Periodontitis and non-alcoholic fatty liver disease: a Mendelian randomisation analysis

Feiyan Yu  
Stomatological Hospital, Shanxi Medical University

Dongchao Wu  
Stomatological Hospital, Shanxi Medical University

Chong Han  
Stomatological Hospital, Shanxi Medical University

Qianqian Wang  
Stomatological Hospital, Shanxi Medical University

Yang Yang  
Stomatological Hospital, Shanxi Medical University

Yi Li  
Stomatological Hospital, Shanxi Medical University

Rao Qin  
Stomatological Hospital, Shanxi Medical University

Linkun Xu  
Stomatological Hospital, Shanxi Medical University

Yue Chen  
Stomatological Hospital, Shanxi Medical University

Dongning He  (kqhdn001@outlook.com)  
Stomatological Hospital, Shanxi Medical University

Research Article

Keywords: periodontitis, non-alcoholic fatty liver disease, Mendelian randomisation, causal relationship

Posted Date: February 9th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2556093/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Previous observational and epidemiological studies have reported the association between periodontitis (PD) and non-alcoholic fatty liver disease (NAFLD). However, evidence from long-term randomized controlled trials (RCTs) is lacking. Therefore, this study aimed to explore the causal relationship between PD and NAFLD.

Methods

Genetic information for individuals of European ancestry with PD (17,353 clinically diagnosed cases and 28,210 controls) and NAFLD (8,434 clinically diagnosed cases and 770,180 controls) were obtained using published genome-wide association study statistics, following which we conducted two-sample bivariate Mendelian randomisation (MR) analyses. Various techniques such as inverse-variance weighted (IVW), weighted median, and MR-Egger regression methods were used to calculate the causal relationship between exposure and the result.

Results

No causal effect of genetically determined PD on NAFLD existed (odds ratio [OR] = 0.99, 95% confidence interval [CI]: 0.90–1.10, P = 0.95). Furthermore, no causal effect of NAFLD on PD was observed in the reverse MR analysis (OR = 1.02, 95% CI: 0.92–1.13, P = 0.63). No heterogeneity was observed between individual single nucleotide polymorphisms pursuant to the heterogeneity assessment (Q > 0.1). Horizontal pleiotropy pursuant to the outcomes of MR-Egger regression and MR Pleiotropy RESidual Sum and Outlier overall tests (P > 0.1) were less likely to distort the causal relationship between PD and NAFLD.

Conclusion

Collectively, we did not find substantial evidence to support a causal association between PD and NAFLD in this bidirectional MR study.

Clinical relevance:

Periodontitis does not seem to be a risk factor for worsening of non-alcoholic fatty liver disease.

Introduction
Periodontitis (PD) is an oral inflammatory disease with a global prevalence of 11% prevalence, and is characterised by gingival redness and bleeding, periodontal pocket formation, tooth loosening, and alveolar bone resorption(1). Local, systemic, environmental, and genetic factors influence the complicated pathogenic mechanism of PD(2). Plaque is generally considered the initiating factor in PD, with genetic variation also playing a crucial role in the disease onset and progression, which might affect subgingival pathogenic bacteria removal or persistence, thereby increasing a person's susceptibility to PD and ultimately resulting in tooth loosening and loss(3, 4). Additionally, PD can result in the large-scale production of pro-inflammatory mediators, which through systemic circulation, allows the conversion of local inflammation to systemic inflammation, thereby impacting other organs(5). Reportedly, PD is a potential risk factor for non-alcoholic fatty liver disease (NAFLD)(6).

NAFLD is a diverse illness characterised by inflammation, intestinal dysbiosis, lipotoxicity, hepatic lipid build-up, and insulin resistance(7, 8). NAFLD affects approximately 25% of individuals worldwide(9). Intricate interactions between genetic variation, environmental factors, metabolic and demographic factors, and gut microbiota influence the pathogenesis of NAFLD(10). An individual's genotype has been linked to the typical expression of NAFLD, and genetic variation affects susceptibility to NAFLD, the progression to fibrosis, and hepatic steatosis(11). Additionally, research has demonstrated that NAFLD is a complicated multisystem illness affecting extrahepatic organs and regulatory pathways, along with the local pathology associated with the liver(12). Reportedly, PD might be an extrahepatic complication of NAFLD(13).

Previous epidemiological studies have revealed parallels between PD and NAFLD. A case-control study showed that patients with NAFLD had considerably more PD-causing bacterial detections than non-NAFLD controls(14). According to a previous systematic analysis, including 12 trials and 53,384 participants, PD might serve as a risk factor for the onset and progression of NAFLD(15). However, not all studies have arrived at the same conclusion, and a significant cross-sectional investigation comprising 11,914 individuals from various Hispanic groups failed to establish a relationship between the clinical characteristics of PD and NAFLD(16). Overall, the findings of observational studies are contradictory, making it challenging to determine whether a causal relationship between PD and NAFLD exists, as these studies have several intrinsic flaws, including reverse causality, confounding variables, and measurement error. Typically, randomised controlled trials (RCTs) could help identify a causal relationship between the two diseases and provide the highest level of evidence for clinical studies. However, no human RCTs on the causal relationship between PD and NAFLD exist due to limitations such as ethical review issues, the prolonged time required, and insufficient funding. Mendelian randomisation (MR) methods based on pooled genome-wide association study (GWAS) data provide us with a new option to assess the causality of the hypothesised exposure-outcome pathway, to better understand the causal relationship between PD and NAFLD and ultimately to develop effective treatment plans for patients.

Mendel's second law states that when parents with two (or more) pairs of relative traits are crossed, genes on non-homologous chromosomes behave as free combinations when the offspring generation produces gametes while alleles segregate. This process is similar to the random grouping of RCTs. The
GWAS data required for current MR methods are relatively easy to obtain compared with RCTs, thereby saving time, effort, research funding, etc. In addition, single nucleotide polymorphisms (SNPs) have been used as instrumental variables (IVs) in MR research to counteract the drawbacks of previous observational studies because genotypes are usually unaffected by disease(17). MR thus overcomes the limitations of observational studies and RCTs and is considered the 'gold standard' for providing medical evidence. Herein, we investigated the causal relationship between PD and NAFLD in a bidirectional double-sample MR investigation using publicly accessible summary statistics from the GWAS.

**Methods**

**Data Sources:**

We use openly available research or accessible datasets for MR studies. Consent paperwork or ethical declarations was not required. Based on the most recent meta-analysis of GWAS gene-lifestyle interactions by the Association for Gene Therapy Endpoints in Dentistry, we obtained summary statistics for PD, whose maximum sample volume to date comprised 17,353 cases with a clinical diagnosis and 28,210 controls(18). The American Academy of Periodontology case definition/Centre for Disease and Control Prevention was used as the basis for clinically diagnosing the PD cases included in this investigation(19). Additionally, based on the findings of a meta-analysis of the GWAS of electronic health records, including a sample size of 17,353 clinically diagnosed cases and 28,210 controls at the time of the analysis, we could acquire information on genetic variants associated with NAFLD(20). All SNPs and their associated pooled data were only accessible from studies on populations with European ancestry to prevent population stratification bias.

**Selection of the Genetic Instruments:**

Our bidirectional two-sample MR study was conducted in a framework shown in Fig. 1. MR must adhere to three presumptions to produce objective results. The three presumptions are as follows: (1) genetic variation is substantially associated with the exposure; (2) genetic variation is independent of certain variables related to the exposure and outcome; (3) genetic variation is allowed to be associated with the outcome solely through exposure (lack of horizontal polymorphism)(21–23).

Many quality control techniques were performed to screen for genetic IVs that complied with the three main assumptions of MR. We only examined genetic instruments with P-values < 1×10^{-5} and those from GWAS that were linked to exposure of interest. In addition, linkage disequilibrium (LD) was assessed between SNPs for exposure of interest, and SNPs with \( R^2 > 0.001 \) or clump distances < 10,000 kb were eliminated based on 1,000 genomic data from European individuals(24). SNPs with the minimum P-value were included when several SNPs were identified at the single locus.

Additionally, we retrieved potential confounding factors that might influence PD and NAFLD by searching the PhenoScanner website. Finally, confounding variables associated with drinking, smoking, and body mass index (BMI) were eliminated(25). Therefore, we needed to search each SNP in PhenoScanner to
check for polymorphisms and eliminate SNPs associated with potential outcome variables or confounders to satisfy the above hypotheses. In addition, the first hypothesis could be further tested by calculating the F-statistic for each SNP using the following equation: 

$$F = \frac{R^2 \times (N-2)}{1-R^2}.$$ 

$R^2$ represents the exposure variance explained by each exposure variance explained by the IV. IVs with an F-statistic < 10 are deemed as weaker instruments and are eliminated from MR analysis(26). We exploited these well-chosen SNPs to serve as the definitive genetic IVs, which were later used for MR analysis.

**Statistical Analysis:**

We used several complementary approaches in this study, comprising the IVW, the MR-Egger regression, and the weighted median methods to quantify the causal effects of exposure on results. The primary analytical method was the IVW approach(27). When all the chosen SNPs are effective IVs, the IVW technique, which is commonly used for basic causal estimates, yields the most accurate outcomes. The IVW uses summary data from each genetic variation, along with risk variables, beta coefficients, and standard errors, to regress the outcomes of each genetic variation individually to evaluate causality. The IVW method calculates a weighted average of the Wald estimates(22).

Data from several genetic instruments can be combined using the weighted median approach to perform consistency analysis by computing a single weighted median estimator. The MR-Egger approach allows each IV to demonstrate pleiotropy, and the procedure is consistent if the instrument strength is unrelated to these pleiotropic effects. The MR-Egger regression can produce estimates after pleiotropic effects have been eliminated and can test for horizontal pleiotropy via intercepts, although it reduces statistical power(28). However, the MR-Egger regression is less accurate and more susceptible to external genetic variation. Internal assumptions are not necessary for the weighted median regression approach, which computes the weighted median of the Wald ratio estimates and is resistant to horizontal pleiotropic bias(23).

Furthermore, to maintain the consistency of SNPs used as IVs in various analyses, we only employed variations of all analysed traits in this study’s analysis and did not replace missing variants with surrogates.

**Sensitivity Analysis:**

Several sensitivity analyses were performed to analyse whether the heterogeneity and pleiotropy of the genetic instruments evaluated would bias the MR results to rule out potential violations of the MR premise.

Horizontal pleiotropy occurs when a genetic variant is linked to more than one trait in various pathways and can contradict the results of MR analysis. We used MR-Egger regression, which allows us to identify confounding variables that could skew the MR results to evaluate and correct for horizontal pleiotropy. We evaluated the MR-Egger regression intercept and ran the MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) global test to determine the presence of pleiotropy(29). Furthermore, funnel plot asymmetry could also be considered a sign of horizontal pleiotropy(30).
The IVW method and MR-Egger regression were used to identify heterogeneity, and heterogeneities were quantified using Cochran’s Q statistic. A leave-one-out analysis was also performed to verify the reliability and consistency of the findings. The “two-sample MR” and “MR-PRESSO” in R (version 4.0.3) packages were used for all studies(30).

Results

Causal Effects of PD on NAFLD

Ninety-seven independent SNPs with significant P-values \(< 1 \times 10^{-5}\) were included as IV SNPs for PD. SNPs with LD were then eliminated. In addition, SNP rs2921075, which was associated with a known confounder BMI, was also eliminated. The mean F-statistic for PD was 21.34, and all the F-statistic values were >10. Finally, 11 significantly related SNPs were carefully selected. Information on IVs for PD is provided in Supplementary Table 1. MR estimations of the various techniques are presented in Table 1. Overall, no causal relationship was observed between the risk of NAFLD and PD predicted by genetics. The IVW method’s findings demonstrated that no statistically significant correlation existed between an increased risk of having PD and an increased risk of developing NAFLD (PD: NAFLD = 0.99, 95% CI: 0.90–1.10, \(P = 0.95\)). Reliable results were also obtained using the MR-Egger and weighted median approaches. Scatter and funnel plots of SNP impact sizes for PD and NAFLD are presented in Figs. 2a and 3a, respectively. No heterogeneity was observed between individual SNPs, as determined by the heterogeneity test and funnel plot, \(Q = 3.0 > 0.1\). The causative association between PD and NAFLD was less likely to be affected by horizontal pleiotropy, based on the MR-Egger regression and MR-PRESSO global test and scatter plot findings, wherein \(P > 0.1\) (Table 2). No single SNP was responsible for the causative estimates of PD, according to the leave-one-out analysis (Fig. 4a).

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Methods</th>
<th>N.SNP</th>
<th>Beta</th>
<th>SE</th>
<th>P.value</th>
<th>OR</th>
<th>95%CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>NAFLD</td>
<td>IVW</td>
<td>11</td>
<td>-0.0028</td>
<td>0.05</td>
<td>0.95</td>
<td>0.99</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MR-Egger</td>
<td>11</td>
<td>-0.0068</td>
<td>0.08</td>
<td>0.93</td>
<td>0.99</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WM</td>
<td>11</td>
<td>-0.017</td>
<td>0.06</td>
<td>0.79</td>
<td>0.98</td>
<td>0.86</td>
</tr>
<tr>
<td>NAFLD</td>
<td>PD</td>
<td>IVW</td>
<td>7</td>
<td>0.024</td>
<td>0.05</td>
<td>0.63</td>
<td>1.02</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MR-Egger</td>
<td>7</td>
<td>0.04</td>
<td>0.10</td>
<td>0.72</td>
<td>1.04</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WM</td>
<td>7</td>
<td>0.039</td>
<td>0.05</td>
<td>0.47</td>
<td>1.03</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Abbreviations: PD: periodontitis; NAFLD: Non-alcoholic fatty liver disease; CI: confidence interval; IVW: inverse variance weighting; N.SNPs: number of SNPs used in MR: OR: odds ratio; MR: Mendelian randomisation.
Table 2
The heterogeneity and horizontal pleiotropy of individual SNPs.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>MR-Egger</th>
<th>IVW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$Q$</td>
<td>$df$</td>
</tr>
<tr>
<td>PD</td>
<td>NAFLD</td>
<td>3.03</td>
<td>9</td>
</tr>
<tr>
<td>NAFLD</td>
<td>PD</td>
<td>6.47</td>
<td>5</td>
</tr>
</tbody>
</table>

MR-Egger test for horizontal pleiotropy

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Outcome</th>
<th>Intercept</th>
<th>SE</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>NAFLD</td>
<td>0.00077</td>
<td>0.013</td>
<td>0.95</td>
</tr>
<tr>
<td>NAFLD</td>
<td>PD</td>
<td>-0.0027</td>
<td>0.0155</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Abbreviations: $df$, degree of freedom; MR, Mendelian randomisation; $Q$, heterogeneity statistic $Q$.

Causal Effects of NAFLD on PD

A total of 407 independent SNPs were considered for NAFLD. SNPs with LD were then eliminated. In addition, SNP rs9922619 and rs429358, which were associated with a known confounder BMI, were also eliminated. The mean F-statistic for NAFLD was 49.97, and all the F-statistic values were > 10. Finally, seven relevant SNPs were carefully selected. Specifics of the NAFLD IVs are provided in Supplementary Table 2. No evidence of a causal link was observed between NAFLD and PD risk using the IVW technique (NAFLD: PD = 1.02, 95% CI: 0.92–1.13, $P = 0.63$) (Table 1). No discernible indication of horizontal pleiotropy or heterogeneity existed in the NAFLD IVs (Table 2, Figs. 2b and 3b). No single SNP was responsible for the causative estimates of NAFLD, according to the leave-one-out analysis (Fig. 4b).

Discussion

Our study is the first to use various complementary MR techniques for exploring the two-way causal relationship between NAFLD and PD. Our two-sample MR analysis found neither a causal relationship between inherited or predicted PD and NAFLD nor a causal relationship between inherited or predicted NAFLD and PD in individuals of European ancestry.

Previous epidemiological studies have reported a correlation between PD and NAFLD(31). A previous case-control study reported that NAFLD’s clinical or biochemical characteristics and PD are correlated(32). A similar conclusion was drawn by another cohort study, wherein patients with PD had higher rates of NAFLD, suggesting that PD might be a modifiable risk factor for NAFLD(33). Additionally, several case-control studies have reported that non-surgical periodontal therapy positively impacted liver function indices (alanine aminotransferase and aspartate aminotransferase) in patients with NAFLD(14,
Nevertheless, not all observational studies have drawn the same conclusions. A cross-sectional study of Hispanics/Latinos reported no correlation between the prevalence of NAFLD and the clinical parameters of PD(16). Furthermore, a cross-sectional study of Japanese adults reported no correlation between PD and the biochemical parameters of NAFLD(35). Although there have been meta-analyses on the association between PD and NAFLD, they have contradictory findings(15, 36, 37). These two meta-analyses suggested that PD might be a potential risk for NAFLD onset and progression(15, 37). Five studies comprising 27,703 participants were included in another meta-analysis, which revealed that PD was not a risk factor for NAFLD when various metabolic parameters were adjusted(36). It is noteworthy that a correlation between the two is different from having a causal relationship. Previous cross-sectional studies have been controversial because they measure exposure and outcomes after they exist or have occurred; therefore, their understanding of the association between exposure and outcome is erratic and, at times, leads to reverse causality. Cohort studies can screen for the temporal order of exposure and outcome; however, they cannot completely exclude the effects of residual confounding and could lead to unstable causal relationships. Most studies included in these meta-analyses were cross-sectional in design, with an overall inclusion of a small number of studies. Significant heterogeneity was observed among these studies. The incidence of NAFLD might have been overestimated or underestimated in the included studies owing to the different diagnostic methods used for PD and NAFLD, which could also raise confounding bias. Additionally, since most of the included studies were only performed in a few nations, it is possible that the findings cannot be applied to populations of other continents. In summary, current studies fail to establish reliable inferences of causality for the two diseases. MR, on the other hand, is based on the most fundamental Mendelian laws of inheritance, according to which parental alleles are randomly assigned to the offspring, and the genotype determines the phenotype (and disease is certainly a phenotype). As the genotype is innate, it is ahead of the time at which outcomes occur and is not influenced by various confounding factors, such as the environment in later life. Thus, it could serve as a powerful tool for studying the causal relationship between exposure and outcome.

The exact pathogenesis of the relationship between PD and NAFLD in humans is currently unknown. Several possible factors could explain the relationship between PD and NAFLD in observational studies. One possible biological explanation is that the low-grade inflammatory feature of PD is what causes systemic inflammation. The aetiology of obesity-related insulin resistance, a precursor to NAFLD, which in turn encourages NAFLD onset and progression, depends on low-grade (chronic) inflammation(38). Another biological explanation is the altered gut microbial composition brought on by periodontal infections, particularly Porphyromonas gingivalis(39). By altering the balance between pro- and anti-inflammatory molecules in the liver and hepatic glucose and lipid metabolism, gut microbiota might affect NAFLD(40). However, direct evidence from studies of human RCTs supporting these interpretations lacks evidence. Although PD and NAFLD are associated with systemic low-grade inflammation and alterations in their gut microbes, these reasons are insufficient to suggest a causal relationship between the two. These common pathways might impede causality judgments and bias the observational study results, which in turn might lead to bias in current clinical treatments for NAFLD patients with PD.
The current investigation of bidirectional MR has several advantages. First, an MR design reduces residual confounding and other biases, enhancing the ability to infer causality. Our bidirectional MR study guarantees the inference of causation between PD and NAFLD in both directions. Second, several sensitivity tests were performed. The consistent estimate of several models increased our confidence in the established relationships. Third, we applied the most recent GWAS data of PD and NAFLD in the population of European descent to minimise the effects of population stratification. This provided us with sufficient statistical power to evaluate the potential causal relationship between PD and NAFLD. Since a large sample size was employed for these two outcomes, there is sufficient statistical power to identify even relatively weak causal effects.

This study has certain limitations. First, MR relies on genetic variation to influence the participant's lifetime average risk of a specific characteristic. Hence, it cannot determine if exposure to a particular environment during a given life cycle affects the likelihood of an outcome. Second, most of the research participants included in the MR analysis were Europeans. Consequently, we have to validate whether the findings are indicative of the population as a whole. Third, although it might be a challenge to determine the extent of sample overlap, individuals within exposure and result studies might overlap. Thankfully, this study's robust tools (such as F-statistics considerably > 10) should minimise any potential bias brought on by sample overlap.

**Conclusion**

In conclusion, our results did not demonstrate a causal relationship between genetically anticipated PD and NAFLD or a causal relationship between genetically predicted NAFLD and PD. Updated MR analyses based on wider GWAS pooled data and newer genetic methods are required to corroborate the findings of this study.

**Declarations**

**Acknowledgements**

The authors acknowledge and thank the investigators of the original GWAS studies for sharing summary data used in this study.

**Declaration of funding**

This study was supported by the National Natural Science Foundation of China (Grant No. 82001075), the Natural Science Foundation of the Science and Technology Department of Shanxi Province (Grant No. 202103021224232), and the Science and Technology Innovation Project of Higher Education Institutions of Shanxi Province (Grant No. 2022L167).

**Data availability statement**
The summary statistics used and/or analyzed in the current study are available from the corresponding authors on reasonable request.

**Authors contributions**

H-DN and Y-FY presented the current ideas and were responsible for the design of the study. W-DC, HC, W-QQ, YY, LY, QR, X-LK, and CY accessed all data in the study and were responsible for the accuracy of the data analysis. W-DC and Y-FY performed the statistical analysis and wrote the manuscript. All authors were involved in writing and revising the article, and all agreed to publish the submitted version.

**Conflicts of interest**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethics approval**

This study complied with all relevant ethical regulations and received ethical approval for data collection and analysis from the Shanxi Medical University Dental Hospital.

**References**


Figures

Figure 1

Schematics for the bidirectional Mendelian randomisation design. Three presumptions must be met in order for Mendelian randomisation to be valid.
Figure 2

Scatter plot of the causal relationships between non-alcoholic fatty liver disease (NAFLD) and periodontitis (PD) using different Mendelian randomisation (MR) methods. (A) Causal estimates for PD on NAFLD. (B) Causal estimates for NAFLD on PD. Each line's slope reflects the causal estimates for each approach. The light blue line indicates the inverse-variance weighted estimates, the dark blue indicates the MR-Egger estimates, and the green line indicates the weighted median estimates. In the background, each single nucleotide polymorphism's impact on the result (point and vertical line) was contrasted with its effect on the exposure (point and horizontal line).
Figure 3

Funnel plots were used to present the overall heterogeneity of Mendelian randomisation (MR) estimates. (A) MR estimates for periodontitis (PD) on non-alcoholic fatty liver disease (NAFLD). (B) MR estimates for NAFLD on PD. The light blue line indicates the inverse-variance weighted estimates, and the dark blue line indicates the MR-Egger estimates.

Figure 4
Forrest plot of the causal relationships between non-alcoholic fatty liver disease (NAFLD) and periodontitis (PD). (A) Mendelian randomisation (MR) estimates for PD on NAFLD. (B) MR estimates for NAFLD on PD. MR estimates using the inverse-variance weighted fixed effect technique are presented in circles if each single nucleotide polymorphism was excluded one at a time.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)