristol, Bristol, UK
2. Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, UK
3. Division of Pediatric and Public Health, Adams School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

Corresponding author:

Simon Haworth, Bristol Dental School, 1 Trinity Quay, Avon Street, Bristol, BS20PT, simon.haworth@bristol.ac.uk

Abstract:

Mendelian randomization (MR) is a flexible analytical tool which has been widely applied to strengthen causal inference in observational epidemiology and is now gaining attention in many areas including periodontal research. The interpretation of results drawn from MR is based on a series of assumptions which can be unrealistic or difficult to meet faithfully in some settings. However, we argue that with care this does not necessarily prevent valuable deployment of the approach. We argue that clarity of presentation and careful assessment of specific analytical conditions is a fundamental part of all MR analyses. To that end, awareness of its limitations should also guide the design of MR investigations and the presentation of results rather than rule out use all together. Notably, considerations similar to those known to be important in conventional epidemiological settings apply to MR. While MR studies are valuable in their contrast to other study limitations, the application of this technique must be carefully cross-examined. Specific considerations include possible confounders, recruitment strategy and phenotypic measurement, and differential analysis properties across studies. In the case of periodontal research, current MR applications are limited by the available evidence base for genetic contributions to periodontitis, however this sets a specific scene for the strategic use of MR and shines a light on a need for greater research emphasis on the genetics of the condition and intermediaries. This article provides a perspective on the uses and inherent limitations of Mendelian randomization studies and the importance of adhering to basic epidemiological principles when designing them.

Clinical relevance statement

Scientific Rationale for Study:

Mendelian randomization (MR) is a flexible analytical tool which is increasingly being used in periodontal research.

Principal Findings:

The key assumptions underlying MR are difficult to satisfy perfectly. Results of MR studies can be influenced by multiple factors which are discussed, including confounding, study recruitment and measurement and differential analysis properties across studies. This does not prevent use of the method, but calls for careful examination of each application.

Practical Implications:

Results from MR experiments must be carefully examined and ideally compared against other sources of evidence. The main features to consider when interpreting an MR study are discussed.

Summary

Mendelian randomization (MR) is a useful analytical tool which is now being widely applied to periodontal research to examine the effects of possible risk factors on periodontitis. The key feature of MR is that host genetic variation (e.g., genetic data from human genome-wide association studies) can be used as a proxy for a risk factor of interest, to infer the effects of that risk factor on the outcome of interest (e.g., periodontitis). This approach is attractive because MR studies are less prone to confounding and reverse causation than classical observational epidemiologic investigations and can be done without direct measurement of the putative risk factor in the test data. Conversely, MR studies have their own natural limitations and make strong underlying assumptions which are difficult to satisfy perfectly. Since the strengths and limitations of MR studies differ from other study designs, they can help provide causal inference when used as an adjunct to other study designs and are supported by biological and clinical rationale.

For a meaningful MR, it is crucial to consider the nature of the putative risk factor, because the technique can only be used for risk factors which can be proxied by genetic variation. This makes the technique inherently easier to apply when the risk factor is under tight host genetic control such as a metabolite or biomarker. It is inherently difficult to use the technique when the putative risk factor is only tenuously influenced by host genetics.

MR can theoretically be applied to examine the effects of periodontitis on other traits. In practice this is extremely difficult given the limited number of consensus genetic risk loci which have been identified for periodontitis and the small amount of variation in periodontitis liability which these explain. The scope is expected to increase as understanding of the genetic basis of periodontitis improves and valid genetic instruments for periodontitis are developed.

MR is most useful where there is an existing clear hypothesis and where MR estimates can complement evidence from other sources. MR can currently help test some, but not all, hypotheses in periodontal research and should be deployed only in suitable situations after careful consideration of the research question. Like any other method, MR is capable of producing biased and invalid results, meaning each application should be carefully examined.

Background

Oral and dental diseases are persistent global health problems with substantial consequences on those affected, their communities, and health systems (Peres et al., 2019). As part of efforts to improve oral health, considerable efforts have been invested in identifying proximal and distal causes of oral diseases, including risk factors that are shared with systemic diseases over the last decades (Jin et al., 2016). It is important to acknowledge that the major non-communicable oral diseases, namely periodontitis and dental caries, are likely to be causally and undeniably non-causally associated with a wide array of modifiable and non-modifiable risk factors (Darveau, 2010; Heaton & Dietrich, 2012). Indeed, depending on context and intended use, it is not always necessary to distinguish between associations with causal versus non-causal factors. For example, disease prediction approaches (Pylypchuk et al., 2018) can leverage both non-modifiable risk factors and non-causal associations. In other contexts,

however, it becomes important to distinguish between different origins of association. Specifically, causal associations with modifiable risk factors are potentially important for clinical practice and public health because they may inform possible interventions to prevent disease or reduce its burden. Unfortunately, it is often not possible to distinguish causal from non-causal associations in classical observational studies since there are other potential ways for associations to occur. These are summarized in **Table 1** and include chance, reverse-causal association, confounding and multiple forms of bias. Therefore, evidence of association emanating from classical observational epidemiology must be triangulated (Lawlor, Tilling, & Davey Smith, 2016) with data from other study designs to help strengthen its inferential potential.

Well undertaken randomized controlled trials offer the highest level of internal validity (and often available evidence) and help strengthen evidence for causal association. Depending on the exposure in question however (e.g., smoking as a causal risk factor for periodontitis), it is not always ethical or practically possible to carry out a trial. Longitudinal observational studies can be less prone to detecting reverse-causal associations, but are otherwise susceptible to similar problems to cross-sectional observational studies. In such cases, an alternative way to strengthen inference is to employ variables that correlate with the exposure and no other variables which could confound the observed associations, i.e., instrumental variables (Davies, Smith, Windmeijer, & Martin, 2013; Iwashyna & Kennedy, 2013; Lousdal, 2018) and then examine the association between this proxy exposure and the outcome.

Instrumental variables are arguably hypothetical because it is virtually impossible to demonstrate their independence from unknown or unmeasured confounders. In practice, however, variables which approximate the characteristics of instrumental variables may still be useful to help strengthen inference. Human genetic data provide one potential source of such variables. Allocation of alleles at conception is essentially random within a family, meaning alleles are not expected to associate with environmental confounders of the exposure-outcome association. Moreover, germline genotype does not vary with time, with the implication that reverse causation is less of a problem when dealing with genetic exposure variables.

Recognizing the potential utility of genetic data, a form of instrumental variable analysis using genetic data was proposed in 1986 (Katan, 1986) and was subsequently described as 'Mendelian randomization' (MR) in 1991 (Gray & Wheatley, 1991). However, it was in the early 21st century, when genetic data from large-scale studies started to become available, that the approach gained widespread attention (Smith & Ebrahim, 2003). The key feature of MR is that genetic variables are used to proxy the effects of a putative exposure when examining the effects of an exposure on an outcome. MR, therefore, is a form of causal inference analysis which uses genetic data as a source of instrumental variables. The name refers to the pseudo-random allocation of alleles at meiosis under Mendelian inheritance, which can be considered analogous to randomization in a clinical trial. The three main assumptions underlying all MR experiments as detailed in **Figure 1** include relevance (the genetic proxy is strongly associated with the exposure), independence (the genetic proxy must be independent of confounders), and exclusion restriction (the genetic proxy for the exposure cannot have any effect on the outcome except through its effect on the exposure). Provided these assumptions are met, several of the explanations given in **Table 1** can be ruled out, meaning that associations obtained using MR methods are generally interpreted as causal associations.

Adoption and evolution of study design

Early MR experiments were carried out in studies where genotype, exposure and outcome data were available for the same participants, now termed one-sample MR. These studies were carried out in an era where genetic studies were smaller and simpler in terms of the genetic association signals they uncovered. Consequently, the genetic proxies included in these studies tended to be more simple in underpinning. For example, the examination of C-reactive protein's association with metabolic syndrome (Timpson et al., 2005) where genetic variation in the regulatory regions of the protein encoding gene can lead to isomorphic changes in circulating protein level independent of other factors. The methods used in one-sample MR remain largely unchanged and it remains the preferred design if it can be undertaken. Its main advantages are that it allows some of the MR assumptions (i.e. instrument properties of independence from confounders) to be tested directly and it enables comparisons between the derived causal effect estimates and those obtained from conventional observational studies.

Over time, the nature of these experiments has changed and while many changes have been positive, the greater complexity of modern experiments has also created new challenges for interpretation. A major development was the emergence of two-sample methods (Pierce & Burgess, 2013) where estimates of genotype-exposure and genotype-outcome association are obtained from different studies. This greatly expanded the scope of possible experiments (Hemani et al., 2018) but makes it harder to test assumptions about confounding. Moreover, these approaches also introduce additional assumptions such as that the two studies are drawn from *the same underlying population*, with *no participant overlap.*

During the last 2 decades, there has been an increase in the power and availability of genome-wide association studies (GWAS) leading to a far greater arena of potential genetic variables to include in an MR application. This is a generally positive evolvement, allowing for a greater range of possible experiments, but also has the potential to complicate inference. The great power of modern GWAS allows detection of variants which have weak effects on a trait. The inclusion of weakly-associated variants can mean that subsequent analyses suffer from weak instrument bias (Burgess, Thompson, & Collaboration, 2011). One possible approach to reducing weak instrument bias involves using polygenic scores to aggregate the effects of multiple weak instruments into a single strong instrument, but these scores are themselves difficult to interpret (Tan & Timpson, 2022). Further, variants can become associated with a trait in a GWAS through a number of different pathways including some that involve no direct effect on the trait (**Figure 2**) (Holmes & Davey Smith, 2019). When so called 'non-primary' genetic associations lead to a variant being selected as a proxy, reverse-causal effects may become visible as an apparently forward-causal effect in MR (Burgess, Swanson, & Labrecque, 2021). This is a particular concern when there is imprecise measurement (Hemani, Tilling, & Davey Smith, 2017) (**Figure 2)**.

In addition to these complications, there has been a general move away from clearly defined clinical questions mirroring those in a randomized controlled trial towards more broad and exploratory applications. This is being challenged by authors who advocate for clear clinical questions (Gagliano Taliun & Evans, 2021). Overall, the changing approaches to study design have greatly increased the scope and potential of MR experiments, but at the cost of greater complexity and less clear interpretation.

Challenging the assumptions of MR experiments

Since MR was first adopted, there have been marked improvements in the understanding of the properties of human genetic data which challenges the assumptions of the methods. For example, while alleles may be randomly allocated *within a family*, they can be non-randomly distributed within a population (Haworth et al., 2019; Koellinger & de Vlaming, 2019) due to various possible reasons including ancestry, assortative mating and gene-environment correlation, explaining why *within-family* Mendelian randomization produces different estimates from *within-population* Mendelian randomization (Brumpton et al., 2020). These complicating factors were always present even in early studies, but only became visible with the greater power of modern GWAS.

Genetic variants can have effects on multiple traits and outcomes, either through downstream effects on a causal node (termed vertical pleiotropy, and the target of MR experiments) or through acting on multiple causal pathways (termed horizontal pleiotropy). Since genotypes are inherited from haplotypes containing multiple polymorphisms, other variants within a same haplotype can also create apparent horizontal pleiotropy. Horizontal pleiotropy is now recognized to be widespread (Gratten & Visscher, 2016) and some traits (e.g., human height) are associated with a substantial portion of the entire genome (Yengo et al., 2022). The existence of horizontal pleiotropy does not necessarily create bias in MR experiments; however, if the pleiotropic pathways have unbalanced effects on the outcome, then this can distort the MR estimate or mimic a causal effect of the exposure on outcome (**Figure 3**). Unfortunately, there is evidence that unaddressed horizontal pleiotropy distorts many reportedly casual associations seen in MR (Verbanck, Chen, Neale, & Do, 2018).

The assumption in two-sample study MR studies that both samples are drawn from the same underlying population is also potentially problematic and violations of this assumption produce biased causal estimates. The causal effect is inflated in situations where the genotype-exposure association is stronger in the population with the measured genotype-outcome association and deflated in the reciprocal situation. Because the genotype-exposure association is only known in one population, it must be assumed to be the same in both. However, even the same trait can have apparently different genetic association signals when examined in different groups of people (Ioannidis, Patsopoulos, & Evangelou, 2007) due to a combination of possible explanations including, among others, demographics and study protocols. Genetic effects can vary in a trait-specific manner even between populations of broadly similar ancestry (Huang et al., 2022) suggesting that great care is needed when assuming identical genetic effects across two different samples. This is not a problem in one-sample designs.

Finally, induced complexities in the properties of genetic associations being used to instrument exposures of interest, are being recognized and reported. As an example, collider bias has received little attention as a cause of incorrect inference in MR, however there is now growing awareness that bias introduced during study recruitment (e.g., healthy volunteer bias) is carried over to all downstream analyses including GWAS (Y. V. Sun et al., 2023) and MR (Gkatzionis & Burgess, 2019; Schoeler et al., 2023). Misclassification bias in phenotype assignment can also produce bias in MR estimates (Clayton et al., 2023). Taken together, there are multiple explanations for apparently causal associations in MR highlighting the need for careful design and interpretation of these studies.

Methodological developments

There have been a wide range of methods proposed to detect and account for violations of the basic MR assumptions, sometimes termed 'robust' estimators. We suggest that the use of one or more 'robust' approaches should be an essential part of any sensitivity analysis (Burgess, Bowden, Fall, Ingelsson, & Thompson, 2017). These methods continue to be updated regularly and some of the more common methods are summarized in recent reviews by de Leeuw et al. (de Leeuw, Savage, Bucur, Heskes, & Posthuma, 2022) and Burgess et al. (Burgess et al., 2019). Different methods can be employed to detect or account for variants with complicating characteristics, such as identifying genetic variants whose effects on the outcome are disproportionate to their effects on the exposure, which could be indicative of unbalanced horizontal pleiotropy (Bowden, Davey Smith, & Burgess, 2015). It should be noted, however, that these methods typically introduce one or more new assumptions which may themselves be difficult to test and potentially invalid (Burgess & Thompson, 2017). Given this limitation, it has been suggested that a panel of potential approaches should be used, and their effects should be compared (Slob & Burgess, 2020). Comparison of effect estimates obtained from different methods can provide valuable information into the degree of violation of the underlying assumptions and help the reader to understand how far the experiment likely strayed from the ideal conditions.

We argue that comparison of naïve and robust MR estimates can be thought of as analogous to comparing minimally adjusted and fully adjusted estimates in a classical observational study. While fully adjusted estimate cannot be considered 'true', the degree of agreement or disagreement across these estimates helps guide a reader as to the likely impact of confounding in a classical observational study, or the potential impact of (detectable) assumption violations in the MR setting. Comparison of estimates derived from different analytical conditions is therefore suggested as a key part of overall interpretation. As a cautionary note, over-reliance on advanced and emerging statistical methods in MR can cause problems since these methods contain many implicit assumptions which are easily misunderstood, even by expert authors and reviewers. A recent example of this involves an attempt to test for non-linear effects of 25-hydroxyvitamin D on all-cause mortality. This MR was initially published and then retracted when serious methodological problems leading to incorrect inference were highlighted by a methodological expert writing in a different journal (Burgess, 2023), and some of the specific methodological challenges have since been discussed further by Wade and colleagues (Wade et al., 2023).

Rather than trying to correct for bias in the statistical analysis stage, it is reasonable to place more emphasis on carefully designing experiments that violate fewer assumptions to begin with. Within-family MR (Brumpton et al., 2020) inherently avoids some, but not all, sources of bias. The use of negative controls (Sanderson, Richardson, Hemani, & Davey Smith, 2021) remains a useful tool to detect problems with the MR context.

Interpretation of MR studies

Overall, there is now sufficient evidence that the MR assumptions are theoretical and are unlikely to fully satisfied in real-world applications except for specific settings such as within-family MR. Importantly, and as is the case for any applied theoretical method, this does not exclude use of the approach, rather calls for transparency in the presentation of methods, data and results, and careful interpretation. In the case of MR, studies must therefore rely on a judgment of how far the experiment has strayed from theoretical conditions, bearing in mind the possible sources of bias and their likely impact on results. In this sense, interpretation of MR studies is not dissimilar to interpretation of any

other epidemiological study, the important and valuable difference being that these complications are not the same as those in other analysis–hence enhancing the value in comparison of results and the persistence of an association signal.

The main factors to consider when interpreting a MR study are summarized in **Table 2**. Minor violations of the underlying assumptions can be partially accounted for with statistical methods and may not change the overall interpretation, analogous to weak confounding in an observational study. Poorly designed experiments, underpowered studies, and strong violations of the assumptions of the test used will result in uninterpretable effect estimates, just as in a classical observational study. A recent review article by Davey Smith and Ebrahim commented on the explosion of MR studies with highly implausible findings, claiming to find effects of exposures which cannot reasonably be proxied by genetic variation (Smith & Ebrahim, 2024). In our experience, endogenous biological variables (for example protein expression) often have satisfactory genetic proxies, while exogenous variables (for example lifestyle traits) are difficult to proxy with host genetic variation.

Applications in periodontal research

In principle, Mendelian randomization can be applied to test for postulated effects of a possible risk factor on periodontitis and to test for postulated effects of periodontitis on other traits. As of early 2024, there are now many examples in the published and grey literature of studies testing effects of possible risk factors on periodontitis. Illustrative examples include studies examining the possible effects of total adiposity, vitamin D, tobacco smoking, alcohol consumption, educational attainment, fasting glucose and depression (Baumeister, Freuer, et al., 2022; Baumeister et al., 2021; Baumeister, Reckelkamm, et al., 2022; Shungin et al., 2015; Wang et al., 2022; Zhang et al., 2023), with varying results. Other studies have used MR to explore whether therapeutic manipulation of target molecules might have effects on periodontitis (Alayash et al., 2024; Nolde et al., 2023). These studies (termed "drug target MR") differ from others only in the choice of potentially druggable exposures as the exposure variables, and are otherwise similar to other MR studies. There have been few one-sample applications, with the majority using two-sample designs. A general limitation of these studies is the underlying genomics evidence base of periodontitis—there are few available studies with genetic data and clearly defined clinical endpoints of periodontitis. Many studies use summary statistic data from the Gene-Lifestyle Interactions in Dental Endpoints (GLIDE) consortium which remains a useful resource, but one with natural limitations in statistical power and phenotypic definition (i.e., adoption of 'any periodontitis' as the GWAS analytical endpoint) as discussed by the authors (Shungin et al., 2019).

The lack of well-powered genome-wide association studies for periodontitis is an even greater problem when attempting to test for downstream effects of periodontitis, since this can only be accomplished if there are valid genetic proxies for periodontitis. However, very few valid association signals have emerged from the literature to date. MR authors have attempted to address this by defining genotype-exposure association at a weak threshold (Ma et al., 2023; Y. Q. Sun, Richmond, Chen, & Mai, 2020) or by compiling lists of variants which are reportedly associated with periodontitis in individual studies, even if these were not associated in larger studies or meta-analyses (Bell, Gibson, Harshfield, & Markus, 2020; Corlin et al., 2021; Czesnikiewicz-Guzik et al., 2019). All these approaches are potentially problematic due to the issue of weak instrument bias and the need for consistency in genotype-exposure estimates discussed above. Moreover, this issue highlights the need for concerted efforts to

better understand the genetic basis of periodontitis using high-fidelity clinical disease endpoints and large sample sizes.

Aside from the practical problem arising from lack of genetic proxies for periodontitis, there are non-trivial problems with interpreting MR that aim to test for causal effects of a binary variable such as periodontitis. These issues are discussed in detail by Burgess and Labrecque (Burgess & Labrecque, 2018), but the practical implication is that MR using periodontitis as an exposure should be considered in terms of an underlying continuous liability scale to periodontitis—a distinction not often considered in the literature. Periodontitis-associated genotypes are assigned at conception but are not penetrant until some point later in life, meaning any interpretation of the effects of those genotypes on other outcomes needs to consider age, especially in the context of two-sample designs. To date, there is little consensus about what possible clinical or population-level intervention is proxied by these studies.

Importantly, alternatives to MR exist and an overall attitude of combination in the face of all analysis specific limitations should be embraced. While MR is often regarded as a distinct analytical paradigm, it can also be thought of as a test for genetic correlation between two traits, under the special case that the correlation is induced only by vertical pleiotropy (Figure 3). As such, it exists on a continuum of possible approaches which exploit genetic correlation for casual inference (Walker, Zheng, Gaunt, & Smith, 2022). One exciting and relatively recent approach is the latent causal variable method (O'Connor & Price, 2018) which allows causal inference when the genetic correlation arises from a combination of both horizontal and vertical pleiotropy. To date this has not been widely applied to periodontitis (Haworth et al., 2021).

Summary

MR is an extremely useful and flexible tool whose value and interpretation depends entirely on the study design and application. We argue the basic MR assumptions are best regarded as theoretical and are difficult to satisfy in real-word applications. Interpretation of MR studies therefore relies on a judgment of how far the experiment has strayed from the theoretical ideals. Minor violations of the underlying assumptions can be partially accounted for with statistical methods and may not change the overall interpretation. However, strong violations lead to uninterpretable and potentially misleading estimates. Steps should be taken to minimize bias during experimental design, rather than in the statistical phase—the same basic considerations regarding sampling frame and phenotypic assessment which are considered in classical observational studies should also be considered in MR.

The most useful experiments are those which target a clear clinical question using carefully selected genetic variables as proxies for the exposure, well-defined outcomes, appropriate populations and have adequate statistical power. To-date, few experiments in periodontitis have achieved this, partially due to the scarcity of consensus genetic evidence for periodontitis. Well-powered GWAS based of high-fidelity periodontitis measures are needed to support the development of robust genetic instruments and generally advance our understanding of the genomic basis of periodontal diseases.

Figures

Figure 1 Illustration of the three classical MR assumptions

a) the genetic proxy for the exposure (G) must strongly associate with the exposure ('relevance')

b) the genetic proxy for the exposure must be independent of confounders of the exposure-outcome association ('independence')

c) the genetic proxy for the exposure cannot have any effect on the outcome except through its effect on the exposure ('exclusion restriction')

Figure 2 Impact of phenotypic measurement.

The biological processes leading to periodontitis are not measured directly in large genetic association studies but must be inferred from clinical endpoints or other data (a). Even with perfect measurement, a genetic proxy for periodontitis may in fact proxy a heritable risk factor on the causal pathway leading to periodontitis (b). In this specific example, it might lead to harmful effects of smoking being mis-attributed to periodontitis and violation of the independence assumption; in other examples it can lead to incorrect inference of causal direction. With imperfect measurement the phenotype may also capture heritable factors which influence measurement without being on the causal pathway to periodontitis, e.g heritable factors may influence self-reports (c).

Figure 3 Illustration of different potential configurations of pleiotropy

In an ideal scenario, the genotype only affects the exposure variable without pleiotropic effects (a). If the genotype affects more than one pathway (termed horizontal pleiotropy) that does not automatically violate the MR assumptions, since those pathways may have no effect on the outcome (b) or may have effects on the outcome which are balanced (c). Unbalanced horizontal pleiotropy (d) violates the exclusion restriction assumption but may be detectable using robust estimators. Vertical pleiotropy (e) does not violate the MR assumptions and is indeed the target of MR experiments. In practice, the pleiotropic configuration is unknown and may include a combination of both horizontal and vertical pathways.

Disclosure statement

Dr. Haworth has nothing to disclose

Dr. Timpson has nothing to disclose

Dr. Divaris has nothing to disclose

Funding statement

No specific funding was obtained for this work. SH is supported by Wellcome through grant 227534/Z/23/Z

References

Alayash, Z., Baumeister, S. E., Holtfreter, B., Kocher, T., Baurecht, H., Ehmke, B., . . . Reckelkamm, S. L. (2024). Complement C3 as a potential drug target in periodontitis: Evidence from the cis-Mendelian randomization approach. *J Clin Periodontol, 51*(2), 127-134. doi:10.1111/jcpe.13894

Baumeister, S. E., Freuer, D., Baurecht, H., Reckelkamm, S. L., Ehmke, B., Holtfreter, B., & Nolde, M. (2022). Understanding the consequences of educational inequalities on periodontitis: A Mendelian randomization study. *J Clin Periodontol, 49*(3), 200-209. doi:10.1111/jcpe.13581

Baumeister, S. E., Freuer, D., Nolde, M., Kocher, T., Baurecht, H., Khazaei, Y., . . . Holtfreter, B. (2021). Testing the association between tobacco smoking, alcohol consumption, and risk of periodontitis: A Mendelian randomization study. *J Clin Periodontol, 48*(11), 1414-1420. doi:10.1111/jcpe.13544

Baumeister, S. E., Reckelkamm, S. L., Baurecht, H., Nolde, M., Kocher, T., Holtfreter, B., . . . Hannemann, A. (2022). A Mendelian randomization study on the effect of 25-hydroxyvitamin D levels on periodontitis. *J Periodontol, 93*(8), 1243-1249. doi:10.1002/JPER.21-0463

Bell, S., Gibson, J. T., Harshfield, E. L., & Markus, H. S. (2020). Is periodontitis a risk factor for ischaemic stroke, coronary artery disease and subclinical atherosclerosis? A Mendelian randomization study. *Atherosclerosis, 313*, 111-117. doi:10.1016/j.atherosclerosis.2020.09.029

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol, 44*(2), 512-525. doi:10.1093/ije/dyv080

Brumpton, B., Sanderson, E., Heilbron, K., Hartwig, F. P., Harrison, S., Vie, G. A., . . . Davies, N. M. (2020). Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat Commun, 11*(1), 3519. doi:10.1038/s41467-020-17117-4

Burgess, S. (2023). Violation of the Constant Genetic Effect Assumption Can Result in Biased Estimates for Non-Linear Mendelian Randomization. *Hum Hered, 88*(1), 79-90. doi:10.1159/000531659

Burgess, S., Bowden, J., Fall, T., Ingelsson, E., & Thompson, S. G. (2017). Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology, 28*(1), 30-42. doi:10.1097/EDE.0000000000000559

Burgess, S., Davey Smith, G., Davies, N. M., Dudbridge, F., Gill, D., Glymour, M. M., . . . Theodoratou, E. (2019). Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res, 4*, 186. doi:10.12688/wellcomeopenres.15555.3

Burgess, S., & Labrecque, J. A. (2018). Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol, 33*(10), 947-952. doi:10.1007/s10654-018-0424-6

Burgess, S., Swanson, S. A., & Labrecque, J. A. (2021). Are Mendelian randomization investigations immune from bias due to reverse causation? *Eur J Epidemiol, 36*(3), 253-257. doi:10.1007/s10654-021-00726-8

Burgess, S., & Thompson, S. G. (2017). Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol, 32*(5), 377-389. doi:10.1007/s10654-017-0255-x

Burgess, S., Thompson, S. G., & Collaboration, C. C. G. (2011). Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol, 40*(3), 755-764. doi:10.1093/ije/dyr036

Clayton, G. L., Goncalves, A., Soares, Goulding, N., Borges, M. C., Holmes, M. V., . . . Carter, A. R. (2023). A framework for assessing selection and misclassification bias in mendelian randomisation studies: an illustrative example between body mass index and covid-19. *BMJ, 381*, e072148. doi:10.1136/bmj-2022-072148

Corlin, L., Ruan, M., Tsilidis, K. K., Bouras, E., Yu, Y. H., Stolzenberg-Solomon, R., . . . Michaud, D. S. (2021). Two-Sample Mendelian Randomization Analysis of Associations Between Periodontal Disease and Risk of Cancer. *JNCI Cancer Spectr, 5*(3). doi:10.1093/jncics/pkab037

Czesnikiewicz-Guzik, M., Osmenda, G., Siedlinski, M., Nosalski, R., Pelka, P., Nowakowski, D., . . . Guzik, T. J. (2019). Causal association between periodontitis and hypertension: evidence from Mendelian randomization and a randomized controlled trial of non-surgical periodontal therapy. *Eur Heart J, 40*(42), 3459-3470. doi:10.1093/eurheartj/ehz646

Darveau, R. P. (2010). Periodontitis: a polymicrobial disruption of host homeostasis. *Nat Rev Microbiol, 8*(7), 481-490. doi:10.1038/nrmicro2337

Davies, N. M., Smith, G. D., Windmeijer, F., & Martin, R. M. (2013). Issues in the reporting and conduct of instrumental variable studies: a systematic review. *Epidemiology, 24*(3), 363-369. doi:10.1097/EDE.0b013e31828abafb

de Leeuw, C., Savage, J., Bucur, I. G., Heskes, T., & Posthuma, D. (2022). Understanding the assumptions underlying Mendelian randomization. *Eur J Hum Genet, 30*(6), 653-660. doi:10.1038/s41431-022-01038-5

Gagliano Taliun, S. A., & Evans, D. M. (2021). Ten simple rules for conducting a mendelian randomization study. *PLoS Comput Biol, 17*(8), e1009238. doi:10.1371/journal.pcbi.1009238

Gkatzionis, A., & Burgess, S. (2019). Contextualizing selection bias in Mendelian randomization: how bad is it likely to be? *Int J Epidemiol, 48*(3), 691-701. doi:10.1093/ije/dyy202

Gratten, J., & Visscher, P. M. (2016). Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Med, 8*(1), 78. doi:10.1186/s13073-016-0332-x

Gray, R., & Wheatley, K. (1991). How to avoid bias when comparing bone marrow transplantation with chemotherapy. *Bone Marrow Transplant, 7 Suppl 3*, 9-12.

Haworth, S., Kho, P. F., Holgerson, P. L., Hwang, L. D., Timpson, N. J., Renteria, M. E., . . . Cuellar-Partida, G. (2021). Assessment and visualization of phenome-wide causal relationships using genetic data: an application to dental caries and periodontitis. *Eur J Hum Genet, 29*(2), 300-308. doi:10.1038/s41431-020-00734-4

Haworth, S., Mitchell, R., Corbin, L., Wade, K. H., Dudding, T., Budu-Aggrey, A., . . . N, J. T. (2019). Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nat Commun, 10*(1), 333. doi:10.1038/s41467-018-08219-1

Heaton, B., & Dietrich, T. (2012). Causal theory and the etiology of periodontal diseases. *Periodontol 2000, 58*(1), 26-36. doi:10.1111/j.1600-0757.2011.00414.x

Hemani, G., Tilling, K., & Davey Smith, G. (2017). Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet, 13*(11), e1007081. doi:10.1371/journal.pgen.1007081

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., . . . Haycock, P. C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *eLife, 7*. doi:10.7554/eLife.34408

Holmes, M. V., & Davey Smith, G. (2019). Can Mendelian Randomization Shift into Reverse Gear? *Clin Chem, 65*(3), 363-366. doi:10.1373/clinchem.2018.296806

Huang, Q. Q., Sallah, N., Dunca, D., Trivedi, B., Hunt, K. A., Hodgson, S., . . . Kuchenbaecker, K. (2022). Transferability of genetic loci and polygenic scores for cardiometabolic traits in British Pakistani and Bangladeshi individuals. *Nat Commun, 13*(1), 4664. doi:10.1038/s41467-022-32095-5

Ioannidis, J. P., Patsopoulos, N. A., & Evangelou, E. (2007). Heterogeneity in meta-analyses of genome-wide association investigations. *Plos One, 2*(9), e841. doi:10.1371/journal.pone.0000841

Iwashyna, T. J., & Kennedy, E. H. (2013). Instrumental variable analyses. Exploiting natural randomness to understand causal mechanisms. *Ann Am Thorac Soc, 10*(3), 255-260. doi:10.1513/AnnalsATS.201303-054FR

Jin, L. J., Lamster, I. B., Greenspan, J. S., Pitts, N. B., Scully, C., & Warnakulasuriya, S. (2016). Global burden of oral diseases: emerging concepts, management and interplay with systemic health. *Oral Dis, 22*(7), 609-619. doi:10.1111/odi.12428

Katan, M. B. (1986). Apolipoprotein E isoforms, serum cholesterol, and cancer. *Lancet, 1*(8479), 507-508. doi:10.1016/s0140-6736(86)92972-7

Koellinger, P. D., & de Vlaming, R. (2019). Mendelian randomization: the challenge of unobserved environmental confounds. *Int J Epidemiol, 48*(3), 665-671. doi:10.1093/ije/dyz138

Lawlor, D. A., Tilling, K., & Davey Smith, G. (2016). Triangulation in aetiological epidemiology. *International journal of epidemiology, 45*(6), 1866-1886. doi:10.1093/ije/dyw314

Lousdal, M. L. (2018). An introduction to instrumental variable assumptions, validation and estimation. *Emerg Themes Epidemiol, 15*, 1. doi:10.1186/s12982-018-0069-7

Ma, C., Wu, M., Gao, J., Liu, C., Xie, Y., Lv, Q., & Zhang, X. (2023). Periodontitis and stroke: A Mendelian randomization study. *Brain Behav, 13*(2), e2888. doi:10.1002/brb3.2888

Nolde, M., Alayash, Z., Reckelkamm, S. L., Kocher, T., Ehmke, B., Holtfreter, B., . . . Baumeister, S. E. (2023). Downregulation of interleukin 6 signaling might reduce the risk of periodontitis: a drug target Mendelian randomization study. *Front Immunol, 14*, 1160148. doi:10.3389/fimmu.2023.1160148

O'Connor, L. J., & Price, A. L. (2018). Distinguishing genetic correlation from causation across 52 diseases and complex traits. *Nat Genet, 50*(12), 1728-1734. doi:10.1038/s41588-018-0255-0

Peres, M. A., Macpherson, L. M. D., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., . . . Watt, R. G. (2019). Oral diseases: a global public health challenge. *Lancet, 394*(10194), 249-260. doi:10.1016/S0140-6736(19)31146-8

Pierce, B. L., & Burgess, S. (2013). Efficient design for Mendelian randomization studies: subsample and 2-sample instrumental variable estimators. *Am J Epidemiol, 178*(7), 1177-1184. doi:10.1093/aje/kwt084

Pylypchuk, R., Wells, S., Kerr, A., Poppe, K., Riddell, T., Harwood, M., . . . Jackson, R. (2018). Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. *Lancet, 391*(10133), 1897-1907. doi:10.1016/S0140-6736(18)30664-0

Sanderson, E., Richardson, T. G., Hemani, G., & Davey Smith, G. (2021). The use of negative control outcomes in Mendelian randomization to detect potential population stratification. *Int J Epidemiol, 50*(4), 1350-1361. doi:10.1093/ije/dyaa288

Schoeler, T., Speed, D., Porcu, E., Pirastu, N., Pingault, J. B., & Kutalik, Z. (2023). Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nat Hum Behav, 7*(7), 1216-1227. doi:10.1038/s41562-023-01579-9

Shungin, D., Cornelis, M. C., Divaris, K., Holtfreter, B., Shaffer, J. R., Yu, Y. H., . . . Franks, P. W. (2015). Using genetics to test the causal relationship of total adiposity and periodontitis: Mendelian randomization analyses in the Gene-Lifestyle Interactions and Dental Endpoints (GLIDE) Consortium. *Int J Epidemiol, 44*(2), 638-650. doi:10.1093/ije/dyv075

Shungin, D., Haworth, S., Divaris, K., Agler, C. S., Kamatani, Y., Keun Lee, M., . . . Johansson, I. (2019). Genome-wide analysis of dental caries and periodontitis combining clinical and self-reported data. *Nat Commun, 10*(1), 2773. doi:10.1038/s41467-019-10630-1

Slob, E. A. W., & Burgess, S. (2020). A comparison of robust Mendelian randomization methods using summary data. *Genet Epidemiol, 44*(4), 313-329. doi:10.1002/gepi.22295

Smith, G. D., & Ebrahim, S. (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol, 32*(1), 1-22. doi:10.1093/ije/dyg070

Smith, G. D., & Ebrahim, S. (2024). Mendelian randomisation at 20 years: how can it avoid hubris, while achieving more? *Lancet Diabetes Endocrinol, 12*(1), 14-17. doi:10.1016/S2213-8587(23)00348-0

Sun, Y. Q., Richmond, R. C., Chen, Y., & Mai, X. M. (2020). Mixed evidence for the relationship between periodontitis and Alzheimer's disease: A bidirectional Mendelian randomization study. *Plos One, 15*(1), e0228206. doi:10.1371/journal.pone.0228206

Sun, Y. V., Liu, C., Hui, Q., Zhou, J. J., Gaziano, J. M., Wilson, P. W., . . . Phillips, L. S. (2023). Correction for Collider Bias in the Genome-wide Association Study of Diabetes-Related Heart Failure due to Bidirectional Relationship between Heart Failure and Type 2 Diabetes. *medRxiv*. doi:10.1101/2023.09.22.23295915

Tan, V. Y., & Timpson, N. J. (2022). The UK Biobank: A Shining Example of Genome-Wide Association Study Science with the Power to Detect the Murky Complications of Real-World Epidemiology. *Annu Rev Genomics Hum Genet, 23*, 569-589. doi:10.1146/annurev-genom-121321-093606

Timpson, N. J., Lawlor, D. A., Harbord, R. M., Gaunt, T. R., Day, I. N., Palmer, L. J., . . . Davey Smith, G. (2005). C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *Lancet, 366*(9501), 1954-1959. doi:10.1016/S0140-6736(05)67786-0

Verbanck, M., Chen, C. Y., Neale, B., & Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet, 50*(5), 693-698. doi:10.1038/s41588-018-0099-7

Wade, K. H., Hamilton, F. W., Carslake, D., Sattar, N., Davey Smith, G., & Timpson, N. J. (2023). Challenges in undertaking nonlinear Mendelian randomization. *Obesity (Silver Spring), 31*(12), 2887-2890. doi:10.1002/oby.23927

Walker, V. M., Zheng, J., Gaunt, T. R., & Smith, G. D. (2022). Phenotypic Causal Inference Using Genome-Wide Association Study Data: Mendelian Randomization and Beyond. *Annu Rev Biomed Data Sci, 5*, 1-17. doi:10.1146/annurev-biodatasci-122120-024910

Wang, Y., Chu, T., Gong, Y., Li, S., Wu, L., Jin, L., . . . Deng, H. (2022). Mendelian randomization supports the causal role of fasting glucose on periodontitis. *Front Endocrinol (Lausanne), 13*, 860274. doi:10.3389/fendo.2022.860274

Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., . . . Hirschhorn, J. N. (2022). A saturated map of common genetic variants associated with human height. *Nature, 610*(7933), 704-712. doi:10.1038/s41586-022-05275-y

Zhang, X., Jiang, H., Zhang, L., Li, C., Chen, C., Xing, M., . . . Ma, Y. (2023). Potential Causal Association between Depression and Oral Diseases: A Mendelian Randomization Study. *Genes (Basel), 14*(12). doi:10.3390/genes14122191