Investigating the causal link between periodontal diseases and poorer cognitive outcomes using a Mendelian Randomisation approach

Thesis by alternative format

Chenyi Gao

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Medicine and Health

School of Dentistry

September 2024

Intellectual property and publications

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others. This thesis is constructed of chapters comprising of one or more related publications or works still under preparation for publication:

Gao, C., & Kang, J. (2024). Oral disease and its association with cognitive decline and dementia. In D. Deng, N. Dame-Teixeira, & T. Do (Eds.), Oral microbiome and the host: Symbiosis, Dysbiosis and Microbiome interventions for maintaining oral and systemic health. Springer. (under editing).

I contributed to conception, literature review, interpretation and draft write up. Dr. Jing Kang contributed to the conception, interpretation and critically revised the manuscript

Gao, C., Larvin, H., Bishop, D. T., Bunce, D., Pavitt, S., Wu, J., & Kang, J. (2023). Oral diseases are associated with cognitive function in adults over 60 years old. Oral diseases. https://doi.org/10.1111/odi.14757

I and Dr. Jing Kang contributed to data analysis, interpretation, and manuscript write-up. Dr. Harriet Larvin contributed to data analysis and interpretation. Dr. Jing Kang, Prof. Jianhua Wu, Prof. Susan Pavitt, Prof. David Timothy Bishop and Prof. David Bunce contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Gao, C., Iles, M., Larvin, H., Bishop, D. T., Bunce, D., Ide, M., Sun, F., Pavitt, S., Wu, J., & Kang, J. (2024). Genome-Wide Association Studies on Periodontitis: A Systematic Review. PLoS One, 19(9), e0306983. https://doi.org/10.1371/journal.pone.0306983

I contributed to systematic search, paper screening, quality assessment, data extraction, data analysis, interpretation, and manuscript write-up. Dr. Harriet Larvin and Miss Fanyiwen Sun contributed to paper screening, quality assessment, data analysis. Dr. Jing Kang contributes to the paper screening, quality assessment, data extraction, data analysis and interpretation. Dr. Jing Kang, Prof. Jianhua Wu, Prof. Susan Pavitt, Prof. David Timothy Bishop and Prof. David Bunce contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Gao, C., Iles, M. M., Bishop, D. T., Larvin, H., Bunce, D., Wu, B., Luo, H., Nibali, L., Pavitt, S., Wu, J., & Kang, J. (2024). Genetic risk factors for periodontitis: a genome-wide association study using UK Biobank. (under review on Journal of Clinical Oral Investigation).

I contributed to data analysis, interpretation, and manuscript write-up. Dr. Jing Kang and Dr. Harriet Larvin contributed to data analysis and interpretation. Dr. Jing Kang, Prof. David Timothy Bishop and Dr. Mark Iles, contributed to study design, data analysis and interpretation. Dr. Jing Kang, Prof. Jianhua Wu, Prof. Susan Pavitt, Prof. David Timothy Bishop and Prof. David Bunce contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Gao, C., Larvin, H., Zhou, K., Bishop, D. T., Iles, M. M., Bunce, D., Wu, B., Luo, H., Pavitt, S., Wu, J., & Kang, D. J. (2024). Investigating the causation between periodontitis and dementia: a Two-sample Mendelian Randomisation study using UK Biobank data. (ready to submit to Journal)

I contributed to conception, data analysis and interpretation, and drafted the manuscript. Prof. David Timothy Bishop, Dr. Mark Iles, Dr. Jing Kang contributed to study design, data analysis and interpretation. Dr. Jing Kang, Dr. Jianhua Wu, Dr. Harriet Larvin contributed to data acquisition. Dr. Jing Kang, Prof. Jianhua Wu, Prof. Susan Pavitt, Prof. David Timothy Bishop and Prof. David Bunce contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Acknowledgements

This thesis would not have been possible without the guidance, support, and encouragement of many individuals and organisations/institutes who have helped me throughout this academic journey.

First and foremost, I would like to express my deepest gratitude to all my supervisors, Dr. Jing Kang, Prof. Jianhua Wu, Prof. David Bunce, Prof. David Timothy Bishop, Prof. Susan Pavitt. Their unwavering support and patience, insightful guidance, valuable feedback, and warm encouragement were instrumental in the development of this work and my personal development. It has been a great honour to work with such a brilliant supervisory team. My deepest thanks go to my main supervisor for see my potential and offering this precious PhD opportunities, standing by my side, providing unwavering support and encouragement at every step—not only in my research but also in my career and professional development. Your commitment to my growth through teaching opportunities, constant information sharing, and tireless feedback during manuscript development, on my CV and job applications has been truly invaluable. Prof. Wu, your meticulous and constructive suggestions and feedback during PhD research and manuscript development, and continuous supports in my professional development, have been pivotal in shaping my academic journey. Prof. Bishop, your expertise in genetic epidemiology and your patient, hands-on guidance in data analysis and interpretation have been indispensable. Prof. Bunce, your rigorous feedback and comprehensive manuscript reviews, coupled with your psychological insights into result interpretation, have greatly enriched my work. Prof. Pavitt, your insightful advice on my career development and your willingness to step in as my main supervisor when Dr. Kang transitioned to KCL, have been nothing short of remarkable. Having you all as my supervisors has made me incredibly fortunate, and your collective support has made my PhD journey not just successful, but profoundly rewarding. Having you all to be my supervisor making me the luckiest PhD study ever and bring me an ever enjoyable PhD study.

I am extremely grateful to my funder, Alzheimer's Society, for their exceptional support in both my research and professional development. I would like to particularly thank Mr. David Hughes and Ms. Heather Corrie—words cannot express my gratitude for your kind and generous donation for supporting dementia research and PhD students. This work would not have been possible without your great support.

I would like to extend my special thanks to my collaborators. Dr. Mark Iles, your patience with my email inquiries, hands-on guidance in genetic data analysis, and comprehensive feedbacks on manuscript have been invaluable and greatly help me improve my research quality. Prof. Bei Wu, your enthusiasm for our research and supports on building research network, and hospitality

during my placement visit at New York University have enabled our data analysis in RDC data centre and greatly enriched my experience too. Dr. Huabin Luo, I am deeply grateful for your kindness in sharing your expertise with NHANES data analysis and for your willingness to perform data analysis in RDC data centre to facilitate this PhD project. Your expertise, input, and support have been instrumental in producing high-quality work during my PhD study.

I also want to express my sincere thanks to Dr. Harriet Larvin and Mr. Ke Zhou for your excellent contributions and wonderful collaborative efforts. It is my pleasure to work with both of you.

I would also like to extend my sincere thanks to the anonymous reviewers who offered valuable comments and constructive feedback during the publication process, as well as to the participants who provided data to the UK Biobank and the National Health and Nutrition Examination Survey.

Lastly, I would like to express my deepest thanks to my beloved family and friends. To my mother, father, and Chengni, thank you for your endless support and belief in my ability to achieve excellence. To my dear friends Yang Zhao, Xiaoxuan Fu, Miao Gao, Nuan Shangguan, and Qianyu Wang, your emotional support and companionship during stressful times have been a source of happiness and strength. Thank you to my little Persica and Taro for loving me and bringing so much joy with your adorable presence, which truly reduced my stress during my PhD journey.

I acknowledge the use of ChatGPT in the proof-reading of the thesis before submission. I confirm that the proof-reading undertaken by ChatGPT was in accordance with the Postgraduate Researcher Proof-reading Policy.

Abstract

Dementia, a leading cause of mortality and disability, affects a growing number of older adults and is increasingly seen in younger populations. Despite heightened public health concern, no treatment exists to reverse dementia's progression, making prevention crucial. Periodontitis has been suggested as a potential causal risk factor for dementia, but inconsistent study results, confounding factors, reverse causation, and limited biomedical research complicate our understanding of this relationship. Mendelian Randomisation (MR) offers a cost- and time-efficient method for investigating causal links through genetic epidemiology.

The aim of this thesis is to explore the causal relationship between periodontitis and cognitive decline, including dementia, using the MR approach with data from the UK Biobank (UKB) and National Health and Nutrition Examination Survey (NHANES).

To address this, both epidemiological and genetic epidemiological methods were employed. First, a systematic review of the literature revealed discrepancies in the association between periodontitis and dementia, varying by gender, region, and disease severity. A cross-sectional study using NHANES data confirmed a significant positive association between periodontitis and poorer cognitive function. This led to the next step—causality investigation via MR. After reviewing existing periodontitis genome-wide association studies (GWASs), an original periodontitis GWAS on periodontitis using UKB data were performed to identify the genetic risk variants of periodontitis that may be able to use as genetic instruments for periodontitis. A two-sample MR analysis using UKB data found partial evidence of causality, but results were inconsistent across models and require cautious interpretation.

In conclusion, this thesis confirms a positive association between periodontitis and poorer cognitive function but suggests that periodontitis is unlikely to be a direct causal risk factor for dementia. These findings provide preliminary evidence and valuable insights for researchers, clinicians, and public health professionals for understanding this causal relationship, as well as potentially guiding future research on the underlying mechanisms and encouraging a holistic approach to patient care.

Table of Content

List of abbreviations

Abbreviation	Full Term
A1	Effect Allele
A1 Freq	A1 Frequency
A2	Non-Effect Allele
AF	Allele Frequency
AgP	Aggressive Periodontitis
AL	Attachment Loss
AP	Apical Periodontitis
APOE	Apolipoprotein E4
ASH1L	ASH1 Like Histone Lysine Methyltransferase
AUC	Area Under Curve
BBB	Blood-Brain Barrier
BMI	Body Mass Index
CD33	CD33 Molecule
CDC-AAP	The Centres For Disease Control And Prevention And American Academy Of Periodontology
CERAD W-L	The Consortium To Establish A Registry For Alzheimer's Disease Word Learning Subtest
CF	Cognitive Function
CFI	Comparative Fit Index
Chr: Pos	Chromosome: Position
CI	Confidence Interval
CNS	Central Nervous System
CP	Chronic Periodontitis
CR1	Complement Receptor 1
CRACR2A	Calcium Release Activated Channel Regulator 2A (CRACR2A)
df	Degree Of Freedom
DMFT	The Number Of Decayed, Missing And Filled Teeth
DR15	DQA1*01:02~HLA-DQB1*06:02~HLA-DRB1*15:01
DSST	The Digit Symbol Substitution Test
EAF	Effect Allele Frequency
ERB	Ethics Review Board
ERG	E-26 Transformation–Specific TF-Related Gene
FCER1G	Fc Epsilon Receptor Ig
GC	Genomic Control
GCF	Global Cognitive Function

GLIDE	Gene-Lifestyle Interactions In Dental Endpoints
GPR141	G Protein-Coupled Receptor 141
GRC	Genome Reference Consortium
GRCH37	Genome Reference Consortium Released "Build 37" Of The Human Genome
Gtex	Genotype-Tissue Expression
GWA	Genome-Wide Association
GWAS	Genome-Wide Association Study
HLA	Human Leukocyte Antigen
HRC	Haplotype Reference Consortium
HWE	Hardy-Weinberg Equilibrium
IBS	Pair-Wise Identity-By-State
ICD10	International Classification Of Disease (10th Edition)
ICD9	International Classification Of Disease (9th Edition)
IL	Interleukin
INFO	Imputation Information
IVW	Inverse-Variance Weighted
KCNQ5	Potassium Voltage-Gated Channel Subfamily Q Member 5
LD	Linkage Disequilibrium
LDSCR-I	LDSR Intercept
LDSR	Linkage Disequilibrium Score Regression
MAF	Minor Allele Frequency
MDS	Multimentional Scaling
MDSC	Multimentional Scaling Components
MR	Mendelian Randomisation
MS4A	Membrane-Spanning 4-Domains Subfamily A
NHANES	National Health And Nutrition Examination Survey
OH	Oral Health
OR	Odds Ratio
P. Gingivalis	Porphyromonas Gingivalis
PC	Principal Components
PCA	Principal Components Analysis
PEO	The Population Of This Study, The Exposure Of This Study, The Outcome Of This Study
PIR	Poverty Index Ration
PRISMA	Preferred Reporting Items For Systematic Reviews And Meta-Analyses
PRS	Polygenic Risk Score

QC	Quality Control
RMSEA	Root-Mean-Square Estimate
ROC	Receiver Operator Characteristic
RR	Relative Risk
RS ID	SNP Reference Identifier
SD	Standard Deviation
SE	Standard Error;
SEM	Structure Equation Model
SIGLEC5	Sialic Acid Binding Ig Like Lectin 5
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
TREM2	Triggering Receptor Expressed On Myeloid Cells 2
UKB	UK Biobank

Chapter 1: Introduction

Poorer cognitive outcomes such as dementia are major concerns for older adults, impacting health status and quality of life (e.g., disability and dependency), and placing burden on care givers and health care systems (Park et al., 2003; Prince et al., 2013). Meanwhile, oral health problems in older adults are also widespread. Periodontal diseases are common oral health problems where severe cases affect around 19% of the adult population globally (WHO, 2023b). It has been associated with various systemic health problems and has been suggested to be a risk factor for dementia (Demmer et al., 2020; Kamer et al., 2008). However, despite a growing body of epidemiological work exploring the association between periodontal diseases and cognitive outcomes in the past two decades using both longitudinal and cross-sectional designs, inconsistent results (Larvin et al., 2023), concerns of confounding effects (Davies et al., 2018; Sekula et al., 2016), and lack of biomedical studies exploring potential underlying mechanisms (Thomson and Barak, 2021) have led to controversy and debate concerning any causal relationship between periodontal diseases and dementia. One methodological approach that may shed light on this relationship is Mendelian Randomisation (MR). This is an advanced epidemiological research approach that has been used in the past decade to investigate causal relations between modifiable risk factor and disease (Emdin et al., 2017; Sanderson et al., 2022). The approach has considerable potential to provide important insights into the causal association between periodontal diseases and dementia, having the advantages of limiting confounding effects and addressing reverse causation concerns (Emdin et al., 2017; Sanderson et al., 2022) if a suitable genetic predictor of periodontal diseases status can be identified. Given this background, this thesis will use MR to attempt to investigate the relationship between periodontal diseases and cognition and dementia, and also present a research output that contributes to MR analysis and results.

This introductory chapter considers the disease definition, the disease burden for both poor cognitive outcomes at older age and dementia, and periodontal diseases. The research objectives and rational for the research design and main research methods, are then described. The subsequent chapters present the research findings, and the final chapter discusses the interpretation of the findings and the strengths and limitations of present research.

This chapter includes content from my published book chapter *"Oral disease and its association with cognitive decline and dementia." (Gao and Kang, 2024)*, from the upcoming Springer Book "***Oral Microbiome and host***" edited by Naile Dame Teixeira, Thuy Do and Dongmei Deng. The chapter provides a more comprehensive review of common oral diseases, tooth loss, periodontal diseases and dental caries whilst this introductory chapter focuses specifically on material that aligns with the thesis focus. That is, periodontal diseases, poor cognitive outcomes including dementia.

Chapter 1.1 Study background

1.1.1 Cognitive function and dementia

Disease definition

Cognitive function is obviously central to our daily lives and includes a wide range of abilities such as attention, memory, perception, language, executive function, motor skills and construction (Harvey, 2019). Cognitive function changes throughout our lifespan (Bialystok and Craik, 2006) and can vary depending on multiple factors such as education, lifestyle behaviours (e.g., smoking, drinking, physical activity), mental health, comorbidities (e.g., traumatic brain injury, diabetes), and age (Baumgart et al., 2015). For example, age is significantly associated with cognitive function and negatively impacts multiple cognitive domains including attention and memory (Murman, 2015). In addition, age-related brain structural changes such as reduced brain size and changes in neuronal networks are associated with observed age-related decline in cognitive function (Murman, 2015). Although changes in cognitive function with age do not necessarily indicate underlying neuropathology, they can still be an alarming part of aging as they can influence financial, personal and societal burden (Deary et al., 2009).

In addition to normal age-related decline in cognitive function, neuropathological cognitive decline or cognitive impairment in later life is of considerable concern to the public where the most common form of cognitive impairment at late life stage is degenerative dementia (Murman, 2015). Dementia refers to a general loss or impairment in a range of cognitive domains that may severely affect daily living (Arvanitakis et al., 2019). For example, memory loss is the most "well-known" symptom, while mood (e.g., anxiety related to memory loss) and behavioural changes (e.g., personality change and inappropriate behaviours) may manifest before memory issues occur (WHO, 2023a).

Diseases causes

The causes of dementia includes a variety of diseases or injury directly or indirectly causing damage or alteration to the neuron or neuron network in the brain which can result in cognitive impairment and cognitive decline such as stroke and alcohol misuse (Murman, 2015). Genetic factors also play fundamental roles in cognitive decline and dementia. The heritability of late onset dementia was found between 60%-80% from twins studies and the SNP-based heritability estimated from different cohorts (i.e., the UKB, The KRONOS/Tgen dataset, The GR@ACE data, ROSMAP, The Amsterdam Dementia cohort data) was found between 38%-66% (Baker et al., 2023). Increasing heritability with age (i.e. before age of 60 years old vs. after age of 60 years old) was reported for age-sensitive cognitive domains including general memory (heritability

estimated between 23%-46%), working memory (heritability estimated between 26%--38%) and executive function (heritability estimated between 27%-77%) (Reynolds and Finkel, 2015).

Despite various causes and pathological mechanisms for neurodegeneration diseases (e.g., Alzheimer's disease and Frontotemporal dementia), chronic neuroinflammation has been found as a shared feature among neurodegeneration diseases (Zhang et al., 2023). The neuroinflammation was previously thought to be the consequence of protein aggregation in the central nervous system (CNS), but later, neuroinflammation was also found to contribute to the aggregation in neurodegeneration diseases progression at an early stage (Zhang et al., 2023) and the immune response in the brain has since been recognized as one of the core pathologies in the disease (Heppner et al., 2015; Kinney et al., 2018). It has been suggested that brain inflammation have a dual function: (1) a neuroprotective effect in acute-phase, and (2) an adverse reaction in chronic phase. In the adverse reaction phase, many pro-inflammation and toxic products such as reactive oxygen species, nitric oxide, and cytokines will be released by microglia, the resident immune cells within the central nervous system (Kinney et al., 2018). In most cases the immune response is thought to be beneficial but if the stimulus is persistent and unresolved the immune response will become neurotoxic (Zhang et al., 2023). Variety of factors may be recognised as the cause of a persistent stimulus such as the environmental factors (e.g., systemic infections) and genetic factors (e.g., *apolipoprotein E4 (APOE4*) mutations) (Zhang et al., 2023).

In addition to the genetic risk factor *APOE* in dementia, the effects of inflammation and involvement of immune system on dementia were also supported by the immune-related genes discovered in dementia. For example, previous genetic meta-analysis study also found that there is shared genetic architecture between late onset Alzheimer's disease and immune-mediated traits such as immune-mediated disease *haplotype human leukocyte antigen (HLA) DQA1*01:02~HLA-DQB1*06:02~HLA-DRB1*15:01 (DR15)* haplotype (Kunkle et al., 2019). A previous study suggested the HLA variety determined by *HLA* genes is critical to human adaptive immunity with a range of immune responses activities (Wang et al., 2020). This includes inflammation mediated by the immune response, T-cell transendothelial migration, infection and brain development and plasticity in Alzheimer's diseases pathogenesis etc. Specifically, on *HLA-DR15* was reported associated with human autoimmune disease via interleukin (IL)-17 regulation (Wang et al., 2020) where ILs are a type of cytokines critical to the pro-inflammation and anti-inflammation response and are significantly associated with the risk of Alzheimer's disease (Kinney et al., 2018; Rainero et al., 2004). Many other Genome-wide Association studies (GWAS) found inflammation-related genes (e.g., *CR1 (complement receptor 1), MS4A (membrane-spanning 4-domains subfamily A), TREM2 (triggering receptor expressed on myeloid cells 2), and CD33 molecule (CD33)*) related to the risk of Alzheimer's diseases (Karch and Goate, 2015) and variants associated with Alzheimer's disease are enriched in

immune-related tissues across the body, including whole blood, spleen, liver, and key immune cells of the brain (i.e., microglia) (Jansen et al., 2019). These evidence suggest the fundamental role of the immune response and inflammation in dementia development.

In addition to the inflammation in the CNS, there are increasing studies suggesting the importance of the immune response in the peripheral system in promoting neurodegeneration, cognitive decline and dementia pathology, specifically Alzheimer's disease (Bettcher et al., 2021; Walker et al., 2019). Investigating the systemic inflammation or disease trigger the systemic inflammation may contribute to the discovery of new treatment targets for dementia.

Diseases burden

Alzheimer's disease is the most prevalent form of dementia (Arvanitakis et al., 2019), accounting for around 60%-70% of dementia cases. Alzheimer's disease and other dementias are also a leading cause of mortality worldwide and also disability and dependency in late life (Prince et al., 2016; WHO, 2023a). The impact of dementia is also multidimensional with physical, psychological and social impacts on not only patients but also carers and health care systems. Due to limited awareness and understanding of dementia, stigmatization (i.e., negative attitude and behaviours direct towards dementia patients) and diagnosis barriers may also result (WHO, 2023a). For example, it has been estimated that 75% of dementia patients are undiagnosed worldwide, especially in lower income countries where the rate increases to 90% (ADI, 2022).

It has been estimated that there are 55 million people in the world living with dementia, and the number is forecast to increase to 139 million by 2050 (ADI, 2022). This increase in dementia cases around the world is primarily due to global population growth and population aging (Nichols et al., 2022). Although dementia is strongly associated with age and more often observed in old adults (>=65 years old), the prevalence of early-onset dementia (i.e., dementia before 65 years old) is also high. Previous studies estimate approximately 3.9 million people around the world to live with early onset dementia (Hendriks et al., 2021). Since patients often receive diagnosis at a late disease stage (Prince et al., 2013) and current treatment methods do little to slow or reverse the course of dementia (Livingston et al., 2020), it is of pressing importance that research investigates potential modifiable risk factors for dementia prevention and intervention.

Oral health factors, particularly periodontal diseases status, is worth consideration in this respect as evidenced by suggestive association studies and a few biomedical studies. In the next section of this chapter, the definition and burden of periodontal diseases will be introduced and then evidence on its potential of being a modifiable risk factor will also be comprehensively discussed.

1.1.2 Periodontal disease

Diseases definition

Periodontal diseases refer to a range of conditions that affects the tissue surrounding or supporting the teeth, including gingiva or gums, bone and ligament (Kinane et al., 2017; Pihlstrom et al., 2005). Periodontal diseases can result from multiple causes such as developmental, inflammatory, traumatic, neoplastic, genetic and metabolic factors (Pihlstrom et al., 2005).

Gingivitis, the inflammation at gingiva, is the early stage or mildest form of periodontal diseases and results from harmful microorganisms being present in biofilm or dental plaque (Kinane et al., 2017; Pihlstrom et al., 2005). Gingivitis is characterised by bleeding while brushing and occasionally reported pain (Kinane et al., 2017). At this stage, the inflammation is limited to the soft tissues, epithelium and connective tissue (Kurgan and Kantarci, 2018) and periodontal diseases can be easily reversed by good oral hygiene (Kinane et al., 2017; Pihlstrom et al., 2005).

Periodontitis is a more severe form of periodontal diseases, in which case the inflammation progresses to the supporting structures (Kurgan and Kantarci, 2018). This leads to the loss of surrounding gingiva, supporting bone, ligaments, which result in deepened periodontal pockets and accompanying attachment loss (Kinane et al., 2017). The chronic periodontitis present with clinical characteristics such as redness, bleeding of gingival pocket area on probing, deepened periodontal pocket, alteration in the texture and swelling of marginal gingiva, recession of the marginal gingiva, increased tooth mobility and tooth loss at end stage (Kinane et al., 2017) The structure loss associated periodontitis is irreversible. There are also different forms of periodontitis with chronic and aggressive periodontitis as two most common presentations (Stabholz et al., 2010). Chronic periodontitis is more often observed in older adults and characterised by a slow progression rate in terms of periodontal tissue disruption and bone loss regards to age; whereas aggressive periodontitis is characterised by a more rapid rate of progression and is more often observed in young people (Armitage et al., 2010; Cardoso et al., 2018). Genetics play a stronger role in aggressive periodontitis than chronic periodontitis (Van der Velden, 2017).

According to the World Workshop of Periodontology 2018, the periodontitis has been further summarised to different stage and grade of periodontitis (Tonetti et al., 2018). This current classification consists of four stages with consideration of severity and complexity. The stage I periodontitis is characterised by an early stage of attachment loss which develops from the persistent gingival inflammation and biofilm dysbiosis. The stage II periodontitis is characterised by greater damage to the tooth support. Stage III periodontitis is characterised by significant attachment apparatus damage and may accompanied with tooth loss in absence of advanced

treatment. Stage IV periodontitis involves significant periodontal support damage, leading to substantial tooth loss. The grade of periodontitis adds one more dimension, which considers the rate of disease progression (grade A—slow rate of progression, grade B—moderate rate progression, grade C—rapid rate of progression). Indirect evidence includes the percentage of bone loss relative to age, as well as the level of destruction and biofilm deposits. Meanwhile, smoking and metabolic control of diabetes have been considered grade modifiers.

Diseases causes

Gingivitis is a risk factor for periodontitis as previous studies have found that inflammation in the gingival tissue is a prerequisite for periodontitis (Kurgan and Kantarci, 2018). However, despite increased risks, not all sites or individuals with gingivitis will develop into periodontitis over time; some remain stable instead (Kurgan and Kantarci, 2018). Multiple variables may play a role in this progression, including lifestyle behaviours such as smoking and oral hygiene, nutrition, obesity, diabetes, genetic risk factors (Genco and Borgnakke, 2013; Kinane et al., 2017; Tonetti et al., 2018). According to previous meta-analysis of periodontitis genetic studies, the heritability of periodontitis from GWAS was 7% to 38% as estimated from GWAS and family studies with increase by diseases severity, while self-reported gingivitis has a heritability of 29% (Nibali et al., 2019).

More than 65 genes have been found to play a role in periodontitis in genetic studies (e.g., candidate gene study and genome-wide association study (GWAS)) and the majority of these genes account for a small proportion of total risk of periodontitis (Loos and Van Dyke, 2020). Most of the genes involved in periodontal diseases are associated with the immune response (Shaddox et al., 2021). For example, the *IL* genes play not only fundamental role in inflammatory response but also in tissue destruction in periodontitis. The pro-inflammatory IL molecules previously found in periodontitis include IL-1, IL-1α, IL-1β, IL-2, IL-5, IL-6, IL-7, IL-9, IL-12, IL-17, IL-18, IL-23, and IL-33 (Liu., 2022; Liu and Li, 2022; Martínez-García and Hernández-Lemus, 2021). These pro-inflammatory mediators, such as cytokines, were released by leukocytes of the innate immune system and stimulated by periodontitis pathogens (Martínez-García and Hernández-Lemus, 2021). Genes or single nucleotide polymorphisms (SNPs) that modulate IL molecules, such as IL-1 release, have been reported in periodontitis (Karimbux et al., 2012; Shaddox et al., 2021). However, not all ethnicities replicate these SNPs and their association with periodontitis (Shaddox et al., 2021).

Diseases burden

Despite the simple reversibility of gingivitis, periodontitis is still prevalent in around 10%-19% of populations globally (Disparity), 2021; WHO, 2023b), and affects almost half of UK adults,

and more than half of US adults. Periodontitis can lead to pain or discomfort and difficulty in chewing. The end stage of periodontal diseases is tooth loss which has a severe impact on function, and an adverse psychological and social impact (Emami et al., 2013; WHO, 2023b), especially for those unable to replace with dentures. WHO estimate around 7% of adults (>=20 years old) worldwide to experience complete tooth loss. This prevalence increases to 23% for people aged 60 year or older (WHO, 2023b).

Periodontal diseases as common peripheral inflammatory conditions have been suggested to also contribute to whole body inflammatory burden (Kinane et al., 2017) worsening many conditions such as diabetes mellitus (Genco et al., 2020), cardiovascular disease (Larvin et al., 2021), and also cognitive outcomes including dementia where neuroinflammation plays a critical role in the pathogenesis (Heneka et al., 2015). In the next section, I discuss current understanding of the association between periodontal diseases and poorer cognitive outcomes and dementia, as well as potential underlying mechanisms.

1.1.3 Association evidence from observational studies

In the past two decades, there have been increasing numbers of population studies investigating whether there is an association between periodontal diseases and cognitive function, as this may contribute to developing prevention approaches and/or delaying the onset of dementia. The earliest paper reporting the association between oral disease and cognitive function that we were able to identify is a Japanese paper published in 1994 (Kondo et al., 1994) which has been discussed by many later studies (e.g., (Thomson and Barak, 2021). This study used a case-control design where sex and age were matched in 60 Alzheimer's disease patients with two controls for each case (120 controls in total) to investigate the lifestyle risk factors for Alzheimer's disease (Kondo et al., 1994). Five risk factors were found, and the results suggested participants who have all five risk factors were more than nearly 1,000 times likely to develop Alzheimer's disease compare with those with none of these risk factors. The five risk factors were tooth loss, psychosocial inactivity, physical inactivity, head injury and low education. This research provides an initial indication that oral disease may contribute to Alzheimer's disease, prompting many later studies to investigate the potential of oral diseases as risk factors for poorer cognitive outcomes including dementia and potential mechanisms underlying this association further. However, the sample size of this preliminary evidence is relatively small which may affect the results' validity and generalisation, additional evidence is needed to further validate whether tooth loss is a risk factor for cognitive decline. Three years later, an animal model study found neurological impairment following loss of molar teeth suggesting that tooth loss might be a risk factor for dementia (Kato et al., 1997). By contrast however, a prospective longitudinal study published in

the 2001 did not find significant associations between baseline dentition status and incidence dementia (Shimazaki et al., 2001).

More recently, many longitudinal and cross-sectional studies suggest there is positive association between periodontal diseases and dementia or cognitive impairment or decline (Choi et al., 2019; Demmer et al., 2020; Kang et al., 2020; Kang et al., 2019; Winning et al., 2022). There are also some studies that have reviewed potential underlying mechanisms underlining increasing interest in this field (Kamer et al., 2008; Stein et al., 2006). There are also some studies that have reviewed potential underlying mechanisms underlying the increasing interest in this field (Kamer et al., 2008; Stein et al., 2006). These mechanisms included: (1) the inflammatory response to periodontal diseases may potentially harm the brain; (2) the oral bacteria may be potentially transported to brain and lead to infection; (3) malnutrition associated with periodontal diseases might increase the risk of Alzheimer's disease; (4) periodontal diseases might increase the risk of Alzheimer's disease via third factors (e.g., stroke or cerebrovascular injury); (5) shared genetic risk factors (e.g., *IL-1* gene) between periodontal diseases and Alzheimer's disease contribute to the association observed between them. These mechanisms are discussed in detail in later sections.

However, the systematic review including 56 cross-sectional and longitudinal studies (searched between 1993-2013) found that significant associations between periodontal diseases and poor cognitive outcomes were not consistently established across studies (Wu et al., 2016). Similar findings from systematic review including 16 longitudinal and case-control designed studies (searched until 2016) were also reported (Tonsekar et al., 2017). By focusing on case-control studies only, a more recent systematic review and meta-analysis (search until 2018) found one out of 12 included studies failed to identify a significant result, but their meta-analysis did yield significant results suggesting periodontal diseases increases the risk of dementia incidence (Nadim et al., 2020). The relative risk ratio of dementia based on periodontal diseases status in this study was 1.38 with 95% confidence interval (1.01-1.90).

Since there has been no recent systematic review and meta-analysis performed for both cross-sectional and longitudinal studies, as well as considering the impact of study factors on the results of the association, prior to this thesis, a systematic review and meta-analysis study from our research team was conducted to understand current observational studies on the association between periodontal diseases and cognitive decline and dementia. There are 49 observational studies, including 21 cross-sectional and 28 longitudinal studies, investigating the association between periodontitis and cognitive function in the literature up until 2022 (Larvin et al., 2023). However, this systematic review found that the quality of the included studies was mixed, and findings were also mixed. By using ROBINS-I checklist, most of the studies (n=39) included had selection bias or a risk of confounding which leads to serious risk of bias. Meanwhile, the meta-

analysis results also suggested that the association between periodontitis, cognitive function, and dementia varied based on gender, the measurements method of periodontitis and the severity of periodontitis etc. More importantly, a significant association result is not consistently reported across studies.

The lack of consistent findings leads to questions concerning whether having periodontal diseases is a risk factor for poorer cognitive outcome, especially dementia. For example, reverse causation is possible with evidence showing that the presence of poor cognitive function might be a risk factor for subsequent oral health problems. Dementia may result in limited ability to conduct adequate daily activities including oral health care. Therefore, specific oral health care interventions may need to be designed by dental health professionals and implemented by care givers in order to maintain good oral health and prevent the development of oral diseases. Much of this research addresses the oral health concerns in the people with impaired cognitive ability. There are also several early studies that have addressed oral health need that attempted to create guidance for dentists when treating dementia patients (e.g., (Ghezzi and Ship, 2000; Niessen and Jones, 1987)). For example, dementia disease severity needs to be considered, as well as the involvement of care givers, instructions to caregivers, and consideration given on use of sedation during treatment for managing uncooperative behaviours. Although numerous early studies (1993-2013) have been conducted to test the effect of cognitive function on oral health, the effect is still not fully understood or well explained because of methodological issues (e.g., sample size, inconsistent assessment tools for dental and cognitive function/dementia) in early studies as pointed out by the systematic review (Wu et al., 2016).

1.1.4 Potential mechanisms

Life-course hypothesis

With developments in the field (e.g., publication of consensus standardized disease definition, larger sample studies), higher-quality studies have been published recently and provide further evidence for the effect of cognitive function in association to oral diseases. In particular, a newly emerged “life-course” model provides an explanation of the mechanisms underlying the association observed between oral disease, tooth loss specifically, and poorer cognitive outcomes with consideration of cumulative effects throughout life course(Thomson and Barak, 2021). This model could also be used to explain the impact of cognitive function on the periodontal diseases progression towards end stage of the disease (i.e., tooth loss). In the previous definition section, I touched upon the disease progression for both periodontal diseases and dementia and noted numerous common factors throughout the lifespan that could contribute to or prevent both diseases such as education, smoking and alcohol. The Life-course model emphasises the

importance of considering life-long exposure to the risk factors when looking at oral health outcomes due to the chronic and cumulative nature of common oral diseases (Thomson and Barak, 2021). In particular, tooth loss is a cumulative and final state of many oral diseases including periodontal diseases. It suggests the impact of cognitive function on tooth loss could start at an early life stage, as it contributes to better dental care (e.g., routine dentist visits, better dental hygiene) along life span. Meanwhile, better cognitive function contributes to better life-style behaviours (e.g., less smoking, alcohol drinking, higher education) which further enhance the effect on oral health outcomes through the life course (Thomson and Barak, 2021). Those with better cognitive function at early life stage are also likely to maintain a better cognitive function and delay the progress of dementia or cognitive impairment at later life stage. During the aging process, the combined effect from better cognitive function cumulate and lead to a better oral health condition at later life stages. On the contrast, the adverse effect also cumulated from early life stage along the life course and leads to a poorer oral health status at late life stage. Therefore, the association can be observed between oral health and cognitive function.

Based on the current research findings and life-course model, it is also likely that the association found between periodontal diseases and poorer cognitive outcomes including dementia may be due to a reverse causation (i.e., poorer cognitive function leads to poorer oral outcome.). The association between periodontal diseases and poorer cognitive outcomes including dementia is not yet well understood. In the next section, some mechanisms potentially explaining how periodontal diseases contribute to dementia development are discussed.

Shared inflammation pathway hypothesis

There are also several mechanisms that have been proposed to explain the association between periodontal diseases and poorer cognitive outcomes including dementia. In the next subsection, mechanisms suggesting periodontal diseases as causal risk factor for poor cognitive outcomes, dementia in particular, are discussed and evaluated.

Prior to association found between periodontal diseases and poorer cognitive outcomes, there were already studies suggesting that systemic inflammation may increase the risk of dementia. For example, a meta-analysis of a population-based study assessing the association between inflammatory markers, including C-reactive protein, IL-6, α1-antichymotrypsin, lipoprotein-associated phospholipase A2 activity, and fibrinogen, and dementia found an increased risk of dementia (Darweesh et al., 2018). However, the exact effects of these inflammatory markers involved in Alzheimer's disease's pathology remain to be understood.

It has been suggested that systemic inflammation, characterised by increased levels of circulating proinflammatory cytokines and chemokines in the blood, can lead to various molecular changes that may harm the CNS and initiate or promote neurodegeneration (Walker et al., 2019). For instance, For instance, systemic inflammation can trigger reactive and proinflammatory microglial and astrocytic phenotypes in the CNS. This may result in harmful effects such as tau hyperphosphorylation, β-amyloid oligomerization, complement activation, and the breakdown of neurotransmitters into toxic metabolites (Walker et al., 2019). These effects occur as the CNS detects inflammatory proteins transported through various pathways to the brainstem. These pathways include the transport of proinflammatory cytokines across the blood-brain barrier (BBB), the transport of cytokines through interactions with endothelial cells and circumventricular organs, and the stimulation of the vagus nerve.(Walker et al., 2019).

As mentioned in section 1.1.1, inflammation and immune response are part of the core pathology in both dementia (Heppner et al., 2015; Kinney et al., 2018) and periodontal diseases (Hajishengallis and Chavakis, 2021; Kinane et al., 2017). Periodontal diseases have also been found associated with systemic inflammation and many comorbidities characterised by chronic inflammation (Hajishengallis and Chavakis, 2021) such as diabetes (Preshaw et al., 2011), cardiovascular diseases (Larvin et al., 2021) and COVID-19 (Larvin et al., 2020). Similar to dementia patients, the increased blood level of pro-inflammatory mediators/cytokines (e.g., IL-1, IL-6, C-reactive protein and fibrinogen) as well as neutrophil numbers observed in patients with periodontitis compared to healthy controls (D'Aiuto et al., 2013; H. Bokhari et al., 2012; Hajishengallis and Chavakis, 2021; Schenkein et al., 2020). Meanwhile, reduced systemic inflammatory markers have also been observed after patients received periodontal treatments (D'Aiuto et al., 2013; H. Bokhari et al., 2012; Hajishengallis and Chavakis, 2021; Schenkein et al., 2020). It has been suggested that these increased inflammatory markers may be caused by the haematogenous spread of periodontal bacteria or the overflow of inflammatory mediators from periodontal tissues into the bloodstream (Hajishengallis and Chavakis, 2021). The association between periodontal diseases and systemic inflammation has been observed and contributes to several systemic diseases, including cognitive decline and dementia.

Through this shared inflammation pathway, periodontal diseases may also contribute to cognitive decline and dementia by initiating or promoting neuroinflammation in the CNS (Kamer et al., 2008). This effect of periodontal diseases may be achieved by an increase in serum levels of inflammatory markers such as C-reactive protein which could potentially reach the brain through systemic circulation and neuropathways (Kamer et al., 2008). In addition to the common increases in increased inflammatory cytokines from both diseases, one of the supporting evidence for the shared inflammation pathway is that there some shared genes and inflammation markers between periodontal diseases and dementia such as *IL genes*. A recent study analysed the common

molecular mechanisms between periodontitis and Alzheimer's disease and found 21 shared genes (e.g., *IL-6, IL-17A, IL-1α, IL-1β, IL-10, C-reactive protein*) between periodontitis and Alzheimer's diseases were enriched in cell cytokine activity, inflammatory response, and extracellular membrane (Ge et al., 2024). This finding indicates the link between periodontitis and Alzheimer's disease pathogenesis and support the shared inflammatory pathway mechanism between both diseases. However, it should be noted that these shared genetic pathways for periodontal diseases and Alzheimer's diseases not necessarily support the causal relationship between them and may also indicate that genetics may be third factors contributing to both diseases.

The primary supporting evidence for the shared inflammatory pathway hypothesis is that the key oral pathogen, *Porphyromonas gingivalis (P. gingivalis)*, in the chronic periodontitis has been observed in the brain of Alzheimer's patients (Dominy et al., 2019). The presence of *P. gingivalis* in brains of mice has also been associated with increased levels of amyloid-β, complement activation, and neuroinflammation. Meanwhile, gingipains, the toxic proteases produced by *P. gingivalis*, were also observed and the level of gingipain was associated with tau and ubiquitin, the markers of Alzheimer's diseases pathology (Dominy et al., 2019). As suggested by a previous review (Hajishengallis and Chavakis, 2021), *P. gingivalis* may transport across the brain-blood barrier by increasing the vascular endothelial permeability (Farrugia et al., 2021; Hajishengallis and Chavakis, 2021), thereby affecting neurological function in Alzheimer's patients. However, more investigation on the way the pathogen is transported to the brain is still needed for better and clearer understanding of this potential mechanism (Hajishengallis and Chavakis, 2021).

There are several ways that the *P. gingivalis* may exert its impact on the brain of Alzheimer's diseases patient, such as: (1) triggering immune response and increases the blood level inflammatory cytokine/markers and then affecting brain as discussed above; (2) causing neuroinflammatory response directly by causing brain infection as evidence by above observation of *P. gingivalis* in the brain, and (2) interacting with the *APOE,* which is encoded by the strongest genetic risk factor for late-onset Alzheimer's diseases (Hajishengallis and Chavakis, 2021). *P. gingivalis* gingipains could break down *APOE* at a specific site (Arg residues) (Lönn et al., 2018), whilst the *APOE* fragment may be neurotoxic (Muñoz et al., 2019) and may contribute to Alzheimer's diseases (Hajishengallis and Chavakis, 2021). A previous review suggested that systemic inflammation contributes to the alteration in BBB which may enable the *P. gingivalis* transport into brain and causing further brain infections and stimulate the inflammation response in the CNS (Varatharaj and Galea, 2017). Another evidence that supports the impact of periodontal diseases on Alzheimer's diseases via *P. gingivalis* gingipains found that gingipain inhibitor reduced neuroinflammation and prevented further neurodegeneration in a mouse model (Dominy et al., 2019). However, as this clinical evidence is from a mouse model, the effect of

gingipains inhibitor in an Alzheimer's diseases patient is still unknown. Meanwhile, evidence directly supporting the periodontal disease as a cause of dementia is still scarce, and more high-quality evidence is still needed to uncover whether periodontal diseases can cause poorer cognitive outcomes and dementia. It would also be interesting and beneficial for future study to understand how *P. gingivalis* can interact with other oral pathogens and affect cognitive function.

Other possible mechanisms hypotheses

Another possible mechanism concerns tooth loss and pain when chewing resulting from periodontal diseases affecting cognitive function is via nutrition pathways. The nutrition model suggests that impaired chewing ability caused by tooth loss could lead to reduced food choice that may contribute to malnutrition (Thomson and Barak, 2021). A recent systematic review and meta-analysis found a 21% increased risk of malnutrition in edentulous older adults or older adults without functional dentition compared to those with functionally adequate dentition (Zelig et al., 2022). In consequence, normal cognitive function might be affected due to lack of nutrition (Yu et al., 2021). Additionally, the dietary choices made by tooth loss patients may result in a high fat diet that contributes to obesity (Nascimento et al., 2016). As obesity is a risk factor for dementia (Beydoun et al., 2008), the increased risk of obesity due to tooth loss related food choice may increase the risk of dementia.

Pain, induced by periodontitis, supressing normal cognitive functioning (Moriarty et al., 2011) may also serve as a simple and direct mechanism underlying the observation of relatively lower cognitive performance in patients with periodontitis. However, the pain pathway would not be able to lead to dementia if pain killers managed the pain.

Although there are many studies and hypothesised mechanisms proposed to explain the causal effect of periodontal diseases and dementia, limitations and lack of evidence makes it difficult to draw clear conclusions concerning the underlying causal relationship. Association due to a common covariates/confounding factor and without a causal relationship is also highly possible with many shared risk factors (e.g., smoking) which are not well adjusted in the statistical analysis in many existing observational studies (Thomson and Barak, 2021). Reverse causation as suggested in the Life-course model is also possible. Meanwhile, the biomedical study which trying to explore the mechanisms of the effect of periodontal diseases on poorer cognitive outcomes especially dementia is limited and far from confirming the causal impact. Therefore, unlike more promising and well evidenced effects from poor cognitive function on periodontal disease, the effect of periodontal disease on cognitive function, especially dementia, still need further investigation with particular focus on exploring its causal relationship with considerations on other common covariates effect and understanding the biological underpinnings.

No causal link underlying

In addition to the aforementioned underlying mechanisms, it is also possible that no causal relationship exists between periodontal diseases and poor cognitive outcomes, including dementia. The observed association in observational studies may simply reflect the high prevalence of both conditions in older adults. Consequently, more robust evidence is needed to establish a causal relationship, alongside further investigation into each hypothesised mechanism.

Chapter 1.2 Investigating the causal link

Taken together, existing studies suggest an association between periodontal diseases and poor cognitive outcomes, including dementia. Several hypotheses regarding underlying mechanisms have been proposed, including the life-course hypothesis, which posits that poor cognitive outcomes contribute to the development of periodontal diseases, the shared inflammatory hypothesis, and other mechanistic hypotheses suggesting that periodontal diseases contribute to poor cognitive outcomes. It is also possible that no causal link exists. However, direct evidence (e.g., biomedical studies, randomised controlled trials, or clinical trials) for each hypothesis, and to directly address causal questions, remains scarce. Whether periodontal diseases lead to poor cognitive outcomes, particularly dementia, is a question that remains unanswered. In this thesis, I conducted a Mendelian Randomisation analysis to investigate the causal relationship, as it offers the advantages of time efficiency and lower cost compared to other research methods for addressing causal questions.

1.2.1 Mendelian Randomisation

Mendelian Randomisation is a method which use genetic variants as an instrument to assess the causal relationship. It is generally based on the statistical methodology called "instrumental variable analysis" by using genetic variants (e.g., single-nucleotide polymorphisms) fulfilling three key assumptions as genetic instrument of exposure of interest (Sanderson et al., 2022). The three essential assumptions are:

(1) Assumption 1: relevance (the genetic instrument should be strongly associated with the exposure of interest);

(2) Assumption 2: exchangeability (the genetic instrument should not be associated with other confounding factors);

(3) Assumption 3: exclusion restriction (the genetic instrument should only affect the outcome of interest via exposure of interest).

Similar to the randomised control trial, MR also considers participants as being "randomised" but based on the natural, random assortment of genes during meiosis proposed in Mendel's second

law to achieve the random allocation of genetic variants in a population (Emdin et al., 2017; Smith and Ebrahim, 2003) instead of randomly allocating participants into treatment and placebo group in a randomised control trial. Since the genetic variants' "randomisation" occur before birth during gamete formation and conception, investigating the association using MR is not susceptible to a reverse causation or confounding factors and is considered as the same level of evidence as randomised control trial and systematic review (Davies et al., 2018; Smith and Ebrahim, 2003) if the three assumptions are justified.

Given the difficulties, the time and the high cost of implementing a large-scale randomised control trials or biomedical research exploring the causal mechanisms, MR is considered as a good alternative because it allows researchers to use existing databases containing genetic information of participants, exposure and outcome of interest leading to lower cost and are more time efficient in the situations where an instrument can be implemented. Since the establishment of many large genetic data resources (e.g., NHANEs, UK Biobank, FinnGen consortium, Gene-lifestyle interactions in dental endpoints [GLIDE] etc.) and increasing number of genetic association studies (e.g., candidate gene study, genome-wide association study) available in the literature, MR has been increasingly used in recent years to examine the casual relationship for a wide range of diseases (e.g., cancer, cardiovascular disease), traits and phenotypes. However, MR studies investigating the causal link between periodontal diseases and dementia are limited leaving the question of this potential causal link unanswered.

Chapter 1.3 Research approach

1.3.1 Research question

Is there an association between periodontal diseases and poorer cognitive function? Especially, is periodontitis a causal risk factor for dementia?

1.3.2 Aims and objectives

The overall aim of this thesis is to investigate the causal relationship between periodontal diseases and poor cognitive outcomes including dementia using method combining traditional epidemiology approach and more advanced genetic epidemiology approach – MR. The research hypothesis is that having periodontal diseases can increase the risk of dementia. The main data resources are National health and Nutrition Examination Survey (NHANES) and UK Biobank (UKB)

The thesis delivered by achieving the following objectives:

1. Quantify the association between severe periodontal disease and cognitive function using NHANES data and multivariable regression analysis (Chapter 2)
2. Review existing literature of genome-wide association study (GWAS) for genetic risk variants of periodontitis and explore the genetic risk variants for periodontitis using UKB data (Chapter 3)
3. Investigate the relationship between periodontitis and dementia through two-sample MR analyses using UKB data (Chapter 4)

1.3.3 Outline of thesis

Prior to the application of the MR approach, it is necessary to review and assess the exiting evidence for a comprehensive understanding and evaluation on the current evidence for the association between periodontal diseases and dementia. Therefore, in Chapter2, I include a review study that systematically examines the published observational studies exploring the association between periodontal diseases and dementia. The overall findings suggest that there is high heterogeneity between studies and association results differ for a variety of reasons including gender, study region, periodontal disease severity, thus recommendations on further investigation on the association using more homogenous study design is proposed.

Chapter 2 includes a published observational study we conducted using the high quality cross-sectional NHANES 2011-2014 data which incorporate the golden standard of periodontal measurements and reliable cognitive function tests to further validate the association. Previous reviews have noted that smoking and nutrition were not well-controlled in earlier observational studies (Larvin et al., 2023), therefore, we made attempts to use structural equation modelling and adjust for various potential confounders. The results suggest a significant association between periodontal diseases and cognitive function in participants aged 60 or older, and future studies should aim to explore the causal relationship between periodontal diseases and dementia.

This leads to chapter 3, investigating the genetic instruments for periodontitis. Valid genetic instruments of the exposure (i.e., periodontitis) is an important step prior to the MR. Although various types of genetic association analyses are available, in this thesis, I chose to use genome-wide association study (GWAS) to select our genetic instruments, due to its advantages in comprehensively scanning the genome for genetic-phenotype associations, high reproducibility of findings, and reduced bias compared to hypothesis-driven candidate gene studies (Duncan et al., 2019). The first study published in Chapter 3 is a systematic review of periodontitis GWAS, aimed at understanding the genetic risk variants of periodontitis identified by GWAS. So far, there is no systematic review of the genetic risk variants identified by Genome-wide association study (GWAS) for periodontitis, despite some other types of reviews on the genetic aspects of

periodontal diseases (Loos et al., 2020; Masumoto et al., 2019; Shaddox et al., 2021) and one systematic review on the heritability of periodontitis (Nibali et al., 2019). This study also systematically evaluated the current issues in the field of periodontitis GWAS. Among the included studies, the review noted that no common genetic risk variants have been identified across the GWAS which possibly due to small sample size in many periodontitis GWAS, variability in participants ethnicities which implicates a different genetic structure, heterogeneity in the periodontitis measurements and definitions. Since the summary statistics of GWAS on periodontitis could not be obtained from other investigators, a meta-analysis combining all previous studies to give a more robust result was not possible. It is highly recommended for future GWAS studies to use a consistent periodontal diseases definition and obtain standardised clinical measurements of periodontal diseases plus share the GWAS summary statistics to enable further downstream analysis. The summary of genome-wide significant genetic risk variants of periodontitis from this systematic review is used as one set of potential genetic instruments in the next step of MR. The second manuscript in Chapter 3 is an original study using the genome-wide Association approach on UK Biobank to assess the genetic variation of periodontitis. From this GWAS using UKB, I found four significant genetic risk variants for periodontitis that had not been identified before. This study serves another potential set of genetic instrument for the next MR analysis.

Chapter 4 presents manuscript which use the mendelian randomisation analysis to explore the causal relationship between periodontitis and all-cause dementia using UKB data. Because of the risk of overfitting in one sample MR, I chose to perform two sample MR by splitting UKB European sample into two independent groups: a periodontal disease group (contains only periodontitis patient and healthy controls) and a dementia group (containing only dementia patients and controls that are independent from the periodontal disease group to avoid sample overlap. This method ensured both groups has sufficient sample size and eliminate bias induced from data harmonisation process by using individual level of data. The results of this study did not show significant causal relationship between periodontitis and dementia.

Chapter 1.4 Reference

ADI AsDI. 2022. World alzheimer report 2022:Life after diagnosis: Navigating treatment, care and support. https://www.alzint.org/resource/world-alzheimer-report-2022/.

Armitage GC, Cullinan MP, Seymour GJ. 2010. Comparative biology of chronic and aggressive periodontitis: Introduction. Periodontology 2000. 53(1):7-11. doi: 10.1111/j.1600-0757.2010.00359.x.

Arvanitakis Z, Shah RC, Bennett DA. 2019. Diagnosis and management of dementia: Review. JAMA : the journal of the American Medical Association. 322(16):1589-1599. doi: 10.1001/jama.2019.4782.

Baker E, Leonenko G, Schmidt KM, Hill M, Myers AJ, Shoai M, de Rojas I, Tesi N, Holstege H, van der Flier WM et al. 2023. What does heritability of alzheimer's disease represent? PloS one. 18(4):e0281440. doi: 10.1371/journal.pone.0281440.

Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. 2015. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: A population-based perspective. Alzheimer's & Dementia. 11(6):718-726. doi.

Bettcher BM, Tansey MG, Dorothée G, Heneka MT. 2021. Peripheral and central immune system crosstalk in alzheimer disease — a research prospectus. Nature reviews Neurology. 17(11):689-701. doi: 10.1038/s41582-021-00549-x.

Beydoun MA, Beydoun HA, Wang Y. 2008. Obesity and central obesity as risk factors for incident dementia and its subtypes: A systematic review and meta-analysis. Obesity reviews. 9(3):204-218. doi: 10.1111/j.1467-789X.2008.00473.x.

Bialystok E, Craik FIM. 2006. Lifespan cognition : Mechanisms of change. Oxford: Oxford University Press.

Caetano AJ, D'Agostino EM, Sharpe P, Nibali L. 2022. Expression of periodontitis susceptibility genes in human gingiva using single-cell rna sequencing. Journal of periodontal research. 57(6):1210-1218. doi: 10.1111/jre.13057.

Cardoso EM, Reis C, Manzanares-Céspedes MC. 2018. Chronic periodontitis, inflammatory cytokines, and interrelationship with other chronic diseases. Postgraduate medicine. 130(1):98-104. doi: 10.1080/00325481.2018.1396876.

Choi S, Kim K, Chang J, Kim SM, Kim SJ, Cho HJ, Park SM. 2019. Association of chronic periodontitis on alzheimer's disease or vascular dementia. Journal of the American Geriatrics Society (JAGS). 67(6):1234-1239. doi: 10.1111/jgs.15828.

D'Aiuto F, Orlandi M, Gunsolley JC. 2013. Evidence that periodontal treatment improves biomarkers and cvd outcomes. Journal of clinical periodontology. 40(s14):S85-S105. doi: 10.1111/jcpe.12061.

Darweesh SKL, Wolters FJ, Ikram MA, de Wolf F, Bos D, Hofman A. 2018. Inflammatory markers and the risk of dementia and alzheimer's disease: A meta-analysis. Alzheimer's & dementia. 14(11):1450-1459. doi: 10.1016/j.jalz.2018.02.014.

Davies NM, Holmes MV, Davey Smith G. 2018. Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. BMJ (Online). 362:k601-k601. doi: 10.1136/bmj.k601.

Deary IJ, Corley J, Gow AJ, Harris SE, Houlihan LM, Marioni RE, Penke L, Rafnsson SB, Starr JM. 2009. Age-associated cognitive decline. British medical bulletin. 92(1):135-152. doi: 10.1093/bmb/ldp033.

Demmer RT, Norby FL, Lakshminarayan K, Walker KA, Pankow JS, Folsom AR, Mosley T, Beck J, Lutsey PL. 2020. Periodontal disease and incident dementia: The atherosclerosis risk in communities study (aric). Neurology. 95(12):e1660-e1671. doi: 10.1212/WNL.0000000000010312.

Guidance chapter 5: Periodontal diseases. 2021. UK.GOV; [accessed 2024 16/02]. https://www.gov.uk/government/publications/delivering-better-oral-health-an-evidence-based-toolkit-for-prevention/chapter-5-periodontal-diseases.

Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, Nguyen M, Haditsch U, Raha D, Griffin C et al. 2019. Porphyromonas gingivalis in alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. Science advances. 5(1):eaau3333-eaau3333. doi: 10.1126/sciadv.aau3333.

Duncan LE, Ostacher M, Ballon J. 2019. How genome-wide association studies (gwas) made traditional candidate gene studies obsolete. Neuropsychopharmacology (New York, NY). 44(9):1518-1523. doi: 10.1038/s41386-019-0389-5.

Emami E, de Souza RF, Kabawat M, Feine JS. 2013. The impact of edentulism on oral and general health. International Journal of Dentistry. 2013:498305-498307. doi: 10.1155/2013/498305.

Emdin CA, Khera AV, Kathiresan S. 2017. Mendelian randomization. JAMA : the journal of the American Medical Association. 318(19):1925-1926. doi: 10.1001/jama.2017.17219.

Farrugia C, Stafford GP, Potempa J, Wilkinson RN, Chen Y, Murdoch C, Widziolek M. 2021. Mechanisms of vascular damage by systemic dissemination of the oral pathogen porphyromonas gingivalis. The FEBS journal. 288(5):1479-1495. doi: 10.1111/febs.15486.

Gao C, Kang J. 2024. Oral disease and its association with cognitive decline and dementia. In: Deng D, Dame-Teixeira N, Do T, editors. Oral microbiome and the host: Symbiosis, dysbiosis and microbiome interventions for maintaining oral and systemic health. Springer. (under editing).

Ge F, Zhao Y, Zheng J, Xiang Q, Luo P, Zhu L, He H. 2024. Discovering common pathogenetic processes between periodontitis and alzheimer's disease by bioinformatics and system biology approach. BMC oral health. 24(1):1074-1013. doi: 10.1186/s12903-024-04775-9.

Genco RJ, Borgnakke WS. 2013. Risk factors for periodontal disease. Periodontology 2000. 62(1):59-94. doi: 10.1111/j.1600-0757.2012.00457.x.

Genco RJ, Graziani F, Hasturk H. 2020. Effects of periodontal disease on glycemic control, complications, and incidence of diabetes mellitus. Periodontology 2000. 83(1):59-65. doi: 10.1111/prd.12271.

Ghezzi EM, Ship JA. 2000. Dementia and oral health. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology. 89(1):2-5. doi: https://doi.org/10.1016/S1079-2104(00)80003-7.

H. Bokhari SA, Khan AA, Butt AK, Azhar M, Hanif M, Izhar M, Tatakis DN. 2012. Non-surgical periodontal therapy reduces coronary heart disease risk markers: A randomized controlled trial. Journal of clinical periodontology. 39(11):1065-1074. doi: 10.1111/j.1600-051X.2012.01942.x.

Hajishengallis G, Chavakis T. 2021. Local and systemic mechanisms linking periodontal disease and inflammatory comorbidities. Nature reviews Immunology. 21(7):426-440. doi: 10.1038/s41577-020-00488-6.

Harvey PD. 2019. Domains of cognition and their assessment. Dialogues in clinical neuroscience. 21(3):227-237. doi.

Hendriks S, Peetoom K, Bakker C, van der Flier WM, Papma JM, Koopmans R, Verhey FRJ, de Vugt M, Koehler S. 2021. Global prevalence of young-onset dementia a systematic review and meta-analysis. JAMA neurology. 78(9):1080-1090. doi: 10.1001/jamaneurol.2021.2161.

Heneka MTP, Carson MJP, Khoury JEP, Landreth GEP, Brosseron FP, Feinstein DLP, Jacobs AHP, Wyss-Coray TP, Vitorica JP, Ransohoff RMP et al. 2015. Neuroinflammation in alzheimer's disease. Lancet neurology. 14(4):388-405. doi: 10.1016/S1474-4422(15)70016-5.

Heppner FL, Ransohoff RM, Becher B. 2015. Immune attack: The role of inflammation in alzheimer disease. Nature reviews Neuroscience. 16(6):358-372. doi: 10.1038/nrn3880.

Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, Sealock J, Karlsson IK, Hägg S, Athanasiu L et al. 2019. Genome-wide meta-analysis identifies new loci and functional pathways influencing alzheimer’s disease risk. Nature genetics. 51(3):404-413. doi: 10.1038/s41588-018-0311-9.

Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ. 2008. Inflammation and alzheimer's disease: Possible role of periodontal diseases. Alzheimer's & dementia. 4(4):242-250. doi: 10.1016/j.jalz.2007.08.004.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

Kang J, Wu B, Bunce D, Ide M, Pavitt S, Wu J. 2019. Cognitive function and oral health among ageing adults. Community dentistry and oral epidemiology. 47(3):259-266. doi: 10.1111/cdoe.12452.

Karch CM, Goate AM. 2015. Alzheimer’s disease risk genes and mechanisms of disease pathogenesis. Biological psychiatry (1969). 77(1):43-51. doi: 10.1016/j.biopsych.2014.05.006.

Karimbux NY, Saraiya VM, Elangovan S, Allareddy V, Kinnunen T, Kornman KS, Duff GW. 2012. Interleukin-1 gene polymorphisms and chronic periodontitis in adult whites: A systematic review and meta-analysis. Journal of periodontology (1970). 83(11):1407-1419. doi: 10.1902/jop.2012.110655.

Kato T, Usami T, Noda Y, Hasegawa M, Ueda M, Nabeshima T. 1997. The effect of the loss of molar teeth on spatial memory and acetylcholine release from the parietal cortex in aged rats. Behavioural brain research. 83(1-2):239-242. doi: 10.1016/S0166-4328(97)86078-0.

Kinane DF, Stathopoulou PG, Papapanou PN. 2017. Periodontal diseases. Nature reviews Disease primers. 3(1):17038-17038. doi: 10.1038/nrdp.2017.38.

Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT. 2018. Inflammation as a central mechanism in alzheimer's disease. Alzheimer's & dementia : translational research & clinical interventions. 4(1):575-590. doi: 10.1016/j.trci.2018.06.014.

Kondo K, Niino M, Shido K. 1994. A case-control study of alzheimer's disease in japan – significance of life-styles. Dementia and geriatric cognitive disorders. 5(6):314-326. doi: 10.1159/000106741.

Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, Boland A, Vronskaya M, van der Lee SJ, Amlie-Wolf A et al. 2019. Genetic meta-analysis of diagnosed alzheimer's disease identifies new risk loci and implicates aβ, tau, immunity and lipid processing. Nature genetics. 51(3):414-430. doi.

Kurgan S, Kantarci A. 2018. Molecular basis for immunohistochemical and inflammatory changes during progression of gingivitis to periodontitis. Periodontology 2000. 76(1):51-67. doi: 10.1111/prd.12146.

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2023. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. Age and ageing. 52(2). doi: 10.1093/ageing/afad015.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021. Risk of incident cardiovascular disease in people with periodontal disease: A systematic review and meta-analysis. Clinical and experimental dental research. 7(1):109-122. doi: 10.1002/cre2.336.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Frontiers in medicine. 7:604980-604980. doi: 10.3389/fmed.2020.604980.

Liu X, Li H. 2022. A systematic review and meta-analysis on multiple cytokine gene polymorphisms in the pathogenesis of periodontitis. Frontiers in immunology. 12:713198-713198. doi: 10.3389/fimmu.2021.713198.

Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, Brayne C, Burns A, Cohen-Mansfield J, Cooper C et al. 2020. Dementia prevention, intervention, and care: 2020 report of the lancet commission. The Lancet (British edition). 396(10248):413-446. doi: 10.1016/S0140-6736(20)30367-6.

Lönn J, Ljunggren S, Klarström-Engström K, Demirel I, Bengtsson T, Karlsson H. 2018. Lipoprotein modifications by gingipains of porphyromonas gingivalis. Journal of periodontal research. 53(3):403-413. doi: 10.1111/jre.12527.

Loos, G. B, Van Dyke, E. T. 2020. The role of inflammation and genetics in periodontal disease. Periodontol 2000. 83(1):26-39. doi: 10.1111/prd.12297.

Loos BG, Van Dyke TE. 2020. The role of inflammation and genetics in periodontal disease. Periodontology 2000. 83(1):26-39. doi: 10.1111/prd.12297.

Martínez-García M, Hernández-Lemus E. 2021. Periodontal inflammation and systemic diseases: An overview. Frontiers in physiology. 12:709438-709438. doi: 10.3389/fphys.2021.709438.

Masumoto R, Kitagaki J, Fujihara C, Matsumoto M, Miyauchi S, Asano Y, Imai A, Kobayashi K, Nakaya A, Yamashita M et al. 2019. Identification of genetic risk factors of aggressive periodontitis using genomewide association studies in association with those of chronic periodontitis. Journal of periodontal research. 54(3):199-206. doi: 10.1111/jre.12620.

Moriarty O, McGuire BE, Finn DP. 2011. The effect of pain on cognitive function: A review of clinical and preclinical research. Progress in neurobiology. 93(3):385-404. doi: 10.1016/j.pneurobio.2011.01.002.

Muñoz SS, Garner B, Ooi L. 2019. Understanding the role of apoe fragments in alzheimer’s disease. Neurochemical research. 44(6):1297-1305. doi: 10.1007/s11064-018-2629-1.

Murman DL. 2015. The impact of age on cognition. Seminars in hearing. 36(3):111-121. doi: 10.1055/s-0035-1555115.

Nadim R, Tang J, Dilmohamed A, Yuan S, Wu C, Bakre AT, Partridge M, Ni J, Copeland JR, Anstey KJ, Chen R. 2020. Influence of periodontal disease on risk of dementia: A systematic literature review and a meta-analysis. doi.

Nascimento GG, Leite FRM, Conceição DA, Ferrúa CP, Singh A, Demarco FF. 2016. Is there a relationship between obesity and tooth loss and edentulism? A systematic review and meta-analysis. Obesity reviews. 17(7):587-598. doi: 10.1111/obr.12418.

Nibali L, Bayliss-Chapman J, Almofareh SA, Zhou Y, Divaris K, Vieira AR. 2019. What is the heritability of periodontitis? A systematic review. Journal of Dental Research. 98(6):632-641. doi: 10.1177/0022034519842510.

Nichols E, Steinmetz JD, Vollset SE, Abd-Allah F, Abdoli A, Abu-Gharbieh E, Alipour V, Almustanyir S, Amu H, Arabloo J et al. 2022. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019. The Lancet Public health. 7(2):e105-e125. doi: 10.1016/S2468-2667(21)00249-8.

Niessen LC, Jones JA. 1987. Professional dental care for patients with dementia. Gerodontology. 6(2):67-71. doi: 10.1111/j.1741-2358.1987.tb00391.x.

Park HL, O'Connell JE, Thomson RG. 2003. A systematic review of cognitive decline in the general elderly population. International journal of geriatric psychiatry. 18(12):1121-1134. doi: 10.1002/gps.1023.

Pihlstrom BL, Michalowicz BS, Johnson NW. 2005. Periodontal diseases. Lancet. 366(9499):1809-1820. doi: 10.1016/S0140-6736(05)67728-8.

Preshaw PM, Alba AL, Herrera D, Jepsen S, Konstantinidis A, Makrilakis K, Taylor R. 2011. Periodontitis and diabetes: A two-way relationship. Diabetologia. 55(1):21-31. doi: 10.1007/s00125-011-2342-y.

Prince M, Ali GC, Guerchet M, Prina AM, Albanese E, Wu YT. 2016. Recent global trends in the prevalence and incidence of dementia, and survival with dementia. Alzheimer's research & therapy. 8(1):23-23. doi: 10.1186/s13195-016-0188-8.

Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. 2013. The global prevalence of dementia: A systematic review and metaanalysis. Alzheimer's & dementia. 9(1):63-75.e62. doi: 10.1016/j.jalz.2012.11.007.

Rainero I, Bo M, Ferrero M, Valfrè W, Vaula G, Pinessi L. 2004. Association between the interleukin-1alpha gene and alzheimer's disease: A meta-analysis. Neurobiology of aging. 25(10):1293-1298. doi.

Reynolds CA, Finkel D. 2015. A meta-analysis of heritability of cognitive aging: Minding the “missing heritability” gap. Neuropsychology Review. 25(1):97-112. doi.

Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, Palmer T, Schooling CM, Wallace C, Zhao Q, Davey Smith G. 2022. Mendelian randomization. Nature Reviews Methods Primers. 2(1):6. doi: 10.1038/s43586-021-00092-5.

Schenkein HA, Papapanou PN, Genco R, Sanz M. 2020. Mechanisms underlying the association between periodontitis and atherosclerotic disease. Periodontology 2000. 83(1):90-106. doi: 10.1111/prd.12304.

Sekula P, Del Greco M F, Pattaro C, Köttgen A. 2016. Mendelian randomization as an approach to assess causality using observational data. Journal of the American Society of Nephrology. 27(11):3253-3265. doi: 10.1681/ASN.2016010098.

Shaddox LM, Morford LA, Nibali L. 2021. Periodontal health and disease: The contribution of genetics. Periodontology 2000. 85(1):161-181. doi: 10.1111/prd.12357.

Shimazaki Y, Soh, Saito T, Yamashita Y, Koga T, Miyazaki H, Takehara T. 2001. Influence of dentition status on physical disability mental impairment, and mortality in institutionalized elderly people. Journal of dental research. 80(1):340-345. doi: 10.1177/00220345010800010801.

Smith GD, Ebrahim S. 2003. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? International journal of epidemiology. 32(1):1-22. doi: 10.1093/ije/dyg070.

Stabholz A, Soskolne WA, Shapira L. 2010. Genetic and environmental risk factors for chronic periodontitis and aggressive periodontitis. Periodontology 2000. 53(1):138-153. doi: 10.1111/j.1600-0757.2010.00340.x.

Stein PS, Scheff S, Dawson D. 2006. Alzheimer's disease and periodontal disease: Mechanisms underlying a potential bidirectional relationship. Grand Rounds in Oral Systemic Medicine. 3(1):14-24. doi.

Thomson WM, Barak Y. 2021. Tooth loss and dementia: A critical examination. Journal of Dental Research. 100(3):226-231. doi: 10.1177/0022034520957233.

Tonetti MS, Greenwell H, Kornman KS. 2018. Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. Journal of clinical periodontology. 45(S20):S149-S161. doi: 10.1111/jcpe.12945.

Tonsekar PP, Jiang SS, Yue G. 2017. Periodontal disease, tooth loss and dementia: Is there a link? A systematic review. Gerodontology. 34(2):151-163. doi: 10.1111/ger.12261.

Van der Velden U. 2017. What exactly distinguishes aggressive from chronic periodontitis: Is it mainly a difference in the degree of bacterial invasiveness? Periodontology 2000. 75(1):24-44. doi: 10.1111/prd.12202.

Varatharaj A, Galea I. 2017. The blood-brain barrier in systemic inflammation. Brain, behavior, and immunity. 60:1-12. doi: 10.1016/j.bbi.2016.03.010.

Walker KA, Ficek BN, Westbrook R. 2019. Understanding the role of systemic inflammation in alzheimer's disease. ACS chemical neuroscience. 10(8):3340-3342. doi: 10.1021/acschemneuro.9b00333.

Wang Z-X, Wan Q, Xing A. 2020. Hla in alzheimer's disease: Genetic association and possible pathogenic roles. NeuroMolecular Medicine. 22(4):464-473. doi.

Dementia. 2023a. World Heath Organisation; [accessed 2024 2/14]. https://www.who.int/news-room/fact-sheets/detail/dementia.

Oral health. 2023b. https://www.who.int/news-room/fact-sheets/detail/oral-health: World Health Organization; [accessed 2024 5th February].

Winning L, Naseer A, De Looze C, Knight SP, Kenny RA, O'Connell B. 2022. Tooth loss and cognitive decline in community dwelling older irish adults: A cross-sectional cohort study. Journal of dentistry. 119:104077-104077. doi: 10.1016/j.jdent.2022.104077.

Wu B, Fillenbaum GG, Plassman BL, Guo L. 2016. Association between oral health and cognitive status: A systematic review. Journal of the American Geriatrics Society (JAGS). 64(8):1752-1752. doi: 10.1111/jgs.14572.

Yu W, Yu W, Liu X, Wan T, Chen C, Xiong L, Zhang W, Lü Y. 2021. Associations between malnutrition and cognitive impairment in an elderly chinese population: An analysis based on a 7-year database. Psychogeriatrics. 21(1):80-88. doi: 10.1111/psyg.12631.

Zelig R, Goldstein S, Touger-Decker R, Firestone E, Golden A, Johnson Z, Kaseta A, Sackey J, Tomesko J, Parrott JS. 2022. Tooth loss and nutritional status in older adults: A systematic review and meta-analysis. JDR Clinical & Translational Research. 7(1):4-15. doi: 10.1177/2380084420981016.

Zhang W, Xiao D, Mao Q, Xia H. 2023. Role of neuroinflammation in neurodegeneration development. Signal transduction and targeted therapy. 8(1):267-267. doi: 10.1038/s41392-023-01486-5.

Chapter 2: Quantifying the association between periodontitis and poorer cognitive outcomes

Cross-sectional study

Title: Oral diseases are associated with cognitive function in adults over 60 years old

Published in *Oral Diseases*, 30, 3480–3488.
https://doi.org/10.1111/odi.14757

Authors:
Chenyi Gao[1] , Harriet Larvin[2], David Timothy Bishop[3], David Bunce[4], Susan Pavitt[1], Jianhua Wu[1, 2,5], Jing Kang[1]*.
[1] School of Dentistry, University of Leeds
[2] Leeds Institute for Data Analytics, University of Leeds
[3] Leeds Institute of Medical Research, School of Medicine, University of Leeds
[4] School of Psychology, University of Leeds
[5] Centre of Primary Care, Wolfson Institute of Population Health, Queen Mary, University of London, UK

Metrics
Journal Impact factor : 2.9
Citation : 5

Data availability statement

This study has used data from the National Health and Nutrition Examination Survey conducted by the Centers for Disease Control and Prevention in the United States. The data is free and anyone can access it from: https://wwwn.cdc.gov/nchs/nhanes/Default.aspx

Funding statement

This project is funded by the Alzheimer's Society. Chenyi Gao is supported by the Alzheimer's Society Heather Corrie Ph.D. studentship. The views expressed are those of the author(s) and not necessarily those of the Alzheimer's Society. The Alzheimer's Society did not involve in design and conduct this study, as well as data collection or preparation, review, approval, and decision to submit the manuscript.

Conflict of interest disclosure

There is no conflict of interest for all authors.

Ethics approval and participant consent statement

Our study performed secondary data analyses using NHANES 2011-14. The ethnic approval for NHANES was conducted by NCHS Ethics Review Board (ERB) priory to the data collection: https://www.cdc.gov/nchs/nhanes/irba98.htm. Participants consents were obtained by NHANES and the related documentation can be found via the link https://www.cdc.gov/nchs/nhanes/genetics/genetic_participants.htm

ABSTRACT

Objective: To investigate the bi-directional association between oral diseases and cognitive function comprehensively.

Subjects and Methods: This cross-sectional study utilized data from the National Health and Nutrition Examination Survey. Oral diseases include periodontitis, dental caries, and tooth loss (end point of oral disease resulting in tooth extraction). Cognitive function included three domains: memory, processing speed, and executive function. A global cognitive score was then derived from sum of the three cognitive domains. Oral-cognition associations were examined using various statistical models: 1) Regress oral disease on cognitive function; 2) Regress cognitive function on oral disease; 3) Structural equation modelling treating cognition and oral disease as latent variables.

Results: There were 2508 participants aged 60+ who had both oral and cognitive information. Associations between various oral disease and global cognitive score were observed (Odds ratio $OR_{cog->periodontitis}$ 0.95, 95% Confidence Interval [0.92, 0.99]; $\beta_{cog->caries}$ -0.13, [-0.23, -0.04]; $\beta_{cog->tooth\ loss}$ -0.03 [-0.04, -0.01]; $\beta_{tooth\ loss->cog}$ -0.04 [-0.06,-0.02]; $\beta_{caries->cog}$ -0.03 [-0.06,-0.01]; $\beta_{periodontitis->cog}$ -0.39 [-0.69,-0.10]). Significant correlation was also found between these oral disease and cognitive function using structural equation model (r -0.22, [-0.34, -0.10]).

Conclusions: This study found robust bi-directional associations between oral disease and cognitive function using various modelling approaches among the ageing population.

INTRODUCTION

Oral diseases include dental caries, periodontitis, and tooth loss (which is the result of oral disease that leads to tooth extraction). In recent decades, some evidence has emerged supporting the association between oral disease and cognitive decline/dementia (Kang et al., 2020; Kang et al., 2019b). It has been suggested that the association is possibly due to shared inflammation pathways between oral diseases and cognitive decline (Kamer et al., 2008; Noble et al., 2013), where the inflammation might be caused by chronic oral bacterial infection and dysbiosis (Dominy et al., 2019; Singhrao et al., 2014). On the other hand, cognitive decline could influence oral hygiene behaviours, therefore, reciprocal direction of associations may be observed. However, few studies have explored both directions of such associations, as shown in several systematic reviews (Lin, 2018; Oh et al., 2018; Shen et al., 2016; Tonsekar et al., 2017). Mixed results and conclusions from literature might be due to inaccurate measures in oral health or cognitive functions, study design, population differences, methodological variation, and sample selection issues (e.g., sample size and bias) (Larvin et al., 2023; Wu et al., 2016). Therefore, more high-quality original studies using well-designed and clinically examined oral and cognitive data, with detailed information on participants' demographic characteristics, lifestyle, and medical history, are valuable contributing to the knowledge development in this area.

While biological/clinical experiments, or long-term followed up life course studies, would be the ideal ones for assessing the association between oral disease and cognitive function, these studies are usually costly, time consuming and sometimes unethical to conduct. Hence, large-scale epidemiological studies are the alternatives, which are more feasible, cheaper, practical, and efficient to carry out. In this cross-sectional study, we used data from the National Health and Nutrition Examination Survey (NHANES), a nationally representative health survey in the United States, to comprehensively investigate the association between various oral diseases and various cognitive function domains.

METHODS

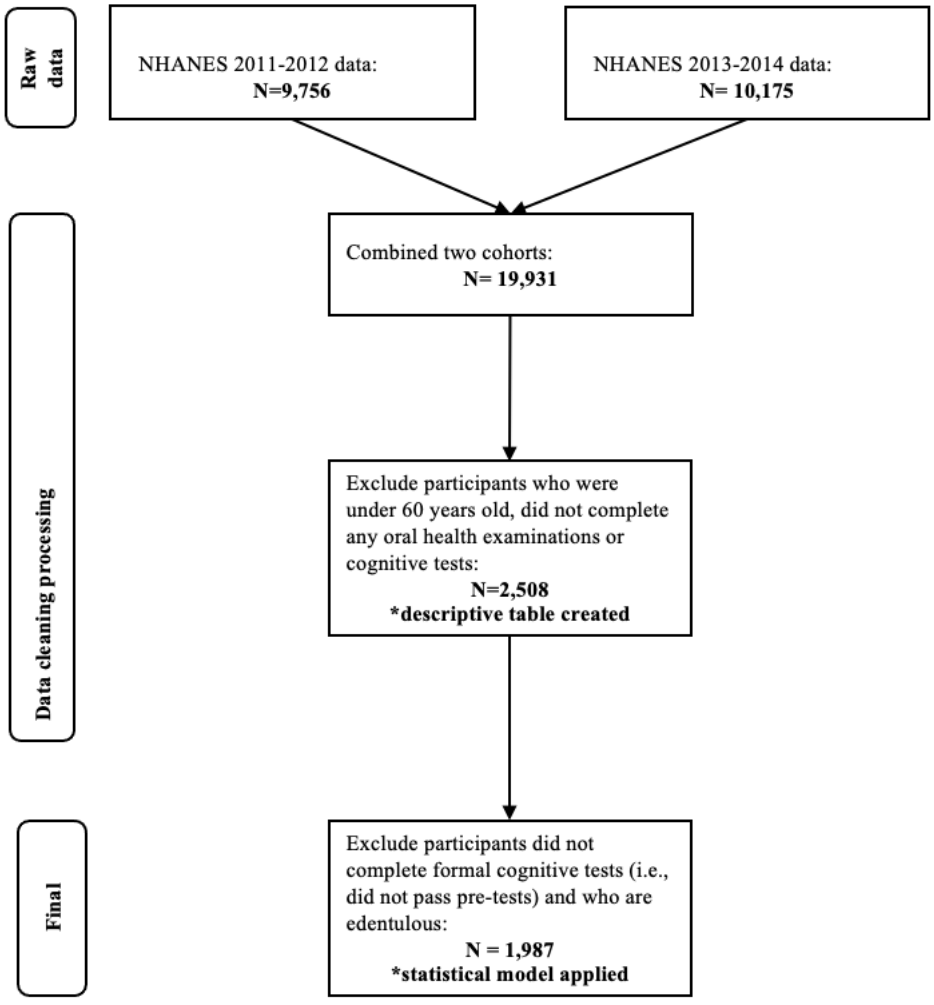
Study design

The study followed the STROBE guidelines (Von Elm et al., 2007). NHANES is a biennial national survey designed to examine and collect the information on both health and nutritional status across the non-institutionalized United States population using a stratified, multistage, probability sampling design. Data has been collected using various methods including questionnaires, interviews, and physical examinations ((NCHS), 2022). The methods and design for the survey are available via the link: https://www.cdc.gov/nchs/nhanes/index.htm

Study participants

Participants aged 60 or over from NHANES 2011-12 and 2013-14 were included in the analysis as these two cycles contain detailed dental clinical examination (6 sites per tooth, 28 permanent teeth without wisdom teeth) and participants aged 60 or over received cognitive tests in three different domains (memory, processing speed, and executive function). Participants under 60 years old, or with incomplete cognitive tests or oral examination were excluded from this study. The flow chart of participants selection is presented in Figure 1.

Figure 1. Flow chart of sample selection from raw data in NHANES (2011-2014) to the final sample



N, number of participants

Participants' consents were obtained by NHANES and the related documentation can be found via the link (https://www.cdc.gov/nchs/nhanes/genetics/genetic_participants.htm).

Cognitive function

Cognitive function include memory measured by The Consortium to Establish a Registry for Alzheimer's Disease Word Learning subtest (CERAD W-L, raw score range 0-40) (Morris et al., 1989), processing speed measured by the Digit Symbol Substitution test (DSST, raw score range 0-105) (Bienias et al., 2003; Plassman et al., 2007; Proust-Lima et al., 2006), and executive function measured by the animal fluency test (raw score range 0-40) (Strauss et al., 2006). Participants were required to pass a pre-test before the formal tests for the processing speed and executive function. The scores obtained for each cognitive domain were then all scaled to 0-10. Then, the global cognitive function (GCF) score (0-30) is derived based on the sum of scores form all cognitive domains with equal weight. The detailed methodology was documented in NHANES (2011/2012: https://wwwn.cdc.gov/Nchs/Nhanes/2011-2012/CFQ_G.htm; 2013/2014: https://wwwn.cdc.gov/Nchs/Nhanes/2013-2014/CFQ_H.htm).

Oral diseases

Oral diseases were reflected by periodontitis, dental caries, and tooth loss. Periodontitis was clinically assessed and classified as 'none, mild, moderate, severe' by using a standard case definition for surveillance of periodontitis (Eke et al., 2012; Holtfreter et al., 2015). Dental caries is reported as the number of decayed, missing and filled teeth (DMFT, scale 0-28). Tooth loss is measured by the number of remaining teeth count between 0 (edentulous) and 28 (all teeth present without wisdom teeth). The detailed description of oral disease and classification can be found elsewhere (Kang et al., 2019a).

Covariates

Based on previous studies, we included 'common risk factors' for oral disease and cognitive decline as covariates (Hong et al., 2018; Kang et al., 2022; Kang et al., 2020; Kang et al., 2019b; Larvin et al., 2023; Larvin et al., 2021a; 2021b; 2022a; Larvin et al., 2022b; Larvin et al., 2021c; Larvin et al., 2020): demographic variables (age, sex, ethnicity, education qualification, marital status, and poverty index ratio), anthropometric measures (body mass index (BMI, kg/m^2), waist measurement), lifestyle factors (smoking status, cigarette number in the past 30 days, at least 12 alcohol drinks consumed in the past year, substance misuse, physical activity, intake of sugar, carbohydrate, and energy); comorbidities (cardiovascular diseases (including Congestive heart failure, Coronary heart disease, Angina, Heart attack), diabetes, liver disease, arthritis, depression and sleep disorder); dental hygiene behavior (time since last dental visit, tooth brushing frequency per day, use dental floss). Data for both "time since last dental visit" and "dental floss" was

available in NHANES 2011-2014 and toothbrush frequency data was only available in 2013-2014.

Statistical analyses

Descriptive statistics were performed presenting cognitive function in quartiles among participants' characteristics. Continuous variables were presented as mean (SD) or median (interquartile range) and categorical variables were reported as frequency (%).

Participants who did not passed cognitive function pretests (pretest for animal fluency test and DSST test) or those who were edentulous were excluded at modelling stage, because no data was collected for them in formal cognitive tests and no periodontitis presented in edentulous. If a variable had more than 50% missing data, it will be excluded in the modelling. Moreover, 184 participants with recorded diastolic blood pressure <50Hmm were considered as outlier and coded as missing data. Periodontitis was further grouped as binary (yes/no) in modelling stage for simplicity. Due to the significant amount of missing data in smoking status (n=1230) and drug use (n=1168), both was grouped into three categories (non-smoker, ever smoker, unknown), where "unknown" stand for missing response. Cigarette numbers (missing response=2169) and tooth-brushing frequency (not measured in participants aged 60+) were excluded due to insufficient data (missing response=2169).

To test the relationship between each oral disease (periodontitis, dental caries, tooth loss) and each cognitive function domain (memory, processing speed, executive function), appropriate regression models were performed when treating cognitive function and oral disease as outcomes, respectively. Null models were tested first, and gradually adjusted for demography, lifestyle, anthropometry, presence of comorbidities, and dental hygiene behaviors in the fully adjusted models.

Finally, since the cognitive function consists of multiple domains and there are also multiple oral diseases, a structural equation model (SEM) was applied to understand the overall associations by assuming two latent variables: cognition and oral disease, in which model the variables consist of the latent variables was considered with equal weight without bias. Memory, executive function, and processing speed were the three observed variables for latent variable cognition, and dental caries and periodontitis were the two observed variables for latent oral disease variable (tooth loss was highly correlated with dental caries therefore not used as another observed variable). All the covariates adjusted in the regression models were also adjusted in the SEM. The comparative fit index (CFI more than 0.9 would indicate good model fit) and root-mean-square estimate (RMSEA, less than 0.06 as an indication of a good fit) were used to examine the model fit (Bentler, 1990).

The variables which contained small proportion of missing data were handled via multiple imputation methods and coefficients were combined using Rubin's rule(Rubin, 2004). Data was also analyzed on complete cases (those with missing data) and sensitivity analysis was conducted to assess the impact of missing data. Statistical analyses were performed in R version 4.1.1 (https://cran.r-project.org/) with various packages. Significance level were set at 0.05.

RESULTS

Overall, 2508 (50.1% male; mean age 69.3, SD 6.7) participants aged over 60 fulfilled the inclusion criteria (Figure 1). Full descriptive tables are in Appendix Table a.1-3. Excluding edentulous participants and those who did not pass the cognitive pre-test, 1,987 participants remained. The average global cognitive score was 14.1 (SD 3.9) out of 30, and 49.4% participants had periodontitis (Table 1).

Table 1. Overall sample characteristics and characteristics in highest and lowest cognitive function quartiles: NHANES, 2011-2014.

	Overall	Cognitive function	
		<25%	>=75%
N	2508	605	606
Demographic			
Gender, male (%)	1257 (50.1)	352 (58.2)	235 (38.8)
Age, mean (SD)	69.3 (6.7)	71.8 (7.0)	66.3 (5.5)
Race (%)			
Mexican American	225 (9.0)	63 (10.4)	39 (6.4)
Other Hispanic	264 (10.5)	97 (16.0)	35 (5.8)
Non-Hispanic White	1149 (45.8)	200 (33.1)	393 (64.9)
Non-Hispanic Black	624 (24.9)	193 (31.9)	87 (14.4)
Other	246 (9.8)	52 (8.6)	52 (8.6)
Education, College or Above (%)	1241 (49.5)	142 (23.5)	466 (76.9)
PIR, Mean (SD)	2.6 (1.6)	1.9 (1.4)	3.4 (1.6)
Cognitive Function			
Memory, Mean (SD)	6.0 (1.8)	4.1 (1.3)	7.7 (1.0)
Executive Function, Mean (SD)	3.6 (1.5)	2.3 (1.0)	5.20 (1.3)
Processing Speed, Mean (SD)	4.4 (1.6)	2.7 (1.0)	6.2 (1.0)
Global Cognitive Function, Mean (SD)	14.1 (3.9)	9.1 (1.8)	19.1 (1.8)
Oral Health			

Missing Teeth, Mean (SD)	11.9 (10.2)	16.1 (9.9)	7.0 (8.6)
DMFT, Mean (SD)	18.4 (7.3)	20.3 (7.5)	16.0 (6.6)
Periodontitis Severity (%)			
None	1268 (50.6)	291 (48.1)	354 (58.4)
Mild-Moderate	978 (39.0)	239 (39.5)	206 (34.0)
Severe	262 (10.4)	75 (12.4)	46 (7.6)

N, number of participants; SD, standard deviation; PIR, poverty index ration, which is defined as the ratio of total family income to the US poverty level; <25%, the participants with global cognitive function scores that are the lowest 25% in the overall sample size; >=75%, participants with global cognitive function scores that are the highest 25% in the overall sample size. There were missing data in the following variables: education (<0.1%), income ratio (8.2%), processing speed (0.4%), executive function (3.1%)

For cognitive function as outcome, fully adjusted models showed that the participants with periodontitis (beta -0.39, 95%CI [-0.69, -0.10]), more missing teeth (beta -0.04, 95%CI [-0.06, -0.02]) or more caries (DMFT beta -0.03, 95%CI [-0.06, -0.01]) were likely to have lower global cognitive score. In particular, the higher number of missing teeth was associated with lower executive function (beta -0.01, 95%CI [-0.02, -0.00]) and processing speed (beta -0.02, 95%CI [-0.03, -0.01]). The higher number of DMFT is related to lower processing speed (beta -0.01, 95%CI [-0.02, -0.01]) only. Having periodontitis was associated with poorer memory (beta -0.15, 95%CI [-0.30,0.00] and processing speed (beta -0.16, 95%CI [-0.28, -0.04]) (Table 2). Tables summaries results from each stage of adjustments can be view in Appendix Table a.4.

For oral disease as outcome, three different regression models were applied based on the distribution of oral health outcomes: linear model for the number of DMFT, negative binomial model for the number of missing teeth, binomial logistic regression for periodontitis (Histogram is available in Appendix Figure a.1). In the fully adjusted models, the effect of global cognitive score on periodontitis (OR 0.95, 95%CI [0.92,0.99]), the number of missing teeth (beta -0.03, 95%CI [-0.04, -0.01]) and DMFT (beta -0.13, 95%CI [-0.23, -0.04]) was demonstrated. Periodontitis was associated with poorer memory (OR 0.93, 95%CI [0.87,1.00] and processing speed (OR 0.88, 95%CI [0.81,0.96]). A higher number of DMFT was negatively associated with processing speed (beta -0.36, 95%CI [-0.58, -0.14]); while executive function (OR -0.06, 95%CI [-0.09, -0.03]) and poorer processing speed (beta -0.07, 95%CI [-0.11, -0.04]) were strongly associated with the higher number of missing teeth (Table 3). Tables summaries results from each stage of adjustments can be view in Appendix Table a.5.

Table 2. Oral health's association with cognitive function outcomes: NHANES 2011-2014 (n = 1,987)

		Cognitive outcomes (β [Δ], 95% CI)			
	Model	**Memory**	**Executive function**	**Processing speed**	**Global cognitive score**
Missing Teeth	Unadjusted	-0.04 [-0.05,-0.03] ***	-0.05 [-0.06,-0.04] ***	-0.08 [-0.09,-0.07] ***	-0.17 [-0.19,-0.15] ***
	Adjusted[a]	-0.00 [-0.01,0.01]	-0.01 [-0.02,-0.00] **	-0.02 [-0.03,-0.01] ***	-0.04 [-0.06,-0.02] ***
DMFT	Unadjusted	-0.03 [-0.04,-0.02] ***	-0.04 [-0.03,-0.01] ***	-0.04 [-0.05,-0.02] ***	-0.09 [-0.11,-0.06] ***
	Adjusted[a]	-0.01 [-0.02,0.00]	-0.01 [-0.02,0.00]	-0.01 [-0.02,-0.01] **	-0.03 [-0.06,-0.01] **
Periodontitis	Unadjusted	-0.57 [-0.72,-0.41] ***	-0.79 [-0.53,-0.26] ***	-0.79 [-0.94,-0.65] ***	-1.75 [-2.10,-1.41] ***
	Adjusted[a]	-0.15 [-0.30,0.00] *	-0.08 [-0.22,0.05]	-0.16 [-0.28,-0.04] **	-0.39 [-0.69,-0.10] **

[Δ] Linear regression models were used to obtain coefficients: oral health factor as predictor, cognitive function as outcome. The beta represents the coefficient estimate of per tooth increase or per unit of DMFT increase or having periodontitis on each cognitive function. [a]Adjusted for demographic variables (age, sex, ethnicity, education qualification, marital status, and poverty index ratio), anthropometric measures (BMI, waist measurement), lifestyle factors (smoking status, alcohol intake, substance misuse, physical activity, intake of sugar, carbohydrate, and energy); comorbidities (cardiovascular diseases, diabetes, liver disease, arthritis, depression and sleep disorder); dental hygiene behavior (time since last dental visit, use dental floss)

N, number of participants; CI, confidence interval; DMFT, the number of missing, decayed, filled tooth.

***p<.001, **p<.01, *p<.05

Table 3. Cognitive function's association with oral health outcomes. NHANES (n = 1,987), 2011-2014.

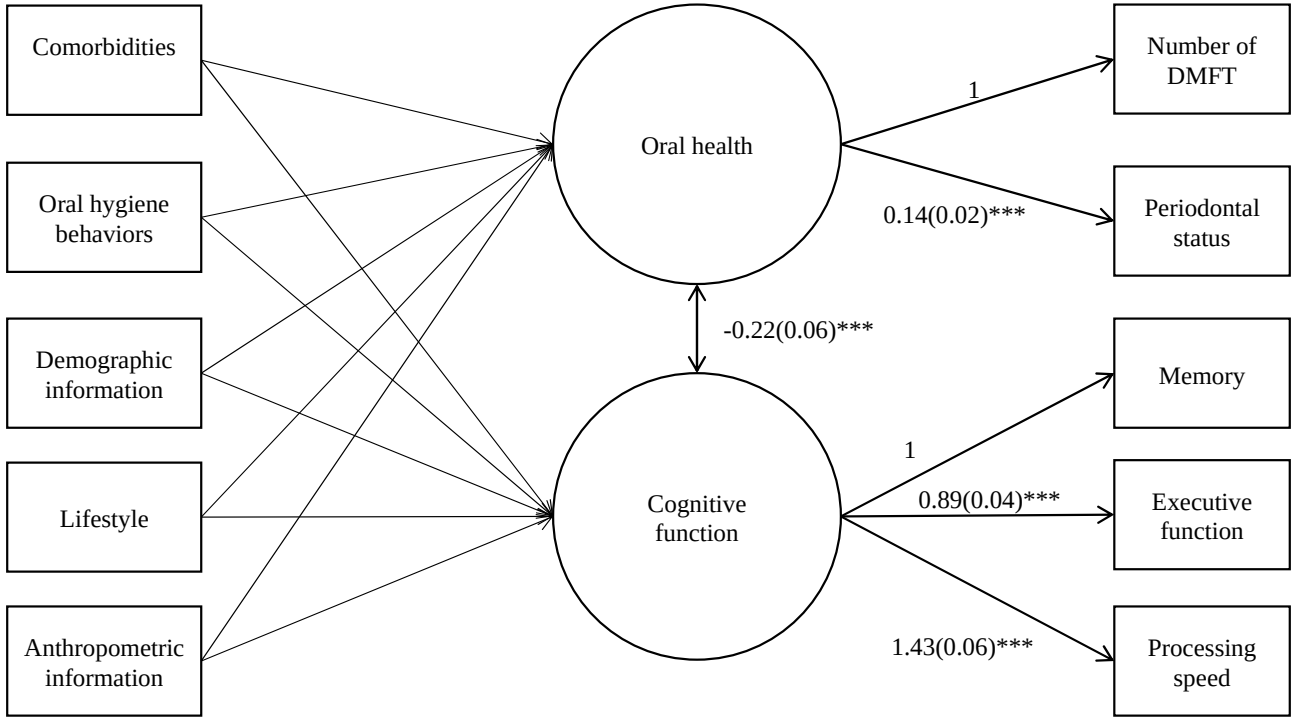
		Periodontitis	DMFT	Missing teeth
		OR, 95% CI	Beta, 95%CI	Beta, 95%CI
	Models	Logistic	Linear	Negative binomial
Memory	Unadjusted	0.82 [0.77,0.86]***	-0.42 [-0.58,-0.25]***	-0.10 [-0.12,-0.07]***
	Adjusted[a]	0.93 [0.87,1.00]*	-0.15 [-0.33,0.02]	-0.01 [-0.04,0.02]
Executive Function	Unadjusted	0.94 [0.89,1.00]*	-0.35 [-0.54,-0.16]***	-0.15 [-0.18,-0.12]***
	Adjusted[a]	0.95 [0.88,1.03]	-0.17 [-0.38,0.03]	-0.06 [-0.09,-0.03]***
Processing Speed	Unadjusted	0.91 [0.86,0.95]***	-0.54 [-0.71,-0.36]***	-0.22 [-0.25,-0.19]***
	Adjusted[a]	0.88 [0.81,0.96]**	-0.36 [-0.58,-0.14]**	-0.07 [-0.11,-0.04]***
Global cognitive score	Unadjusted	0.88 [0.86,0.91]***	-0.23 [-0.30,-0.16]***	-0.08 [-0.09,-0.07]***
	Adjusted[a]	0.95 [0.92,0.99]**	-0.13 [-0.23,-0.04]**	-0.03 [-0.04,-0.01]***

[a]Adjusted for demographic variables (age, sex, ethnicity, education qualification, marital status, and poverty index ratio), anthropometric measures (BMI, waist measurement), lifestyle factors (smoking status, alcohol intake, substance misuse, physical activity, intake of sugar, carbohydrate, and energy); comorbidities (cardiovascular diseases, diabetes, liver disease, arthritis, depression and sleep disorder); dental hygiene behavior (time since last dental visit, use dental floss)

N, number of participants; CI, confidence interval; DMFT, the number of missing, decayed, filled tooth.

***p<.001, **p<.01, *p<.05

Figure 2. Structure Equation model considering oral health and cognitive function as two latent variables.



The cycle represented the latent variable; the rectangle represented observed values. The coefficient estimate and standardised error were marked next to the solid line, which indicated the association. Covariates were summarised and represented by its category in this diagram: demographic variables (age, sex, ethnicity, education qualification, marital status, and poverty index ratio), anthropometric measures (BMI, waist measurement), lifestyle factors (smoking status, alcohol intake, substance misuse, physical activity, intake of sugar, carbohydrate, and energy); comorbidities (cardiovascular diseases, diabetes, liver disease, arthritis, depression and sleep disorder); dental hygiene behaviour (time since last dental visit, use dental floss)

*Note. ***p<0.001*

Finally, SEM showed association between the two latent variables: oral disease and cognitive function, after full adjustment (correlation r= –0.22, 95%CI [-0.34, -0.10]) (i.e., better cognitive function associated with better oral health—high value in oral disease indicate worse oral health) (Figure 2). The results showed that the data was well fitted in the model (CFI=0.99, RMSEA =0.05, SRMR=0.02).

Sensitivity analysis using complete cases was performed and results remain robust. Please see the Appendix Table a.6 -7 for detailed regression results and Appendix Figure a.2 for SEM results of sensitivity analysis.

DISCUSSION

Our study revealed robust associations between oral diseases (dental caries, periodontitis, and tooth loss) and poor global cognitive function in individuals over 60 years of age. Interestingly, we found that processing speed is associated with every oral disease, but memory or executive function is only associated with certain oral disease.

To date, hundreds of original studies have examined the association between oral diseases and cognitive function/decline. Our results is supported by some of them: previous findings showed the negative effect of poor OH on the CF (Demmer et al., 2020; Kang et al., 2019b; Winning et al., 2022), especially the bidirectional relationship over a six-year follow-up (Kang et al., 2020). The studies successfully found the association including the current study are similar on large sample size (more than 1000 participants), well-adjusted statistical model (e.g., with consideration on essential health behaviours and health status), and participants aged 45-60 or over (mid to old age). However, there are still many other previous studies failed to find such association. For example, some found no differences on dementia outcome between periodontitis group and non-periodontitis group (Holmer et al., 2022). Compare to our study, this study did not consider the effect from smoking and sample size in final follow-up and analysis despite of large sample size at beginning. The definition of periodontitis is also different where they consider deep pocket depth only but our study and some other study use both clinical attachment loss and pocket depth simultaneously. This could also lead to different conclusions. Varied study quality and mixed evidence have also been noted from a recent systematic review, where they discussed how study design factors can influence the prevalence and risk estimates of cognitive disorders in relation to oral diseases. These factors include cross-sectional vs longitudinal study designs, inconsistent definitions of oral diseases (especially periodontal disease), self-reported vs clinical assessed oral diseases/cognitive decline or dementia, varying sample sizes, confounding effects, and data quality (Larvin et al., 2023). Our study will contribute to the expanding knowledge pool on the association between oral diseases and cognitive health by providing reliable, and robust evidence with adjustments on essential covariates.

The impact of oral diseases, from dental caries or periodontitis to its end stage—tooth loss, on cognitive function might be explained through the ‘nutrition pathway’: oral disease or tooth loss can cause chewing disability, leading to limited diet selection and lack of nutrition intake (Kossioni, 2018). This can contribute to poorer cognitive function due to malnourishment (Shatenstein et al., 2012). However, this mechanism may only explain the association between oral diseases and cognitive impairment in those with severe tooth loss, especially those who do not use dentures (Witter et al., 1990). Hence, the ‘commonly inflammatory pathway’ theory is more supported by existing evidence (Beydoun et al., 2020; Matsushita et al., 2020; Tonsekar et al., 2017). This theory suggested that oral pathogens such as *Porphyromonas gingivalis*. which is responsible for chronic oral disease like periodontitis, may travel through blood vessels to the body and brain and contribute to neurodegeneration (Dominy et al., 2019). However, this theory is not fully convincing too, as there are still studies arguing that the exiting evidence supporting this mechanism is still limited and mixed in study quality (Thomson and Barak, 2021). Future high quality biological studies are needed to further support this hypothesis. Other ‘life course’ studies have proposed that it is the poorer health behaviour and lifestyle choice (e.g. smoking and high sugar consumption) of people with lower cognitive function at a young age that result in worse oral health in older age. These individuals with lower cognitive function at a young age are also at higher risk of cognitive decline and dementia in older age (Thomson and Barak, 2021; Thomson et al., 2019). True, better cognitive function may indicate better awareness and capability of dental care, whereas poor oral hygiene behaviors due to cognitive impairment could worsen the oral health conditions (Thomson et al., 2018; Wu et al., 2007). Our observed reciprocal relationship between oral health and cognitive function among people over 60 years old provided a snapshot evidence supporting the "oral-cognition” association, but it is still a long way to go to find the actual mechanism.

Furthermore, each cognitive domain showed different levels of association with each oral disease. Processing speed is related to all oral diseases, but periodontitis was associated with memory but not executive function. Tooth loss and dental carries were associated with executive function but not memory. However, this interpretation should be approached with caution in the absence of a plausible biological explanation. Future studies should explore specific cognitive domains and oral diseases to understand their shared pathogenicity and identify targeted populations for intervention.

Our study has several strengths. First, our study benefited from the nationally representative, high quality, large-scale NHANES data in the aging population with minimal selection bias. Clinical measurements of oral features and professional cognitive assessment contribute to a more reliable result. Confounding factors which ranged from demographic information to oral hygiene

behaviors, multiple methods, questionnaire, interview, physical measurements were also used and guaranteed by the quality of the data ((NCHS), 2022). Secondly, our study utilized comprehensive and well-adjusted statistical modelling approaches. Especially, the use of SEM considered oral health and cognitive function as latent variables commonly with less bias from measurement (Tomarken and Waller, 2005). This further ensure the validity and reliability of the robustness in the results, and sensitivity analysis further demonstrated our result's robustness.

Our study also has some limitations. First, NHANES is cross-sectional survey, thus we cannot draw causal inference; the underlying mechanism and causality still needs investigation. Second, residual selection bias may occur by excluding edentulous participants or those who failed cognitive pre-test. Third, there were different levels of data missing in several covariates (e.g., 46.57% data missing in drug use) which might be due to unwillingness of sharing this experience. Therefore, reporter bias might occur. While smoking was an important mediator/moderator in the association between oral disease and cognitive function, in our study high proportion of missing data in smoking status made such insightful analyses not possible. Finally, it is worth noting that the current study did not include the edentulous participants, but the residual damage of periodontitis could still remain in those patients. It is not fully understood biologically yet how the residual damage influence on the cognitive function status, but evidence from many studies have shown edentulism contributes to greater risks of dementia (Stein et al., 2010) and cognitive decline (Naorungroj et al., 2015).

Oral health is undoubtable an important part of healthy ageing, and oral care in the ageing population should not be neglected. While we cannot conclude whether the association between oral disease and poor cognitive function is causal, prevention of oral disease is crucial for a better quality of life.

CONCLUSIONS

Our study showed the associations between oral diseases and cognitive function in the ageing population, emphasizing the importance of maintaining a good cognitive function and oral health. Our study also provides insight into the association between specific oral diseases and cognitive domains. Further studies are required to explore whether a causal association exists and to investigate the biological mechanism.

ACKNOWLEDGEMENT:

We would like to acknowledge the funders for this project: Alzheimer's Society and Barts Charity. Chenyi Gao is supported by the Alzheimer's Society Heather Corrie Ph.D. studentship. J. Wu is supported by Barts Charity (MGU0504).

Author contributions: CG and JK contributed to data analysis, interpretation, and manuscript write-up. HL contributed to data analysis and interpretation. JK, JW, SP, DTB and DB contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Reference list

(NCHS) NCfHS. 2022. National health and nutrition examination survey.

Bentler PM. 1990. Comparative fit indexes in structural models. Psychological bulletin. 107(2):238-246. doi: 10.1037/0033-2909.107.2.238.

Beydoun MA, Beydoun HA, Hossain S, El-Hajj ZW, Weiss J, Zonderman AB. 2020. Clinical and bacterial markers of periodontitis and their association with incident all-cause and alzheimer's disease dementia in a large national survey. J Alzheimers Dis. 75(1):157-172. doi: 10.3233/JAD-200064.

Bienias JL, Beckett LA, Bennett DA, Wilson RS, Evans DA. 2003. Design of the chicago health and aging project (chap). Journal of Alzheimer's disease. 5(5):349-355. doi: 10.3233/JAD-2003-5501.

Demmer RT, Norby FL, Lakshminarayan K, Walker KA, Pankow JS, Folsom AR, Mosley T, Beck J, Lutsey PL. 2020. Periodontal disease and incident dementia: The atherosclerosis risk in communities study (aric). Neurology. 95(12):e1660-e1671. doi: 10.1212/WNL.0000000000010312.

Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, Nguyen M, Haditsch U, Raha D, Griffin C et al. 2019. Porphyromonas gingivalis in alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. Science advances. 5(1):eaau3333-eaau3333. doi: 10.1126/sciadv.aau3333.

Eke PI, Page RC, Wei L, Thornton-Evans G, Genco RJ. 2012. Update of the case definitions for population-based surveillance of periodontitis. Journal of periodontology (1970). 83(12):1449-1454. doi: 10.1902/jop.2012.110664.

Holmer J, Eriksdotter M, Häbel H, Hed Myrberg I, Jonsson A, Pussinen PJ, Garcia-Ptacek S, Jansson L, Sandborgh-Englund G, Buhlin K. 2022. Periodontal conditions and incident dementia: A nationwide swedish cohort study. Journal of periodontology (1970). doi: 10.1002/JPER.21-0518.

Holtfreter B, Albandar JM, Dietrich T, Dye BA, Eaton KA, Eke PI, Papapanou PN, Kocher T. 2015. Standards for reporting chronic periodontitis prevalence and severity in epidemiologic studies: Proposed standards from the joint eu/USA periodontal epidemiology working group. Journal of clinical periodontology. 42(5):407-412. doi: 10.1111/jcpe.12392.

Hong J, Whelton H, Douglas G, Kang J. 2018. Consumption frequency of added sugars and uk children's dental caries. Community Dent Oral Epidemiol. 46(5):457-464. doi: 10.1111/cdoe.12413.

Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ. 2008. Inflammation and alzheimer's disease: Possible role of periodontal diseases. Alzheimer's & dementia. 4(4):242-250. doi: 10.1016/j.jalz.2007.08.004.

Kang J, Palmier-Claus J, Wu J, Shiers D, Larvin H, Doran T, Aggarwal VR. 2022. Periodontal disease in people with a history of psychosis: Results from the uk biobank population-based study. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12798.

Kang J, Smith S, Pavitt S, Wu J. 2019a. Association between central obesity and tooth loss in the non-obese people: Results from the continuous national health and nutrition examination survey (nhanes) 1999-2012. J Clin Periodontol. 46(4):430-437. doi: 10.1111/jcpe.13091.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

Kang J, Wu B, Bunce D, Ide M, Pavitt S, Wu J. 2019b. Cognitive function and oral health among ageing adults. Community dentistry and oral epidemiology. 47(3):259-266. doi: 10.1111/cdoe.12452.

Kossioni AE. 2018. The association of poor oral health parameters with malnutrition in older adults: A review considering the potential implications for cognitive impairment. Nutrients. 10(11):1709. doi: 10.3390/nu10111709.

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2023. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. Age and ageing. 52(2). doi: 10.1093/ageing/afad015.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021a. Multimorbid disease trajectories for people with periodontitis. Journal of clinical periodontology. 48(12):1587-1596. doi: 10.1111/jcpe.13536.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021b. Risk of incident cardiovascular disease in people with periodontal disease: A systematic review and meta-analysis. Clinical and experimental dental research. 7(1):109-122. doi: 10.1002/cre2.336.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2022a. The additive effect of periodontitis with hypertension on risk of systemic disease and mortality. Journal of periodontology (1970). 93(7):1024-1035. doi: 10.1002/JPER.21-0621.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2022b. Periodontitis and risk of immune-mediated systemic conditions: A systematic review and meta-analysis. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12812.

Larvin H, Wilmott S, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021c. Additive effect of periodontal disease and obesity on covid-19 outcomes. Journal of dental research. 100(11):1228-1235. doi: 10.1177/00220345211029638.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Frontiers in medicine. 7:604980-604980. doi: 10.3389/fmed.2020.604980.

Lin CS. 2018. Revisiting the link between cognitive decline and masticatory dysfunction. BMC Geriatr. 18(1):5. doi: 10.1186/s12877-017-0693-z.

Matsushita K, Yamada-Furukawa M, Kurosawa M, Shikama Y. 2020. Periodontal disease and periodontal disease-related bacteria involved in the pathogenesis of alzheimer's disease. J Inflamm Res. 13:275-283. doi: 10.2147/JIR.S255309.

Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. 1989. The consortium to establish a registry for alzheimer's disease (cerad). Part i. Clinical and neuropsychological assessment of alzheimer's disease. Neurology. 39(9):1159-1165. doi.

Naorungroj S, Schoenbach VJ, Wruck L, Mosley TH, Gottesman RF, Alonso A, Heiss G, Beck J, Slade GD. 2015. Tooth loss, periodontal disease, and cognitive decline in the atherosclerosis risk in communities (aric) study. Community dentistry and oral epidemiology. 43(1):47-57. doi: 10.1111/cdoe.12128.

Noble JM, Scarmeas N, Papapanou PN. 2013. Poor oral health as a chronic, potentially modifiable dementia risk factor: Review of the literature. Current neurology and neuroscience reports. 13(10):1-8. doi: 10.1007/s11910-013-0384-x.

Oh B, Han DH, Han KT, Liu X, Ukken J, Chang C, Dounis K, Yoo JW. 2018. Association between residual teeth number in later life and incidence of dementia: A systematic review and meta-analysis. BMC Geriatr. 18(1):48. doi: 10.1186/s12877-018-0729-z.

Plassman BL, Langa KM, Fisher GG, Heeringa SG, Weir DR, Ofstedal MB, Burke JR, Hurd MD, Potter GG, Rodgers WL et al. 2007. Prevalence of dementia in the united states: The aging, demographics, and memory study. Neuroepidemiology. 29(1-2):125-132. doi: 10.1159/000109998.

Proust-Lima C, Amieva H, Dartigues J-F, Jacqmin-Gadda H. 2006. Sensitivity of four psychometric tests to measure cognitive changes in brain aging-population–based studies. American journal of epidemiology. 165(3):344-350. doi: 10.1093/aje/kwk017.

Rubin DB. 2004. Multiple imputation for nonresponse in surveys. Hoboken, N.J. ;: Wiley-Interscience.

Shatenstein B, Ferland G, Belleville S, Gray-Donald K, Kergoat M-J, Morais J, Gaudreau P, Payette H, Greenwood C. 2012. Diet quality and cognition among older adults from the nuage study. Experimental gerontology. 47(5):353-360. doi: 10.1016/j.exger.2012.02.002.

Shen T, Lv J, Wang L, Wang W, Zhang D. 2016. Association between tooth loss and dementia among older people: A meta-analysis. Int J Geriatr Psychiatry. 31(8):953-955. doi: 10.1002/gps.4396.

Singhrao SK, Harding A, Simmons T, Robinson S, Kesavalu L, Crean S. 2014. Oral inflammation, tooth loss, risk factors, and association with progression of alzheimer's disease. Journal of Alzheimer's disease. 42(3):723-737. doi: 10.3233/JAD-140387.

Stein PS, Kryscio RJ, Desrosiers M, Donegan SJ, Gibbs MB. 2010. Tooth loss, apolipoprotein e, and decline in delayed word recall. Journal of dental research. 89(5):473-477. doi: 10.1177/0022034509357881.

Strauss E, Sherman EMS, Spreen O. 2006. A compendium of neuropsychological tests : Administration, norms, and commentary. Oxford: Oxford University Press.

Thomson WM, Barak Y. 2021. Tooth loss and dementia: A critical examination. Journal of Dental Research. 100(3):226-231. doi: 10.1177/0022034520957233.

Thomson WM, Broadbent JM, Caspi A, Poulton R, Moffitt TE. 2019. Childhood iq predicts age-38 oral disease experience and service-use. Community Dent Oral Epidemiol. 47(3):252-258. doi: 10.1111/cdoe.12451.

Thomson WM, Smith MB, Ferguson CA, Kerse NM, Peri K, Gribben B. 2018. Oral status, cognitive function and dependency among new zealand nursing home residents. Gerodontology. 35(3):185-191. doi: 10.1111/ger.12337.

Tomarken AJ, Waller NG. 2005. Structural equation modeling: Strengths, limitations, and misconceptions. Annual review of clinical psychology. 1(1):31-65. doi: 10.1146/annurev.clinpsy.1.102803.144239.

Tonsekar PP, Jiang SS, Yue G. 2017. Periodontal disease, tooth loss and dementia: Is there a link? A systematic review. Gerodontology. 34(2):151-163. doi: 10.1111/ger.12261.

Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, Initiative S. 2007. The strengthening the reporting of observational studies in epidemiology (strobe) statement: Guidelines for reporting observational studies. Annals of internal medicine. 147(8):573-577. doi.

Winning L, Naseer A, De Looze C, Knight SP, Kenny RA, O'Connell B. 2022. Tooth loss and cognitive decline in community dwelling older irish adults: A cross-sectional cohort study. Journal of dentistry. 119:104077-104077. doi: 10.1016/j.jdent.2022.104077.

Witter DJ, Cramwinckel AB, van Rossum GMJM, Käyser AF. 1990. Shortened dental arches and masticatory ability. Journal of Dentistry. 18(4):185-189. doi: 10.1016/0300-5712(90)90107-P.

Wu B, Fillenbaum GG, Plassman BL, Guo L. 2016. Association between oral health and cognitive status: A systematic review. Journal of the American Geriatrics Society (JAGS). 64(8):1752-1752. doi: 10.1111/jgs.14572.

Wu B, Plassman BL, Liang J, Wei L. 2007. Cognitive function and dental care utilization among community-dwelling older adults. American journal of public health (1971). 97(12):2216-2221. doi: 10.2105/AJPH.2007.109934.

Chapter 3: Genetic instruments of periodontitis

Chapter 3.1 GWAS systematic review

Title: Genome-Wide Association Studies on Periodontitis: A Systematic Review

Published in *PlosOne, 19(9)*
doi: 10.1371/journal.pone.0306983

Chenyi Gao[1], Mark Iles[2], Harriet Larvin[3], David Timothy Bishop[4], David Bunce[5], Mark Ide[6], Fanyiwen Sun[2], Susan Pavitt[1], Jianhua Wu [3], Jing Kang[7]*.

[1] School of Dentistry, University of Leeds, UK
[2] School of Medicine, University of Leeds, UK
[3] Wolfson Institute of Population Health, Queen Mary, University of London, UK
[4] Leeds Institute of Medical Research, School of Medicine, University of Leeds, UK
[5] School of Psychology, University of Leeds, UK
[6]Centre for Host Microbial Interactions, Faculty of Dentistry Oral and Craniofacial Sciences, King's College London, UK
[7]Oral Clinical Research Unit, Faculty of Dentistry Oral and Craniofacial Sciences, King's College London, UK

Metrics

Journal Impact factor : 3.7
Citation : 0

Abstract

Objectives: This study aims to systematically review the existing literature and critically appraise the evidence of genome-wide association studies (GWAS) on periodontitis. This study also aims to synthesise the findings of genetic risk variants of periodontitis from included GWAS.

Methods: A systematic search was conducted on PubMed, GWAS Catalog, MEDLINE, GLOBAL HEALTH and EMBASE via Ovid for GWAS on periodontitis. Only studies exploring single-nucleotide polymorphisms (SNPs) associated with periodontitis were eligible for inclusion. The quality of the GWAS was assessed using the Q-genie tool. Information such as study population, ethnicity, genomic data source, phenotypic characteristics(definition of periodontitis), and GWAS methods(quality control, analysis stages) were extracted. SNPs that reached conventional or suggestive GWAS significance level(5e-8 or 5e-06) were extracted and synthesized.

Results: A total of 15 good-quality GWAS on periodontitis were included (Q-genie scores ranged from 38-50). There were huge heterogeneities among studies. There were 11 identified risk SNPs (rs242016, rs242014, rs10491972, rs242002, rs2978951, rs2738058, rs4284742, rs729876, rs149133391, rs1537415, rs12461706) at conventional GWAS significant level ($p<5x10^{-8}$), and 41 at suggestive level ($p<5x10^{-6}$), but no common SNPs were found between studies. Three SNPs (rs4284742 [G], rs11084095 [A], rs12461706 [T]) from three large studies were from the same gene region–*SIGLEC5*.

Conclusion: GWAS of periodontitis showed high heterogeneity of methodology used and provided limited SNPs statistics, making identifying reliable risk SNPs challenging. A clear guidance in dental research with requirement of expectation to make GWAS statistics available to other investigators are needed.

Key words: Genome-Wide Association Study, Periodontitis, Periodontal Diseases, Systematic Review

Introduction

Periodontitis is a common disease affecting the tissues supporting and surrounding the teeth (Pihlstrom et al., 2005). This disease is a major cause of tooth loss and is associated with a range of multiple long-term chronic conditions (e.g., diabetes (Larvin et al., 2022; Preshaw et al., 2011), cardiovascular diseases (Larvin et al., 2021; Sanz et al., 2020), and cognitive impairments (Gao et al., 2023; Kang et al., 2020; Kang et al., 2019; Larvin et al., 2023), which can negatively impact quality of life and increase the risk of mortality (Larvin et al., 2020). Because of the effect of periodontitis, periodontal treatment is critical to the patients and several recent periodontal treatment approaches has found more effectively reduce periodontitis (Isola et al., 2024b) and also the inflammation mediator (C-reactive protein) that participated in both periodontitis and other systemic disease (i.e., cardiovascular disease) (Isola et al., 2024a). Despite of evolving periodontal treatments, how do we define the periodontitis is vitally important in either investigating the pathology but also in following treatment approaches. Although there have been consensus reports on universally accepted periodontitis definition such as for the 1999 classification of periodontal diseases (Lang et al., 1999a; Lang et al., 1999b; Lindhe et al., 1999); and then later the 2018 AAP/EFP classification of periodontal and peri-implant diseases and conditions (Caton et al., 2018)), the actual definition of periodontitis employed in dental studies shows considerable heterogeneity.

Understanding of the pathology and aetiology of periodontitis is fundamental to improve he treatment approaches of periodontitis. Based on the current understanding, the pathology and aetiology of periodontitis is complex and multifactorial, including microorganism pathogens, environmental factors, lifestyle behaviours (such as nutrition, oral hygiene, and smoking), (Preshaw et al., 2011), epigenetic factors (Loos and Van Dyke, 2020) and genetic factors (Genco and Borgnakke, 2013; Preshaw et al., 2011). By focusing on the genetic impact on periodontitis, a recent systematic review on heritability of periodontitis suggested 7%-38% across different study designs (e.g., twin study, other family study, or genome-wide association study (GWAS)) (Nibali et al., 2019). GWAS comprehensively investigates the association between a trait or diseases and hundreds of thousands of genetic variants (most commonly single nucleotide polymorphisms, SNPs) across the genome (Uffelmann et al., 2021); the technique is considered agnostic in terms of not depending on any aspect of the disease biology. GWAS as a technique aims to identify SNPs statistically associated with the trait or disease of interest, so called genetic risk variants or loci. The statistical technique involves comparing the allele frequency differences between cases (with the trait in question) and controls (persons without the trait of interest). The GWA approach has been applied to most common diseases since technology allowed the conduct of such studies in 2005 (Klein et al., 2005). Since then, this GWA technique has been applied to oral diseases including periodontitis, contributing to finding more SNPs/genes associated with

periodontitis, complementing genetic studies based on biological mechanisms proposed to influence the likelihood of periodontitis.

Current GWAS of periodontitis face many challenges, most notably limited sample size, population stratification, variation in methodologies applied, and use of non-consensus definitions of periodontitis, that complicate interpretation of the consistency of the results (Uffelmann et al., 2021). Even though there are some reviews (Loos and Van Dyke, 2020; Masumoto et al., 2019; Shaddox et al., 2021) on the genetic aspects of periodontitis, there have been limited attempts to systematically evaluate GWAS studies of periodontitis. To date, there is only one systematic review on the heritability of periodontitis (Nibali et al., 2019), and one descriptive review of periodontitis(Shaddox et al., 2021).

Since more GWAS on periodontitis have been published recently, a systematic evaluation and, ideally meta-analysis, of the available evidence would improve the understanding of the genetic mechanisms of periodontitis, address the current research gap, and contribute to better design and analysis of future GWAS. The aim of this systematic review on periodontitis is to: 1) critically appraise the evidence of GWAS of periodontitis. 2) Synthesise the findings and results by summarizing SNPs identified by high quality GWAS, and 3) meta-analyse appropriate studies/SNPs if summary statistics of GWAS were available.

Materials and methods

The study was to systematically review the GWAS that explored the genetic risk factors associated with periodontitis in the general population. This study is registered at the PROSPERO platform (ID: CRD42023456388). The main amendment in this manuscript compares to the protocol is that the meta-analysis was not performed due to appropriate data's unavailability.

Search strategy

This systematic review was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement 2020 (Page et al., 2021), as well as the guideline for performing systematic review for gene association studies (Little et al., 2006). The search was conducted on PubMed, the GWAS catalog, the ScienceDirect, the Embase, Global Health, Medline via Ovid to ensure the broad coverage of GWAS. The details search strategies in each database were summarised in Appendix Table b.1—4. The search term including periodontitis (e.g., periodontal disease, periodontitis, periodon*) and GWAS terms (e.g., GWA, GWAS, genome wide association, whole genome association, WGA, WGAS). The "*" sign is to indicate any words start with "periodon".

PEO: The Population of this study is human participants. The Exposure of this study is the genetic variants (SNPs). The Outcome of this study is 'periodontitis'.

The search was restricted to the period 2005-2023 to limit the irrelevant genetic study, since GWAS study were not conducted prior to 2006. The search was completed on 31 Aug 2023.

Study selection

Strict eligibility criteria were developed by the four authors through detailed discussion to ensure relevant study inclusion:

1. Study design is a genome-wide association study on human population including all ethnicities.
2. Study disease/primary outcome/disease phenotype is of any form of periodontitis, including clinically diagnosed periodontitis with any diagnosis criteria, or self-reported. We do not specify the heritability, power, allele frequency at the stage of the search but will synthesize such information at the data extraction.

Exclusion criteria are as below:

1. The study is a genetic/candidate gene study but not a whole genome wide scan for risk/protective SNPs of periodontitis.
2. The study used oral pathogen or gingivitis as a proxy for periodontitis. Oral pathogens and gingivitis were not considered as periodontitis as they did not directly imply the presence of periodontitis.

Search results were downloaded and imported to Covidence (link: https://get.covidence.org/literature-review?campaignid=18165361407&adgroupid=138405766537&gclid=EAIaIQobChMIwfHnroSs_AIVTLDtCh10zAAEEAAYASAAEgJP_fD_BwE) for study screening and eligibility assessment. First, duplication screening and removal is automatically done by COVIDENCE. Then, a three stage of eligibility assessment performed, and this consists of: initial abstract screening, full text screening and conflict resolve. Two authors (CG, JK) processed the abstract screening to narrow down the number of studies to be include in full text screening and eligibility assessment where all study might potentially include a GWA analysis on periodontitis is marked as “include” or “Maybe”. All “include” and “Maybe” marked study were screened for its full text to confirm its eligibility where GWAS on periodontitis or study include GWA analysis on periodontitis were marked as “include” only. Papers where the two authors had differing opinions regarding inclusion or exclusion were then screened by two more authors (HL, FS) and discussed by all four authors for the final decision. The eligibility of papers was confirmed before data extraction and quality assessment. A data extraction form was developed and based on the study (Vukovic et al., 2018). The information of SNP location and chromosome, and the nearest gene was recorded from dbSNP (https://www.ncbi.nlm.nih.gov/snp/) based on the Genome Reference

Consortium (GRC) released "build 37" of the human genome (GRCh37) version to keep the consistency of the SNP location.

Data extraction

The steps for identifying genome-wide associated significant SNPs vary across studies, and one or two or three steps were applied among the included GWAS: 1) Conducting statistical test for association on one dataset only—discovery stage; 2) Conducting meta-analysis on (an)other dataset(s)—meta-analysis stage, optional; or 3) Further seeking an independent replication / validation using other dataset(s), optional. Sometimes step 2 and 3 can swop (Uffelmann et al., 2021).

In this systematic review, no study was excluded by including replication or meta-analysis stage or not, but SNPs with $p<5x10^{-6}$ were extracted from the final stage of each identified study. SNPs with conventional GWAS significance ($p<5x10^{-8}$) were also highlighted. For example, some studies contained discovery, validation/replication, and meta-analysis stage, so we extracted only the SNPs that were significant at the meta-analysis stage. If a study performed GWAS association analysis at discovery stage only, significant SNPs from discovery stage were extracted. The essential statistics and information were extracted on Excel sheets, including SNP ID, Chromosome and position, Odds Ratio or Coefficient Beta, Standard error or confidence interval, Nearest Gene, risk allele, ethnicity, sample size, quality control procedure, periodontitis definition etc. However, due to the unavailability of full list summary statistics from almost all studies (except Shungin et al 2019), the further meta-analysis cannot be performed. Additional information (i.e, participants number, age, ethnicity, participants inclusion/exclusion criteria, data resource, type of periodontitis, clinical phenotype: measurements & definitions, study design: GWAS stage included, quality control during analysis, GWAS significant threshold, statistical model in GWAS) on methodology were also extracted.

Quality assessment

The Q-Genie tool, a validated tool developed by McMaster University for rating the quality of genetic association studies (Sohani et al., 2015; Sohani et al., 2016), for GWAS quality assessment was used. This tool includes 11 assessment areas including rationale, outcome classification, comparison groups, exposure, source of bias, power, analysis, statistical methods used, test of assumptions and inferences, and conclusion, with each area scored max 5, making the total highest score 55. All papers were assessed by two authors in parallel, and any deviation between individual ratings was discussed and validated by a third author.

In addition to the Q-Genie tool, we also checked the quality control procedure in each included paper following the guidance of GWAS quality control (i.e., call rate, HWE, MAF, relatedness,

population stratification, heterozygosity rate, sex mismatch) (Agler et al., 2019; Marees et al., 2018; Uffelmann et al., 2021). We have also extracted the genomic inflation, reference alleles and allele frequency from the 1000 Genomes allele frequency table from dbSNP for checking the inflation reported and comparing the reported alleles and allele frequency(Winkler et al., 2014). Additional information such as covariates controlled in the association analysis model and imputation quality check were also extracted from each study.

Results

Study selection

After the systematic search in GWAS catalog and PubMed, there were 16 papers from GWAS catalog, 202 papers from PubMed, 491 papers from EMBASE, Global Health and Medline Via Ovid, and 38 papers from ScienceDirect retrieved. After removal of duplicate studies and screening the abstract, 88 papers' full text were assessed for eligibility. Of these, 15 papers were deemed eligible (Bevilacqua et al., 2018; de Coo et al., 2021; Divaris et al., 2013; Feng et al., 2014; Hong et al., 2015; Munz et al., 2019; Munz et al., 2017; Petty et al., 2023; Sanders et al., 2017; Schaefer et al., 2010; Shaffer et al., 2014; Shimizu et al., 2015; Shungin et al., 2019; Tegelberg et al., 2021; Teumer et al., 2013). All included studies were published between 2010 and 2023. Figure 1 displays the PRISMA flow chart, and PRISMA 2020 checklist is provided as Appendix Table b.5 and b.6.

Quality assessment

The Q-Genie tool was applied to assess the quality of each eligible study. Overall, the studies have satisfactory quality (scores ranged from 38 –50, Table 1). However, the quality varies regarding the classification of the outcome, description of comparison groups, whether the study is adequately powered, statistical methods, description of the test and inferences (scores varied from 2 to 5 in each area among included GWAS). It is worth noting that many studies have insufficient sample size, and only five studies (Munz et al., 2019; Munz et al., 2017; Sanders et al., 2017; Shimizu et al., 2015; Shungin et al., 2019) each had a total sample over 10,000 (Table 2).

For quality control procedure performed, all included studies have had or reported quality control steps taken but varied from 3 steps reported to 11 steps reported (Table 3). There were 5 studies did not report the genomic inflation score or linkage disequilibrium score regression intercept for inflation check. Most included SNPs have similar effect allele or minor allele frequency reported but few SNPs with mismatch allele as the 1000 Genomes project were also noted (e.g., rs4242220, rs12969041, rs2027756) and effect allele or minor allele from one study was not reported (Table 4).

Figure 1. The PRISMA flow chart of the study inclusion process

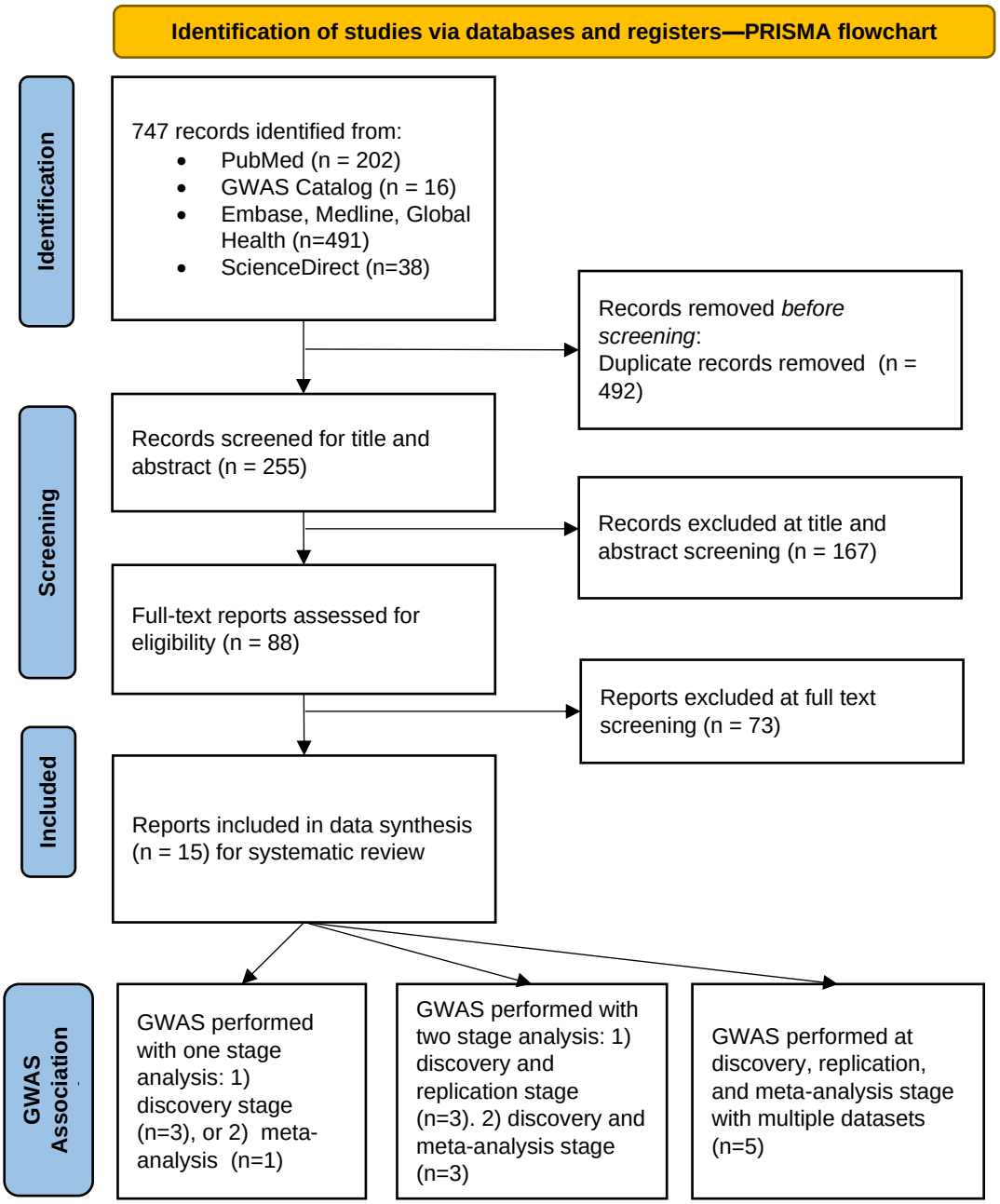


Table 1. Quality assessment rating of all 15 included studies using Q-Genie Tool

	Paper	Adequacy of the presented hypothesis and rationale	Classification of the outcome	Description of comparison groups	Technical classification of the exposure	Non-technical classification of the exposure	Disclosure and discussion of sources of bias	Study adequately powered	Description of planned analyses	Statistical methods	Description and test of all assumptions and inferences	Conclusions drawn by the authors were supported by the results and appropriate methods	Overall
1	Bevilacqua et al. (2017)	4	5	3	5	3	4	1	4	4	1	4	38
2	deCoo et al. (2021)	5	5	3	6	5	4	2	4	4	4	4	46
3	Divaris et al. (2013)	5	5	3	6	6	4	3	4	4	5	5	50
4	Feng et al. (2014)	4	4	3	4	3	2	2	4	4	4	3	37
5	Hong et al. (2015)	4	5	3	5	4	3	2	4	4	4	5	43
6	Munz et al. (2017)	4	4	4	5	4	2	4	5	5	4	5	46
7	Munz et al. (2019)	4	4	4	5	4	2	5	5	5	4	5	47
8	Sanders et al. (2017)	4	4	2	4	4	3	5	4	5	4	5	44
9	Schaefer et al. (2010)	5	3	4	4	3	2	2	4	4	4	4	39
10	Shaffer et al. (2014)	4	5	2	4	4	3	2	4	4	4	5	41
11	Shimizu et al. (2015)	5	3	3	4	4	5	5	5	5	4	5	48

12	Shungin et al. (2019)	5	2	3	5	4	5	5	4	5	4	5	47
13	Tegelberg et al. (2021)	2	2	2	3	3	4	2	4	2	4	5	33
14	Teumer et al. (2013)	4	4	2	3	5	4	2	3	4	4	4	39
15	Petty et al., (2023)	4	5	3	5	3	4	2	4	4	4	4	42

Table 2. Summary of all the 15 included studies.

Study	participants number (in analysis)		age	ethnicity	data resource	type of periodontitis	definition of periodontitis	GWAS stage included
	case	control						
Bevilacqua et al., (2018)	442	160	18-89 years old	Italian (isolated sample)	'Friuli Venezia Giulia Genetic Park'project	CP	the criteria of the AAP (Armitage, 1999)	discovery stage
de Coo et al., (2021)	441	1141	14-35 years old	Spanish	The PerioGEN study	AgP	the classification of periodontal diseases (Papapanou et al., 2018; Tonetti et al., 2018)	discovery stage
Divaris et al., (2013)	3251	1909	53-79 years old	European American	(1) the Atherosclerosis Risk in Communities (ARIC) cohort. (2) the Health, Aging and Body Composition (Health ABC) Study	CP	the CDC-AAP consensus	(1) discovery stage. (2) replication stage. (3) meta-analysis stage
Feng et al., (2014)	712	1973	25-89 years old	White, Black, other	(1) the Dental Registry and DNA Repository (DRDR) of the University of Pittsburgh School of Dental Medicine. (2)two independent population-based cohorts from Brazil (Porto Alegre & Rio de Janeiro)	CP	30% or more teeth with sites of clinical AL of five millimeters or more	(1) discovery stage. (2) replication stage. (3) meta-analysis stage
Hong et al., (2015)	414	263	44-88 years olds	Korean	the Korean Genome and Epidemiologic Study (KoGES) cohort	periodontitis	the CDC-AAP criteria (Eke et al. 2012).	Discovery stage: (1) GWAS overall population. (2) validate the SNPs in the subpopulation: moderate case, severe cases vs. healthy control.

Munz et al., (2017)	2067	8819	mean age 32-74 years	Dutch, German, Turkish	(1) the biobank Popgen. (2) the Competence Network "FoCus - Food Chain Plus" & the Dortmunder Gesundheitsstudie - DOGS. (3) recruited Turkish participants described in other study (Shaefer et al., 2015; Renauer et al., 2015)	AgP & CP	(1) CP: the number of proximal sites with AL>=4 mm. (2) AP: ≥2 teeth with 30% alveolar bone loss under the age of 35 and no diabetes (Shaefer et al., 2015)	(1) Two GWAS performed in discovery stage . (2) replication stage. (3)two meta-analysis performed (1. discovery sample meta-analysis, all sample meta-analysis)
Munz et al., (2019)	5095	9908		Dutch, German, European American	(1) the biobank Popgen. (2) the Competence Network "FoCus–Food Chain Plus", the Dortmunder Gesundheitsstudie–DOGS and the Heinz Nixdorf Recall Studies. (3) the ACTA (Academisch Centrum Tandheelkunde Amsterdam) and Rotterdam and Wageningen by the B-Proof Study. (4)the Atherosclerosis Risk in Communities (ARIC) Study. (5) a meta-analysis of SHIP and SHIPTREND cohorts	AgP & CP	(1)Dutch AgP: were ≥2 affected teeth with ≥ 30% bone loss in patients <36 years of age. (2) European American CP: the CDC-AAP (Page & Eke., 2007). (3) German CP: subjects within the first and the third tertile of proportion of proximal sites with AL ≥ 4 mm were contrasted after stratification by sex and 10-year age groups Age-specific tertiles were defined to include severely diseased cases within each age stratum.	meta-analysis stage
Sanders et al., (2017)	(1) discovery stage: 10935. (2) replication stage: 4402		18-76 years old	Hispanic/Latino, European American, African American	(1) The Hispanic Community Health Study / Study of Latinos (HCHS/SOL). (2)The ARIC cohort (1987 to present)	CP	4 case definitions from the CDC-AAP	(1) discovery stage. (2) replication stage.

Schaefer et al., (2010)	438	1320	<=35 years old	Dutch, German	(1) the cases in discovery stage weas recruited from Germany from 2003-2008. (2) The control population in the discovery stage was recruited from the popgen biobank. (3) the replication sample was recruited from Netherlands between 2003-2006	AgP	(1) localized AgP: >=50% bone loss at 2–6 teeth. (2) generalized AgP was characterized by >=50% bone loss at >=7 teeth. (Shaefer et al., 2009)	(1) two GWAS performed at discovery stage (GWAS1 & GWAS2). (2) replication stage. (3) meta-analysis stage
Shaffer et al., (2014)	(1) PD1: 176 (2) PD2: 93	(1) PD1: 497 (2) PD2: 529	18-49 years old	non- Hispanic Caucasian	the Centre for Oral Health Research in Appalachia	CP	Two or more sextants with probing depths of >=5.5mm: (1) PD1 assumed edentulous sextants had not been affected by CP; (2) PD2 assumed the missing teeth in edentulous sextants had been affected by CP.	two GWAs performed at discovery stage (i.e., PD1, PD2)
Shimizu et al., (2015)	2760	15158	17-98 years old	Japanese	the Biobank Japan and participants recruited from Health Sciences University of Hokkaido and Tokyo Medical and Dental University, the Rotary Club of Osaka-Midosuji District, the PharmaSNP Consortium.	Periodontitis	criteria of the AAP (Armitage 2004; Armitage and Cullinan 2004, 2010; Page and Eke 2007)	(1) discovery stage. (2) validation stage. (3) meta-analysis stage
Shungin et al., (2019)	36332	470262	UK Biobank 40-69 years old	Mixed ethnicity	GLIDE & UK biobank	Periodontitis and loose tooth	(1) UK biobank: self-reported oral health questionnaire. (2) GLIDE: ARIC, SHIP, SHIP-Trend and HCHS/SOL (criteria published by the CDC-AAP (Page & Eke, 2007));	(1) discovery stage. (2) meta-analysis stage

							COHRA1 (>=2 sextants had probing depth of >= 5.5 mm, or ever having 'gum surgery'); TwinGene and MDC (at least two tooth surfaces had probing depth of >=5 mm, or at least four tooth surfaces had probing depth of >=4 mm); BBJ (clinical diagnosis by physicians at participating hospitals); TMDUAGP (the 1999 international workshop for a classification of periodontal disease and conditions (Armitage, 1999));WGHS (self-reported periodontitis or not)	
Tegelberg et al., (2021)	3245		30-65 years old34	Finish	(1) the national Health 2000 Survey. (2) the Northern Finland Birth Cohort 1966 Study	periodontal pocketing	The number of teeth with≥4 mm deep periodontal pockets	(1) discovery stage (two GWAS performed) (2) meta-analysis stage
Teumer et al., (2013)	(1) SHIP: 3365 (44.8% controls). (2) SHIP-TREND: 667 (46.4% controls)		20-79 years old	German	(1) The Study of Health in Pomerania (SHIP). (2) SHIP-TREND	CP	four different periodontitis case definitions included: (1) mean proximal AL (mean PAL), (2) proportion of proximal sites with AL ≥ 4 mm, (3) the CDC-AAP case definition (Page & Eke 2007)	(1) discovery stage: four GWAS were performed for four phenotype definitions. (2) meta-analysis stage: combined analysis of results from GWAS of PAL and Severe &moderate periodontitis
Petty et al., (2023)	333	546	>18 years old	Mixed ethnicity: African, European, East Asian,	patients received treatment in at the UTHealth School of Dentistry or at the University	AP	An AP lesion was characterized radiographically as a rarefaction lesion with the	(1) discovery stage.

				Hispanic/Latino, south Asian, Other	of Pittsburgh School of Dental Medicine		disappearance of the periodontal ligament space and discontinuity of the lamina dura.	(2) validation stage

AL, attachment loss; AP, Apical periodontitis; AgP, Aggressive periodontitis; CP, Chronic periodontitis; CDC-AAP, the Centers of Disease Control American; Academy of Periodontology; AAP, American Academy of Periodontology.

Table 3. Summary of quality control steps included in the 15 studies.

Study	Sample Call rate	SNP call rate	HWE	MAF	Relatedness	population stratification	Heterozygosity rate	Sex Mismatch	Imputation quality check	adjusted covariates in the statistics model	inflation check (LDSC intercept/GC lambda)
Bevilacqua et al., (2018)					√	√				sex, age, smoking	
de Coo et al., (2021)	√	√	√		√	√				gender, age and the first PC	
Divaris et al., (2013)	√	√	√	√	√	√		√	√	GWAS use ARIC data at discovery stage adjusted for age, sex, examination centre and first 10PCs)	GWAS in ARIC cohort (Discovery stage):λ = 1.019 for moderate and 1.024 for severe CP.
Feng et al., (2014)		√	√	√						1) discovery stage: age, sex, diabetes status, and smoking status, first five PCs adjusted in the complete dataset for managing population structure.2) replication stage: age, sex, ethnicity, diabetes status, and smoking status. BMI additionally adjusted in the sample from Porto Alegre	
Hong et al., (2015)		√	√	√	√	√	√	√		age, sex, smoking, drinking, education and BMI.	
Munz et al., (2017)	√	√	√	√	√	√		√	√	sex, smoking status and 6 MDSCs (German only).	1)before adjustment: λ = 1.08 for German and λ = 1.01 for Netherland. 2) after adjusting 1-6 MDSC in German: λ = 1.04
Munz et al., (2019)			√	√					√		the P-value correction method used for inflation management

Sanders et al., (2017)	√	√	√	√	√	√		√	√	1) discovery stage adjusted for the fixed effects of sex, age, field centre, cigarette use, sampling weights, genetic subgroup and 5 genetic principal components representing ancestry, and the random effects of census block group, household, and kinship. 2) replication stage adjusted for age, sex, smoking, and 10 ancestry principal components	$\lambda = 0.990$
Schaefer et al., (2010)	√	√	√	√							$\lambda = 1.04$ in stage 1 and 1.12 in stage 2.
Shaffer et al., (2014)			√	√	√	√				age	$\lambda =$ 0.997 and 0.991 for PD1 and PD2, respectively.
Shimizu et al., (2015)		√	√	√	√	√					$\lambda = 1.024$
Shungin et al., (2019)	√	√	√	√	√	√	√	√	√	1) UK Biobank GWAS: Age, age squared, sex and genotyping array were as covariates in association testing. 2) GLIDE GWAS for the binary trait of periodontitis, age, age-squared and other study-specific covariates were instead included as covariates in association tests.	1) Combined: $\lambda_{GC} = 1.09$, h2 LDSR = 0.046, LDSR-i = 1.00 for periodontitis/loose teeth). 2) GLIDE: Association analysis results corrected by $\lambda = 1.07$
Tegelberg et al., (2021)	√	√	√	√						age, sex, smoking, plaque, and the first ten principal components	1) age, sex, and the first ten PCs adjusted: $\lambda = 0.99$ for NFBC66 and 1.01 for T2000 data. 2) age, sex, smoking, and the first ten

											PCs adjusted: λ=0.98 for NFBC66 and 0.99 for T2000 data. 3) age, sex, smoking, plaque, and the first ten PCs: λ=0.98 for NFBC66 and 0.99 for T2000 data.
Teumer et al., (2013)		√	√	√					√		The pvalues and SEs of both the individual cohort results and combined analysis results were corrected for genomic control if kGC was >1.
Petty et al., (2023)	√		√	√	√	√	√	√	√		

SD, Standard deviation; PCA, principal components analysis; HWE, Hardy–Weinberg equilibrium; LDSR, linkage disequilibrium score regression; LDSCR-I, LDSR intercept; GC, genomic control; IBS, pair-wise identity-by-state, MDSC, Multimentional scaling components; MDS, multimentional scaling; λ, GC lambda/genomic inflation factor; SE, standard error; AF, allele frequency.

Table 4. The top signal significant SNPs at 5e-6 significant level of periodontal disease extracted from the 13 included papers.

Study	SNP	Chromosome: Position	OR (95% CI)	effect allele	EAF (or MAF)	other allele	p	Nearest Gene	1000Genomes alleles and allele frequency	Reference data ethnicity
Bevilacqua et al. (2017)	rs242016	12:3788260	3.7 (2.32 -- 6.29)	A	0.21	G	1.50E-08	CRACR2A	G=0.8469 A=0.1531	Europe
	rs242014	12:3789135	3.7 (2.32 -- 6.29)	T	0.2	C	1.60E-08	CRACR2A	C=0.8469 T=0.1531	Europe
	rs10491972	12:3789949	3.7 (2.32 -- 6.29)	G	0.2	A	1.70E-08	CRACR2A	A=0.8469 G=0.1531	Europe
	rs242002	12:3807915	3.6 (2.28 -- 6.13)	T	0.21	G	2.73E-08	CRACR2A	G=0.8479 T=0.1521	Europe
deCoo et al. (2021)	rs35709256	11:926122823	2.07 (1.55-- 2.77)	A	Cases 0.12 / Controls 0.07		9.48E-07	FAT3	G=0.9056 A=0.0944	Europe
	rs4807188	19:1975217	2.8 (1.83 -- 4.27)	A	Cases 0.05/ Controls 0.02		1.81E-06	CSNK1G2	G=0.9841 A=0.0159	Europe
	rs2074872	17:10222061	2.12 (1.55 -- 2.91)	A	Cases 0.41/ Controls 0.33		2.84E-06	MYH13, LOC107985004	G=0.6054 A=0.3946	Europe
	rs116611488	1:205005227	4.22 (2.31 -- 7.72)	T	Cases 0.03/ Controls 0.009		2.90E-06	none	C=0.9861 T=0.0139	Europe
	rs4854545	2:69313773	4.34 (2.34 -- 8.02)	G	Cases 0.03/ Controls 0.009		2.97E-06	ANTXR1	G=0.0099 T=0.9901	Europe
	rs78672540	8:5247045	2.32 (1.62 -- 3.31)	C	Cases 0.08/ Controls 0.04		3.78E-06	**none**	T=0.9334 C=0.0666	Europe
	rs13439823	8:108440377	2.01 (1.49 -- 2.7)	A	Cases 0.43/Controls 0.37		4.22E-06	ANGPT1	G=0.6054 A=0.3946	Europe
	rs11993287	8:144966243	0.58 (0.46 -- 0.73)	A	Cases 0.31/ Controls 0.38		4.31E-06	none	G=0.6750 A=0.3250	Europe
Divaris et al. (2013)	rs2521634	7:24378040	1.49 (1.28 -- 1.73)	G	Cases 0.80/ Controls 0.74	A	3.50E-07	LOC107986777	G=0.830 A=0.170	American
	rs7762544	6:41379315	1.4 (1.24 -- 1.59)	G	Cases 0.21/ Controls 0.16	A	7.50E-08	none	G=0.171 A=0.829	American
	rs3826782	19:6887736	2.01 (1.52 -- 2.65)	A	Cases 0.05/ Controls 0.04	G	8.20E-07	ADGRE1	G=0.840 A=0.160	American

Hong et al. (2015)	rs4242220	5:166744741	0.53 (0.41 -- 0.69)	minor allele: C	0.24	A	2.84E-06	TENM2	T=0.7381 G=0.2619	East Asian
	rs12969041	18:13191184	2.86 (1.92 -- 4.27)	minor allele: A	0.28	G	2.79E-07	none	C=0.7431 T=0.2569	East Asian
	rs2027756	18:13190107	2.86 (1.92 -- 4.27)	minor allele: A	0.28	G	2.79E-07	none	C=0.753 A=0.000, T=0.247	East Asian (1000 Genomes_ 3x)
Munz et al. (2017)	rs2978951	8:6823295	1.25 (1.16 -- 1.35)	A	0.41	G	2.06E-08	none	A=0.3966 G=0.6034	Europe
	rs2738058	8:6821617	1.28 (1.18 -- 1.38)	T	0.43	C	6.78E-10	none	T=0.4175 C=0.5825	Europe
	rs4284742	19:52131733	1.34 (1.21 -- 1.48)	G	0.76	A	1.34E-08	SIGLEC5	A=0.2584 G=0.7416	Europe
	rs4970469	1:27312816	1.52 (1.29 -- 1.81)	G	0.9	A	1.20E-06	none	G=0.8598 A=0.1402	Europe
	rs1122900	5:36689181	1.27 (1.16 -- 1.4)	A	0.4	C	8.00E-07	none	A=0.4414 C=0.5586	Europe
	rs2070901	1:161185058	1.29 (1.16 -- 1.44)	T	0.24	G	4.36E-06	FCER1G	G=0.7306 T=0.2694	Europe
Munz et al. (2019)	rs729876	16:13388778	1.23 (1.15 -- 1.32)	T	1) German (Aggressive periodontitis): Cases 0.84/ Controls 0.80. 2) Netherlands: Cases 0.81/ Controls 0.79. 3) European American (severe): Cases 0.84 / Controls 0.80. 4) European American (moderate): Cases 0.82 /Controls 0.80.	C	1.21E-08	LOC107984137	T=0.8990 C=0.1010	Global
	rs11084095	19:52127030	1.17 (1.11 -- 1.24)	A	1) German (Aggressive	G	5.09E-08	SIGLEC5 - AC018755.18	G=0.7829 A=0.2171	Global

					periodontitis): Cases 0.46/ Controls 0.41. 2) Netherlands: Cases 0.47/ Controls 0.42. 3) European American (severe): Cases 0.41 / Controls 0.39. 4)European American (moderate): Cases 0.42 /Controls 0.39. 5) German (Chronic Periodontitis): Cases 0.83/Controls 0.80					
	rs9982623	21:47691216	1.23 (1.13 -- 1.33)	C	1) German (Aggressive periodontitis): Cases 0.88/ Controls 0.86. 2) Netherlands: Cases 0.88/ Controls 0.87. 3) European American (severe): Cases 0.89 / Controls 0.86. 4)European American (moderate): Cases 0.89 /Controls 0.86. 5) German	T	8.65E-07	MCM3AP	C=0.8660 T=0.1340	Global

					(Chronic Periodontitis): Cases 0.86/Controls 0.86					
Sanders et al. (2017)	rs149133391	1:231716531	beta: -0.139 (-0.09 -- -0.19)	T	MAF= 0.011	C	7.90E-09	TSNAX-DISC1	T=0.9888 C=0.0112	Global
	rs75715012	11:21649150	beta: 0.045 (0.03 -- 0.06)	G	MAF= 0.089	A	1.10E-07	none	G=0.9119 A=0.0881	Global
	rs186066047	5:3875774	beta: 0.23 (0.14 -- 0.31)	G	MAF= 0.003	A	1.70E-07	none	G=0.9966 A=0.0034	Global
	rs10456847	6:18955171	beta: -0.03(-0.04 -- -0.02)	C	MAF= 0.330	G	2.60E-07	none	C=0.6316 G=0.3684	Global
	rs79308117	1:155479179	beta: 0.18 (0.11 -- 0.25)	A	MAF= 0.005	C	2.80E-07	ASH1L	A=0.9870 C=0.0130	Global
Schaefer et al. (2010)	rs1537415	9:138529722	1.59 (1.36 --1.86)	G	1) GWAS 1: MAF = Cases 0.5/ Controls 0.375. 2) GWAS 2: MAF= Cases 0.493/ Controls 0.371. 3) Validation: MAF= Cases 0.49/ Controls 0.396	C	5.51E-09	GLT6D1	G=0.5845 C=0.4155	Europe
Shaffer et al. (2014)	rs733048	4:13242797	2.4 ()	not found	1) PD 1: MAF=0.22, 2) PD2: MAF=0.22		1.00E-06	LOC105374494	G=0.745 A=0.255	American
	rs10457525	6:129872966	2.33 ()	not found	1) PD 1: MAF=0.80		3.50E-06	LOC102723409	G=0.744 T=0.256	American
	rs7749983	6:129874355	2.39 ()	not found	1) PD 1: MAF=0.19		2.40E-06	LOC102723409	T=0.768 A=0.232	American
Shimizu et al. (2015)	rs9446777	6:73581051	0.86 (0.80 -- 0.93)	A	1)GWAS: MAF= Cases 0.12 /Controls 0.14, 2)replication:	G	4.83E-06	KCNQ5	A=0.8333 G=0.1667	East Asian

					MAF= Cases0.13/ Controls 0.15					
	rs2392510	7:37746569	0.87 (0.82 -- 0.92)	C	1)GWAS: MAF= Cases 0.37 /Controls 0.41, 2)replication: MAF= Cases0.37/ Controls 0.40	T	4.17E-06	GPR141	C=0.4712 T=0.5288	East Asian
Shungin et al. (2019)	rs12461706	19:52121235	1.05 ()	T	0.4		3.90E-09	SIGLEC5	A=0.7821 T=0.2179	Global
Tegelberg et al. (2021)	rs200392355	6:7452510	beta: 0.16 (0.09 -- 0.22)	*CT*	0.46		1.22E-06	LOC102724234	CT=0.0005	European (ALFA)
	rs2409703	8:10958526	beta: 0.28 (0.16 -- 0.39)	*C*	0.079		1.61E-06	*XKR6*	T=0.9066 C=0.0934	European
	rs11630851	15:76021782	beta: 0.30 (0.18 -- 0.42)	*T*	0.067		9.39E-07	DNM1P35	C=0.9374 T=0.0626	European
	rs4444613	20:13340138	beta: -0.28 (-0.38 -- -0.18)	*A*	0.087		1.35E-07	TASP1	G=0.8807 A=0.1193	European
	rs2003705	20:37763743	beta: -0.16 (-0.23 -- -0.10)	*T*	0.2		1.68E-06	LOC107985448	C=0.7863 T=0.2137	European
Petty et al., (2023)	rs12036106	1:21998838		T	0.14	C	5.07E-07	RAP1GAP,USP48	C=0.8604 T=0.1396	Global
	rs13031512	2:188670717		A	0.11	C	2.60E-06	TFPI,LINC01090	C=0.9485 A=0.0515	Global
	rs72870126	4:88859319		G	0.16	A	4.80E-06	MEPE,SPP1	A=0.7472 G=0.2528	Global
	rs369717575	4:169522766		TA	0.02	T	1.96E-06	PALLD	(A)8=0.9773 delA=0.0227	Global
	rs36793	5:109328537		T	0.94	C	2.07E-06	LINC01848,TMEM232	C=0.0946 T=0.9054	Global
	rs148550758	6:168124558		C	0.02	T	2.04E-06	LINC02487,LINC01558	T=0.9846 C=0.0154	Global
	rs7835237	8:23877060		G	0.03	C	3.62E-06	STC1,ADAM28	C=0.9663 G=0.0337	Global

	rs12800372	11:68871815		C	0.25	G	4.34E-06	TPCN2,LOC338694	G=0.8189 C=0.1811	Global

OR, Odds Ratio; CI, confidence Interval; MAF, minor allele frequency; EAF, effect allele frequency. Note. The included studies not listed in this SNP tables as they do not have SNPs that reach our criteria (5E-06). The highlighted 3 SNPs are from gene *SIGLEC5*.

Data synthesis

Table 2 summarized every study. The majority (n=13) of studies included populations from Europe (German, Finish, Italian, Spanish, Dutch, Finnish, Turkish) and America (European American, American African, Caucasian, non -Hispanic Caucasian, Hispanic and Latino), with seven studies included mixed ethnicity population and six studies included only European or non-Hispanic Caucasian. There were also two studies included samples from East Asia only (Korean and Japanese (Hong et al., 2015; Shimizu et al., 2015).

Out of these 15 studies, there are six studies focused on the chronic periodontitis, two studies focused on aggressive periodontitis, one study focused on apical periodontitis and one focused on periodontal pocketing. The rest of 5 studies were interested in periodontitis regardless of periodontitis type or including multiple types of periodontitis. The definition and measurement of periodontitis were also varied with most of 15 studies employing full-mouth dental examination performed by trained examiner or dentists and some studies also including radiographs. Eight studies utilised or incorporate different versions of criteria from the Centres for Disease Control and Prevention and American Academy of Periodontology (CDC-AAP) definition (Armitage, 1999; Armitage, 2004; Armitage and Cullinan, 2010; Eke et al., 2012; Page and Eke, 2007) in which one study measured two sites per tooth (Hong et al., 2015), two studies measured four sites per tooth (Bevilacqua et al., 2018; Teumer et al., 2013), two studies measured six sites per tooth (Divaris et al., 2013; Sanders et al., 2017) and three studies did not specify. Some studies also included radiographs such as x ray. The detailed study characteristic can be viewed in Table 2.

The majority of studies analysed periodontitis as a binary phenotype (case/control), while five studies analysed periodontitis as a continuous variable or included linear regression as part of their analysis to investigate the risk SNPs for periodontitis related traits (Bevilacqua et al., 2018; Sanders et al., 2017; Shungin et al., 2019; Tegelberg et al., 2021; Teumer et al., 2013). Two studies had no GWAS significant SNPs found nor the SNPs reaching our lowered suggestive significance threshold for SNP inclusion (Feng et al., 2014; Teumer et al., 2013). Therefore, these two studies were excluded in the next step of SNPs extraction in Table 4. Except for Shungin et al 2019, no studies have provided a *full list of summary statistic* (total sample size, number of cases, number of controls, odds ratios, 95% confidence interval or standard errors, p-value) but only reported top signal SNPs. It is also noted that some datasets were used in multiple studies (e.g. ARIC data was used in (Divaris et al., 2013; Sanders et al., 2017; Shungin et al., 2019).

Table 4 reported all the top signal SNPs extracted from the remaining 13 studies with full details, where 11 risk SNPs (rs242016, rs242014, rs10491972, rs242002, rs2978951, rs2738058, rs4284742, rs729876, rs149133391, rs1537415, rs12461706) were associated with periodontitis

at conventional GWAS significance level ($p<5x10^{-8}$), plus 41 SNPs that reached the suggestive level of significance ($p<5x10^{-6}$). Although there are no common SNPs reported across study, three large-scale study reported three genome-wide significant SNPs (i.e., rs4284742 effect allele [G], rs11084095 [A], rs12461706 [T]) from the gene *SIGLEC5* (Munz et al., 2019; Munz et al., 2017; Shungin et al., 2019).

Discussion

This systematic review has identified, critically evaluated, and synthesized genetics evidence from 15 publicly available GWAS studies on periodontitis, published between 2010 and 2023. The majority of studies had good quality, but 10 out of 15 were small studies with a sample size of less than 10,000. A total of 11 SNPs (rs242016, rs242014, rs10491972, rs242002, rs2978951, rs2738058, rs4284742, rs729876, rs149133391, rs1537415, rs12461706) at genome-wide conventional significant level ($p<5x10^{-8}$) and 41 SNPs at suggestive level ($p<5x10^{-6}$) were associated with the risk of having periodontitis. In which, three SNPs from three large studies (i.e., rs4284742 [G], rs11084095 [A], rs12461706 [T]) were reported in a same gene --- *SIGLEC5*.

This systematic review has identified several risk variants associated with periodontitis from existing GWAS studies. However, there was huge heterogeneity among study designs and methodologies: sample sizes varied from hundreds to hundred thousand (with the proportion of cases varied from 7.2% to 73.4%, Table 2), ethnicity and population differences, especially, the periodontitis measurements and definitions which may have impact on the GWAS results. The periodontal measurements in the included studies varied from self-reported questionnaire, clinical examination to radiographs, full month examination to half mouth examination, and even studies using same criteria (e.g., CDC-AAP) employed measurements varied from on six sites per tooth to two sites per tooth. Although half mouth examination and fewer sites measurement could lead to more efficient measurement processes, risk of misclassification remains. By utilising questionnaire in measurements, self-reporting bias may also contribute to either underestimation or overestimation of the number of cases. For example, Shimizu et al. 2015 recruited and measured controls separately using self-reported health questionnaire, which may cause underestimation of cases from control participants and may contribute to reduced power to identify genome-wide significant SNPs. Although it is unclear to what extent these heterogeneities in periodontitis definition and measurements would lead to heterogeneity in the GWAS results across studies, future GWAS on periodontitis could benefit from more detailed measurements on periodontal conditions and use of standardised classification criteria. The 2018 periodontal status classification is advocated for use, and proposals have been made to further refine the use of current periodontitis classification to enhance epidemiological data collection and analysis (Holtfreter et al., 2024). In addition, more GWAS studies in different ethnic groups, with larger sample sizes and considering covariates are also important for observing the potential

for ethnic differences on GWAS results or susceptibility of periodontitis and observe genome-wide significant SNPs.

Heterogeneity in the methods and results reporting was also noted. The current 15 studies incorporated multiple study designs such as inclusion of multiple GWA stages (e.g., 8 studies included discovery and replications). Meanwhile it has been noted that the replication stage had relatively small sample size than discovery stage with five studies included replication stage that have had more than 1,000 participants. However, insufficient sample size in the replication stage could lead to both type I and type II error in the replication results (McGuire et al., 2021). Meanwhile, it is also important to replicate the results from the combined analysis stage (meta-analysis of discovery and validation stage), which is also missing in the included studies. Future study could utilise better approaches to assess the reproducibility such as the meta-analysis model-based assessment (McGuire et al., 2021). In terms of results reporting, several included studies did not report the coefficient beta or odds ratio, the standard error or confidence interval, and effect allele of the reported SNPs. This missing information leads to difficulty during data synthesis. Genomic inflation scores were also missing in few studies and not all studies reported or not conducted all quality control steps selected from GWAS quality control guidance. A standardized guideline and consensus on GWAS reporting may be needed to uniform the GWAS report.

To date, there is no guideline in PRISMA on how to perform systematic review on GWAS studies, except Winkler (2014) suggested protocol for genome-wide association meta-analyses in terms of the quality control and conduction of such analyses, requiring the availability of full SNPs and associated statistics reported, allele frequencies, and population stratification (Winkler et al., 2014) . Analytic tools like METAL might estimate the pooled effect for overlapping samples, but it needs genome-wide data to perform appropriate estimation. In addition, a standard and more detailed and widely acceptable quality assessment tool for GWAS is also needed. Although Q-genie tool has been used to assess validity and reliability here, it was designed for genetic association studies, not particularly for GWAS. An assessment tool specifically for GWAS with clear guidance on scoring would be beneficial not only for the assessors but also for the readers to better understand quality assessment criteria.

Form the included 15 studies, three studies based on populations of German, Dutch, European American, Turkish and Asian with sample size > 10,000 participants were commonly discovered three unique SNPs in the gene *SIGLEC5*, where the effect alleles of all three SNPs has been reported for their protective effect on periodontitis (Munz et al., 2019; Munz et al., 2017; Shungin et al., 2019). A recent study has investigated the three genome-wide significant SNPs in the region of SIGLEC5 and shown an impact on SIGLEC5 expression indicating that *SIGLEC5* is indeed

the target gene for the signal. (Mueller et al., 2022). *SIGLEC5* codes for sialic acid-binding Ig-like lectins as a transmembrane inhibitory receptor and is responsible for binding sialic acids and sialic acid-containing glycan ligands. It is expressed in cells in the innate immune system and plays a role in inflammation regulation, both in infection and wound healing (Pillai et al., 2012). They observed SNP rs11084095 at *SIGLEC5* can influence ERG binding and enhancer activity (Mueller et al., 2022), where ERG is important for endothelial homeostasis, including acute response to injury and repair of the endothelium (Heiss et al., 2015). Meanwhile, the SNP rs12461706 was found in complete linkage disequilibrium with rs11084095. In addition, SNP rs4284742 has been shown to affect MAFB binding affinity, with the common allele enhancing the binding affinity compared to the alternative allele (Mueller et al., 2022). MAFB is suggested to be associated with the activation of *SIGLEC5* expression and contribute to early-onset periodontitis. Further investigation of *SIGLEC5* in periodontitis pathologies and intervention targeting the biological pathway underpinned by *SIGLEC5* may contributes to both aetiology understanding and disease treatment (Harper et al., 2015). In addition to these three SNPs from the same gene region, the rest of the 8 SNPs may also contribute to the periodontitis, however, further GWAS replication with larger sample size may be needed.

The three SNPs in *SIGLEC5*, two of them were found from both aggressive and chronic periodontitis and one were found from mixed definition defined periodontitis (i.e., including both self-reported and also clinical definition defined periodontitis). This suggested that the *SIGLEC5* and the three SNPs may play fundamental roles in all types of periodontitis instead of some specific periodontitis and investigation on *SIGLEC5* may contributes to all type of periodontitis treatment and common pathology understanding. According to the GWAS catalog, most of GWAS significant SNPs found on 15 studies were reported only for periodontitis, except rs2738058 was also found in kidney diseases in Chinese population (i.e., IGA glomerulonephritis) (Li et al., 2015). Meanwhile, several mapped genes of these GWAS significant SNPs were also found significant in IGA glomerulonephritis (e.g., *DEFA9P* and *DEFA10P* (Li et al., 2015)), neuropsychological conditions (e.g., *TMF1P1* (Neale et al., 2014)), despite of reporting uncommon SNP. These may suggest somewhat genetic similarity between periodontitis and these conditions but further investigation on their relationship with periodontitis still needed.

In comparison to the previous systematic review on the heritability of periodontitis (Nibali et al., 2019), our focus lies more heavily on the methodology utilised in GWAS and synthesis of results. Moreover, comparing with the review article of genetics of periodontitis by Shaddox et al 2021, we employed a systematic approach and included a greater number of studies than the prior review. It is noted that 6 out of 8 studies that used chronic periodontitis as disease phenotype did not meet the required sample size of >10K cases for such disease with low heritability, except the studies Munz et al (2017) and Munz et al (2019). For the 8 studies that used aggressive

periodontitis with high heritability as disease definition, smaller sample size is usually acceptable but no studies showed any common risk variants, indicating a potential of false positive results.

The findings that NO common SNPs were consistently reported through all the included studies highlighted a significant level of heterogeneity in the results obtained from GWAS of periodontitis. Given this lack of repeatability in GWAS finding, any identified genetic variants must be interpretated with caution. Furthermore, we observed a reluctance within the dental research community to share GWAS results. Only one study (Shungin et al 2019) provided a comprehensive list of SNPs statistics from their GWAS of periodontitis, while the remaining studies offered only a limited number of top-signaled SNPs, with some providing incomplete statistics such as lacking odds ratios or 95% confidence intervals. When we attempted to obtain full SNP statistics from these studies, they either declined or did not respond, underscoring a significant transparency issue in current dental research practices. In many common diseases within the medical field, guidelines exist mandating GWAS data sharing as a standard practice expected by funders and publishers. We call for similar guidelines to be established in dental research, requiring the sharing of GWAS statistics from published work to facilitate advancement within the field.

The current study has several strengths, such as summarizing existing risk variants in the literature and discussed the study design of current GWAS on periodontitis, and presenting each study clearly with its database used, population, and GWAS testing method used. However, there are also some limitations to consider. Publication bias should be kept in mind while interpreting the results, as only publicly available GWAS studies were included in the review (Vukovic et al., 2018). Additionally, potential biases from the selected studies may exist, such as those from the periodontitis definition, and variation of methodology applied (e.g. number of covariates adjusted in the association analysis). The current study failed to obtain access to full GWAS summary statistics to perform meta-analysis which could contribute to resolve small sample size in many studies and detect risk variants in combined sample. For example, SNPs with consistently border-line association with periodontitis could have been missed, which may have in theory become statistically significant in meta-analysis if most studies had provided a full list of statistics for all SNPs. In addition, since the majority of included studies sampled were from white population (European, European American, Caucasian etc.) and only few of them included Latino, Black, Asian or mixed ethnicity, it is difficult to draw conclusion based on each ethnicity group. More studies investigating Black and Asian populations could help to understand the underlying ethnicity difference of periodontitis, and future GWAS should also provide summary statistics in repository such as GWAS catalog to facilitate further meta-analysis.

Conclusions

To conclude, our systematic review of 15 GWAS studies on periodontitis identified 11 SNPs were at genome-wide significance $p<5x10^{-8}$ level. Variants near or in the gene region *SIGLEC5* were reported most frequently (i.e. in three large scale studies) for its potential role on periodontitis. These results imply potential therapeutic targets pathway underlined by the *SIGLEC5*. Further investigation on this gene could contributed to the periodontitis treatment approach design. However, the heterogeneity on study design, study sample size and target population between studies has been noted. To improve our understanding of periodontitis and support the development of effective treatment options, more high-quality and homogeneous methodology used in GWAS studies are needed. These studies should use standardized periodontitis definitions and assessment tools, have larger sample sizes, and include different ethnicities. Data repository of GWAS results should be made available so that further meta-analysis can be possible, especially in dental research, to ensure research transparency and reproducibility. These efforts will contribute to greater understanding of this oral disease and ultimately benefit public health.

Data availability: the data included in the current study can be found in the included studies and their supplementary files.

Acknowledgment

Funding statement: Chenyi Gao is supported by external fund from the Alzheimer's Society (Alzheimer's Society Heather Corrie Ph.D. studentship: 546 (AS-PhD-19b-012)). The views expressed are those of the author(s) and not necessarily those of the Alzheimer's Society. The Alzheimer's Society did not involve in design and conduct this study, as well as data collection or preparation, review, approval, and decision to submit the manuscript. Jianhua Wu and Harriet Larvin are supported by internal fund from Barts Charity (MGU0504). There was no additional external funding received for this study. . The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions: CG contributed to systematic search, paper screening, quality assessment, data extraction, data analysis, interpretation, and manuscript write-up. HL, FS contributed to paper screening, quality assessment, data analysis. JK contributes to the paper screening, quality assessment, data extraction, data analysis and interpretation. JK, JW, SP, DTB and DB contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Ethics approval statement: as a systematic review only used data publicly available in previously published paper (main report or supplements), ethics approval is not considered necessary.

Competing interest disclosure: The authors declare no conflict of interest nor competing interest.

Reference

Agler CS, Shungin D, Ferreira Zandoná AG, Schmadeke P, Basta PV, Luo J, Cantrell J, Pahel TD, Jr., Meyer BD, Shaffer JR et al. 2019. Protocols, methods, and tools for genome-wide association studies (gwas) of dental traits. Methods Mol Biol. 1922:493-509. doi: 10.1007/978-1-4939-9012-2_38.

Armitage GC. 1999. Development of a classification system for periodontal diseases and conditions. Ann Periodontol. 4(1):1-6. doi: 10.1902/annals.1999.4.1.1.

Armitage GC. 2004. Periodontal diagnoses and classification of periodontal diseases. Periodontology 2000. 34(1):9-21. doi: 10.1046/j.0906-6713.2002.003421.x.

Armitage GC, Cullinan MP. 2010. Comparison of the clinical features of chronic and aggressive periodontitis. Periodontology 2000. 53(1):12-27. doi: 10.1111/j.1600-0757.2010.00353.x.

Bevilacqua L, Navarra CO, Pirastu N, Lenarda RD, Gasparini P, Robino A. 2018. A genome-wide association study identifies an association between variants in efcab4b gene and periodontal disease in an italian isolated population. Journal of periodontal research. 53(6):992-998. doi: 10.1111/jre.12598.

Caton JG, Armitage G, Berglundh T, Chapple ILC, Jepsen S, Kornman KS, Mealey BL, Papapanou PN, Sanz M, Tonetti MS. 2018. A new classification scheme for periodontal and peri-implant diseases and conditions – introduction and key changes from the 1999 classification. Journal of clinical periodontology. 89(S1):S1-S8. doi: 10.1002/JPER.18-0157.

de Coo A, Cruz R, Quintela I, Herrera D, Sanz M, Diz P, Rodriguez Grandio S, Vallcorba N, Ramos I, Oteo A et al. 2021. Genome-wide association study of stage iii/iv grade c periodontitis (former aggressive periodontitis) in a spanish population. J Clin Periodontol. 48(7):896-906. doi: 10.1111/jcpe.13460.

Divaris K, Monda KL, North KE, Olshan AF, Reynolds LM, Hsueh WC, Lange EM, Moss K, Barros SP, Weyant RJ et al. 2013. Exploring the genetic basis of chronic periodontitis: A genome-wide association study. Hum Mol Genet. 22(11):2312-2324. doi: 10.1093/hmg/ddt065.

Eke PI, Page RC, Wei L, Thornton-Evans G, Genco RJ. 2012. Update of the case definitions for population-based surveillance of periodontitis. Journal of periodontology (1970). 83(12):1449-1454. doi: 10.1902/jop.2012.110664.

Feng P, Wang X, Casado PL, Kuchler EC, Deeley K, Noel J, Kimm H, Kim JH, Haas AN, Quinelato V et al. 2014. Genome wide association scan for chronic periodontitis implicates novel locus. BMC Oral Health. 14:84. doi: 10.1186/1472-6831-14-84.

Gao C, Larvin H, Bishop DT, Bunce D, Pavitt S, Wu J, Kang J. 2023. Oral diseases are associated with cognitive function in adults over 60 years old. Oral diseases. doi: 10.1111/odi.14757.

Genco RJ, Borgnakke WS. 2013. Risk factors for periodontal disease: Risk factors for periodontal diseases. Periodontology 2000. 62(1):59-94. doi: 10.1111/j.1600-0757.2012.00457.x.

Harper AR, Nayee S, Topol EJ. 2015. Protective alleles and modifier variants in human health and disease. Nat Rev Genet. 16(12):689-701. doi: 10.1038/nrg4017.

Heiss C, Rodriguez-Mateos A, Kelm M. 2015. Central role of enos in the maintenance of endothelial homeostasis. Antioxid Redox Signal. 22(14):1230-1242. doi: 10.1089/ars.2014.6158.

Holtfreter B, Kuhr K, Borof K, Tonetti MS, Sanz M, Kornman K, Jepsen S, Aarabi G, Volzke H, Kocher T et al. 2024. Aces: A new framework for the application of the 2018

periodontal status classification scheme to epidemiological survey data. J Clin Periodontol. doi: 10.1111/jcpe.13965.

Hong K-W, Shin M-S, Ahn Y-B, Lee H-J, Kim H-D. 2015. Genomewide association study on chronic periodontitis in korean population: Results from the yangpyeong health cohort. Journal of clinical periodontology. 42(8):703-710. doi: 10.1111/jcpe.12437.

Isola G, Pesce P, Polizzi A, Lo Giudice A, Cicciù M, Scannapieco FA. 2024a. Effects of minimally invasive non-surgical therapy on c-reactive protein, lipoprotein-associated phospholipase a2, and clinical outcomes in periodontitis patients: A 1-year randomized, controlled clinical trial. Journal of periodontology (1970). doi: 10.1002/JPER.23-0518.

Isola G, Polizzi A, Santonocito S, Alibrandi A, Pesce P, Kocher T. 2024b. Effect of quadrantwise versus full-mouth subgingival instrumentation on clinical and microbiological parameters in periodontitis patients: A randomized clinical trial. Journal of periodontal research. doi: 10.1111/jre.13279.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

Kang J, Wu B, Bunce D, Ide M, Pavitt S, Wu J. 2019. Cognitive function and oral health among ageing adults. Community dentistry and oral epidemiology. 47(3):259-266. doi: 10.1111/cdoe.12452.

Klein RJ, Zeiss C, Chew EY, Tsai J-Y, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST et al. 2005. Complement factor h polymorphism in age-related macular degeneration. Science (American Association for the Advancement of Science). 308(5720):385-389. doi: 10.1126/science.1109557.

Lang N, Bartold PM, Cullinan M, Jeffcoat M, Mombelli A, Murakami S, Page R, Papapanou P, Tonetti M, Van Dyke T. 1999a. Consensus report: Aggressive periodontitis. Annals of Periodontology. 4(1):53-53. doi: https://doi.org/10.1902/annals.1999.4.1.53.

Lang NP, Soskolne WA, Greenstein GR, Cochran DL, Corbet EF, Meng H-x, Newman MG, Novak MJ, Tenenbaum H. 1999b. Consensus report: Necrotizing periodontal diseases. Annals of Periodontology. 4:78-78. doi.

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2023. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. Age Ageing. 52(2). doi: 10.1093/ageing/afad015.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021. Risk of incident cardiovascular disease in people with periodontal disease: A systematic review and meta-analysis. Clinical and experimental dental research. 7(1):109-122. doi: 10.1002/cre2.336.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2022. Periodontitis and risk of immune-mediated systemic conditions: A systematic review and meta-analysis. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12812.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Front Med (Lausanne). 7:604980. doi: 10.3389/fmed.2020.604980.

Li M, Foo JN, Wang JQ, Low HQ, Tang XQ, Toh KY, Yin PR, Khor CC, Goh YF, Irwan ID et al. 2015. Identification of new susceptibility loci for iga nephropathy in han chinese. Nature communications. 6(1):7270-7270. doi: 10.1038/ncomms8270.

Lindhe J, Ranney R, Lamster I, Charles A, Chung C-P, Flemmig T, Kinane D, Listgarten M, Löe H, Schoor R et al. 1999. Consensus report: Chronic periodontitis. Annals of Periodontology. 4(1):38-38. doi: https://doi.org/10.1902/annals.1999.4.1.38.

Little J, Higgins J, Bray M, Ioannidis J, Khoury M, Manolio T, Smeeth L, Sterne J. 2006. The hugenet™ huge review handbook, version 1.0. doi.

Loos BG, Van Dyke TE. 2020. The role of inflammation and genetics in periodontal disease. Periodontology 2000. 83(1):26-39. doi: 10.1111/prd.12297.

Marees AT, Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM. 2018. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. International journal of methods in psychiatric research. 27(2):e1608-n/a. doi: 10.1002/mpr.1608.

Masumoto R, Kitagaki J, Fujihara C, Matsumoto M, Miyauchi S, Asano Y, Imai A, Kobayashi K, Nakaya A, Yamashita M et al. 2019. Identification of genetic risk factors of

aggressive periodontitis using genomewide association studies in association with those of chronic periodontitis. Journal of periodontal research. 54(3):199-206. doi: 10.1111/jre.12620.

McGuire D, Jiang Y, Liu M, Weissenkampen JD, Eckert S, Yang L, Chen F, Berg A, Vrieze S, Jiang B et al. 2021. Model-based assessment of replicability for genome-wide association meta-analysis. Nature communications. 12(1):1964-1964. doi: 10.1038/s41467-021-21226-z.

Mueller R, Chopra A, Dommisch H, Schaefer AS. 2022. Periodontitis risk variants at siglec5 impair erg and mafb binding. J Dent Res. 101(5):551-558. doi: 10.1177/00220345211049984.

Munz M, Richter GM, Loos BG, Jepsen Sr, Divaris K, Offenbacher S, Teumer A, Holtfreter B, Kocher T, Bruckmann C et al. 2019. Meta-analysis of genome-wide association studies of aggressive and chronic periodontitis identifies two novel risk loci. European journal of human genetics : EJHG. 27(1):102-113. doi: 10.1038/s41431-018-0265-5.

Munz M, Willenborg C, Richter GM, Jockel-Schneider Y, Graetz C, Staufenbiel I, Wellmann J, Berger K, Krone B, Hoffmann P et al. 2017. A genome-wide association study identifies nucleotide variants at siglec5 and defa1a3 as risk loci for periodontitis. Hum Mol Genet. 26(13):2577-2588. doi: 10.1093/hmg/ddx151.

Neale BM, Corvin A, Huang H, Agartz I, Amin F, Begemann M, Buccola NG, Cantor RM, Chen RYL, Cheng W et al. 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature (London). 511(7510):421-427. doi: 10.1038/nature13595.

Nibali L, Bayliss-Chapman J, Almofareh SA, Zhou Y, Divaris K, Vieira AR. 2019. What is the heritability of periodontitis? A systematic review. J Dent Res. 98(6):632-641. doi: 10.1177/0022034519842510.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE et al. 2021. The prisma 2020 statement : An updated guideline for reporting systematic reviews. doi.

Page RC, Eke PI. 2007. Case definitions for use in population-based surveillance of periodontitis. Journal of periodontology. 78(7):1387-1399. doi: 10.1902/jop.2007.060264.

Petty LE, Silva R, Chaves de Souza L, Vieira AR, Shaw DM, Below JE, Letra A. 2023. Genome-wide association study identifies novel risk loci for apical periodontitis. Journal of endodontics. doi: 10.1016/j.joen.2023.07.018.

Pihlstrom BL, Michalowicz BS, Johnson NW. 2005. Periodontal diseases. Lancet. 366(9499):1809-1820. doi: 10.1016/S0140-6736(05)67728-8.

Pillai S, Netravali IA, Cariappa A, Mattoo H. 2012. Siglecs and immune regulation. Annu Rev Immunol. 30:357-392. doi: 10.1146/annurev-immunol-020711-075018.

Preshaw PM, Alba AL, Herrera D, Jepsen S, Konstantinidis A, Makrilakis K, Taylor R. 2011. Periodontitis and diabetes: A two-way relationship. Diabetologia. 55(1):21-31. doi: 10.1007/s00125-011-2342-y.

Sanders AE, Sofer T, Wong Q, Kerr KF, Agler C, Shaffer JR, Beck JD, Offenbacher S, Salazar CR, North KE et al. 2017. Chronic periodontitis genome-wide association study in the hispanic community health study / study of latinos. J Dent Res. 96(1):64-72. doi: 10.1177/0022034516664509.

Sanz M, Marco del Castillo A, Jepsen S, Gonzalez-Juanatey JR, D'Aiuto F, Bouchard P, Chapple I, Dietrich T, Gotsman I, Graziani F et al. 2020. Periodontitis and cardiovascular diseases: Consensus report. Journal of clinical periodontology. 47(3):268-288. doi: 10.1111/jcpe.13189.

Schaefer AS, Richter GM, Nothnagel M, Manke T, Dommisch H, Jacobs G, Arlt A, Rosenstiel P, Noack B, Groessner-Schreiber B et al. 2010. A genome-wide association study identifies glt6d1 as a susceptibility locus for periodontitis. Human molecular genetics. 19(3):553-562. doi: 10.1093/hmg/ddp508.

Shaddox LM, Morford LA, Nibali L. 2021. Periodontal health and disease: The contribution of genetics. Periodontology 2000. 85(1):161-181. doi: 10.1111/prd.12357.

Shaffer JR, Polk DE, Wang X, Feingold E, Weeks DE, Lee M-K, Cuenco KT, Weyant RJ, Crout RJ, McNeil DW, Marazita ML. 2014. Genome-wide association study of

periodontal health measured by probing depth in adults ages 18-49 years. G3 : genes - genomes - genetics. 4(2):307-314. doi: 10.1534/g3.113.008755.

Shimizu S, Momozawa Y, Takahashi A, Nagasawa T, Ashikawa K, Terada Y, Izumi Y, Kobayashi H, Tsuji M, Kubo M, Furuichi Y. 2015. A genome-wide association study of periodontitis in a japanese population. Journal of dental research. 94(4):555-561. doi: 10.1177/0022034515570315.

Shungin D, Haworth S, Divaris K, Agler CS, Kamatani Y, Keun Lee M, Grinde K, Hindy G, Alaraudanjoki V, Pesonen P et al. 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.

Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, Norman G, Anand SS. 2015. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (q-genie) tool. BMC genetics. 16(1):50-50. doi: 10.1186/s12863-015-0211-2.

Sohani ZN, Sarma S, Alyass A, de Souza RJ, Robiou-du-Pont S, Li A, Mayhew A, Yazdi F, Reddon H, Lamri A et al. 2016. Empirical evaluation of the q-genie tool: A protocol for assessment of effectiveness. BMJ Open. 6(6):e010403-e010403. doi: 10.1136/bmjopen-2015-010403.

Tegelberg P, Leppilahti JM, Ylöstalo A, Tervonen T, Kettunen J, Suominen AL, Ylöstalo P. 2021. Genome-wide association study of periodontal pocketing in finnish adults. BMC oral health. 21(1):1-611. doi: 10.1186/s12903-021-01964-8.

Teumer A, Holtfreter B, Volker U, Petersmann A, Nauck M, Biffar R, Volzke H, Kroemer HK, Meisel P, Homuth G, Kocher T. 2013. Genome-wide association study of chronic periodontitis in a general german population. J Clin Periodontol. 40(11):977-985. doi: 10.1111/jcpe.12154.

Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, Martin HC, Lappalainen T, Posthuma D. 2021. Genome-wide association studies. Nature Reviews Methods Primers. 1(1). doi: 10.1038/s43586-021-00056-9.

Vukovic V, Stojanovic J, Vecchioni A, Pastorino R, Boccia S. 2018. Systematic review and meta-analysis of snps from genome-wide association studies of head and neck cancer. Otolaryngol Head Neck Surg. 159(4):615-624. doi: 10.1177/0194599818792262.

Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, Ferreira T, Fall T, Graff M, Justice AE et al. 2014. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc. 9(5):1192-1212. doi: 10.1038/nprot.2014.071.

Chapter 3.2 GWAS

Title: Genetic risk factors for periodontitis: a genome-wide association study using UK Biobank

Manuscript under review at *Journal of Clinical Oral Investigation (IF=3.1)*

Authors:

Chenyi Gao[1], Mark M. Iles[2,3], David Timothy Bishop[4], Harriet Larvin[5], David Bunce[6], Bei Wu[7], Huabin Luo[8], Luigi Nibali[10], Susan Pavitt[1], Jianhua Wu[5], Jing Kang[9]*.

Affiliations:

[1] School of Dentistry, University of Leeds, Leeds, UK

[2] Leeds Institute for Data Analytics, University of Leeds, Leeds, UK

[3] NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

[4] Leeds Institute of Medical Research, School of Medicine, University of Leeds, Leeds, UK

[5] Wolfson Institute of Population Health, Queen Mary University of London, London, UK

[6] School of Psychology, University of Leeds, Leeds, UK

[7] Rory Meyers College of Nursing, New York University, New York, US

[8] Department of Public Health, East Carolina University, Greenville, US

[9] Oral Clinical Research Unit, Faculty of Dentistry Oral Craniofacial Sciences, King's College London

[10] Periodontology Unit, Centre for Host Microbiome Interactions, Faculty of Dentistry, Oral & Craniofacial Sciences, King's College, London, UK.

Abstract

Objective: Periodontitis is associated with many health conditions, but its genetic basis is not yet fully understood. This genome-wide association study (GWAS) aimed to investigate the genetic variants associated with periodontitis.

Method: This study involved UK Biobank participants of European descent. Individuals were categorised as "having periodontitis" if they self-reported having 'painful gums', 'bleeding gums' or 'loose teeth' (n=68,482). Those without these symptoms (n=307,342) were classified as controls. We performed GWAS of this binary periodontitis phenotype using logistic regression models with PLINK2.0 adjusting for age, sex, and the first 15 principal components to account for population stratification.

Results: A total of 376,611 participants (mean baseline age = 57 ± 7.9 SD) were included in the GWAS, and four significant loci were identified: rs775476621 on chromosome 11 (Odds Ratio, OR[T]: 3.08, p=1.01 $x10^{-8}$), rs751014048 on chromosome 11 (OR[G]: 3.07, p=1.04 $x10^{-8}$), rs149922301 on chromosome 4 near gene *RP11-61G19.1* (OR[A]: 1.18, p=2.71 $x10^{-8}$), and rs368467810 on chromosome 6 near genes *HIST1H3L* (OR[TTTA]: 0.96, p=3.88 $x10^{-8}$).

Conclusion: Four loci associated with periodontitis were identified, none of which had been previously linked to this condition. Further exploration of the functions of these loci may enhance our understanding of periodontitis aetiology and subsequent drug development.

Introduction

Periodontitis is a common inflammatory condition caused by bacterial infection, affecting the tissues surrounding and supporting the teeth (Eke et al., 2012). Periodontitis is recognised as a major cause of tooth loss (Eke et al., 2012) and is associated with various long-term chronic conditions such as diabetes (Larvin et al., 2022b), cognitive impairment and dementia (Kang et al., 2020; Larvin et al., 2022a), and cardiovascular disease (Larvin et al., 2021), resulting in worsening quality of life (Mao et al., 2023) and an increased risk of mortality (Larvin et al., 2024; Larvin et al., 2020). Periodontitis affects 20--50% of the global population (Nazir, 2017), with severe cases impacting more than 1 billion adults worldwide (WHO, 2023). The pathology and aetiology of periodontitis are not yet fully understood due to its multifaceted nature. This includes the complex interplay of microorganisms, pathogens, environmental factors such as nutritional intake and smoking, and genetic factors (Page et al., 1997; Preshaw et al., 2011). While there have been numerous experimental studies into the roles of microorganisms, pathogens and environmental factors in the development of periodontitis, genetic influences are still not well understood. Improved knowledge of the genetics of periodontitis could potentially benefit not only our understanding of disease aetiology, but also risk assessment and personalised treatment plans (Shungin et al., 2019).

Recent study estimated the heritability of periodontitis, ranging from 7% to 38% (Nibali et al., 2019). To date, 15 genome-wide association studies (GWAS) of various forms of periodontitis have been conducted as identified in a GWAS systematic review (Gao et al., 2024). Here, 11 single nucleotide polymorphisms (SNP) were collectively found that reached the conventional level for genome-wide statistical significance ($p<5x10^{-8}$), and achieved a significance level of 41 SNPs at $5x10^{-6}$. However, no common loci were identified across these studies, even among those of similar ethnicity (Gao et al., 2024). This may be partially due to heterogeneity across studies and small sample size.

Further GWAS investigations using larger sample sizes are warranted. The aim of this study, therefore, was to use data from the UK Biobank (UKB) to identify genetic risk factors for periodontitis in a population of European descent.

Method

Sample and data resources

UKB (Bycroft et al., 2018) is a prospective cohort study involving over 500,000 participants aged 40 to 69 years at initial recruitment (2006--2010). A wide range of demographic and health-related, variables and lifestyle data was recorded for participants through questionnaires, physical examinations, imaging and genotyping at baseline. UKB has continued to follow-up participants'

health-related outcomes through continued contact and access to medical records (Sudlow et al., 2015) (https://www.ukbiobank.ac.uk/).

In the current study, the sample was drawn from the final release of UK Biobank genetic data (sample size n=488,377), of which 409,548 (84% of total genotyped sample) persons were classified as of European ancestry by integration of self-reported ethnicities and principal components analysis (PCA) results of population structure (Bycroft et al., 2018). At the time of analysis, 139 UKB recruits had withdrawn their consent and so were excluded from this study.

Phenotype: Periodontitis

The UKB variable "self-reported dental health" (variable code: 100538) was used as a surrogate for periodontitis. Participants who reported "painful gums", "bleeding gums", or "loose teeth" were classified as having periodontitis; whereas, those reporting other or no dental health issue were classified as controls. This self-reported measure of periodontitis has previously been validated as a reliable surrogate for periodontitis (Abbood et al., 2016) and is suggested to be characteristic of moderate to severe periodontitis (Papapanou et al., 2018). Participants who chose "prefer not to say" were coded as missing for this variable. The same definition has been used in previous UKB studies (Kang et al., 2022; Larvin et al., 2020) in the absence of a clinical diagnosis.

Genotyping and imputation

Details of DNA extraction and quality filtering are described elsewhere (https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/ukb_dna_processing.pdf).

Two genotyping arrays were employed: (1) the Affymetrix UK Biobank Axiom Array (sample size n=438,427) and (2) the Affymetrix UK BiLEVE Axiom Array (sample size n=49,950). The genotyped samples shared 95% common single nucleotide variants (SNV) and produced 825,927 SNVs for genotyping in total. The imputation of missing genotypes was implemented in IMPUTE4 (https://jmarchini.org/software/), which was developed from IMPUTE2 undertaken by UK Biobank using genotype reference data from the Haplotype Reference Consortium (HRC). The UK10K and 1000G genotype reference panels were employed to impute SNVs not included in the HRC. This study performed quality control filtering and genome-wide analysis on Version 3 of the imputed data, containing 487,409 participants for 93,095,623 autosomal SNVs. Details of DNA extraction, genotyping, and imputation are documented elsewhere (Bycroft et al., 2018).

UKB also provides the values of principal components for each participant to adjust for population stratification; as suggested by UKB, the first 15 are used in most analyses.

Statistical analysis

Sample and variant exclusions were based on the quality control metrics provided by UK Biobank. Only participants of European ancestry were included in this study. For sample quality control, UKB performed heterozygosity checks by fitting a linear regression model with the first six principal components and outlined 968 outliers with an unusually high heterozygosity rate and missing call rate > 5%. The heterozygosity rate and missing call rate outliers were removed from the current study. Participants were also excluded for a discrepancy in biological sex and self-reported sex. Related pairs of participants were randomly removed one from every pair of participants with a kinship coefficient >0.08838835 (Manichaikul et al., 2010) which was calculated and provided by UKB. Based on the imputation information (INFO) scores and minor allele frequency information provided by UKB, an exclusion list containing SNPs with INFO score ≤ 0.5 was created. Duplicated SNVs were also removed before GWA analyses, resulting in 58,443,190 SNVs included in the final GWAS.

GWA analyses were conducted on autosomal chromosomes using a logistic regression model that assumed an additive mode of inheritance, implemented in PLINK 2.0 (Linux) (https://www.cog-genomics.org/plink/2.0/) (Chang et al., 2015). Two distinct association analyses were performed: (1) an association analysis without any covariates, and (2) an association analysis with age, sex, and the first 15 principal components (PC1-15) included as covariates to adjust for baseline age, sex and population stratification. In the analysis, we defined the threshold for genome-wide significance as the conventional level of $p<5x10^{-8}$ and a suggestive level of significance as $p<5x10^{-6}$. Manhattan plots and Q-Q plots were created in R 4.2.2 (https://www.r-project.org/about.html), using the packages "qqman" (Stephen, 2014) and "fastqq" (https://github.com/gumeo/fastqq). In addition, SNVs that were significant in previous GWASs reviewed by Gao et al. (2024) were also examined in our study to confirm replication of the results. The efficacy of genome-wide significant SNVs ($p<5x10^{-8}$) in predicting periodontitis were evaluated using Receiver Operating Characteristic (ROC) curves (Fawcett, 2006) in the package "ROCR" in R4.2.2 (Sing et al., 2005).

Functional annotation using FUMA 1.6.0

SNVs with minor allele frequency (MAF) >= 0.001 from the model including covariates were followed up for functional annotation in FUMA v.1.6.1 (Watanabe et al., 2017) using the function SNP2GENE. The default parameter in the SNP2GENE function was used with 1000 Genome Phase 3 as the reference panel for its largest number of reference SNVs.

SNVs were declared as independent significant loci if they reached $p < 5x10^{-8}$ and were in no more than weak linkage disequilibrium ($r^2<0.6$). Lead SNVs at a locus were further defined as the most significant among those with $r^2>0.1$. ANNOVAR (2017-07-17) was applied to estimate

the functional consequence of SNVs. MAGMA (v1.08) was implemented to conduct gene-based tests, gene-set tests and gene-property tests with the Genotype-Tissue Expression (GTEx) v8 53 tissue types. Details of the function of FUMA and analysis of SNP2GENE are available in (Watanabe et al., 2017).

Results

Following initial quality control, 376,611 European participants (mean age =57.6 years old, 53.76% women=68,482, controls=307,342) remained, of which 787 participants were dropped from the association analysis due to missing phenotype data. The sample characteristics are presented in Table 1.

Table 1. Sample characteristics

	Overall	Periodontitis	
	(n=376,611)	Yes (n = 68,482)	No (n = 307, 342)
Age, mean (SD)	57 (7.9)	56 (7.8)	58 (8.0)
Sex, n=Female (%)	202460 (53.8)	40834 (59.6)	161197 (52.4)
Missing data in phenotype (%)	787 (0.2)		

SNPs—Periodontitis Association: null model and covariates model

In the null model (no covariates), 8 SNVs reached genome-wide significance ($p<5x10^{-8}$) in relation to periodontitis (Table 2). After adjusting for 17 covariates (age, sex, PC1-15), three of the 8 SNVs remained genome-wide significant: rs775476621 (Odds $Ratio_{[T]}=3.08$, $p=1.01x10^{-8}$) and rs751014048 ($OR_{[G]}=3.07$, $p=1.04x10^{-8}$) on chromosome 11; rs149922301 ($OR_{[A]}=1.18$, $p=2.71x10^{-8}$) on chromosome 4. An additional SNV also reached significance in the covariates model: rs368467810 ($OR_{[TTTA]}=0.96$, $p=3.88x10^{-8}$) on chromosome 6. Overall, 674 SNVs (including four GWAS significant SNVs) reached our suggestive level of significance ($p<5x10^{-6}$) in the covariates model, and may also contribute to periodontitis.

Manhattan plots are presented in Figure 1 to show the association signals across the genome. Additionally, Q-Q plots compared the observed p-value with the expected p-value under the null hypothesis providing an indication of inflation in the GWAS results. A genomic inflation factor lambda = 0.98 for both models indicated no inflation in the results.

Table 2. Association results for SNVs with $p<5x10^{-8}$ for both models with and without covariates adjusted.

RS ID	Chr: Pos	A1	A2	A1 freq	Model	OR	95%CI	p
rs775476621	11: 34892994	T	A	0.00	no covariate model	3.05	(2.08, 4.47)	$1.04x10^{-8}$
					covariates model	3.08	(2.10, 4.52)	$1.01x10^{-8}$
rs751014048	11: 34991008	G	A	0.00	no covariate model	3.04	(2.08, 4.45)	$1.09x10^{-8}$
					covariates model	3.07	(2.09, 4.51)	$1.04x10^{-8}$
rs185910248	5: 124199710	A	G	0.01	no covariate model	0.86	(0.82, 0.91)	$1.78x10^{-8}$
					covariates model	0.86	(0.82, 0.91)	$6.37x10^{-8}$
rs28381639	6: 31837774	G	A	0.04	no covariate model	1.08	(1.05, 1.11)	$2.12x10^{-8}$
					covariates model	1.08	(1.05, 1.11)	$5.47x10^{-8}$
rs149922301	4: 10809690	A	G	0.01	no covariate model	1.18	(1.12, 1.26)	$2.62x10^{-8}$
					covariates model	1.18	(1.12, 1.26)	$2.71x10^{-8}$
rs34465217	6: 31880807	C	CA	0.06	no covariate model	1.07	(1.05, 1.1)	$4.02x10^{-8}$
					covariates model	1.07	(1.05, 1.10)	$5.99x10^{-8}$
rs35664695	6: 31929808	T	C	0.04	no covariate model	1.08	(1.05, 1.11)	$4.31x10^{-8}$
					covariates model	1.08	(1.05, 1.11)	$1.15x10^{-7}$
rs41315812	6: 31943232	T	C	0.04	no covariate model	1.08	(1.05, 1.11)	$4.76x10^{-8}$
					covariates model	1.08	(1.05, 1.11)	$1.25x10^{-7}$
rs368467810	6: 27849721	TTTA	T	0.36	no covariate model	0.96	(0.95, 0.98)	$1.59x10^{-7}$
					covariates model	0.96	(0.95, 0.98)	$3.88x10^{-8}$

RS ID, SNP reference identifier; Chr: Pos, Chromosome: Position; A1, effect allele; A2, non-effect allele; A1 freq, A1 frequency; OR, Odds Ratio; CI, Confidence Interval; p, p-value. Note. The SNVs significant in the covariates model were highlighted where dark green represents the SNVs with MAF >0.001, and light green represents MAF<0.00

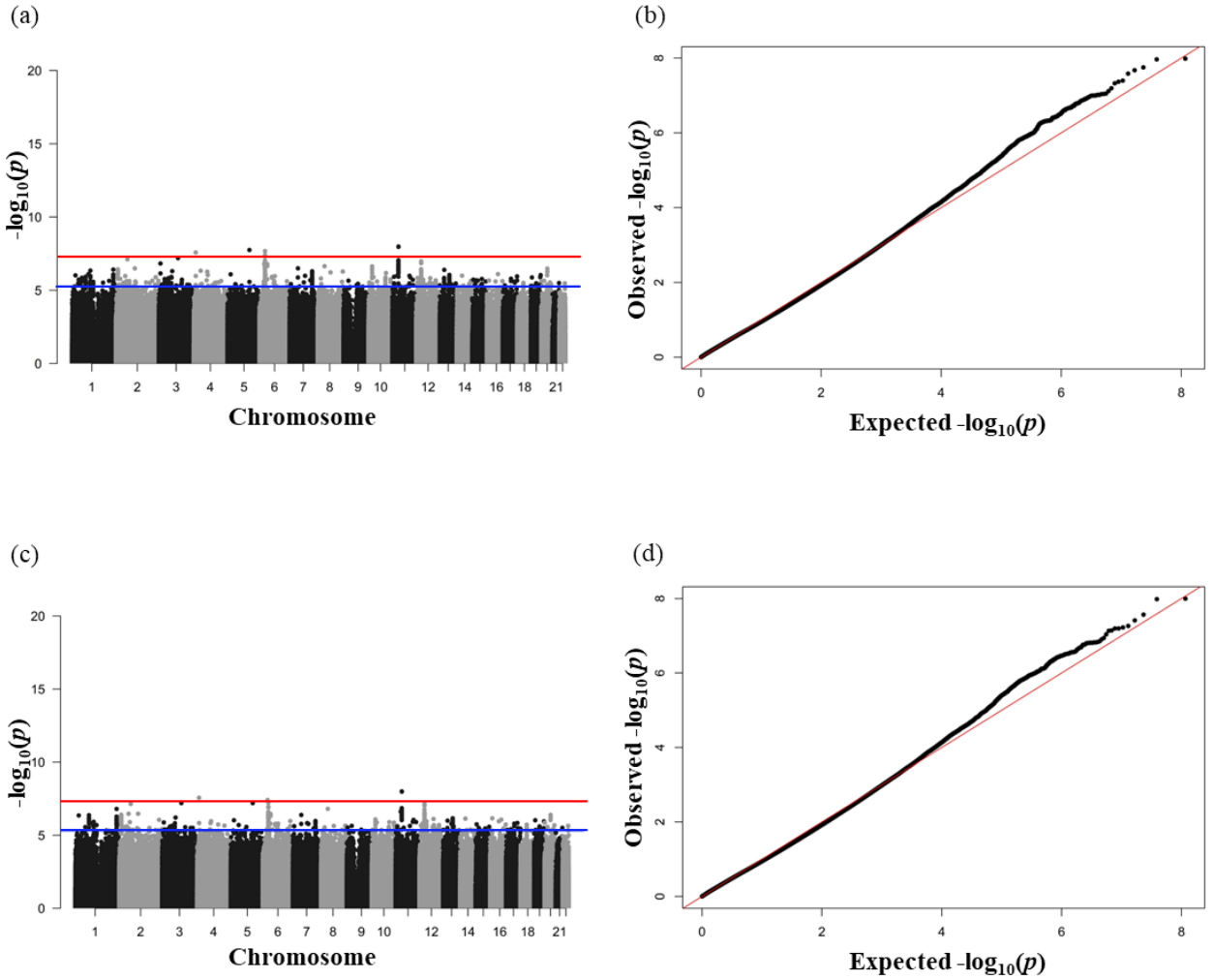


Figure 1. (a) Manhattan plot of the association analysis model without adjusting for any covariates, and (b) corresponding Q-Q plot. (c) Association analysis model with adjustment for age, sex and the first 1-15 PCs, and (d) Q-Q plots of the corresponding model (genomic inflation factor lambda = 0.98). Note: The red line in (a) and (c) Manhattan plot represents the genome-wide significant threshold ($p<5x10^{-8}$); while the blue line represents the suggestive line of significance ($p<5x10^{-6}$).

To cross-validate the accuracy and performance of our GWAS models, significant SNVs from both models without and with covariates (age, sex, first 15 PCs) were entered into the ROC curve estimate respectively. Better performance and model accuracy were detected for the covariates model (area under curve (AUC)= 0.572) compared to the no covariates model (AUC =0.503). Figure 2 shows the ROC curve and AUC for both association models.

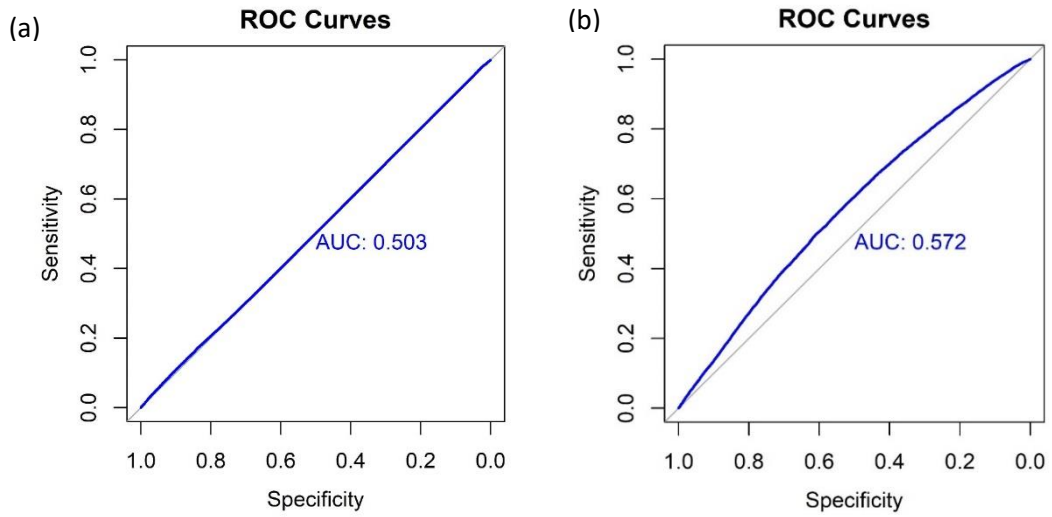


Figure 2. (a) ROC analysis including the significant SNVs from the association analysis model without adjustment for covariates. (b) ROC analysis including the four significant SNVs from the association analysis model with adjustment for age, sex, and the first 1-15 principal components. The grey line represents the reference line (AUC=0.500), and the area under curve (AUC) is marked along the reference line of both ROC curves.

In addition to the results from the current study, SNPs reported as associated with periodontitis in 13 previous studies reviewed by Gao et al. (2024), were also investigated. Based on the results from our GWAS with covariates, none of the previously reported SNPs reached $p<5x10^{-8}$ in our study (Appendix Table d.1).

FUMA analysis

In the covariates model, there were 18,849,429 SNVs with (MAF > = 0.001) imported into FUMA for further downstream functional analysis. Two leading and independent genome-wide significant SNVs (rs149922301, rs368467810) were identified, both located in intergenic regions, where 132 genes mapped to the leading SNVs, 31 of which are pseudogenes. Four genes were annotated to the two leading SNVs: (1) rs149922301*: RP11-61G19.1* (distance=110405) and *MIR572* (distance=560761); and (2) rs368467810: *HIST1H4L* (distance= 8432) and *HIST1H3L* (distance= 8472). The regional plot can be viewed in Appendix Figure c.1.

In MAGMA gene-based analysis, the input SNVs were mapped to 19,174 protein coding genes in which gene *CCDC91* was the only significant gene after Bonferroni correction ($p <= 2.608x10^{-6}$) (Appendix Figure c.2). In the gene set analysis, no gene was found to be significant after Bonferroni correction, with $p = 1$ after Bonferroni correction for all gene sets (Appendix Table c.2).

The results of MAGMA tissue-expression analysis suggested that the current periodontitis GWAS were not significantly enriched for expression in any tissue specificities. (Appendix Figure c.3).

Discussion

The current study used UKB data to investigate the genetic risk variants for self-reported periodontitis and found four significant loci associated with the self-reported periodontitis phenotype: rs775476621[T] ($p=1.01 x10^{-8}$) and rs751014048[G] ($p=1.04 x10^{-8}$) on chromosome 11, rs149922301[A] on chromosome 4 near the gene *RP11-61G19.1* ($p=2.71 x10^{-8}$) and rs368467810[TTTA] on chromosome 6 near the genes *HIST1H3L* ($p=3.88 x10^{-8}$). The gene analysis of GWAS results prioritised 132 genes. The tissue expression analysis did not show any significant expression.

When compared to prior GWAS investigations of periodontitis, only two SNPs reported in previous studies reached nominal significance ($p < 0.05/52 = 0.001$) after Bonferroni correction (see Appendix Table c.1) in our results: rs12461706 [T] identified in Shungin et al. (2019) (our study $p = 5.34x10^{-4}$), and rs11084095 [A] identified in Munz et al. (2019) ($p = 4.58x10^{-4}$). None of the genome-wide significant ($p < 5x10^{-8}$) SNPs in our study were previously associated with periodontitis. Additionally, most of the previously reported SNPs did not achieve statistical significance ($p < 0.05$) in our GWAS.

There are several reasons for the lack of significance of the previously reported SNPs and why the SNVs highlighted here have not been reported before. One reason is the differences in the ethnic groups studied. For example, SNPs (rs2392520[C] $p=4.17x10^{-6}$) discovered in a Japanese sample (Shimizu et al., 2015) were not statistically significant ($p<0.05$) in our sample ($p=0.52$). Similarly, most of the genome-wide significant ($p<5x10^{-8}$) and suggestive significant SNPs ($p<5x10^{-6}$) discovered in a Korean sample from Hong et al. (2015) were not significant in this study. The two replicated SNPs were both from the study using the European sample or using the European sample as part of overall sample (Munz et al., 2019; Shungin et al., 2019).

A second reason is the limited sample size in most prior studies ($n < 10,000$), which may not be sufficient for a GWAS to find “true risk/protective SNPs”. For example, Petty et al. (2023)

included only 879 mixed ethnicity participants (333 periodontitis cases) and did not find any genome-wide significant SNPs. Their suggestive SNPs (rs12800372 [C], $p<5x10^{-6}$) did not reach statistical significance in our study. To the best of our knowledge, there is only one study (Shungin et al., 2019) with a larger sample size than the current study and their finding of a genome-wide significant SNP (rs12461706 [T]) for periodontitis was consistent with the results of our study ($p= 5.34x10^{-4}$). Since Shungin et al. (2019) used UKB data as part of the overall sample, replicating their genome-wide significant SNPs is unsurprising. Another SNP we replicated was rs11084095 [A] from Munz et al. (2019) (with European participants (n =15,003).

The third reason for the previously reported SNPs not reaching the same level of significance or non-significant in this study might be due to differences in phenotype definition. For Shungin et al. (2019), Munz et al. (2019), and the present study, although we all included European samples and had sample sizes larger than 10,000, different measures were used for periodontitis and varying case definitions employed. Shungin et al. (2019) used mixed periodontitis measures and definitions as multiple datasets were included. Periodontitis measures included self-reported periodontitis, clinically diagnosed periodontitis and clinical examination measures, defined using criteria such as the CDC-AAP definition (Page and Eke, 2007) and the 1999 international workshop for the classification of periodontal diseases (Armitage, 1999). Munz et al. (2019) used mixed definitions as well which included the CDC-AAP definition (Page and Eke, 2007) and definition based on bone loss (>= 30%) or loss of attachment (>=4mm) solely. Our study utilised a self-reported periodontitis definition encompassing bleeding gums, painful gums and loose teeth. This definition varied from Shungin et al. (2019), who identified UKB participants with tooth loss only as a surrogate for periodontitis. These differences in case definitions are likely to have affected case and control classification in the final study population, thereby impacting the final results.

Despite the differences, the significance ($p<10^{-3}$) observed at the rs12461706 [T] locus from Shungin et al. (2019), and rs11084095 [A] from Munz et al. (2019) partially support the important role of gene *SIGLEC5*. Both rs12461706 [T] and rs11084095 [A] were annotated on *SIGLEC5* which functions innate immune systems and contributes to periodontitis (Mueller et al., 2022).

The functions of the SNPs identified in this study are not fully understood, especially SNPs rs775476621 and rs751014048 which had no functionality associated with them using FUMA. There are four genes annotated for the other two identified SNPs: (1) rs149922301: *RP11-61G19.1* (distance=110405) and *MicroRNA 572 (MIR572*) (distance=560761), and (2) rs368467810*: HIST1H4L* (distance= 8432) and *HIST1H3L* (distance= 8472). Gene *MIR572* is attributed to microRNA class which are short non-coding regulatory RNAs involved in the expression of more than 60% of human genes (Seven et al., 2014; Suer et al., 2019). Previous

studies have found that *MIR57*2 is upregulated in several types of cancer/malignancy, such as non-small cell lung cancer (Sun et al., 2022), ovarian cancer (Zhang et al., 2015), renal cell carcinoma (Pan et al., 2018), contributing to malignant development, poor prognosis, and shortened survival time. However, the regulatory role of *MIR572* in periodontitis is not clear and requires further investigation to establish how changes in expression of *MIR572* are related to periodontitis progression.

In addition to *MIR572, HIST1H4L* and *HIST1H3L* are histone genes within the family of protein-coding genes that regulate the DNA binding. Histone modification is essential to many biological processes including brain development and mental illness (Armitage, 1999) such as schizophrenia, bipolar disorder, autism spectrum disorder and depression (Wu et al., 2020). This corresponds to previous findings on associations between periodontitis and mental illness (Kang et al., 2022), and raising the possibility that the mental illness and periodontitis share similar pathological pathways. However, further exploration of the exact biological impact these genes have on periodontitis development is still necessary and, especially, little is known about the function of *RP11-61G19.1.* Future studies on the function of these four genes could improve understanding of the pathology and underlying mechanisms of periodontitis.

In the MAGMA gene analysis, the top gene from the MAGMA gene-based test—*CCDC91, Coiled-Coil Domain Containing 91 (CCDC91)*, is a protein coding gene that has been predicted to be involved in protein binding activities and Golgi to lysosome transportation. *CCDC91* has not been annotated to SNPs reported in previous periodontitis GWAS investigations and it is unclear the role it plays in periodontitis susceptibility. According to the GWAS catalog SNVs in *CCDC91* are associated with traits such as Body Mass Index (Justice et al., 2017). Interestingly, epidemiological evidence suggests that cardiometabolic disorders (King et al., 2022), as well as increased central obesity, are associated with tooth loss (Kang et al., 2019). The relevance of *CCDC91* for periodontitis and how this gene acts differently between other associated traits (i.e., Body Mass Index) and periodontitis still needs further investigation.

According to the ROC curve, both of our association analysis models provided some prediction. These results suggest the eight SNVs discovered in null models and four SNVs together with covariates (i.e., age, sex and PC1-15) provide modest to poor performance on predicting periodontitis. The no covariate model has AUC=0.503 and covariate model has AUC=0.572 respectively. It is likely that periodontitis is a multifaceted disease involving environmental factors, microorganisms and genetic influences. It is of note that model prediction and performance were improved by adding covariates (i.e., age, sex, PC1-15) which underlines the importance of non-genetic factors in periodontitis.

The current study has several strengths. First, the use of high-quality genetic data with standardised quality control increases the reliability of the results. The sample size was larger than most previous GWAS investigations into periodontitis and drew from an ethnically homogeneous population, to maximise power and minimise the likelihood of false positive results. Second, our functional analysis suggested some promising avenues for further exploration of the genetic causes of periodontitis.

This study also has some limitations. Firstly, participants with missing data were not utilised in the association analysis, and we cannot assess whether they are missing at random. Secondly, this study used self-reported oral health status as a proxy for periodontitis, which is not as ideal as clinical measures. Additionally, our results are based on population with European ancestry, not including ethnicities minorities. Another limitation of this study may impact on the significance of findings is that the current study included an older population where the impact of lifestyle, comorbidities or any other potential risk factors may contribute more than genetic factor, while for younger population genetic variants may play a much stronger role in periodontitis. Therefore, interpretation with caution is needed.

Conclusion

The current study has identified four significant loci associated with periodontitis in a European population using a large, high-quality dataset from the UK Biobank. Future studies are needed to further explore the loci identified here to gain a better understand the pathology of periodontitis.

ACKNOWLEDGEMENT

This article has not been published or being considered for publication elsewhere, in whole, in part, except as an abstract.

Data availability: UK Biobank data were used in the current study and application is required for access.

Conflict of interest: The authors declare no conflict of interest.

Author contributions: CG contributed to data analysis, interpretation, and manuscript write-up. JK, HL, MI, DTB contributed to data analysis. JK, JW, SP, DTB, MI and DB contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Funding statement: Chenyi Gao is supported by the Alzheimer's Society (Alzheimer's Society Heather Corrie Ph.D. studentship: 546 (AS-PhD-19b-012)). The views expressed are those of the author(s) and not necessarily those of the Alzheimer's Society. The Alzheimer's Society did not involve in design and conduct this study, as well as data collection or preparation, review, approval, and decision to submit the manuscript. Jianhua Wu and Harriet Larvin are supported by Barts Charity (MGU0504). Mark Iles is supported in part by the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

References

Abbood HM, Hinz J, Cherukara G, Macfarlane TV. 2016. Validity of self-reported periodontal disease: A systematic review and meta-analysis. Journal of periodontology (1970). 87(12):1474-1483. doi: 10.1902/jop.2016.160196.

Armitage GC. 1999. Development of a classification system for periodontal diseases and conditions. Ann Periodontol. 4(1):1-6. doi: 10.1902/annals.1999.4.1.1.

Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J et al. 2018. The uk biobank resource with deep phenotyping and genomic data. Nature (London). 562(7726):203-209. doi: 10.1038/s41586-018-0579-z.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation plink: Rising to the challenge of larger and richer datasets. Gigascience. 4(1):7-7. doi: 10.1186/s13742-015-0047-8.

Eke PI, Page RC, Wei L, Thornton-Evans G, Genco RJ. 2012. Update of the case definitions for population-based surveillance of periodontitis. Journal of periodontology (1970). 83(12):1449-1454. doi: 10.1902/jop.2012.110664.

Fawcett T. 2006. An introduction to roc analysis. Pattern recognition letters. 27(8):861-874. doi: 10.1016/j.patrec.2005.10.010.

Gao C, Iles M, Larvin H, Bishop DT, Bunce D, Ide M, Sun F, Pavitt S, Wu J, Kang J. 2024. Genome-wide association studies on periodontitis: A systematic review. PLOS ONE. doi.

Hong KW, Shin MS, Ahn YB, Lee HJ, Kim HD. 2015. Genomewide association study on chronic periodontitis in korean population: Results from the yangpyeong health cohort. J Clin Periodontol. 42(8):703-710. doi: 10.1111/jcpe.12437.

Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, Barata L, Deng X, Czajkowski J, Hadley D et al. 2017. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nature Communications. 8(1):14977. doi: 10.1038/ncomms14977.

Kang J, Palmier-Claus J, Wu J, Shiers D, Larvin H, Doran T, Aggarwal VR. 2022. Periodontal disease in people with a history of psychosis: Results from the uk biobank population-based study. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12798.

Kang J, Smith S, Pavitt S, Wu J. 2019. Association between central obesity and tooth loss in the non-obese people: Results from the continuous national health and nutrition examination survey (nhanes) 1999–2012. Journal of clinical periodontology. 46(4):430-437. doi: 10.1111/jcpe.13091.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

King S, Chow CK, Eberhard J. 2022. Oral health and cardiometabolic disease: Understanding the relationship. Internal medicine journal. 52(2):198-205. doi: 10.1111/imj.15685.

Larvin H, Baptiste PJ, Gao C, Muirhead V, Donos N, Pavitt S, Kang J, Wu J. 2024. All-cause and cause-specific mortality in us adults with periodontal diseases: A prospective cohort study. Journal of clinical periodontology. 51(9):1157-1167. doi: 10.1111/jcpe.14002.

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2022a. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. doi.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021. Risk of incident cardiovascular disease in people with periodontal disease: A systematic review and meta-analysis. Clinical and experimental dental research. 7(1):109-122. doi: 10.1002/cre2.336.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2022b. Periodontitis and risk of immune-mediated systemic conditions: A systematic review and meta-analysis. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12812.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Frontiers in medicine. 7:604980-604980. doi: 10.3389/fmed.2020.604980.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics. 26(22):2867-2873. doi: 10.1093/bioinformatics/btq559.

Mao W, Wu B, Chi I, Yang W, Dong X. 2023. Experiences of discrimination and oral health-related quality of life among foreign-born older chinese americans: Does resilience play a mediating role? Community dentistry and oral epidemiology. 51(2):187-193. doi: 10.1111/cdoe.12723.

Mueller R, Chopra A, Dommisch H, Schaefer AS. 2022. Periodontitis risk variants at siglec5 impair erg and mafb binding. Journal of dental research. 101(5):551-558. doi: 10.1177/00220345211049984.

Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, Teumer A, Holtfreter B, Kocher T, Bruckmann C et al. 2019. Meta-analysis of genome-wide association studies of aggressive and chronic periodontitis identifies two novel risk loci. Eur J Hum Genet. 27(1):102-113. doi: 10.1038/s41431-018-0265-5.

Nazir MA. 2017. Prevalence of periodontal disease, its association with systemic diseases and prevention. International journal of health sciences. 11(2):72-80. doi.

Nibali L, Bayliss-Chapman J, Almofareh SA, Zhou Y, Divaris K, Vieira AR. 2019. What is the heritability of periodontitis? A systematic review. J Dent Res. 98(6):632-641. doi: 10.1177/0022034519842510.

Page RC, Eke PI. 2007. Case definitions for use in population-based surveillance of periodontitis. Journal of periodontology. 78(7):1387-1399. doi: 10.1902/jop.2007.060264.

Page RC, Offenbacher S, Schroeder HE, Seymour GJ, Kornman KS. 1997. Advances in the pathogenesis of periodontitis: Summary of developments, clinical implications and future directions. Periodontology 2000. 14(1):216-248. doi: 10.1111/j.1600-0757.1997.tb00199.x.

Pan X, Quan J, Li Z, Zhao L, Zhou L, Jinling X, Weijie X, Guan X, Li H, Yang S et al. 2018. Mir-566 functions as an oncogene and a potential biomarker for prognosis in renal cell carcinoma. Biomedicine & pharmacotherapy. 102:718-727. doi: 10.1016/j.biopha.2018.03.072.

Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, Flemmig TF, Garcia R, Giannobile WV, Graziani F et al. 2018. Periodontitis: Consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. J Clin Periodontol. 45 Suppl 20:S162-s170. doi: 10.1111/jcpe.12946.

Petty LE, Silva R, Chaves de Souza L, Vieira AR, Shaw DM, Below JE, Letra A. 2023. Genome-wide association study identifies novel risk loci for apical periodontitis. Journal of endodontics. doi: 10.1016/j.joen.2023.07.018.

Preshaw PM, Alba AL, Herrera D, Jepsen S, Konstantinidis A, Makrilakis K, Taylor R. 2011. Periodontitis and diabetes: A two-way relationship. Diabetologia. 55(1):21-31. doi: 10.1007/s00125-011-2342-y.

Seven M, Karatas OF, Duz MB, Ozen M. 2014. The role of mirnas in cancer: From pathogenesis to therapeutic implications. Future Oncology. 10(6):1027-1048. doi: 10.2217/fon.13.259.

Shimizu S, Momozawa Y, Takahashi A, Nagasawa T, Ashikawa K, Terada Y, Izumi Y, Kobayashi H, Tsuji M, Kubo M, Furuichi Y. 2015. A genome-wide association study of periodontitis in a japanese population. J Dent Res. 94(4):555-561. doi: 10.1177/0022034515570315.

Shungin D, Haworth S, Divaris K, Agler CS, Kamatani Y, Keun Lee M, Grinde K, Hindy G, Alaraudanjoki V, Pesonen P et al. 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.

Sing T, Sander O, Beerenwinkel N, Lengauer T. 2005. Rocr: Visualizing classifier performance in r. Bioinformatics. 21(20):3940-3941. doi: 10.1093/bioinformatics/bti623.

Stephen DT. 2014. Qqman: An r package for visualizing gwas results using q-q and manhattan plots. bioRxiv.005165. doi: 10.1101/005165.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M et al. 2015. Uk biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS medicine. 12(3):e1001779-e1001779. doi: 10.1371/journal.pmed.1001779.

Suer I, Guzel E, Karatas OF, Creighton CJ, Ittmann M, Ozen M. 2019. Micrornas as prognostic markers in prostate cancer. The Prostate. 79(3):265-271. doi: 10.1002/pros.23731.

Sun B, Zhao J, Shao ZY. 2022. Mir-572 promotes the development of non-small cell lung cancer by targeting klf2. Eur Rev Med Pharmacol Sci. 26(9):3083-3090. doi: 10.26355/eurrev_202205_28725.

Watanabe K, Taskesen E, van Bochoven A, Posthuma D. 2017. Functional mapping and annotation of genetic associations with fuma. Nature communications. 8(1):1826-1811. doi: 10.1038/s41467-017-01261-5.

Oral health-- periodontal (gum) disease. 2023. https://www.who.int/news-room/fact-sheets/detail/oral-health#:~:text=Periodontal%20(gum)%20disease&text=Severe%20periodontal%20diseases%20are%20estimated,than%201%20billion%20cases%20worldwide.: World Health Organisation; [accessed 2023 15/08/2023].

Wu Y, Cao H, Baranova A, Huang H, Li S, Cai L, Rao S, Dai M, Xie M, Dou Y et al. 2020. Multi-trait analysis for genome-wide association study of five psychiatric disorders. Translational psychiatry. 10(1):209-209. doi: 10.1038/s41398-020-00902-6.

Zhang X, Liu J, Zang D, Wu S, Liu A, Zhu J, Wu G, Li J, Jiang L. 2015. Upregulation of mir-572 transcriptionally suppresses socs1 and p21 and contributes to human ovarian cancer progression. Oncotarget. 6(17):15180-15193. doi: 10.18632/oncotarget.3737.

Chapter 4: Causal relationships between periodontitis and dementia

Chapter 4.1 Mendelian Randomisation

Title: Investigating the causation between periodontitis and dementia: a Two-sample Mendelian Randomisation study using UK Biobank data

Authors:

Chenyi Gao[1], Mark M. Iles[2,3] , Harriet Larvin[5], David Bunce[6], Bei Wu[7], Huabin Luo[8], Susan Pavitt[1], Jianhua Wu[1, 5,] David Timothy Bishop[4*,] Jing Kang[1, 9]*

Ready to submit to Journal

Affiliations:

[1] School of Dentistry, University of Leeds, Leeds, UK

[2] Leeds Institute for Data Analytics, University of Leeds, Leeds, UK

[3] NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK

[4] Leeds Institute of Medical Research, School of Medicine, University of Leeds, Leeds, UK

[5] Wolfson Institute of Population Health, Queen Mary, University of London, London, UK

[6] School of Psychology, University of Leeds, Leeds, UK

[7] Rory Meyers College of Nursing, New York University, New York, US

[8] Department of Public Health, East Carolina University, Greenville, US

[9] Oral Clinical Research Unit, Faculty of Dentistry Oral Craniofacial Sciences, King's College London

Abstract

Objective and aims: Numerous studies suggest that periodontitis might be associated with cognitive function and the development of dementia, however, many inconsistency and limitations from observational studies exist and causality is inconclusive. This study aims to explore the potential causal effect of periodontitis on all-cause dementia.

Methods: A two-sample Mendelian Randomisation (MR) analysis based on UK Biobank (UKB) data (n= ~500,000) has been implemented, where participants were divided into two independent groups. The exposure is the self-reported periodontitis, and the outcome is all-cause dementia measured by both clinical diagnoses based on ICD10 and ICD9 codes, and self-reported dementia. Four different sets of genetic instruments were developed based on four thresholds: main approach ($p < 5 \times 10^{-8}$), alternative approach 1 ($p < 5 \times 10^{-6}$), alternative approach 2 ($p < 10^{-4}$), and alternative approach 3 (best-fit p-value threshold calculated by polygenic risk score analysis). These were analysed to cross-check the association and assess the robustness of results (The number of genetic loci included varied from 3 to 1020). The causal association between periodontitis and dementia was assessed using conventional Inverse-Variance Weighted (IVW) MR, along with MR-Egger regression, Weighted Median, and Mode-Based Estimate methods.

Results: Genetic instruments included in all analytical approaches passed the Mendelian Randomisation (MR) assumption check for instrument relevance. The majority of the results suggested no causal association between periodontitis and dementia. However, in the main approach IVW model (coefficient beta: -0.816, 95% confidence interval [-1.617, -0.015]) and in alternative approach 2, the weighted median model (beta: 0.077, 95% CI [0.006, 0.149]) suggested the potential for a causal relationship between periodontitis and dementia.

Conclusion: The results showed inconsistent evidence of a causal link between periodontitis and dementia based on UKB data. Future studies with clinically defined periodontitis, high-quality data, and larger sample sizes are needed to confirm whether a causal link exists between periodontitis and dementia.

Introduction

Dementia, as one of the leading causes of death and major causes of disability and dependency worldwide, has multi-dimensional effects from the impact of physical and psychological changes and the challenges to social interactions faced by patients, their caregivers and the healthcare system (WHO, 2023). Since there is a growing number of patients estimated worldwide, reaching 139 million in 2050 (ADI, 2022), and the treatment for effectively curing dementia remain very limited, exploring modifiable risk factors that contributing to dementia development becomes increasingly important.

Periodontitis is potentially a risk factor for dementia, as suggested by many observational studies, both longitudinal and cross-sectional (Gao et al., 2023; Kang et al., 2019; Larvin et al., 2022), which have indicated correlations between the two. However, with limited direct biomedical evidence and common confounders (e.g., smoking, nutritional intake) affecting observational studies, it remains unclear whether periodontitis can cause dementia. There also remains the possibility that reverse causation which was suggested by the bidirectional association between periodontitis and dementia (Gao et al., 2023; Kang et al., 2020; Kang et al., 2019). Further evidence on the causal association between periodontitis and dementia is needed to inform a more comprehensive dementia prevention strategy.

Mendelian randomisation (MR), is a statistical method that uses genetic variants as instruments to examine causal relationships. It is not prone to reverse causation if all assumptions are valid and is minimally impacted by confounding effects. It can contribute to revealing the causal nature underlying the association between periodontitis and dementia. To date, there have been two MR studies investigated causation between periodontitis and cognitive impairments or Alzheimer's Diseases using public databases, and neither study found evidence of a causal link (Deng et al., 2024; Hu et al., 2024). It is noted that these two studies used publicly available summary statistics for selecting periodontitis genetic instruments provided in the FinnGen database and by the GLIDE consortium, but so far there has been no consensus on the genetic instruments of periodontitis selected between the two MR studies. Our previous study systematically reviewed the genetic variants of periodontitis from all high quality genome-wide association studies (GWAS) of periodontitis, and no common genetic variants were found, indicating the heterogeneity of the genetic instruments in periodontitis (Gao et al., 2024). UKB data (https://www.ukbiobank.ac.uk/) is of high quality, and is an ideal resource to investigate, as it contains information of participants' genetic variants, oral health conditions, and clinically diagnosed or self-reported dementia cases. In addition, selection of the genetic variants should be cautious to avoid potential false positives, and applying various thresholds and novel methods such as polygenic risk score (PRS) should be considered in the selection of genetic variants for periodontitis (Burgess and Thompson, 2013).

In this study, we used UKB data and applied two-sample Mendelian randomisation approach by dividing UKB participants into two independent groups: a periodontitis group (and controls) and a dementia group (plus controls), based on participants' diseases status. We also choose the genetic instruments of periodontitis with two methods: (1) The traditional approach: conventional GWAS significance threshold ($p<5x10^{-8}$), suggestive threshold ($5x10^{-6}$) and (10^{-4}); (2) the PRS approach for selecting genetic instruments at the best-fit p-value threshold, which reduces bias associated with weak instruments (Burgess and Thompson, 2013; Palmer et al., 2012). By utilising different methods, our results could provide a more comprehensive and robust conclusion.

Method

1. Sample and Data Resources

This study included only the European sample (n=409,548) from the final release of UKB genetic data (UK Biobank project ref 54633). The European ancestry was determined by self-reported ethnicities and its similarity with genetic ancestry was computed by principal components analysis (Bycroft et al., 2018). Participants who registered for withdrawal (n=139) at the time of analysis were excluded from this study.

2. Study design

2.1 Diseases definition

Periodontitis was measured by self-reported dental health questions from UKB (variable code: 100538). Participants who had reported "painful gum, bleeding gum or loose tooth" were classified as 'periodontitis cases', while participants who reported "mouth ulcer, toothache, dentures or none of the above" were classified as 'not having periodontitis'. The self-reported measurement of periodontitis is valid and a reliable surrogate for periodontitis (Abbood et al., 2016; Papapanou et al., 2018), and the same UKB periodontitis definition has been used in previous studies (Kang et al., 2022; Larvin et al., 2020). Participants with no data entry or who reported missing data were classified as missing data.

Dementia was indicated by a clinical diagnosis based on UKB ICD9 or ICD10 diagnosis at all data collection time points, primary care Read codes and/or self-reported dementia. Dementia coding is available in UKB released algorithmically-defined outcomes version 2 (https://biobank.ndph.ox.ac.uk/showcase/refer.cgi?id=460). Dementia cases were identified from primary care data using the Read code list from a previous study (Wilkinson et al., 2019). Participants who had reported any forms of dementia were classified as 'dementia cases' and those who have reported other diseases or no other diseases were classified as 'not having

dementia'. Participants with no data entry or who reported missing data were classified as missing data.

Out of the total 376,611 eligible participants classified as European and who passed the pre-GWAS sample QC (described in the next 2.3 GWAS section), there are 891 participants had periodontitis and dementia, 5,133 had dementia only, 67,591 had periodontitis only. There were 302,957 participants with data entry for dementia or periodontitis measurement (i.e., not reporting missing data) but who did not report either periodontitis or dementia considered as controls.

2.2 Sample allocation

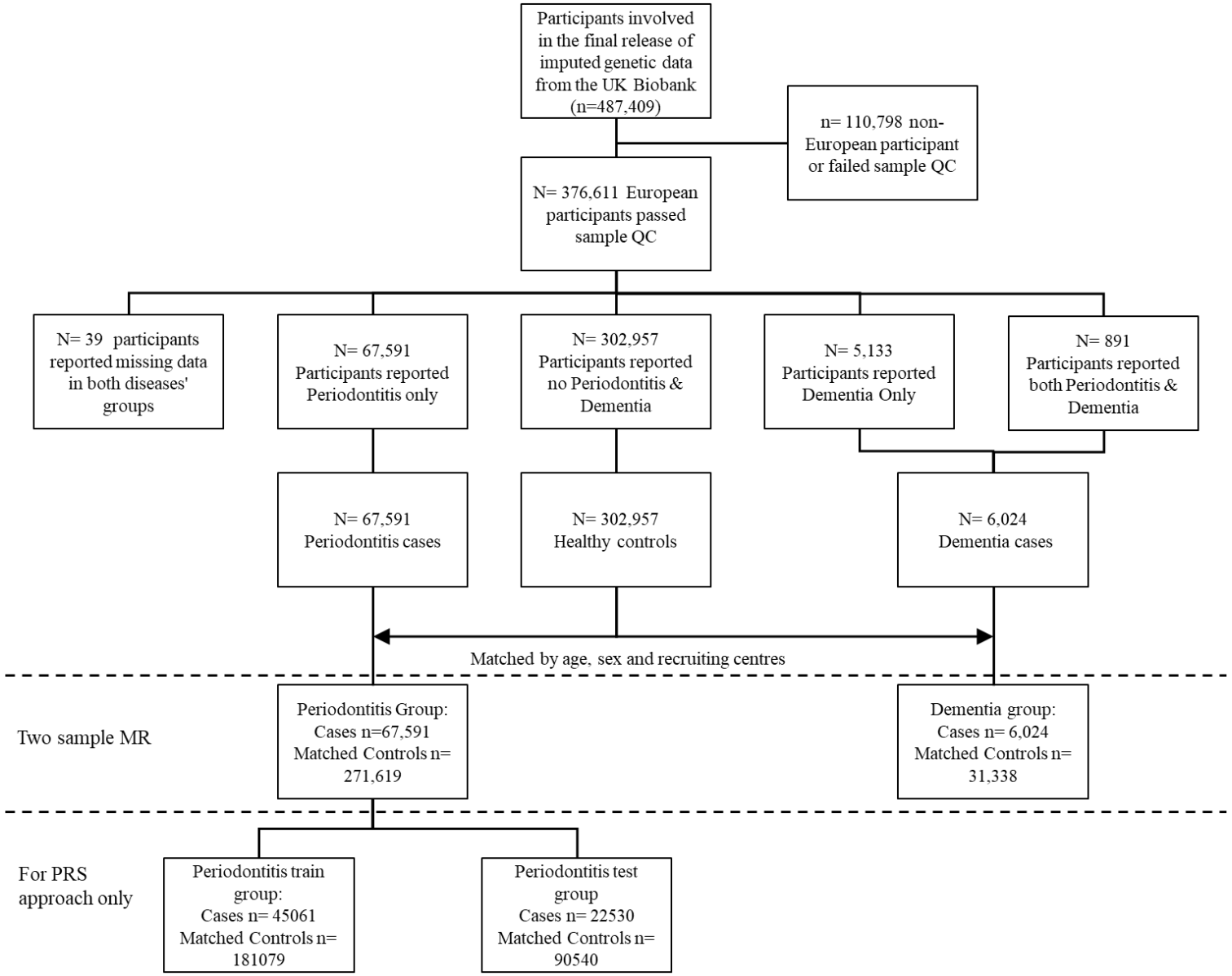
First of all, participants with periodontitis were allocated to the periodontitis group and those with dementia to the dementia group to set up the two sample MR analysis (Figure 1). Participants who reported both diseases were allocated to the dementia group to maximise the dementia group sample size. No participants were allocated to both sample groups to avoid sample overlap and subsequent weak instrument bias (Burgess et al., 2019). Participants who reported as unaffected by both periodontitis and dementia were allocated to the periodontitis or dementia group by matching age, sex and recruiting centre at baseline measures. The number of controls was more than three times that of cases in each group to optimise statistical power (Schlesselman and Stolley, 1982).

To facilitate the PRS analysis, the periodontitis sample was further allocated into a periodontitis "training" group and a periodontitis "test" group by randomly selecting cases and controls. The case -control ratio of 1:3 was retained in the periodontitis training and periodontitis test groups (Figure 1).

2.3. GWAS in UKB

UKB genotyped 487,409 participants using two genotyping arrays and performed imputation resulting in 93,095,623 autosomal variants using UK10K, 1000 Genomes phased and Haplotype Reference Consortium reference panels (Bycroft et al., 2018). We performed pre-GWAS quality control on both the sample and genetic variants in the European ancestry participants.

Figure 1. Sample allocation flow chart



The GWA was performed for periodontitis status in the periodontitis training group and the periodontitis group, as well as for dementia status in the dementia group. The aim of conducting GWAS in the periodontitis group and the periodontitis train group separately is to facilitate the two distinct genetic instruments selection methods: (1) Use of traditional fixed p-value threshold to select the genetic instrument in the periodontitis group; (2) Use of best-fit p-value threshold computed by PRS analysis in the periodontitis training and test groups. A logistic regression model assuming an additive mode of inheritance was assumed across all genome-wide association analysis performed, as well as adjustment for covariates: age, sex, and the first 15 principal components as provided by UKB.

The GWA results of periodontitis from the periodontitis train group were further used to estimate how well these single nucleotide variants (SNV)s can predict periodontitis in the periodontitis test group and also to estimate the best fit p-value threshold for the PRS approach.

2.5 Two Sample Mendelian Randomisation and genetic instruments

Two-sample MR analysis was performed between the periodontitis group and the dementia group (traditional approach), as well as between the periodontitis train and the dementia group (PRS approach). Following GWA, clumping was performed to obtain independently significant SNVs (threshold p value $< 10^{-4}$, Linkage Disequilibrium (LD) $r^2 < 0.1$ and window size 250kb). These SNVs were also potential genetic instruments. To further identify valid genetic instruments, various selection approaches were utilised:

1) Our main approach: Use the genetic instrument based on combining SNVs that reached conventional GWAS significance ($p<5x10^{-8}$) selected from the periodontitis group GWAS.

2) Alternative approach 1: Base the genetic instrument on those SNVs that reached a suggestive level of significance ($p<5x10^{-6}$) within the periodontitis group GWAS.

3) Alternative approach 2: SNVs that reached a suggestive level of significance ($p<10^{-4}$) were selected from the periodontitis group GWAS. This approach was utilised together with alternative approach 1 to cross-check differences with the main approach and the robustness of any suggested causal relationship,

4) Alternative approach 3: Use the genetic instrument that achieved the best fit p-value threshold based on PRS analysis were selected from periodontitis train group GWAS. The best fit p value threshold potentially provided a better explanation of periodontitis variance in the sample, which could again provide an extra sensitivity analysis to check the consistency and robustness of results.

3. statistical analysis

3.1 Mendelian Randomisation

Two-sample MR analyses using these four models with four distinct genetic instruments were utilised in this study (Figure 2).

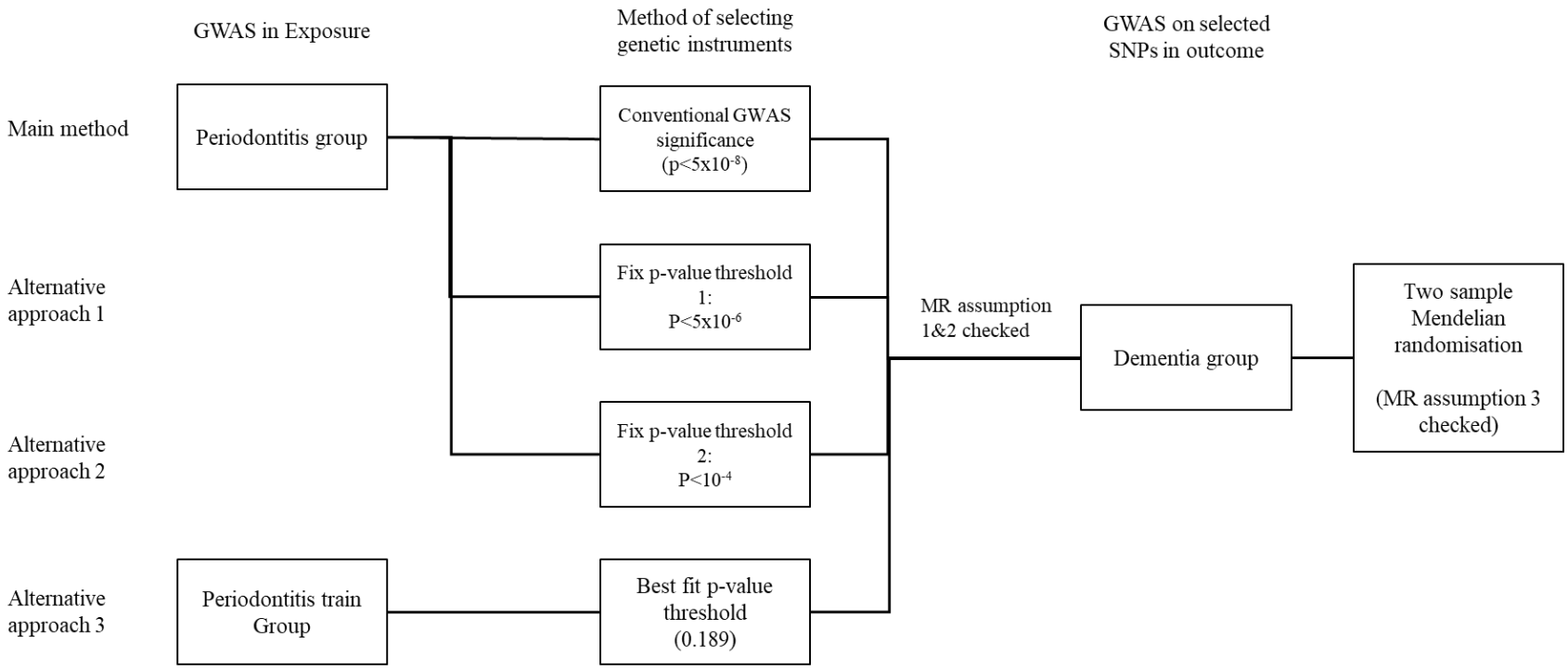
Four MR methods were employed, including Inverse-variance weighted (IVW) MR (Burgess et al., 2013; Burgess et al., 2019), MR-Egger regression (Bowden et al., 2015; Bowden et al., 2016b; Burgess et al., 2019), Weighted median (Bowden et al., 2016a; Burgess et al., 2019), and mode-based estimate (Hartwig et al., 2017) to test and cross-validate the effect of periodontitis on dementia. These different methods reflect a variety of assumptions and scenarios about the MR analysis.

In addition to the methods above, MR on each individual SNVs was performed for the strongest genetic instruments (selected based on $p<5x10^{-8}$) to gain better insight into their effects in the association.

The GWA and clumping were implemented in PLINK 2.0 (Linux) (https://www.cog-genomics.org/plink/2.0/) (Chang et al., 2015). Manhattan plots and Q-Q plots were created in R 4.4.2 (https://www.r-project.org/about.html), using the packages “qqman” (Stephen, 2014), and “fastqq” (https://github.com/gumeo/fastqq). The PRS analysis was performed using PRSice2 (https://choishingwan.github.io/PRSice/) (Choi and O'Reilly, 2019). The MR analysis was conducted in R 4.4.2 using the package “TwoSampleMR” (Hemani et al., 2017; Hemani et al., 2018) and “MendelianRandomization” (Patel et al., 2023; Yavorska and Burgess, 2017)

Figure 2. three main MR analysis method in this study

(a)



(b)

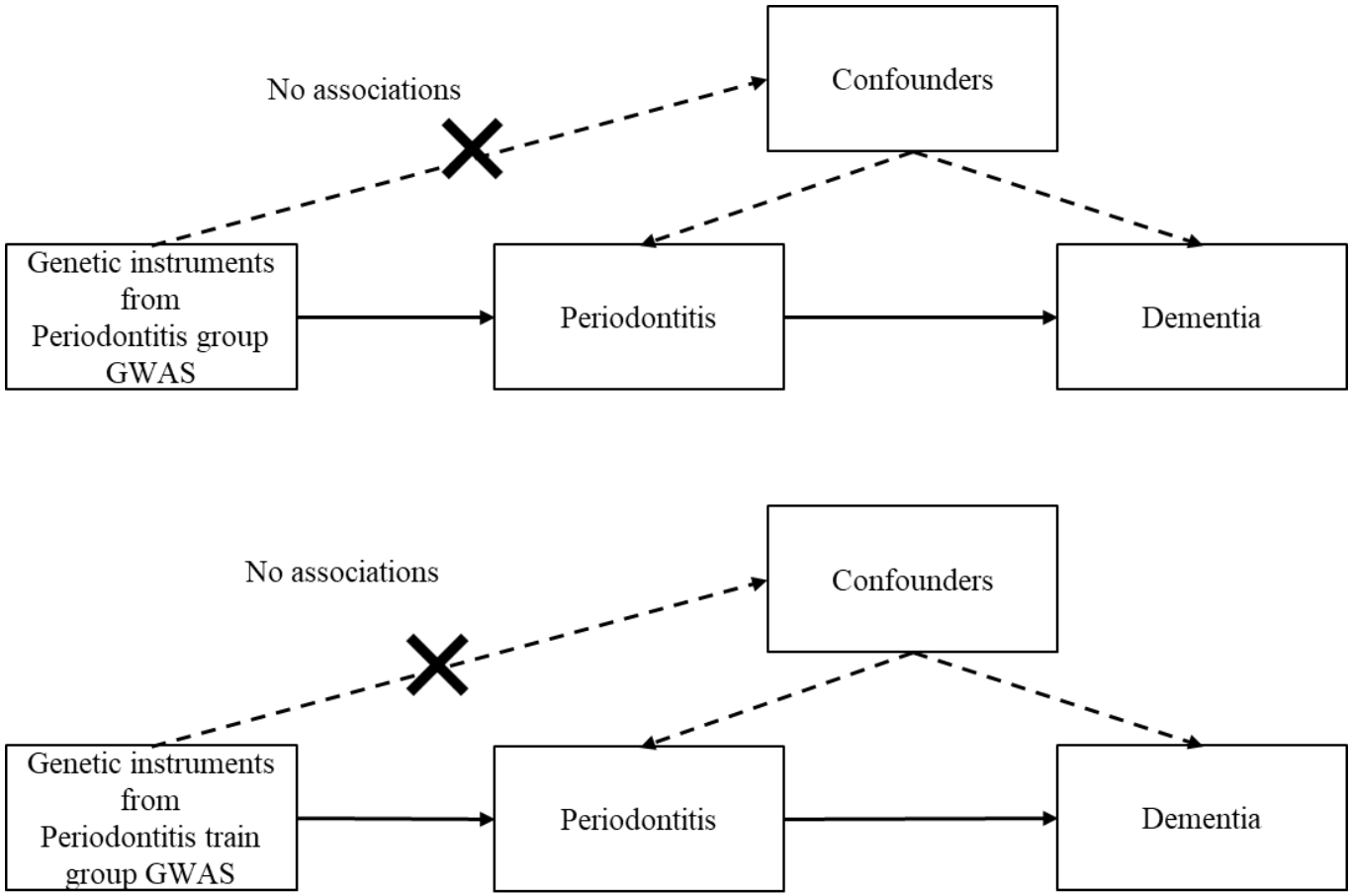


Figure 2. (a) Demonstrates the MR analysis procedures for all approaches utilised in the current study from GWAS in periodontitis, genetic instrument selection, GWAS in dementia, and finally two-sample MR. (b) MR models indicated in our study; the cross sign indicated there is no association allowed between genetic instruments and confounders.

3.2 Check on three main MR assumptions

There are three essential MR assumptions that must not be violated to obtain valid and robust results. We have performed several tests and calculated several statistics to ensure our genetic variants fulfil the assumptions.

1. Assumption 1: The genetic instrument should be relevant (i.e., strong and robustly associated) with the risk factor of interest (Davies et al., 2018). To ensure the relevance of genetic instrument to periodontitis, strict p-value thresholds (i.e., $p<5x10^{-8}$, $p<5x10^{-6}$, and $p<10^{-4}$, best fit p value threshold calculated by PRSice2) were used to choose genetic instruments. For all models, F-statistics were calculated to assess the strength of genetic instruments as well. An F-statistic of at least 10 is regarded as a minimal criterion for limited instrument bias.
2. Assumption 2: the genetic instruments should not be associated with any potential confounding variables (Davies et al., 2018). The association with potential confounders was checked by regressing the selected instrument in each model on several confounders (socioeconomic status, smoking status, alcohol drinking status, blood level of C-reactive protein) with adjustment for age, sex and PC1-15. Significant association with confounders after adjusting for multiple testing using Bonferroni correction indicates a violation of assumption 2 and the SNV(s) were removed from further analysis
3. Assumption 3: The genetic instruments can only influence outcomes via exposure (Davies et al., 2018). MR-Egger regression intercept was used to assess horizontal pleiotropy.

Results

1. Sample characteristics

Sample quality control (QC) included excluding: sex mismatch between biological sex and self-reported sex, 968 participants with high SNV missing rates (>5%) and unusually high heterozygosity rate as documented by UKB, plus related participants (kinship coefficient > 0.08838835 (Manichaikul et al., 2010)), where one participant was dropped randomly from every related pair. The QC on genetic variants included excluding SNVs with minor allele frequency <0.001 and those with imputation information score <= 0.5, resulting in 18,849,429 SNVs from the autosomal chromosomes included in association analysis. As a result of this QC 376,611 European participants from UKB passed sample quality control.

Thirty-nine UKB participants were removed from the analysis due to missing data in both periodontitis and dementia measurements. Among the remaining 376,572 participants (mean age 70.45, standard deviation 7.89), 18.2% of participants had periodontitis and 1.6% of them were diagnosed or self-reported all-cause dementia. Sex was balanced in the remaining sample, with slightly more females present (53.8%). By allocating participants into dementia and periodontitis groups, participants in dementia groups had a higher mean age (the dementia group mean = 76.54 ± 5.3 SD vs. periodontitis group mean = 69.78 ± 7.84 SD) and lower percentage of female participants (46.8% of participants in the dementia group vs. 54.5% of participants in the periodontitis group). 891 participants (14.9% of dementia cases) experienced both periodontitis and dementia and were allocated to dementia group. Sample characteristics for each group's participants are described in Table 1. Sample characteristics of the periodontitis group and the periodontitis train group including selected confounders' summary statistics were included in Appendix Table d.1.

Table 1. Sample characteristics

	Overall	Periodontitis Group		Dementia Group	
		Case	Control	Case	Control
N	376,572	67,591	271,619	6,024	31,338
Age (Mean (SD))	70.45 (7.89)	69.11 (7.75)	69.95 (7.85)	76.48 (5.30)	76.55 (5.30)
Sex = Female (%)	202,440 (53.8)	40,382 (59.7)	144,554 (53.2)	2,843 (47.2)	14,661 (46.8)
Periodontitis = Yes (%)	68,482 (18.2)			891 (14.9)	0 (0.0)
Dementia = Yes (%)	6,024 (1.6)	0 (0)	0 (0)		

SD, standard deviation

2. GWAS results and clumping

To enable MR analysis, GWAS were performed in the periodontitis group and the periodontitis train group.

In the periodontitis group GWAS, five SNVs reached conventional GWAS significance ($p<5x10^{-8}$). Clumping resulted in 1,020 leading significant SNVs (maximum $p=10^{-4}$); three of the GWAS significant SNVs remained. The Manhattan plot and Q-Q plot can be viewed in Appendix Figure

d.1. In the periodontitis training group GWAS, 970 independent significant SNVs were obtained after clumping (Appendix Figure d.2). The F-statistic in all MR approaches > 10 suggests sufficient strength of genetic instruments (assumption 1).

Association between the confounders and the independent SNPs (i.e., potential genetic instruments) revealed that 27 SNVs from periodontitis group GWAS and 12 SNVs from periodontitis train group GWAS violated assumption 2 and have been removed from further analysis.

3. Main approach

The main approach MR was performed on two genome-wide significant SNVs (n=2) (i.e. rs368467810, rs12123266) (F-statistics = 32.2, R^2 = 0.02%) from the periodontitis group GWA study. There was weak significance (Beta: -0.816 95% CI (-1.617, -0.015), p=0.046) observed in the IVW model with no evidence of heterogeneity (Cochran's Q = 0.32 on 1 degree of freedom (df), p-value = 0.57) (Table 2). Results from the mode-based estimate were consistent, if marginally less significant (p= 0.051). MR-Egger regression and weighted median analysis were not performed due to the insufficient number of genetic instruments. Due to the inconsistent results between models, we further performed MR on each individual SNV. The results showed no significance between periodontitis and dementia by using single SNVs.

4. Alternative approach 1 & 2

When lifting the genetic instrument selection criteria to $p<5x10^{-6}$ (SNVs n= 79) (F-statistics = 23.2, R^2=0.54%), the IVW model suggested no association between periodontitis and dementia (Beta: -0.070 (-0.222, 0.083), p>0.5). Insignificant results (p>0.05) were also reported across other MR methods. There is no evidence of horizontal pleiotropy and heterogeneity as indicated by MR-Egger regression intercept (p=0.23) and heterogeneity test (Cochran's Q = 76.73 on df 78, p-value = 0.52) (Table 2).

In the alternative approach 2 ($p<10^{-4}$) (F = 17.4, R^2=6.40%), there were 993 SNVs included in the MR analysis. Similar insignificant results (p>0.05) were found in MR-IVW (Beta: 0.025 (-0.024, 0.075), p>0.5), as well as MR-egger regression and mode-based estimate models. However, MR weighted-median model showed a weak positive significant causal relationship between periodontitis and dementia (Beta: 0.077, 95% CI(0.006, 0.149), p=0.035).

Table2. MR results from all models

	Beta (95%CI)	P	Cochran's Q (df)	$P_{Heterogeneity}$
Main Approach: SNPs (P <5x10^{-8}) From the Periodontitis Group (SNPs N= 2, F=32.2, R2=0.02%)				
MR IVW	-0.816 (-1.617, -0.015)	0.046*	0.32 (1)	0.57
Mode-Based Estimate	-0.820 (-1.643, 0.003)	0.051		
Alternative Approach 1: SNPs (P <5x10^{-6}) From the Periodontitis Group (SNPs N= 79, F=23.2, R2=0.54%)				
MR IVW	-0.070 (-0.222, 0.083)	0.37	76.73 (78)	0.52
MR-Egger Regression	0.048 (-0.197, 0.294)	0.699		
MR-Egger Regression Intercept	-0.008 (-0.020, 0.005)	0.229		
MR Weighted Median	-0.066 (-0.282, 0.150)	0.548		
Mode-Based Estimate	-0.348 (-0.932, 0.237)	0.243		
Alternative Approach 2: SNPs (P <10^{-4}) From the Periodontitis Group (SNPs N=993, F=17.4, R2=6.40%)				
MR IVW	0.025 (-0.024, 0.075)	0.316	954.00 (992)	0.8
MR-Egger Regression	-0.008 (-0.080, 0.063)	0.82		
MR-Egger Regression Intercept	0.003 (-0.001, 0.006)	0.197		
MR Weighted Median	0.077 (0.006, 0.149)	0.035*		
Mode-Based Estimate	0.235 (-0.090, 0.560)	0.157		
Alternative Approach 3: SNPs (P<0.19) From the Periodontitis Train Group (SNPs N= 958, F=17.4, R2=12.39%)				
MR IVW	0.035 (-0.007, 0.076)	0.101	848.24 (957)	0.99
MR-Egger Regression	0.044 (-0.015, 0.103)	0.147		
MR-Egger Regression Intercept	-0.001 (-0.005, 0.003)	0.677		
MR Weighted Median	0.035 (-0.025, 0.095)	0.25		
Mode-Based Estimate	0.039 (-0.231, 0.308)	0.779		
MR: rs12123266 (P<5x10^{-8}) From the Periodontitis Group (F=31.3, R2=0.01%)				
MR IVW	-0.577 (-1.734, 0.581)	0.329		
MR: rs368467810 (P<5x10^{-8}) From the Periodontitis Group (F=33.2, R2=0.01%)				
MR IVW	-1.036 (-2.145, 0.073)	0.067		

CI, confidence interval; df, degree of freedom; Q, Cochran's Q

**p<0.05*

There was also no sign of heterogeneity (Cochran's Q = 954.00 on 992 df, p=0.8) and horizontal pleiotropy (MR-Egger regression intercept p>0.05).

5. Alternative approach 3

As PRS results on the periodontitis test group showed the best-fit p-value threshold is 0.19, all leading SNVs in the periodontitis train group after clumping have been included in the MR analysis (n=958, F=17.4, R^2=12.39%). 20.56% of genetic instruments included in the alternative approach 3 (n=197) were consistent with genetic instruments chosen from the alternative approach 2.

The IVW (Beta: 0.035 (-0.007, 0.076), p>0.05) and all other models based on SNVs from the periodontitis training sample were insignificant, suggested no causal relationship between periodontitis and dementia. Again, there is no evidence of heterogeneity (Cochran's Q = 848.24 on 957 df, p>0.05) and horizontal pleiotropy (MR-Egger regression intercept p>0.05) (Table 2).

Discussion

We have applied various MR methods with differing thresholds to identify valid genetic instruments for periodontitis and to investigate the potential causal relationship between periodontitis and dementia. In keeping with previous analyses, we do not find strong evidence of a causal relationship although a few of our results provide some mild evidence, albeit with inconsistent results regarding the nature of the relationship.

One of the major challenges of performing MR analyses is the difficulty of finding an instrumental variable that is a strong enough predictor of the exposure under consideration. The inconsistency in GWAS analyses of periodontitis, both in our study and others, indicates the challenges of using this approach to understand the association. The majority of previous studies have suggested periodontitis is likely to be associated with a higher risk of dementia. (Choi et al., 2019; Gao et al., 2023; Kang et al., 2020). However, heterogeneity existed in the association results. This includes some studies failing to replicate the association (Holmer et al., 2022) and effect size differencing based on study design factors (e.g., gender, periodontitis measurements, periodontitis severity) (Larvin et al., 2023). In contrast, our findings support previous two-sample MR studies that also find no association between periodontitis and dementia or cognitive impairment (Deng et al., 2024; Hu et al., 2024). Given that observational studies often suffer from uncontrolled confounding effects and reverse causality—issues to which two-sample MR is less prone—the discrepancy between this study and previous observational studies probably suggests

that the previously observed association between periodontitis and dementia might be due to confounding factors or coexistence rather than causation.

A possible explanation for the association observed in epidemiological studies but absent in MR studies is that the association observed is due to shared causes. One of the shared causes hypothesis is a shared inflammatory pathway between periodontitis and dementia. Chronic inflammatory diseases are polygenic, involving genes functioning within the immune response that are shared between various inflammatory diseases (Loos and Van Dyke, 2020). This could potentially be due to common pathogenicity. Based on this concept and missing causation in MR results, periodontitis, a common peripheral inflammatory disease (Kamer et al., 2008), and dementia, where neuroinflammation plays a critical role in diseases progression (Leng and Edison, 2021), potentially shares some inflammatory pathology instead of representing a cause-and-effect relationship. However, due to the complex mechanism underlying inflammatory diseases, further investigation is needed to examine the possibility of a shared inflammatory pathway in between periodontitis and dementia, and how it contributes to different diseases progression. Another possible common cause is lifestyle factors, such as smoking and nutrition intake, which are important causal risk factors for both periodontitis and dementia but were not well adjusted for in all epidemiological studies (Thomson and Barak, 2021). Since both periodontitis and dementia are multifactorial diseases with various factors play important role in the disease's progression, missing causality may be due to shared confounders.

Reverse causation, where dementia increases the risk of periodontitis, is another possibility for the lack of causation in the direction of periodontitis affecting dementia but with the observed significant association between periodontitis and dementia in observational studies. The Life Course hypothesis suggests that poor cognitive function in childhood may contribute to cumulative effects throughout the lifespan (e.g., poorer academic achievement and education level, poorer oral hygiene, and lifestyle behaviours including alcohol consumption and smoking) which ultimately leads to an incremental effect on tooth loss—the final stage of periodontitis—at older age, along with poor cognitive outcomes (Thomson and Barak, 2021). This reverse causation is further supported by previous studies that found an increased risk of periodontitis in participants with decreased cognitive performance (Gao et al., 2023) and in patients with Alzheimer's disease compared to participants without Alzheimer's disease (Ma et al., 2022). Therefore, further investigation into the mechanisms underlying the observed association between periodontitis and dementia in epidemiological studies is still needed.

In addition to the majority of insignificant results, we find weak significant results in our main method ($p<5x10^{-8}$) IVW model and the alternative method 2 ($p<10^{-4}$) MR weighted median model. However, due to the majority of our results suggesting no casual association and inconsistent effect direction observed from these two model results, the significance should be interpreted with caution. It is noteworthy that although the SNVs in the two MR approaches are strong instruments for periodontitis, they explain only (R^2) 0.02% and 6.4% of variance of periodontitis, respectively which potentially decrease the precision of MR estimates (Sanderson et al., 2022). Similarly, only modest variability of periodontitis explained by genetic instruments was also noted in the MR study examining the relationship between periodontitis and biological aging, where an insignificant association was found (Song et al., 2024). This low variance in genetic instruments may be due to the multifactorial nature of periodontitis. The limited variance explained by genetic instruments may decrease the precision of the MR estimates.

On the other hand, the differences between MR models also contribute to the explanation of inconsistent results. For example, IVW provides an efficient causal estimate, other MR methods (i.e., MR-egger, weighted median, mode- based estimate) are more sensitive to outliers, pleiotropy and bias. The significance observed and inconsistent results might be due to potential undetected invalid instruments, or it may also be possible that these SNVs are better predicted by other unadjusted mediators, which could contribute to the observed association. Another possible explanation for the borderline significance detected in IVW model is that the power trade-off in MR-Egger regression, weighted-median and mode-based estimate (Boehm and Zhou, 2022) hindered the observation of significance in other model results. However, this may be able to explain the significance observed in alternative approach 2, as it was observed in weighted median model. The weighted median model benefits from its robustness, even when some invalid instruments are included (Bowden et al., 2016a), including too many invalid or weak instruments may still bias the MR results. It should be noted that the alternative approach 2 included 993 genetic instruments. The risk of invalid instruments can be increased by including too many genetic instruments, even though this increases statistical power and disease representation (Sanderson et al., 2022).

It is also necessary to note that the direction of significance results observed in main approach IVW model and alternative approach 2 weighted median model is inconsistent. There are two possible reasons regarding UKB data quality and disease sample size could explain such inconsistent direction of findings. First, it should be acknowledged that the self-reported measurement of periodontitis is not ideal, although it is a valid approach of periodontitis measurement when clinical measurements absent (Abbood et al., 2016; Papapanou et al., 2018).

To have a better insight, future studies should aim to use clinical measurements and defining periodontitis using standardised periodontitis definition guidelines such as the 2018 periodontal status classification criteria (Holtfreter et al., 2024). The second possible reason is that the sample size for dementia cases is small with only (N= 6,024) from UKB left in analysis and it is way smaller than periodontitis cases (N=67,591). Although by allocating approximate four times more controls in the dementia group could increase the power (Hong and Park, 2012), more dementia cases would provide better insight in dementia phenotype. Dementia subgroups analysis were also not able to be performed due to limited sample size to explore further on each individual type of dementia and their effects on periodontitis. Another limitation is that the three GWAS significant SNVs from the periodontitis group have small R^2, although they were identified as strong instruments by F-statistics. Therefore, the significance observed here should be interpreted with caution.

In addition to the limitation mentioned above, there are also several strengths of the current study. First of all, we utilised various methods to validate the causal association between periodontitis and dementia by using large scale data from UKB. Instead of using one-sample MR directly on UKB, our study split the sample into two to the apply the two-sample MR method. This not only minimised the confounding effect and reverse causation but also reduce the bias due to overfitting which is often experienced in one-sample MR (Burgess et al., 2019). By applying two-sample MR in a single largescale also ensured the power and similarities between sample. Checks on each MR assumption further ensured the validity of current results. However, there are some limitations resulting from the self-reported measurements of periodontitis in UKB. Clinical examination on periodontal status and use of the 2018 periodontal status classification criteria (Holtfreter et al., 2024) is recommended in future studies (Gao et al., 2024).

Conclusion

In conclusion, while there might be a causal relationship between periodontitis and dementia, the evidence remains inconclusive due to inconsistent results obtained from various models and the selection of genetic variants associated with periodontitis. Validated genetic instruments for periodontitis are essential to ensure that the MR approach is appropriate for use in future studies.

ACKNOWLEDGEMENT:
This article has not been published or being considered for publication elsewhere, in whole, in part, except as an abstract.

Data availability: UK Biobank data were used in the current study and application is required for access.

Conflict of interest: The authors declare no conflict of interest.

Author contributions: Chenyi Gao contributed to conception, data analysis and interpretation, and drafted the manuscript. Prof. David Timothy Bishop, Dr. Mark Iles, Dr. Jing Kang contributed to study design, data analysis and interpretation. Dr. Jing Kang, Dr. Jianhua Wu, Dr. Harriet Larvin contributed to data acquisition. Dr. Jing Kang, Prof. Jianhua Wu, Prof. Susan Pavitt, Prof. David Timothy Bishop and Prof. David Bunce contributed to the conceptualizations, funding acquisition, study design and data interpretation. All the author contributed to the critically review the manuscript and final version approval and submission.

Funding statement: Chenyi Gao is supported by the Alzheimer's Society (Alzheimer's Society Heather Corrie Ph.D. studentship: 546 (AS-PhD-19b-012)). The views expressed are those of the author(s) and not necessarily those of the Alzheimer's Society. The Alzheimer's Society did not involve in design and conduct this study, as well as data collection or preparation, review, approval, and decision to submit the manuscript. Jianhua Wu and Harriet Larvin are supported by Barts Charity (MGU0504). Mark Iles is supported in part by the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Reference

Abbood HM, Hinz J, Cherukara G, Macfarlane TV. 2016. Validity of self-reported periodontal disease: A systematic review and meta-analysis. Journal of periodontology (1970). 87(12):1474-1483. doi: 10.1902/jop.2016.160196.

ADI AsDI. 2022. World alzheimer report 2022:Life after diagnosis: Navigating treatment, care and support. https://www.alzint.org/resource/world-alzheimer-report-2022/.

Boehm FJ, Zhou X. 2022. Statistical methods for mendelian randomization in genome-wide association studies: A review. Computational and structural biotechnology journal. 20:2338-2351. doi: 10.1016/j.csbj.2022.05.015.

Bowden J, Davey Smith G, Burgess S. 2015. Mendelian randomization with invalid instruments: Effect estimation and bias detection through egger regression. International journal of epidemiology. 44(2):512-525. doi: 10.1093/ije/dyv080.

Bowden J, Davey Smith G, Haycock PC, Burgess S. 2016a. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genetic epidemiology. 40(4):304-314. doi: 10.1002/gepi.21965.

Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. 2016b. Assessing the suitability of summary data for two-sample mendelian randomization analyses using mr-egger regression: The role of the i2 statistic. International journal of epidemiology. 45(6):1961-1974. doi: 10.1093/ije/dyw220.

Burgess S, Butterworth A, Thompson SG. 2013. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 37(7):658-665. doi: 10.1002/gepi.21758.

Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Kutalik Z, Holmes MV, Minelli C et al. 2019. Guidelines for performing mendelian

randomization investigations: Update for summer 2023. Wellcome open research. 4:186. doi: 10.12688/wellcomeopenres.15555.3.

Burgess S, Thompson SG. 2013. Use of allele scores as instrumental variables for mendelian randomization. International journal of epidemiology. 42(4):1134-1144. doi: 10.1093/ije/dyt093.

Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J et al. 2018. The uk biobank resource with deep phenotyping and genomic data. Nature (London). 562(7726):203-209. doi: 10.1038/s41586-018-0579-z.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. 2015. Second-generation plink: Rising to the challenge of larger and richer datasets. Gigascience. 4(1):7-7. doi: 10.1186/s13742-015-0047-8.

Choi S, Kim K, Chang J, Kim SM, Kim SJ, Cho HJ, Park SM. 2019. Association of chronic periodontitis on alzheimer's disease or vascular dementia. Journal of the American Geriatrics Society (JAGS). 67(6):1234-1239. doi: 10.1111/jgs.15828.

Choi SW, O'Reilly PF. 2019. Prsice-2: Polygenic risk score software for biobank-scale data. Gigascience. 8(7). doi: 10.1093/gigascience/giz082.

Davies NM, Holmes MV, Davey Smith G. 2018. Reading mendelian randomisation studies: A guide, glossary, and checklist for clinicians. BMJ (Online). 362:k601-k601. doi: 10.1136/bmj.k601.

Deng Z, Li J, Zhang Y, Zhang Y. 2024. No genetic causal associations between periodontitis and brain atrophy or cognitive impairment: Evidence from a comprehensive bidirectional mendelian randomization study. BMC oral health. 24(1):571-571. doi: 10.1186/s12903-024-04367-7.

Gao C, Iles M, Larvin H, Bishop DT, Bunce D, Ide M, Sun F, Pavitt S, Wu J, Kang J. 2024. Genome-wide association studies on periodontitis: A systematic review. PLOS ONE. doi.

Gao C, Larvin H, Bishop DT, Bunce D, Pavitt S, Wu J, Kang J. 2023. Oral diseases are associated with cognitive function in adults over 60 years old. Oral diseases. doi: 10.1111/odi.14757.

Hartwig FP, Smith GD, Bowden J. 2017. Robust inference in summary data mendelian randomization via the zero modal pleiotropy assumption. International journal of epidemiology. 46(6):1985-1998. doi: 10.1093/ije/dyx102.

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

Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S, Bowden J, Langdon R et al. 2018. The mr-base platform supports systematic causal inference across the human phenome. eLife. 7. doi: 10.7554/elife.34408.

Holmer J, Eriksdotter M, Häbel H, Hed Myrberg I, Jonsson A, Pussinen PJ, Garcia-Ptacek S, Jansson L, Sandborgh-Englund G, Buhlin K. 2022. Periodontal conditions and incident dementia: A nationwide swedish cohort study. Journal of periodontology (1970). doi: 10.1002/JPER.21-0518.

Holtfreter B, Kuhr K, Borof K, Tonetti MS, Sanz M, Kornman K, Jepsen S, Aarabi G, Völzke H, Kocher T et al. 2024. Aces: A new framework for the application of the 2018 periodontal status classification scheme to epidemiological survey data. Journal of clinical periodontology. doi.

Hong EP, Park JW. 2012. Sample size and statistical power calculation in genetic association studies. Genomics & informatics. 10(2):117-122. doi: 10.5808/gi.2012.10.2.117.

Hu C, Li H, Huang L, Wang R, Wang Z, Ma R, Chang B, Li S, Li H, Li G. 2024. Periodontal disease and risk of alzheimer's disease: A two-sample mendelian randomization. Brain and behavior. 14(4):e3486-n/a. doi: 10.1002/brb3.3486.

Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ. 2008. Inflammation and alzheimer's disease: Possible role of periodontal diseases. Alzheimer's & dementia. 4(4):242-250. doi: 10.1016/j.jalz.2007.08.004.

Kang J, Palmier-Claus J, Wu J, Shiers D, Larvin H, Doran T, Aggarwal VR. 2022. Periodontal disease in people with a history of psychosis: Results from the uk biobank population-based study. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12798.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

Kang J, Wu B, Bunce D, Ide M, Pavitt S, Wu J. 2019. Cognitive function and oral health among ageing adults. Community dentistry and oral epidemiology. 47(3):259-266. doi: 10.1111/cdoe.12452.

Larvin H, Gao C, Kang J, Aggarwal V, Pavitt S, Wu J. 2022. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. submitted for publication to the Age and aging Journal(under peer review).

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2023. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. Age and ageing. 52(2). doi: 10.1093/ageing/afad015.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Frontiers in medicine. 7:604980-604980. doi: 10.3389/fmed.2020.604980.

Leng F, Edison P. 2021. Neuroinflammation and microglial activation in alzheimer disease: Where do we go from here? Nature reviews Neurology. 17(3):157-172. doi: 10.1038/s41582-020-00435-y.

Loos BG, Van Dyke TE. 2020. The role of inflammation and genetics in periodontal disease. Periodontology 2000. 83(1):26-39. doi: 10.1111/prd.12297.

Ma KS, Hasturk H, Carreras I, Dedeoglu A, Veeravalli JJ, Huang JY, Kantarci A, Wei JC. 2022. Dementia and the risk of periodontitis: A population-based cohort study. Journal of dental research. 101(3):270-277. doi: 10.1177/00220345211037220.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. 2010. Robust relationship inference in genome-wide association studies. Bioinformatics. 26(22):2867-2873. doi: 10.1093/bioinformatics/btq559.

Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, Smith GD, Sterne JAC. 2012. Using multiple genetic variants as instrumental variables for modifiable risk factors. Statistical methods in medical research. 21(3):223-242. doi: 10.1177/0962280210394459.

Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, Flemmig TF, Garcia R, Giannobile WV, Graziani F et al. 2018. Periodontitis: Consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. J Clin Periodontol. 45 Suppl 20:S162-s170. doi: 10.1111/jcpe.12946.

Patel A, Ye T, Xue H, Lin Z, Xu S, Woolf B, Mason AM, Burgess S. 2023. Mendelianrandomization v0.9.0: Updates to an r package for performing mendelian randomization analyses using summarized data. Wellcome open research. 8:449. doi: 10.12688/wellcomeopenres.19995.1.

Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, Palmer T, Schooling CM, Wallace C, Zhao Q, Davey Smith G. 2022. Mendelian randomization. Nature Reviews Methods Primers. 2(1):6. doi: 10.1038/s43586-021-00092-5.

Schlesselman JJ, Stolley PD. 1982. Case-control studies : Design, conduct, analysis. New York ;: Oxford University Press.

Song L, Wang Y, Zheng Q, Li W. 2024. Periodontitis prevalence and acceleration of biological aging: Insights from nhanes 2009–2014 and mendelian randomization study. Journal of periodontal research. doi: 10.1111/jre.13345.

Stephen DT. 2014. Qqman: An r package for visualizing gwas results using q-q and manhattan plots. bioRxiv.005165. doi: 10.1101/005165.

Thomson WM, Barak Y. 2021. Tooth loss and dementia: A critical examination. Journal of Dental Research. 100(3):226-231. doi: 10.1177/0022034520957233.

Dementia. 2023. World Heath Organisation; [accessed 2024 2/14]. https://www.who.int/news-room/fact-sheets/detail/dementia.

Wilkinson T, Schnier C, Bush K, Rannikmäe K, Henshall DE, Lerpiniere C, Allen NE, Flaig R, Russ TC, Bathgate D et al. 2019. Identifying dementia outcomes in uk biobank: A validation study of primary care, hospital admissions and mortality data. European journal of epidemiology. 34(6):557-565. doi: 10.1007/s10654-019-00499-1.

Yavorska OO, Burgess S. 2017. Mendelianrandomization: An r package for performing mendelian randomization analyses using summarized data. International journal of epidemiology. 46(6):1734-1739. doi: 10.1093/ije/dyx034.

Chapter 4.2: Extension in chapter 4.1 mendelian randomisation results

This chapter presents the results using the genetic risk variants for periodontitis identified in Chapter 4.1 as genetic instruments. These findings were not included in Chapter 4.1 to maintain the conciseness in the manuscript intended for publication. However, they are described here in Chapter 4.2 to provide a more comprehensive narrative for the thesis. This chapter focuses solely on the methods and results related to the genetic instruments obtained from Chapter 3.1, specifically those genetic instruments that reached significance level $5x10^{-6}$ extracted from the 15 included GWAS studies in the systematic review of GWAS of periodontitis.

Method

In addition to the methods described in Chapter 4.1, I have also conducted Mendelian Randomisation analysis using the potential genetic instruments identified in chapter 3.1 systematic review.

Genetic instruments selection

To summarise, 52 SNPs identified ($p < 5 \times 10^{-6}$) in previous periodontitis GWAS were evaluated for eligibility. Several exclusion criteria were applied to ensure the selection of appropriate genetic instruments and enhance the reliability of MR results:

1. There is sample overlap between our study and the GWAS that identified the SNPs, which could increase bias and the likelihood of type 1 errors (Burgess et al., 2016).
2. Insufficient information was reported for the SNPs. For example, confidence intervals and standard errors were not found in the published paper(s) and corresponding supplementary materials. This lack of information or ambiguous SNP data prevented us from performing the MR analysis.
3. Ambiguous A1 (i.e., effect allele) and A2 (i.e., other allele) were reported for the SNPs.
4. Ambiguous A1 frequency was reported for the SNPs.
5. The SNPs identified in a non-European sample were excluded due to the potential differences in genetic structure between ethnicities (Abdellaoui et al., 2023).
6. SNPs from the same study are in linkage disequilibrium (LD) (r^2>0.1, window size 250kb) were filtered to retain only SNPs with the most significant p-value.

Statistical analysis

Among the selected genetic instruments from each study, SNPs in LD were checked using the web-based tool LDlink (https://ldlink.nih.gov). Two-sample MRs were performed between each GWAS identified genetic instruments and dementia group participants as described in the Chapter 4.1. The summary statistics of the selected genetic instruments for periodontitis were extracted from Table 4 in Chapter 3.2. Association analyses were conducted between the selected SNPs and dementia outcomes in the dementia group using PLINK2. Data harmonisation was carried out using the "TwoSampleMR" package in R. The MR analysis were performed using MR IVW, MR-egger regression, Weighted Median, mode-based method with the "MendelianRandomization" Package in R. Wald ratios were calculated when only one genetic instrument was included from a GWAS. F-statistics and R^2 were calculated to assess the strength of the genetic instruments (Assumption 1). The association between genetic instruments and other confounders were check by searching their association results in the GWAS catalog (https://www.ebi.ac.uk/gwas/) (Assumption 2). SNPs that have been reported in other diseases

or trait that associated with both periodontitis and dementia were excluded from further analysis. The MR-Egger regression intercept was used to assess the horizontal pleiotropy (Assumption 3).

Results and conclusions

Genetic instruments

Based on the exclusion criteria, there are 10 independent significant SNPs in total from three periodontitis GWAS (Bevilacqua et al., 2018; Munz et al., 2017; Tegelberg et al., 2021) were selected as genetic instruments for periodontitis (Table 1). Two SNPs were further excluded due to their potential violation of the MR Assumption 2 by their association with other diseases: (1) rs2738058 was association with IgA glomerulonephritis (Li et al., 2015) which might increase risk of dementia (Li et al., 2018), (2) rs2070901 was associated with several diseases such as asthma (Zhu et al., 2019) which may be linked to both periodontitis (Moraschini et al., 2018) and dementia (Peng et al., 2015).

Table 1. Summary statistics of selected genetic instruments

Study	SNP	Chromosome: Position	OR (95%CI)	A1	EAF (or MAF)	A2	p-value
Bevilacqua et al. (2017) Case n=442 Control n=160	rs242016	12:3788260	3.7 (2.32, 6.29)	A	0.21	G	1.50E-08
Munz et al. (2017) Case n=2067 Control n=8819	rs2738058	8:6821617	1.28 (1.18, 1.38)	T	0.43	C	6.78E-10
	rs42842742	19:52131733	1.34 (1.21, 1.48)	G	0.76	A	1.34E-08
	rs4970469	1:27312816	1.52 (1.29, 1.81)	G	0.9	A	1.20E-06
	rs1122900	5:36689181	1.27 (1.16, 1.4)	A	0.4	C	8.00E-07
	rs2070901	1:161185058	1.29 (1.16, 1.44)	T	0.24	G	4.36E-06
Tegelberg et al. (2021)	rs2409703	8:10958526	beta: 0.28 (0.16, 0.39)	*C*	0.079	T	1.61E-06

Sample n=3245 (continuous phenotype used)	rs11630851	15:76021782	beta: 0.30	(0.18, 0.42)	*T*	0.067	C	9.39E-07
	rs4444613	20:13340138	beta: -0.28	(-0.38, -0.18)	*A*	0.087	G	1.35E-07
	rs2003705	20:37763743	beta: -0.16	(-0.23, -0.1)	*T*	0.2	C	1.68E-06

SNP, Single Nucleotide Polymorphism; Chrom: Pos, Chromosome: Position; OR, Odds Ratio; CI, Confidence interval; A1, effect allele; EAF, effect allele frequency; A2, other allele; beta, coefficient beta

Two-sample MR

Three two-sample MRs were performed (Table 2). Similar to the majority of MR models in Chapter 4.1, most results did not indicate a significant causal association between periodontitis and dementia using genetic instruments derived from previous periodontitis GWASs. The only weak significant association was detected in the MR-Egger regression method, which estimated the casual relationship between genetic instruments from Munz et al. (2017) and dementia group (Beta -0.44, 95%CI (-0.86, -0.03), p=0.04*).

Table 2. MR results summary

	Beta (95%CI)	**P**	**Cochran's Q (df)**	**$P_{Heterogeneity}$**
Bevilacqua et al. (2017) (SNPs N=1, F=26.4, R^2=4.21%)				
MR IVW	-0.01 (-0.05, 0.03)	0.5		
Munz et al. (2017) (SNPs N= 3, F=26.9, R^2=0.74%)				
MR IVW	-0.06 (-0.15, 0.04)	0.24	3.74 (2)	0.15
MR-egger regression	-0.44 (-0.86, -0.03)	0.04*		
MR-egger regression intercept	0.12 (-0.01, 0.24)	0.06		
MR weighted median	-0.07 (-0.20, 0.05)	0.24		
Mode-based estimate	-0.09 (-0.26, 0.07)	0.28		
Tegelberg et al. (2021) (SNPs N=4, F=25, R^2=2.99%)				
MR IVW	0.02 (-0.11, 0.15)	0.75	0.96 (3)	0.8112
MR-egger regression	-0.14 (-0.63, 0.35)	0.58		
MR-egger regression intercept	0.04 (-0.08, 0.16)	0.51		
MR weighted median	0.02 (-0.14, 0.17)	0.83		
Mode-based estimate	0.07 (-0.13, 0.28)	0.49		

CI, confidence interval; df, degree of freedom; Q, Cochran's Q

**p<0.05*

Based on the F-statistics (>10), MR-Egger regression intercept (p> 0.05) and an assessment of associations with other diseases, there was no evidence of violation of MR assumptions in genetic instruments selected from each study.

Conclusions

The mixed results suggest that it is unlikely periodontitis has a causal effect on dementia. A detailed and comprehensive discussion of the results, study limitations and strengths, as well as future directions, can be found in the Chapter 4.1 discussion section.

Reference

Abdellaoui A, Yengo L, Verweij KJH, Visscher PM. 2023. 15 years of gwas discovery: Realizing the promise. American journal of human genetics. 110(2):179-194. doi: 10.1016/j.ajhg.2022.12.011.

Bevilacqua L, Navarra CO, Pirastu N, Lenarda RD, Gasparini P, Robino A. 2018. A genome-wide association study identifies an association between variants in efcab4b gene and periodontal disease in an italian isolated population. J Periodontal Res. 53(6):992-998. doi: 10.1111/jre.12598.

Burgess S, Davies NM, Thompson SG. 2016. Bias due to participant overlap in two-sample mendelian randomization. Genetic epidemiology. 40(7):597-608. doi: 10.1002/gepi.21998.

Li M, Foo J-N, Wang J-Q, Low H-Q, Tang X-Q, Toh K-Y, Yin P-R, Khor C-C, Goh Y-F, Irwan ID et al. 2015. Identification of new susceptibility loci for iga nephropathy in han chinese. Nature communications. 6(1):7270-7270. doi: 10.1038/ncomms8270.

Li X, Sundquist J, Zöller B, Sundquist K. 2018. Dementia and alzheimer's disease risks in patients with autoimmune disorders. Geriatrics & gerontology international. 18(9):1350-1355. doi: 10.1111/ggi.13488.

Moraschini V, Calasans-Maia JdA, Calasans-Maia MD. 2018. Association between asthma and periodontal disease: A systematic review and meta-analysis. Journal of periodontology (1970). 89(4):440-455. doi: 10.1902/jop.2017.170363.

Munz M, Willenborg C, Richter GM, Jockel-Schneider Y, Graetz C, Staufenbiel I, Wellmann J, Berger K, Krone B, Hoffmann P et al. 2017. A genome-wide association study identifies nucleotide variants at siglec5 and defa1a3 as risk loci for periodontitis. Hum Mol Genet. 26(13):2577-2588. doi: 10.1093/hmg/ddx151.

Peng Y-H, Wu B-R, Su C-H, Liao W-C, Muo C-H, Hsia T-C, Kao C-H. 2015. Adult asthma increases dementia risk: A nationwide cohort study. Journal of epidemiology and community health (1979). 69(2):123-128. doi: 10.1136/jech-2014-204445.

Tegelberg P, Leppilahti JM, Ylöstalo A, Tervonen T, Kettunen J, Suominen AL, Ylöstalo P. 2021. Genome-wide association study of periodontal pocketing in finnish adults. BMC Oral Health. 21(1):611. doi: 10.1186/s12903-021-01964-8.

Zhu Z, Zhu X, Liu CL, Shi H, Shen S, Yang Y, Hasegawa K, Camargo CA, Jr., Liang L. 2019. Shared genetics of asthma and mental health disorders: A large-scale genome-wide cross-trait analysis. Eur Respir J. 54(6). doi: 10.1183/13993003.01507-2019.

Chapter 5: Discussion and Conclusions

Chapter 5.1 Summary

My PhD project aims to explore the association between the periodontal disease and poorer cognitive outcomes especially causal relationship between periodontitis and dementia, I used several methods including systematic review and meta-analysis, cross-sectional study with multi-variable regression model and structure equation model and advanced methodology, Mendelian randomisation (MR), to rigorously investigate this association. In this thesis, I found significant association between periodontal diseases and poorer cognitive outcomes including dementia but no significant causal effect of periodontitis on dementia. To conclude, this final chapter summarised the findings from each of the reported studies, and also evaluated the current methodological challenges, the strengths and limitations of the research, and the implications for future research directions are discussed,

5.1.1 Association between periodontitis and poor cognitive outcomes

Chapter 2 provided strong evidence of the association between oral health and cognitive outcomes using multi-variable regression model and structure equation model (SEM) to adjust for a range of covariates that previous evidence linked with both periodontitis and dementia. The overall results from multi-variable regression models and SEM suggested better oral health (including periodontitis, dental caries and tooth loss) is associated with better cognitive functions (including processing speed, memory and executive function). Interestingly, the multi-variable regression analysis showed that different oral health factors were associated with different cognitive domains, where having periodontitis was associated with poorer processing speed and memory in particular. Although, currently, it is unclear which postulated mechanism may underlie the periodontitis effect on these cognitive domains, this finding provided a preliminary insight into the possibility that specific cognitive domains may be linked to periodontitis and suggests a line of future investigation.

I used a combined cognitive function score to represent overall/global cognitive function, which enabled multivariable regression modelling to test the association between periodontal diseases and global cognitive function. In this approach, the combined cognitive function score summed scaled scores from each cognitive domain to ensure that each domain was weighted equally in the global cognitive function score. However, bias may still exist in this manually summed score, which further highlights the importance and benefits of Structural Equation Modelling (SEM) for equally considering each entered variable as latent variables.

However, it should be noted that even though we intended to include smoking and tooth brush frequency in the analysis, high levels of missing data (>50%) meant this variable had to be dropped. Despite of this though, Chapter 2 successfully adjusted for several other covariates

which had not been taken into account in previous studies such as nutritional status, comorbidities and physical activity. This indicate improvements in the statistical modelling in Chapter 2 compared with previous observational studies.

In addition to the adjusted covariates, the level of inflammation is also a key factor in this association. Although the association between periodontal diseases and dementia may be underpinned by shared inflammatory pathways, many other factors may contribute to inflammation in the body, such as genetic predisposition (as discussed in Chapter 1) and age. With age, there is an increase in senescent cells and a reduction in cellular efficacy, leading to the release of pro-inflammatory molecules, such as pro-inflammatory cytokines (e.g., interleukin (IL)-1), by senescent cells, which contribute to a pro-inflammatory milieu in the body (Rea et al., 2018). This suggests an increased inflammatory status with age. These findings also highlight the importance of considering the level of inflammation throughout the body when testing the association between periodontal diseases and poor cognitive outcomes, including dementia, in older age, as well as the shared inflammatory pathway hypothesis. The factors contributing to inflammatory status should be carefully considered in further association studies.

5.1.2 Genetic instruments of periodontitis

In Chapter 3, I presented the first study in the literature that systematically reviewed and critically evaluated the exiting periodontitis genome-wide association study (GWAS) evidence, and identifying 15 periodontitis GWAS publications with satisfactory quality as assessed by the Q-genie tool, which is valid tool assessing the quality and bias in the genetic association study (Sohani et al., 2015; Sohani et al., 2016). Extracted from included studies, 11 single nucleotide polymorphisms (SNPs) reached conventional GWAS level of significance ($p<5x10^{-8}$) and 41 SNPs reached the suggestive level of significance ($p<5x10^{-6}$). Despite three of the SNPs from three largescale studies being from the same gene (i.e., *SIGLEC5*, which primarily functions in the immune system (Mueller et al., 2022), none of the SNPs were commonly reported across studies suggesting low replication of periodontitis GWAS results. High heterogeneity between studies might partly account for the low reproducibility of periodontitis GWAS.

Heterogeneity was noticeable in periodontitis definitions and measurements, sample size and the ethnic composition of study populations. Among the 15 included studies, six investigated chronic periodontitis, three investigated aggressive periodontitis, two investigated both aggressive and chronic periodontitis, one study investigated apical periodontitis, and the remaining studies explored periodontitis or its general surrogates. The types of periodontitis studied may contribute to differences in definitions and the SNPs discovered.

Sample size is especially critical in periodontitis GWAS. A previous meta-analysis of periodontitis genetic studies observed discrepancies in heritability estimates from GWAS (7%) and family studies (29%), although the specific loci contributing to this heritability remain elusive (Nibali et al., 2019). The substantially lower heritability observed in GWAS compared to family studies suggests potential missing heritability in current periodontitis GWAS. A similar phenomenon of missing heritability is commonly seen in other multifactorial diseases, such as late-onset Alzheimer's disease (e.g., 58% estimated from twin studies vs. 7.78% from a GWAS) (Andrews et al., 2023; Génin, 2020). In this context, increasing statistical power in GWAS for diseases with low heritability, including periodontitis, is necessary and may partly enhance the detection of contributing variants, including those with weak effects on the disease of interest (Young and Flint, 2019). Since chronic inflammatory diseases are suggested to be polygenic (Loos and Van Dyke, 2020), periodontitis, especially chronic periodontitis, may also be a polygenic disease, with multiple genetic variants identified across genetic studies (Shaddox et al., 2021). This further underscores the importance of improving statistical power in periodontitis GWAS to identify relevant genetic variants.

Although in Chapter 3 a sample size of 10,000 is mentioned to describe two GWAS with relatively larger sample sizes, this does not necessarily indicate a suitable and well-powered sample size for future periodontitis GWAS. However, performing power/sample size calculations for GWAS differs from non-genetic studies because the heritability of the phenotype or trait must be considered as the effect size (Feng et al., 2011). To perform power calculations, software such as GWAPower considers the number of input SNPs, type I error rate, heritability, expected power, or sample size to determine the power achievable with the existing sample size or the optimal sample size needed to achieve the desired power (Feng et al., 2011). Among the studies included in this systematic review, only five (i.e., (de Coo et al., 2021; Feng et al., 2014; Munz et al., 2017; Shimizu et al., 2013; Teumer et al., 2013)) performed power calculations using different software (e.g., Quanto, the Genetic Power Calculator, the PS Power and Sample Size Program, Purcell's Genetic Power Calculator). These studies all reported statistical power of ≥75% to detect: (1) SNPs at a significance threshold of $p<5x10^{-8}$, with allele frequency of 0.3 and odds ratios (OR) of 0.55 for the protective allele or 1.68 for the risk allele (de Coo et al., 2021); (2) SNPs explaining ≥1.3% variance in mean proximal attachment loss at $p<5x10^{-8}$, and risk variants of periodontitis at $p<5x10^{-8}$ with ORs of 2.1, 1.75, and 1.45 for minor allele frequencies of 5%, 10%, and 30%, respectively (Teumer et al., 2013); (3) a probability of exposure among controls of 20%, and an OR of 1.3 for disease in exposed subjects compared to unexposed subjects (Munz et al., 2017); (4) SNPs at a significance threshold of $p<10^{-7}$ with a relative risk ratio (RR) of 1.8, but power reduced to 12% when RR=1.5 (Feng et al., 2014); (5) SNPs with minor allele frequency (MAF) of 40% and OR of 1.3 (Shimizu et al., 2015), respectively. Heritability of periodontitis was not considered in the above studies. Future periodontitis GWAS may benefit from performing power

calculations, considering the 7% heritability from periodontitis GWAS to determine the minimum sample size required.

In addition to the sample size consideration, less diverse ethnicities explored in periodontitis GWASs literature also limited representativeness of genetic risk variants and discover a common genetic risk variants for each ethnicities respectively. Additionally, the unavailability of the GWAS summary level statistics in some studies prevented a GWAS meta-analysis of synthesised data which would have provided a more robust evaluation across the identified studies.

Previous periodontitis GWAS signals emphasize the importance of the inflammatory response in periodontitis, with several loci identified as regulators of inflammation and contributors to inflammatory diseases. These include protein-coding genes such as *calcium release-activated channel regulator 2A (CRACR2A), sialic acid-binding Ig-like lectin 5 (SIGLEC5), Fc epsilon receptor Ig (FCER1G), ASH1-like histone lysine methyltransferase (ASH1L), potassium voltage-gated channel subfamily Q member 5 (KCNQ5),* and *G protein-coupled receptor 141 (GPR141). CRACR2A* regulates the calcium release-activated channel in key immune cells, particularly T cells (Wang et al., 2019) and biallelic mutations in this gene can lead to immune deficiencies (Wu et al., 2021). *SIGLEC5* is a member of the immunoglobulin lectin family, predominantly expressed by immune cells, and plays a role in regulating suppressive signals for innate immune cells, including lymphoid, myeloid, and T cells (Vuchkovska et al., 2022). *SIGLEC5* has also been shown to interact with Group B Streptococcus β-protein, suppressing T cell activation and reducing proinflammatory cytokine production, such as IFN-γ and IL-22. FCER1G encodes the Fc receptor γ-chain, a molecule crucial for signalling pathways involved in autoimmunity and chronic inflammation (Zhang et al., 2023). This gene is expressed on various immune cells, including neutrophils, macrophages, and eosinophils, where it promotes inflammatory responses. *ASH1L* has been associated with inflammatory cytokine production, specifically suppressing IL-6 production by promoting the ubiquitin-editing enzyme A20 in a mouse model. It has been suggested that *ASH1L* may play a role in preventing or treating inflammatory and autoimmune diseases (Xia et al., 2013). Suppressing the activity of *KCNQ5* has been linked to reduced lipopolysaccharide-induced endothelial cell inflammation and various effects on proinflammatory cytokines, such as IL-1, IL-6, and IL-8, in a Chinese population (Lin et al., 2015). *GPR141* has been suggested to negatively regulate immune responses by modulating the function of two key cells in the innate immune system: monocytes and dendritic cells (Sawabe et al., 2024; Schlitzer et al., 2015). The causal effects of most of these associated genes and their pathways in periodontitis are not yet fully studied.

Among these inflammatory response-related genes, only *SIGLEC5* has been reported across periodontitis genetic association studies, including candidate gene studies and GWAS (Munz et

al., 2018; Shaddox et al., 2021; Shungin et al., 2019; Tong et al., 2019). Although *IL* genes were not discovered in periodontitis GWAS, there are genes affecting key innate immune cells and indirectly impacting the regulation of proinflammatory cytokines, including *ILs*, which again emphasises the role of proinflammatory genes in periodontitis, as described in the above paragraph. As suggested by a previous review, candidate gene studies encountered common problems, including a lack of statistical power, inadequate multiple-testing correction, and false-positive results (Loos and Van Dyke, 2020). This may potentially contribute to the limited overlap between candidate gene studies and GWAS. Although more than 65 genes have been suggested to be associated with periodontitis (Loos and Van Dyke, 2020) and many SNPs have been discovered for periodontitis, much remains to be understood about the individual roles of SNPs and candidate genes in periodontitis.

In addition, despite the identification of these loci and genes in previous GWAS (Table 4 in Chapter 3.1), the causal effect of the identified disease-associated variants should be further investigated. It is common that the genetic variants identified in GWAS may not be causal due to strong linkage disequilibrium between causal variants and other co-inherited variants (Gallagher and Chen-Plotkin, 2018). The less stringent significance threshold used in Chapter 3.1 and the limitations of the included GWAS should also be considered when interpreting the results. A Bayesian approach has been suggested to test the probability that the disease-associated variants are causal (Gallagher and Chen-Plotkin, 2018). Further functional analysis of the causal variants, including the use of publicly available eQTL data to explore gene expression levels in various cell and tissue types, should be pursued to understand how changes in the function or regulation of these genes might contribute to variation in periodontitis risk (Gallagher and Chen-Plotkin, 2018). Moreover, since these genes are involved in inflammatory responses or inflammatory diseases, investigating their role in the association between periodontitis and related systemic diseases, including dementia, is crucial.

To replicate disease-related variants identified in the literature and discover other potential genetic instruments for periodontitis, the second study in Chapter 3 used self-reported periodontal measurements as surrogates to explore genetic risk variants for periodontitis in UK Biobank (UKB). The sample included from the UKB consisted of 376,611 individuals, which contributes to increased statistical power compared to most previous periodontitis GWAS. By applying standardised quality controls and additive models in the association analysis, we replicated two SNPs (i.e., rs12461706, rs11084095) from *SIGLEC5* identified in previous GWASs (Munz et al., 2019; Shungin et al., 2019), providing supporting evidence for the role of *SIGLEC5* in periodontitis. As mentioned in the previous paragraphs, *SIGLEC5* is involved in the innate immune system and may be related to reducing pro-inflammatory cytokines. A previous study using antibody electrophoretic mobility shift assays in peripheral blood mononuclear cells found

that rs11084095 has a regulatory effect on *SIGLEC5* expression. Specifically, the impaired binding affinity of the *E-26 transformation–specific transcription factor-related gene (ERG)* at the minor A-allele of rs11084095 was observed (Mueller et al., 2022). Since *ERG* is primarily expressed in endothelial cells and leukocytes and is an important transcription factor for endothelial homeostasis (Heiss et al., 2015), the impaired binding affinity at *SIGLEC5* may be linked with the disrupted regulatory function of endothelial homeostasis and tissue repair in response to infection (Mueller et al., 2022), which could further contribute to periodontitis susceptibility. Additionally, the SNP rs12461706 has been found to be in strong linkage disequilibrium with rs11084095 ($r^2 = 1$) (Mueller et al., 2022).

We also identified a further four SNVs (i.e., rs775476621, rs751014048, rs149922301, rs368467810) that reached conventional GWAS significance levels in our GWAS. rs149922301 was annotated to the *RP11-61G19.1* gene and *Micro RNA 572*; while rs368467810 is annotated to *HIST1H4L* and *HIST1H3L* by FUMA. However, the role of these genes in periodontitis remains poorly understood. We provided further periodontitis GWAS evidence using a large European sample and provided important insights into the potential genetic underpinnings of periodontitis.

5.1.3 Causal relationship between periodontitis and dementia

The overall results in Chapter 4 suggested that periodontitis is not proven to be a causal risk factor for dementia. However, there were some results in this thesis indicative of a significant causal relationship between periodontitis and dementia. These results should be interpreted with caution due to the conflicting directions of effect (e.g., one MR model find periodontitis increase risk of dementia; whereas the other MR model suggested periodontitis reduced the risk of dementia). This finding is only partially consistent with the prediction based on the significant association results from observational studies.

Several reasons may explain the discrepancy between the MR findings and those for the observational studies. First, there may be a significant causal link between periodontitis and dementia, albeit very small to achieve the same significance level as the association observed in observational studies. Another MR study found significant causal association between periodontitis and brain cortical structure including the medial and lateral orbitofrontal cortex, the inferior temporal cortex also supported the possible causal effect of periodontitis on brain disorders including dementia (Wang et al., 2024). However, due to the inconsistent direction of significant results from Chapter 4, further investigation is still needed. Second, periodontitis and dementia may be indirectly linked due to a shared cause. This might involve common confounding variables such as smoking, or a shared diseases mechanism. The shared

inflammatory pathway hypothesis discussed in Chapter 1 may have value in explaining the association between periodontitis and dementia, although it does not necessarily suggest that periodontitis directly causes dementia. It has been noted that chronic inflammatory diseases often share many common SNPs or genes that function directly in respect to the immune response (Loos et al., 2020). Since both periodontitis and dementia are characterised by inflammatory responses (Kamer et al., 2008; Leng and Edison, 2021), a shared diseases pathway is likely. For example, several shared risk genes involved in the inflammatory response, such as *IL* genes (discussed in Chapter 1), have been identified in previous genetic studies. This is again evidence by shared genetic. A third possibility is that the associations observed may indicate that poorer cognitive outcomes including dementia leads to periodontal diseases rather than vice-versa as suggested by life-course model (Thomson and Barak, 2021). A fourth reason is dementia sample size are not ideal in UKB with only 6000 participants out of approximately 400,000 European participants in the UKB. The final reasons may be that the periodontitis is a multifaceted diseases, which may more strongly affected by other factors such as oral pathogens or environmental factors or the self-report periodontitis measures is less ideal. These is partially evidenced by the low variability of periodontitis (R^2 ranged from 0.01% to 12.39%) can be explained by the genetic instruments selected in this thesis. Heritability estimates the proportion of variance in a particular trait or phenotype that is explained by genetic factors (Barry et al., 2023). The low heritability (7%) estimated from periodontitis GWAS (Nibali et al., 2019) may explain the low variance accounted for by the genetic instruments selected in the current MR analysis. This finding is supported by a previous study that simulated UKB data with varying levels of polygenicity and heritability to assess the magnitude of bias (e.g., weak instrument bias and sample overlap bias) in the MR inverse variance weighted method using traits such as Body Mass Index (BMI) and blood pressure. The study's findings suggested that biases are influenced by genetic architecture and particularly affect polygenic traits or traits with low heritability (Mounier and Kutalik, 2023). Although the F-statistics (>10) indicate that the selected genetic instruments meet the relevance assumption (assumption 1), the low variability (R^2) explained by these instruments should be carefully considered when interpreting the results.

In addition to the lack of significance, it should also be noted that the shared causes (the second reason mentioned above for the lack of significance) may violate MR assumption 3. As discussed in Chapter 1, inflammation has been documented in both periodontal diseases and dementia, and some genes involved in mediating the inflammatory response, such as key inflammatory cytokines (e.g., *IL-1*), have been found to be commonly implicated in both conditions. Although this supports the shared inflammatory response hypothesis (as discussed in Chapter 1), it could also suggest the possibility of shared third causes (violating assumption 2) or the gene contributing directly to dementia, rather than exerting its effect through periodontal disease. Therefore, great care must be taken when selecting genetic instruments to ensure that MR

assumptions are not violated. In our MR analysis, we excluded genetic instruments associated with blood levels of inflammatory cytokines, specifically C-reactive protein. However, as gene mapping was not performed, it remains unclear which mapped genes are associated with the selected genetic instruments, and how these genes might be involved in the inflammatory response, particularly in approaches 2, 3, and 4, where a relatively large number of genetic instruments (ranging from 79 to 993 SNPs) were included. The risk of violating assumption 3 could undermine the validity of our results; however, horizontal pleiotropy was not evident from the MR Egger regression results (Chapter 4.1, Table 2). Future studies investigating the causal relationship between periodontal diseases and dementia would benefit from considering additional shared genes, such as *IL-1*, and exploring how these shared factors might contribute to both conditions.

Taken together, although the research reported in this thesis did not confirm a causal link between periodontal diseases and poorer cognitive outcome, specifically periodontitis and dementia, this thesis still makes a significant contribution to resolve the controversy in the field by making a preliminary attempt to answer casual questions about the association between periodontitis and dementia using advanced research methods. The findings also emphasize the importance of exploring mechanisms underlying the association between periodontitis and dementia in order to inform clinical practice and patient care.

Chapter 5.2 Methodological challenges

Some methodological challenges were identified in Chapters 3 and 4.

While evaluating the genetic instruments for periodontitis in Chapter 3, questions concerning the reproducibility of genetic risk variants were identified in relation to periodontitis GWAS. Two sides of the same issues were apparent: (1) Not all studies included a replication cohort (i.e., a separate and independent dataset to replicate the initial GWAS); and (2) There is no exact replications of genetic risk variants for periodontitis, which means there is no common genetic risk variants reported across periodontitis GWAS. Replication of significant findings is important to address the probability of false positive discovery in all aspects of science area (Ioannidis, 2005; Marigorta et al., 2018) while recognising that with the strict criteria for assessing significance in GWAS, lack of reproducibility has not been an issue for well-define traits (Fatumo et al., 2022). Although several genes (e.g., *IL genes*) were identified in periodontitis by candidate gene study, there was no single nucleotide polymorphisms (SNP) commonly discovered in periodontitis GWAS. A potential false positive discover and publication bias concerns periodontitis GWAS results which impacts on the potential for both false positive and false negative results in MR.

There are several contributing factors that might explain the shortcomings in periodontitis GWAS investigations. First, periodontitis is a complex multi-faceted disease involving interaction between specific bacteria, host response and environmental factors which together play an important role in diseases progression (Lang and Bartold, 2018; Shaddox et al., 2021). Although heritability studies indicate a role of additive genetic factors (Nibali et al., 2019; Shaddox et al., 2021), genetic factors may not stand out at same level of significance as other factors such as oral pathogens or environmental factors. This may contribute to inconsistent discover of genetic risk variants in periodontitis across studies. Besides, the interactions between genetic factors and other factors in diseases progression may be particularly profound in periodontitis.

Second, low heritability estimated from periodontitis GWAS and relatively insufficient sample size as discussed in Chapter 5.1.2 may reduce the representativeness and generalisability of results, contributing to inconsistency between study findings. However, few periodontitis GWAS studies have performed power calculations or considered heritability in these calculations. With this in mind, it is also of note that three largescale periodontitis GWAS investigations have identified two SNPs in Linkage Disequilibrium (LD) annotated on the same gene, *SIGLEC5* (Munz et al., 2019; Munz et al., 2017; Shungin et al., 2019), suggesting that larger sample sizes may contribute to the replication and confirmation of genetic variants in periodontitis. However, this was not the case in the GWAS investigation included in this thesis, since around 400,000 European participants were included in that GWAS investigation (Chapter 4.2). The work in this thesis replicated two genetic risk variants (i.e., rs11084095, rs12461706) around *SIGLEC5* that have been reported in previous studies with relatively larger sample size (i.e., (Munz et al., 2019; Shungin et al., 2019)). However, this is a replication as a SNPs reported in previous periodontitis GWAS achieve nominal significance ($p < 0.05$) in this study after adjusting multiple testing error using Bonferroni correction instead of the more exacting replication (i.e., SNPs reported in previous periodontitis GWAS which achieve again genome-wide significance).

The differences in the significance level between this thesis and previous periodontitis GWAS may lead to the consideration of the final reasons --- the discrepancy of the periodontitis definition and measurements across periodontitis GWAS. This inconsistent periodontitis definition and measurements influence the number of periodontitis patients identified which therefore affects consistency of the genetic variants discovered and complicate the data synthesis. Meanwhile, inconsistent diseases definition and measurements may also pose challenges in estimating disease prevalence and disease susceptibility across epidemiology studies. It has to acknowledge that the GWAS in the current thesis used suboptimal periodontitis definition – self-reported dental health as surrogate. A standardised measurements and diseases definition should be prompted in later studies and not limited to GWAS.

In terms of replication, it is important that future work replicates across different ethnic groups. However, to date, only a few periodontitis GWAS studies have explored populations other than European. This reflects a broader trend across GWAS investigations where approximately 86% of participants are of European ancestry (Abdellaoui et al., 2023; Fatumo et al., 2022). Due to the potential differences in genetic architecture (e.g., allele frequency) (Abdellaoui et al., 2023) and different genetic risk variants for periodontitis (Shaddox et al., 2021), there is a problem not only in results generalisation (Abdellaoui et al., 2023) but also potential differences in disease susceptibility between ethnic groups. Additionally, since the heritability estimation is time and space dependent and is not transferable across different populations (Barry et al., 2023), the differences in terms of geographical region and the environment of participants, and the interaction between gene and environment, may have a bearing on periodontitis. Despite of a growing GWAS research focus on non-European ancestry, greater emphasis on the ethnic diversity in periodontitis GWAS work is still required. This will not only benefit GWAS research but also enable answers to diseases-specific questions across ethnicities such as causation using MR.

Furthermore, inconsistency in essential information (e.g., missing effect allele, effect allele frequency, standard errors, or confidence intervals) reported by studies, and the unavailability of periodontitis GWAS summary statistics, complicate result interpretation and hinder further aggregate analysis in later studies. For example, an intended meta-analysis was not possible for this reason in Chapter 3.1.

In addition to the challenges in conducting periodontitis GWAS research, there are also some methodological challenges in using MR to answer causal questions between periodontitis and dementia. First, despite being used widely to identify modifiable risk factors for disease, MR is not yet a fully developed research method. Currently, methodological limitations may affect the robustness and interpretation of results. As discussed in Chapters 1 and 4, MR relies on the random assortment of alleles to achieve "randomisation", and three key assumptions (described in Chapter 1.2.1) must be met to ensure valid and reliable results. However, appropriate "genetic randomisation" can be compromised by factors such as population stratification (where genetic variants and phenotypes are associated due to the population structure), dynastic effects (where a parent’s phenotype influences the individual’s phenotype), assortative mating (where individuals with similar phenotypes are more likely to mate), and transmission ratio distortion (where alleles are inherited in unequal proportions instead of the expected 50:50 ratio) (Brumpton et al., 2020; Hartwig et al., 2018; Minică et al., 2020; Morris et al., 2020; Sanderson et al., 2022). These sources of confounding introduce correlations between genetic variants and outcomes, potentially

violating MR assumptions. Current MR methods cannot easily address these issues, and biases in MR studies should be carefully considered when interpreting the results (Sanderson et al., 2022).

Selection bias also complicates the interpretation of MR results due to biases in the data sources (Sanderson et al., 2022). In MR studies, selection bias can arise from the influence of an individual's exposure and outcomes on their participation in a study (Hernán et al., 2004; Sanderson et al., 2022). Moreover, recruitment processes can be affected by survivor bias, where only those who survive are observed for outcomes of interest (Sanderson et al., 2022). For example, although the UKB dataset intended to recruit a representative sample in the UK-wide population, achieving a genotyped sample of approximately 500,000 persons, it has been found that participants from UKB are more likely to be older, female, live in less socioeconomically deprived areas, are less likely to have unhealthy risk factors (e.g., obesity, smoking, alcohol consumption) and self-reported health conditions as compared to non-participants (Fry et al., 2017). As such, UKB is held to be more representative of the heathier UK population. Importantly, previous study found 32 genetic variants associated with whether people would like to participate in UKB (Tyrrell et al., 2021). These include loci associated with Alzheimer's diseases which suggested Alzheimer's diseases reduced participation in UKB. This Alzheimer's diseases and UKB participation link also explained why this thesis contains only 6,000 all-cause dementia patients (~1%-1.5% of the European sample) from such large cohort. These biases arise from participants stage creates the correlations between genetic variants and participation (violating assumption 2) and bias MR result interpretation.

Although many of the MR methodological challenges mention cannot be easily managed, there are several new methods being developed to address such challenges that may lead to better standards, validity and reliability in MR research. For example, methods such as multi-variable MR and within-family MR, are still under development or require larger family-based datasets to implement (Sanderson et al., 2022).

More specific to the two-sample approach, it has been noted that there is an increasing number of studies used two-sample MR due to the growing number of publicly available GWAS summary statistics and the ease of applying two-sample MR (e.g., absence of primary data collection and expertise in the researched field) (Hartwig et al., 2016; Sanderson et al., 2022). However, the concerns of overuse, misuse and poor study quality have been raised as well (Sanderson et al., 2022), such as error in data harmonization process (Haycock et al., 2016). There is an urgent need for MR studies to follow guidelines (e.g., (Burgess et al., 2019)) and improve the study quality. Since the large majority of existing GWAS focus on European ancestry, the use of exiting GWAS summary statistic to perform two-sample MR limited the representativeness of population (Sanderson et al., 2022).

This thesis used two-sample MR on a single dataset (i.e., UKB) to explored the causal relationship between a potentially modifiable risk factor, periodontitis, and dementia outcomes. Although there are limitations in UKB data resource, the size of this data resource and access to the individual level of data allowed this thesis to allocate participants into two sample group and run two sample MR. This attempt overcomes the overfitting bias from using one-sample MR and bias arise from data harmonization processing from using just summary statistics of available GWAS. Future study may benefit from using other data resources which contains other ancestry, clinical examinations on periodontal status. Future study would also benefit from including more dementia patients by using diseases specific cohorts to run two-sample MR. These improvements in the MR study testing the casual relationship between periodontitis and dementia may contribute to obtain a more robust causal estimation and promote quality on medical research. In two-sample MR, careful data harmonisation process following two-sample MR guidance (Burgess et al., 2019; Hartwig et al., 2016; Haycock et al., 2016) is needed. Share of all stage of harmonized datasets and clearly documented harmonisation approach for bias assessment by reviewers or other researchers is highly recommended for future studies (Hartwig et al., 2016).

Chapter 5.3 Strengths and limitations

5.3.1 Strengths

The growing body of evidence linking periodontal diseases with poorer cognitive outcome, especially dementia, have led to a comprehensive investigation in this thesis that provided new insights into the causal relationship between periodontal diseases and poorer cognitive outcomes, specifically periodontitis and dementia.

Despite of growing evidence, previous systematic review identified variation in the association and methodological challenges in previous studies (Larvin et al., 2023). Chapter 2 then provided further strong evidence of the association between oral health (including periodontitis) and cognitive outcomes with investigation of specific cognitive domains using NHANES data, adjusting for various confounding factors via multi-variable regression and SEM. Importantly, the clinically examined periodontal status ensures the validity of the results by accurately identifying patients with periodontal disease. The significant associations found in Chapter 2 led to further investigation of the causal relations.

To the best of my knowledge, Chapter 3.1 is the first systematic review of the current GWAS evidence on periodontitis. The chapter comprehensively reviewed the potential genetic instruments identified in previous studies and discussed the potential limitations and research challenges in periodontitis GWAS investigations. Since the majority of GWAS periodontitis

research identified suffered from limited sample size, the second study in Chapter 3 sought to investigate the genetic instruments for periodontitis using UKB data in order to resolve the limitations in many previous GWAS investigations. Standardised quality control procedures were also performed to ensure the validity of the findings (Marees et al., 2018).

Finally, the Chapter 4 provided one of the first MR investigations using UKB data to address causality between periodontitis and dementia while reducing the possibility of confounding effects and reverse causation (Emdin et al., 2017). The MR study in Chapter 4 involved several MR assumption tests (e.g., MR-Egger regression test for horizontal pleiotropy, multivariable regression for testing the association between genetic instruments and potential confounders), a two-sample MR design using single data resource UKB to avoid overfitting bias and error from the data harmonisation process, as well as application of rigorous MR analysis methods (i.e., inverse-variance weighting MR (Burgess et al., 2013; Burgess et al., 2019), MR-egger regression (Bowden et al., 2015; Bowden et al., 2016b; Burgess et al., 2019), weighted median (Bowden et al., 2016a; Burgess et al., 2019) and mode-based estimation (Hartwig et al., 2017)) that cross-validated the results (Boehm and Zhou, 2022; Sanderson et al., 2022).

Another strength worth emphasising is the use of NHANES and UKB. NHANES is a high quality and representative US data resource ((NCHS), 2022; Dye et al., 2019), which included not only detailed and well-designed periodontal status measurement (Dye et al., 2019), but also a wide range of health-related variables which facilitated investigation of cognitive outcomes while adjusting for a range of potential confounding factors (e.g., life-style factors, demographic features, comorbidities). UKB is a large UK-based sample with a high-quality and wide-ranging collection of health measures including comprehensive dementia data from various sources (i.e., primary care, hospital, self-reported data), as well as genetic data which enabled the use of genetic epidemiological methodology (i.e., GWAS and MR) (Bycroft et al., 2018; Sudlow et al., 2015; Wilkinson et al., 2019).

5.3.2 Limitations

However, the limitations of the research also need to be considered when interpreting the results.

The risk of confounding effects from unobserved confounders and cross-sectional data from NHANES led to difficulties in explaining causality and the impact of periodontitis on long-term cognitive outcomes in the Chapter 2. Specifically, the smoking status and toothbrush behaviour variables were dropped in Chapter 2 due to a large amount of missing data. As these are key lifestyle factors strongly associated with periodontal status, failing to adjust for these variables in

the statistical model may introduce potential bias into the results. Exclusion of edentulous participants also left potential the residual effect of periodontitis on edentulous participants and how it affect cognitive outcome unrevealed (Naorungroj et al., 2015; Stein et al., 2010).

Findings from systematic review in Chapter 3 may be subject to publication bias or bias due to inconsistent definitions of periodontitis and variation in methodologies applied across the included studies as discussed in Chapter 3.2. Especially those using incomplete periodontal examinations (e.g., self-reported measurements, probing depth only, or clinical attachment loss only), which may lead to underreported periodontitis cases and subsequently influence their results. Additionally, the results from some included studies may have been biased due to small sample sizes which could possibly have been addressed through meta-analysis. However, this approach was not possible due to GWAS summary statistics not being available in published studies. Another issue concerned the representativeness of included studies as the majority only represented European and Caucasian populations, thereby leading to limited discussion and investigation of genetic instruments from each ethnic group. This is not only due to the small sample sizes but also resulted from limited investigation of other ethnic samples that included GWAS on periodontitis. It is important that genetic differences between ethnicities are investigated (Barbujani et al., 1997; Huang et al., 2015). Finally, it should be noted that a study on apical periodontitis (Petty et al., 2023) has been included in the systematic review in Chapter 3. Since apical periodontitis is caused by deep dental caries and has a different pathophysiology from general periodontitis, including chronic and aggressive periodontitis (Graunaite et al., 2012), the SNPs discovered for apical periodontitis should be considered and discussed separately. This has not been done due to the limited number of GWAS on apical periodontitis.

The following GWAS study in Chapter 3 on periodontitis using UKB data had similar representational concerns as only European participants were included. Importantly, as described in Chapter 1, periodontitis is a complex disease classified by both grade and stage systems. The use of dichotomous categories for periodontitis may not accurately reflect the extent and content of the disease status. Meanwhile, as mentioned in Chapter 1, gingivitis also presents with painful and bleeding gums. Using painful and bleeding gums as a surrogate for periodontitis may lead to the misclassification of participants with gingivitis as periodontitis cases. The lack of sensitivity in the definition used in Chapter 3 is also reflected by the discrepancy between the high prevalence of periodontitis patients reported in previous epidemiological studies (Frencken et al., 2017; Gaffar et al., 2020) and the relatively low prevalence in the GWAS in Chapter 3, where only around 18% of periodontitis cases were identified. On the other hand, the participants from the UKB are relatively older, where the effects of accumulated environmental factors may outweigh the effects of genetic factors. The bias in the results arising from the periodontitis measurements or definitions and the participants' age might explain the lack of significance of previously

discovered genetic risk factors. Another concern is the inclusion of rarer variants. The current study may not have sufficient power to reliably detect signals from rarer variants, despite the large sample size (n=376,611). Tests such as burden tests or sequence kernel association tests could help aggregate the effects of multiple rarer variants (Lee et al., 2014; Wu et al., 2011) .

MR in Chapter 4 and the GWAS investigation in Chapter 3 have the some limitations related to the UKB data. The UKB data contains only self-reported oral health measures and periodontitis could only be defined based on this surrogate measure. Although valid (Abbood et al., 2016; Papapanou et al., 2018) and used widely in previous studies (e.g., (Kang et al., 2022; Larvin et al., 2020)), it is not as sufficiently optimal as clinical examination and measurement, potentially leading to resulting reduction in power. Another specific issue with UKB data concerns the Chapter 4 results is that the limited dementia patients were identified. This may have compromised the robustness of results and further subgroup investigations. Meanwhile, it should be noted that participants who reported both periodontitis and dementia (n=891, 14.8% of total dementia cases) were allocated to the dementia group to increase the sample size of dementia cases. This may potentially confound the association results by increasing the false positive rate. Therefore, although an inconsistent causal relationship between periodontitis and dementia was observed in Chapter 4, given the above limitation, it should be interpreted with caution.

Another limitation concerning the study design is that the current thesis used only GWAS as the method for discovering genetic instruments for periodontitis. As a result, the genetic risk factors identified in candidate gene studies were neglected. Future studies that select genetic instruments from both GWAS and candidate gene studies may provide a more comprehensive representation of periodontitis.

Chapter 5.4 Implications and future research

5.4.1 Clinical perspectives

The partially significant and lacking strong and robust evidence of causation indicates that periodontal diseases may have a weaker effect on cognitive outcomes, specifically the weaker effect of periodontitis on dementia, than previously expected, or indeed, it may not directly affect cognitive outcomes, dementia in particular. These findings have several important implications for clinical practice.

First, these findings may contribute to updating clinical guidelines to ensure that limited clinical resources are prioritised for more effective evidence-based interventions for dementia prevention, rather than focusing too much on oral health care alone. A more holistic approach to patient care and dementia prevention is recommended, in which clinicians should attempt to address a broader

range of risk factors with a focus on overall health. By combining multiple factors, this may lead to a more effective and significant impact on dementia prevention. For example, integrating physical health (e.g., peripheral inflammation, cardiovascular health) (Saeed et al., 2023; Walker et al., 2023), mental health (Kuring et al., 2018), and lifestyle behaviours (e.g., exercise, dietary habits, and oral hygiene) (Brett et al., 2016; Gao et al., 2020; Volkert et al., 2024) may enhance dementia patient care, prevention or intervention.

However, this does not mean that oral health care, including periodontal care, should be neglected. The widely observed significant association when missing direct causal evidence from periodontitis on dementia might be due to that the poorer cognitive outcomes especially dementia has causal impact on oral health status as suggested by life-course model (Thomson and Barak, 2021) discussed in Chapter 1. Importantly, insufficient oral care implemented in the dementia patient may be represented by worsening oral health outcomes in dementia patients (e.g., gingival bleeding, periodontal pockets) (Delwel et al., 2018), and impacting on oral health related quality of life (Kang et al., 2020). Therefore, improvement in oral care management in dementia patients is still needed for improving quality of life for dementia patients but should be considered as part of holistic care on overall health. On the other hand, periodontitis is associated with many other systematic disease (e.g., cardiovascular disease, diabetes etc.) (Genco and Sanz, 2020; Larvin et al., 2021), which may have impact on dementia development (Dove et al., 2021) or worsening overall health status in dementia patients leading to mortality and poorer quality of life. Therefore, consideration on oral care in the holistic approach with improvement in oral care is necessary for clinician when implementing dementia care, prevention or intervention.

5.4.2 Implications for future research

This thesis also has important research implications regarding the causal association between periodontal diseases and poorer cognitive outcomes, especially periodontitis and dementia. Despite of insignificant evidence of causation, future study on the causal relationship between periodontitis and dementia is meaningful for replication purpose. This is also because the MR evidence on this topic is scarce and improvement in methodology to address current thesis limitations (discussed in Chapter 5.3.2) and current method challenges (discussed in Chapter 5.2 “Methodology challenges”) would also contribute to increase robustness of the findings in this thesis. For example, sample size calculation with consideration on the low heritability of diseases, increased dementia patient sample size using disease specific cohorts, use of clinical examination for periodontitis measurement plus use of standardised periodontitis definition criteria would help to increase study robustness and result interpretation. Multivariable MR could be used to tackle the potential effects of selection bias (discussed in Section 5.2) and exploration of the potential mediation effects of other covariates (e.g., smoking, age-related inflammatory markers level) is

also worthy of investigation. Shared genes (e.g., *IL* genes) between both diseases should be considered with cautious while selecting the genetic instruments and test for the assumption 3 of MR. Future studies would also benefit from a more comprehensive consideration of genetic instruments by including strong genetic variants identified in genetic studies, such as candidate gene studies, rather than relying solely on GWAS.

Furthermore, investigating the subgroup interactions such as association between periodontitis and dementia subgroups (i.e., different types of dementia), as well as exploring associations in other ethnicities, such as Asian and Black populations, may provide important insights concerning the role of periodontal diseases in poorer cognitive outcome, especially dementia.

The current research focus, "finding the causal link", could shift towards seeking explanation for observed association between periodontal disease and poorer cognitive outcomes especially dementia. For example, investigations on whether there are indirect effects from periodontitis via other risk factor (e.g., increased risk of comorbidities like diabetes), or effects from shared risk factors (e.g., smoking, diabetes, genetics, nutrition) on both periodontitis and dementia are worth investigating in future epidemiological studies. The shared genes such as *IL* genes and the level of related pro-inflammatory cytokines throughout body from independent causes (e.g., age) are particularly of interest due to their role in inflammation and both diseases. Additionally, biomedical studies are still needed in order to understand the potential mechanisms underlying the association to answer questions such as "why there is association between periodontitis and dementia when there is no causal impact from periodontitis on dementia" which could contribute to improved understanding of disease pathology.

In addition to above future research directions, several challenges in research methods should also be addressed. First, a tool and guidelines for evaluating GWAS, performing and reporting GWAS, and GWAS systematic reviews should be followed for standardised GWAS study design which would further enable data synthesis. Second, shared or publicly accessible GWAS summary statistics should be advocated to enable a range of secondary analyses (e.g., meta-analysis, MR, fine-mapping) without requiring individual level data by other researchers. For example, registration and GWAS summary statistics sharing at GWAS catalog (https://www.ebi.ac.uk/gwas/) and IEU open GWAS project (https://gwas.mrcieu.ac.uk/). The current sharing of GWAS summary statistics remains low and lacks commonly observed standards (MacArthur et al., 2021), despite an increase in recent years in some fields (e.g., the sharing rate of GWAS summary statistics in PubMed molecular epidemiology GWAS publications rose from roughly 3% in 2010 to 23% in 2017) (Thelwall et al., 2020). Recommendations in MacArthur et al. (2021) should be observed by future GWAS investigations in order to advance genomic medicine. MR study design following the guideline of performing

MR (e.g., (Burgess et al., 2019)) and improving two-sample MR study quality is also expected in the future research.

Chapter 5.5 Conclusions

This thesis provides an in-depth investigation into the causal association between periodontal diseases and poorer cognitive outcomes, specifically periodontitis and dementia, using various approaches including Mendelian Randomisation approaches, and found no strong evidence of a significant causal effect between periodontitis and dementia. The thesis started with an cross-sectional study validating and quantified the association between periodontal diseases and poorer cognitive outcomes using advanced statistical approaches (i.e., SEM) using high-quality cross-sectional data from NHANES. The significant association identified led to further investigation of causation using the MR approach. Prior to the MR, genetic instruments of periodontitis were investigated by both systematic review and GWAS. Finally, an MR study was performed highlighting the potential of this advance approach in epidemiological research to reveal causal relationships and identify modifiable risk factors for improving health outcomes thereby aiding disease prevention and intervention.

In conclusion, the evidence presented in the thesis suggests that periodontitis is unlikely to be a direct causal risk factor for dementia, with partial significant and the inconsistent direction of findings suggesting that the results should be carefully interpreted and further investigation is required. Although a significant causal association was not identified, the findings of this thesis help address controversies in the field. The results also highlight the importance of biological mechanisms underlying the observed associations and underline the need for future research to explore how the mechanisms affect the direction of causality between periodontal diseases and poorer cognitive outcome, especially periodontitis and dementia. This work also contributes to clinical practice and interventions by suggesting a more holistic patient care or treatment approach including dental care for dementia patient.

Chapter 5.6 References

(NCHS) NCfHS. 2022. National health and nutrition examination survey.

Abbood HM, Hinz J, Cherukara G, Macfarlane TV. 2016. Validity of self-reported periodontal disease: A systematic review and meta-analysis. Journal of periodontology (1970). 87(12):1474-1483. doi: 10.1902/jop.2016.160196.

Abdellaoui A, Yengo L, Verweij KJH, Visscher PM. 2023. 15 years of gwas discovery: Realizing the promise. American journal of human genetics. 110(2):179-194. doi: 10.1016/j.ajhg.2022.12.011.

Andrews SJ, Renton AE, Fulton-Howard B, Podlesny-Drabiniok A, Marcora E, Goate AM. 2023. The complex genetic architecture of alzheimer's disease: Novel insights and future directions. EBioMedicine. 90:104511. doi: 10.1016/j.ebiom.2023.104511.

Barbujani G, Magagni A, Minch E, Cavalli-Sforza LL. 1997. An apportionment of human DNA diversity. Proceedings of the National Academy of Sciences - PNAS. 94(9):4516-4519. doi: 10.1073/pnas.94.9.4516.

Barry C-JS, Walker VM, Cheesman R, Davey Smith G, Morris TT, Davies NM. 2023. How to estimate heritability: A guide for genetic epidemiologists. International journal of epidemiology. 52(2):624-632. doi: 10.1093/ije/dyac224.

Boehm FJ, Zhou X. 2022. Statistical methods for mendelian randomization in genome-wide association studies: A review. Computational and structural biotechnology journal. 20:2338-2351. doi: 10.1016/j.csbj.2022.05.015.

Bowden J, Davey Smith G, Burgess S. 2015. Mendelian randomization with invalid instruments: Effect estimation and bias detection through egger regression. International journal of epidemiology. 44(2):512-525. doi: 10.1093/ije/dyv080.

Bowden J, Davey Smith G, Haycock PC, Burgess S. 2016a. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. Genetic epidemiology. 40(4):304-314. doi: 10.1002/gepi.21965.

Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. 2016b. Assessing the suitability of summary data for two-sample mendelian randomization analyses using mr-egger regression: The role of the i2 statistic. International journal of epidemiology. 45(6):1961-1974. doi: 10.1093/ije/dyw220.

Brett LBH, Traynor VPBH, Stapley PPMBAH. 2016. Effects of physical exercise on health and well-being of individuals living with a dementia in nursing homes: A systematic review. Journal of the American Medical Directors Association. 17(2):104-116. doi: 10.1016/j.jamda.2015.08.016.

Brumpton B, Sanderson E, Heilbron K, Hartwig FP, Harrison S, Vie GÅ, Cho Y, Howe LD, Hughes A, Boomsma DI et al. 2020. Avoiding dynastic, assortative mating, and population stratification biases in mendelian randomization through within-family analyses. Nature communications. 11(1):3519-3513. doi: 10.1038/s41467-020-17117-4.

Burgess S, Butterworth A, Thompson SG. 2013. Mendelian randomization analysis with multiple genetic variants using summarized data. Genetic epidemiology. 37(7):658-665. doi: 10.1002/gepi.21758.

Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Kutalik Z, Holmes MV, Minelli C et al. 2019. Guidelines for performing mendelian randomization investigations: Update for summer 2023. Wellcome open research. 4:186. doi: 10.12688/wellcomeopenres.15555.3.

Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J et al. 2018. The uk biobank resource with deep phenotyping and genomic data. Nature (London). 562(7726):203-209. doi: 10.1038/s41586-018-0579-z.

de Coo A, Cruz R, Quintela I, Herrera D, Sanz M, Diz P, Rodríguez Grandío S, Vallcorba N, Ramos I, Oteo A et al. 2021. Genome-wide association study of stage iii/iv grade c periodontitis (former aggressive periodontitis) in a spanish population. J Clin Periodontol. 48(7):896-906. doi: 10.1111/jcpe.13460.

Delwel S, Binnekade TT, Perez RSGM, Hertogh CMPM, Scherder EJA, Lobbezoo F. 2018. Oral hygiene and oral health in older people with dementia: A comprehensive review with focus on oral soft tissues. Clinical oral investigations. 22(1):93-108. doi: 10.1007/s00784-017-2264-2.

Dove A, Shang Y, Xu W, Grande G, Laukka EJ, Fratiglioni L, Marseglia A. 2021. The impact of diabetes on cognitive impairment and its progression to dementia. Alzheimer's & dementia. 17(11):1769-1778. doi: 10.1002/alz.12482.

Dye BA, Afful J, Thornton-Evans G, Iafolla T. 2019. Overview and quality assurance for the oral health component of the national health and nutrition examination survey (nhanes), 2011-2014. BMC oral health. 19(1):95-11. doi: 10.1186/s12903-019-0777-6.

Emdin CA, Khera AV, Kathiresan S. 2017. Mendelian randomization. JAMA : the journal of the American Medical Association. 318(19):1925-1926. doi: 10.1001/jama.2017.17219.

Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR, Kuchenbaecker K. 2022. A roadmap to increase diversity in genomic studies. Nature medicine. 28(2):243-250. doi: 10.1038/s41591-021-01672-4.

Feng P, Wang X, Casado PL, Küchler EC, Deeley K, Noel J, Kimm H, Kim JH, Haas AN, Quinelato V et al. 2014. Genome wide association scan for chronic periodontitis implicates novel locus. BMC Oral Health. 14:84. doi: 10.1186/1472-6831-14-84.

Feng S, Wang S, Chen C-C, Lan L. 2011. Gwapower: A statistical power calculation software for genome-wide association studies with quantitative traits. BMC genetics. 12(1):12-12. doi: 10.1186/1471-2156-12-12.

Frencken JE, Sharma P, Stenhouse L, Green D, Laverty D, Dietrich T. 2017. Global epidemiology of dental caries and severe periodontitis – a comprehensive review. Journal of clinical periodontology. 44(S18):S94-S105. doi: 10.1111/jcpe.12677.

Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE. 2017. Comparison of sociodemographic and health-related characteristics of uk biobank participants with those of the general population. American journal of epidemiology. 186(9):1026-1034. doi: 10.1093/aje/kwx246.

Gaffar B, al-Hariki M, Al-Khalifa K, Al-Ansari A, Nazir M, Almas K, Nammour S, Samir N. 2020. Global prevalence of periodontal disease and lack of its surveillance. TheScientificWorld. 2020(2020):1-8. doi: 10.1155/2020/2146160.

Gallagher MD, Chen-Plotkin AS. 2018. The post-gwas era: From association to function. The American Journal of Human Genetics. 102(5):717-730. doi.

Gao SS, Chu CH, Young FYF. 2020. Oral health and care for elderly people with alzheimer's disease. International journal of environmental research and public health. 17(16):5713. doi: 10.3390/ijerph17165713.

Genco RJ, Sanz M. 2020. Clinical and public health implications of periodontal and systemic diseases: An overview. Periodontology 2000. 83(1):7-13. doi: 10.1111/prd.12344.

Génin E. 2020. Missing heritability of complex diseases: Case solved? Human Genetics. 139(1):103-113. doi: 10.1007/s00439-019-02034-4.

Graunaite I, Lodiene G, Maciulskiene V. 2012. Pathogenesis of apical periodontitis: A literature review. Journal of oral & maxillofacial research. 2(4):e1-e1. doi: 10.5037/jomr.2011.2401.

Hartwig FP, Davies NM, Davey Smith G. 2018. Bias in mendelian randomization due to assortative mating. Genetic epidemiology. 42(7):608-620. doi: 10.1002/gepi.22138.

Hartwig FP, Davies NM, Hemani G, Davey Smith G. 2016. Two-sample mendelian randomization: Avoiding the downsides of a powerful, widely applicable but potentially fallible technique. International journal of epidemiology. 45(6):1717-1726. doi: 10.1093/ije/dyx028.

Hartwig FP, Smith GD, Bowden J. 2017. Robust inference in summary data mendelian randomization via the zero modal pleiotropy assumption. International journal of epidemiology. 46(6):1985-1998. doi: 10.1093/ije/dyx102.

Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. 2016. Best (but oft-forgotten) practices: The design, analysis, and interpretation of mendelian randomization studies. The American journal of clinical nutrition. 103(4):965-978. doi: 10.3945/ajcn.115.118216.

Heiss C, Rodriguez-Mateos A, Kelm M. 2015. Central role of enos in the maintenance of endothelial homeostasis. Antioxid Redox Signal. 22(14):1230-1242. doi: 10.1089/ars.2014.6158.

Hernán MA, Hernández-Díaz S, Robins JM. 2004. A structural approach to selection bias. Epidemiology (Cambridge, Mass). 15(5):615-625. doi: 10.1097/01.ede.0000135174.63482.43.

Huang T, Shu Y, Cai Y-D. 2015. Genetic differences among ethnic groups. BMC genomics. 16(1087):1093-1093. doi: 10.1186/s12864-015-2328-0.

Ioannidis JPA. 2005. Why most published research findings are false. PLoS medicine. 2(8):e124-e124. doi: 10.1371/journal.pmed.0020124.

Kamer AR, Craig RG, Dasanayake AP, Brys M, Glodzik-Sobanska L, de Leon MJ. 2008. Inflammation and alzheimer's disease: Possible role of periodontal diseases. Alzheimer's & dementia. 4(4):242-250. doi: 10.1016/j.jalz.2007.08.004.

Kang J, Palmier-Claus J, Wu J, Shiers D, Larvin H, Doran T, Aggarwal VR. 2022. Periodontal disease in people with a history of psychosis: Results from the uk biobank population-based study. Community dentistry and oral epidemiology. doi: 10.1111/cdoe.12798.

Kang J, Wu B, Bunce D, Ide M, Aggarwal VR, Pavitt S, Wu J. 2020. Bidirectional relations between cognitive function and oral health in ageing persons: A longitudinal cohort study. Age and ageing. 49(5):793-799. doi: 10.1093/ageing/afaa025.

Kuring JK, Mathias JL, Ward L. 2018. Prevalence of depression, anxiety and ptsd in people with dementia: A systematic review and meta-analysis. 28(4):393-416. doi.

Lang NP, Bartold PM. 2018. Periodontal health. Journal of clinical periodontology. 45(S20):S9-S16. doi: 10.1111/jcpe.12936.

Larvin H, Gao C, Kang J, Aggarwal VR, Pavitt S, Wu J. 2023. The impact of study factors in the association of periodontal disease and cognitive disorders: Systematic review and meta-analysis. Age and ageing. 52(2). doi: 10.1093/ageing/afad015.

Larvin H, Kang J, Aggarwal VR, Pavitt S, Wu J. 2021. Risk of incident cardiovascular disease in people with periodontal disease: A systematic review and meta-analysis. Clinical and experimental dental research. 7(1):109-122. doi: 10.1002/cre2.336.

Larvin H, Wilmott S, Wu J, Kang J. 2020. The impact of periodontal disease on hospital admission and mortality during covid-19 pandemic. Frontiers in medicine. 7:604980-604980. doi: 10.3389/fmed.2020.604980.

Lee S, Abecasis Gonçalo R, Boehnke M, Lin X. 2014. Rare-variant association analysis: Study designs and statistical tests. American journal of human genetics. 95(1):5-23. doi: 10.1016/j.ajhg.2014.06.009.

Leng F, Edison P. 2021. Neuroinflammation and microglial activation in alzheimer disease: Where do we go from here? Nature reviews Neurology. 17(3):157-172. doi: 10.1038/s41582-020-00435-y.

Lin Y-J, Chang J-S, Liu X, Tsang H, Chien W-K, Chen J-H, Hsieh H-Y, Hsueh K-C, Shiao Y-T, Li J-P et al. 2015. Genetic variants in plcb4/plcb1 as susceptibility loci for coronary artery aneurysm formation in kawasaki disease in han chinese in taiwan. Scientific reports. 5(1):14762-14762. doi: 10.1038/srep14762.

Loos, G. B, Van Dyke, E. T. 2020. The role of inflammation and genetics in periodontal disease. Periodontol 2000. 83(1):26-39. doi: 10.1111/prd.12297.

Loos BG, Van Dyke TE. 2020. The role of inflammation and genetics in periodontal disease. Periodontology 2000. 83(1):26-39. doi: 10.1111/prd.12297.

MacArthur JA, Buniello A, Harris LW, Hayhurst J, McMahon A, Sollis E, Cerezo M, Hall P, Lewis E, Whetzel PL. 2021. Workshop proceedings: Gwas summary statistics standards and sharing. Cell Genomics. 1(1). doi.

Marees AT, Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, Derks EM. 2018. A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. International journal of methods in psychiatric research. 27(2):e1608-n/a. doi: 10.1002/mpr.1608.

Marigorta UM, Rodríguez JA, Gibson G, Navarro A. 2018. Replicability and prediction: Lessons and challenges from gwas. Trends in genetics. 34(7):504-517. doi: 10.1016/j.tig.2018.03.005.

Minică CC, Boomsma DI, Dolan CV, de Geus E, Neale MC. 2020. Empirical comparisons of multiple mendelian randomization approaches in the presence of assortative mating. International journal of epidemiology. 49(4):1185-1193. doi: 10.1093/ije/dyaa013.

Morris TT, Davies NM, Hemani G, Smith GD. 2020. Population phenomena inflate genetic associations of complex social traits. Science advances. 6(16):eaay0328-eaay0328. doi: 10.1126/sciadv.aay0328.

Mounier N, Kutalik Z. 2023. Bias correction for inverse variance weighting mendelian randomization. Genetic epidemiology. 47(4):314-331. doi: 10.1002/gepi.22522.

Mueller R, Chopra A, Dommisch H, Schaefer AS. 2022. Periodontitis risk variants at siglec5 impair erg and mafb binding. Journal of dental research. 101(5):551-558. doi: 10.1177/00220345211049984.

Munz M, Richter GM, Loos BG, Jepsen S, Divaris K, Offenbacher S, Teumer A, Holtfreter B, Kocher T, Bruckmann C et al. 2019. Meta-analysis of genome-wide association studies of aggressive and chronic periodontitis identifies two novel risk loci. Eur J Hum Genet. 27(1):102-113. doi: 10.1038/s41431-018-0265-5.

Munz M, Willenborg C, Richter GM, Jockel-Schneider Y, Graetz C, Staufenbiel I, Wellmann J, Berger K, Krone B, Hoffmann P et al. 2017. A genome-wide association study identifies nucleotide variants at siglec5 and defa1a3 as risk loci for periodontitis. Hum Mol Genet. 26(13):2577-2588. doi: 10.1093/hmg/ddx151.

Naorungroj S, Schoenbach VJ, Wruck L, Mosley TH, Gottesman RF, Alonso A, Heiss G, Beck J, Slade GD. 2015. Tooth loss, periodontal disease, and cognitive decline in the atherosclerosis risk in communities (aric) study. Community dentistry and oral epidemiology. 43(1):47-57. doi: 10.1111/cdoe.12128.

Nibali L, Bayliss-Chapman J, Almofareh SA, Zhou Y, Divaris K, Vieira AR. 2019. What is the heritability of periodontitis? A systematic review. Journal of Dental Research. 98(6):632-641. doi: 10.1177/0022034519842510.

Papapanou PN, Sanz M, Buduneli N, Dietrich T, Feres M, Fine DH, Flemmig TF, Garcia R, Giannobile WV, Graziani F et al. 2018. Periodontitis: Consensus report of workgroup 2 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. J Clin Periodontol. 45 Suppl 20:S162-s170. doi: 10.1111/jcpe.12946.

Petty LE, Silva R, Chaves de Souza L, Vieira AR, Shaw DM, Below JE, Letra A. 2023. Genome-wide association study identifies novel risk loci for apical periodontitis. Journal of endodontics. doi: 10.1016/j.joen.2023.07.018.

Rea IM, Gibson DS, McGilligan V, McNerlan SE, Alexander HD, Ross OA. 2018. Age and age-related diseases: Role of inflammation triggers and cytokines. Frontiers in immunology. 9:586. doi: 10.3389/fimmu.2018.00586.

Saeed A, Lopez O, Cohen A, Reis SE. 2023. Cardiovascular disease and alzheimer's disease: The heart-brain axis. Journal of the American Heart Association. 12(21):e030780-e030780. doi: 10.1161/JAHA.123.030780.

Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, Palmer T, Schooling CM, Wallace C, Zhao Q, Davey Smith G. 2022. Mendelian randomization. Nature Reviews Methods Primers. 2(1):6. doi: 10.1038/s43586-021-00092-5.

Sawabe A, Okazaki S, Nakamura A, Goitsuka R, Kaifu T. 2024. The orphan g protein-coupled receptor 141 expressed in myeloid cells functions as an inflammation suppressor. Journal of leukocyte biology. 115(5):935-945. doi: 10.1093/jleuko/qiae009.

Schlitzer A, McGovern N, Ginhoux F. 2015. Dendritic cells and monocyte-derived cells: Two complementary and integrated functional systems. Seminars in cell & developmental biology. 41:9-22. doi: 10.1016/j.semcdb.2015.03.011.

Shaddox LM, Morford LA, Nibali L. 2021. Periodontal health and disease: The contribution of genetics. Periodontology 2000. 85(1):161-181. doi: 10.1111/prd.12357.

Shimizu S, Momozawa Y, Takahashi A, Nagasawa T, Ashikawa K, Terada Y, Izumi Y, Kobayashi H, Tsuji M, Kubo M, Furuichi Y. 2015. A genome-wide association study of periodontitis in a japanese population. J Dent Res. 94(4):555-561. doi: 10.1177/0022034515570315.

Shimizu T, Deeley K, Briseño-Ruiz J, Faraco IM, Jr., Poletta FA, Brancher JA, Pecharki GD, Küchler EC, Tannure PN, Lips A et al. 2013. Fine-mapping of 5q12.1-13.3 unveils new genetic contributors to caries. Caries Res. 47(4):273-283. doi: 10.1159/000346278.

Shungin D, Haworth S, Divaris K, Agler CS, Kamatani Y, Keun Lee M, Grinde K, Hindy G, Alaraudanjoki V, Pesonen P et al. 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.

Sohani ZN, Meyre D, de Souza RJ, Joseph PG, Gandhi M, Dennis BB, Norman G, Anand SS. 2015. Assessing the quality of published genetic association studies in meta-analyses: The quality of genetic studies (q-genie) tool. BMC genetics. 16(1):50-50. doi: 10.1186/s12863-015-0211-2.

Sohani ZN, Sarma S, Alyass A, de Souza RJ, Robiou-du-Pont S, Li A, Mayhew A, Yazdi F, Reddon H, Lamri A et al. 2016. Empirical evaluation of the q-genie tool: A protocol for assessment of effectiveness. BMJ Open. 6(6):e010403-e010403. doi: 10.1136/bmjopen-2015-010403.

Stein PS, Kryscio RJ, Desrosiers M, Donegan SJ, Gibbs MB. 2010. Tooth loss, apolipoprotein e, and decline in delayed word recall. Journal of dental research. 89(5):473-477. doi: 10.1177/0022034509357881.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M et al. 2015. Uk biobank: An open access resource for identifying the causes

of a wide range of complex diseases of middle and old age. PLoS medicine. 12(3):e1001779-e1001779. doi: 10.1371/journal.pmed.1001779.

Teumer A, Holtfreter B, Völker U, Petersmann A, Nauck M, Biffar R, Völzke H, Kroemer HK, Meisel P, Homuth G, Kocher T. 2013. Genome-wide association study of chronic periodontitis in a general german population. J Clin Periodontol. 40(11):977-985. doi: 10.1111/jcpe.12154.

Thelwall M, Munafò M, Mas-Bleda A, Stuart E, Makita M, Weigert V, Keene C, Khan N, Drax K, Kousha K. 2020. Is useful research data usually shared? An investigation of genome-wide association study summary statistics. Plos one. 15(2):e0229578. doi.

Thomson WM, Barak Y. 2021. Tooth loss and dementia: A critical examination. Journal of Dental Research. 100(3):226-231. doi: 10.1177/0022034520957233.

Tong H, Wei Z, Yin J, Zhang B, Zhang T, Deng C, Huang Y, Zhang N. 2019. Genetic susceptibility of common polymorphisms in nin and siglec5 to chronic periodontitis. Scientific reports. 9(1):2088. doi: 10.1038/s41598-019-38632-5.

Tyrrell J, Zheng J, Beaumont R, Hinton K, Richardson TG, Wood AR, Davey Smith G, Frayling TM, Tilling K. 2021. Genetic predictors of participation in optional components of uk biobank. Nature communications. 12(1):886-886. doi: 10.1038/s41467-021-21073-y.

Volkert D, Beck AM, Faxén-Irving G, Frühwald T, Hooper L, Keller H, Porter J, Rothenberg E, Suominen M, Wirth R, Chourdakis M. 2024. Espen guideline on nutrition and hydration in dementia – update 2024. Clinical nutrition (Edinburgh, Scotland). 43(6):1599-1626. doi: 10.1016/j.clnu.2024.04.039.

Vuchkovska A, Glanville DG, Scurti GM, Nishimura Michael I, White P, Ulijasz AT, Iwashima M. 2022. Siglec-5 is an inhibitory immune checkpoint molecule for human t cells. Immunology. 166(2):238-248. doi: 10.1111/imm.13470.

Walker KA, Le Page LM, Terrando N, Duggan MR, Heneka MT, Bettcher BM. 2023. The role of peripheral inflammatory insults in alzheimer's disease: A review and research roadmap. Molecular neurodegeneration. 18(1):37-37. doi: 10.1186/s13024-023-00627-2.

Wang M, Wang Z, Zhao D, Yu Y, Wei F. 2024. Periodontitis causally affects the brain cortical structure: A mendelian randomization study. Journal of periodontal research. 59(2):381-386. doi: 10.1111/jre.13222.

Wang Y, Huynh W, Skokan TD, Lu W, Weiss A, Vale RD. 2019. Cracr2a is a calcium-activated dynein adaptor protein that regulates endocytic traffic. The Journal of cell biology. 218(5):1619-1633. doi: 10.1083/jcb.201806097.

Wilkinson T, Schnier C, Bush K, Rannikmäe K, Henshall DE, Lerpiniere C, Allen NE, Flaig R, Russ TC, Bathgate D et al. 2019. Identifying dementia outcomes in uk biobank: A validation study of primary care, hospital admissions and mortality data. European journal of epidemiology. 34(6):557-565. doi: 10.1007/s10654-019-00499-1.

Wu B, Rice L, Shrimpton J, Lawless D, Walker K, Carter C, McKeown L, Anwar R, Doody GM, Srikanth S et al. 2021. Biallelic mutations in calcium release activated channel regulator 2a (cracr2a) cause a primary immunodeficiency disorder. eLife. 10. doi: 10.7554/eLife.72559.

Wu Michael C, Lee S, Cai T, Li Y, Boehnke M, Lin X. 2011. Rare-variant association testing for sequencing data with the sequence kernel association test. American journal of human genetics. 89(1):82-93. doi: 10.1016/j.ajhg.2011.05.029.

Xia M, Liu J, Wu X, Liu S, Li G, Han C, Song L, Li Z, Wang Q, Wang J et al. 2013. Histone methyltransferase ash1l suppresses interleukin-6 production and inflammatory autoimmune diseases by inducing the ubiquitin-editing enzyme a20. Immunity. 39(3):470-481. doi: 10.1016/j.immuni.2013.08.016.

Young AI, Flint J. 2019. Solving the missing heritability problem. PLoS genetics. 15(6):e1008222-e1008222. doi: 10.1371/journal.pgen.1008222.

Zhang Z, Zhu Z, Wang X, Liu D, Liu X, Mi Z, Tao H, Fan H. 2023. Comprehensive landscape of immune-based classifier related to early diagnosis and macrophage m1 in spinal cord injury. Aging (Albany, NY). 15(4):1158-1176. doi: 10.18632/aging.204548.

Appendices

Appendix a. Chapter 2.1 supplementary file

Appendix Table a.1. Sample characteristics on cognitive function in quartile: NHANES (n = 2,508), 2011-2014.

	Overall	Cognitive function				
		Q1 (lowest)	Q2	Q3	Q4 (highest)	P value
N	2508	605	605	605	606	
Demographic						
Gender, male (%)	1257 (50.1)	352 (58.2)	335 (55.4)	294 (48.6)	235 (38.8)	<0.001
Age, mean (SD)	69.3 (6.7)	71.8 (7.0)	70.3 (6.8)	68.4 (6.3)	66.3 (5.5)	<0.001
Race (%)						<0.001
Mexican American	225 (9.0)	63 (10.4)	53 (8.8)	54 (8.9)	39 (6.4)	
Other Hispanic	264 (10.5)	97 (16.0)	72 (11.9)	50 (8.3)	35 (5.8)	
Non-Hispanic White	1149 (45.8)	200 (33.1)	253 (41.8)	285 (47.1)	393 (64.9)	
Non-Hispanic Black	624 (24.9)	193 (31.9)	170 (28.1)	141 (23.3)	87 (14.4)	
Other	246 (9.8)	52 (8.6)	57 (9.4)	75 (12.4)	52 (8.6)	
Education, College or Above (%)	1241 (49.5)	142 (23.5)	262 (43.3)	354 (58.5)	466 (76.9)	<0.001
Marital Status, Widowed/Not Married (%)	1070 (42.7)	295 (48.8)	267 (44.2)	242 (40.1)	217 (35.9)	<0.001
Income Ratio, Mean (SD)	2.6 (1.6)	1.9 (1.4)	2.3 (1.5)	2.8 (1.6)	3.4 (1.6)	<0.001
Lifestyle						
At least 12 alcohol drink/ years, Yes (%)	1690 (68.2)	375 (63.1)	406 (67.6)	414 (68.9)	461 (76.5)	<0.001
Number Of Cigarette Smoke in Last 30 Days, Mean (SD)	11.1 (8.5)	9.6 (7.7)	13.6 (9.5)	10.0 (8.1)	10.8 (8.3)	0.006
Smoking Status (%)						0.084
Smoker	343 (13.7)	94 (15.5)	92 (15.2)	73 (12.1)	70 (11.6)	

Non-Smoker	935 (37.3)	234 (38.7)	235 (38.8)	225 (37.2)	217 (35.8)	
Unknown	1230 (49.0)	277 (45.8)	278 (46.0)	307 (50.7)	319 (52.6)	
Drug Use (%)						<0.001
Yes	183 (7.3)	24 (4.0)	32 (5.3)	50 (8.3)	75 (12.4)	
No	1157 (46.1)	207 (34.2)	262 (43.3)	307 (50.7)	364 (60.1)	
Unknown	1168 (46.6)	374 (61.8)	311 (51.4)	248 (41.0)	167 (27.6)	
Sugar Intake (gm/day), Mean (SD)	94.9 (52.1)	87.0 (48.8)	97.5 (55.2)	97.9 (51.0)	97.9 (50.8)	<0.001
Carbohydrate Intake (gm/day), Mean (SD)	221.3 (88.9)	204.4 (83.4)	226.1 (90.1)	227.3 (89.4)	229.1 (87.5)	<0.001
Kcal Intake (kcal/day), Mean (SD)	1812.9 (691.8)	1616.4(654.1)	1835.5 (711.2)	1874.1(689.4)	1944.2 (660.7)	<0.001
Physical Activity (%)	1466 (58.5)	286 (47.3)	333 (55.0)	393 (65.0)	423 (69.8)	<0.001
Anthropometric						
BMI (%)						0.551
Underweight	94 (3.8)	29 (4.9)	22 (3.7)	22 (3.6)	20 (3.3)	
Normal	596 (24.0)	143 (24.1)	155 (26.0)	138 (22.9)	138 (22.8)	
Overweight	886 (35.7)	217 (36.6)	216 (36.2)	215 (35.7)	208 (34.4)	
Obese	906 (36.5)	204 (34.4)	203 (34.1)	228 (37.8)	239 (39.5)	
Systolic Blood Pressure/mmhg, Mean (SD)	133.1 (19.4)	137.1 (20.1)	134.1 (20.5)	131.8 (18.0)	128.6 (16.6)	<0.001
Diastolic Blood Pressure /mmhg, Mean (SD)	70.5 (10.4)	69.9 (10.8)	69.8 (10.8)	70.7 (10.1)	71.7 (9.6)	0.008
Waist (%)						0.008
Low	375 (15.7)	105 (18.7)	101 (17.6)	85 (14.3)	75 (12.6)	
High	459 (19.2)	111 (19.8)	124 (21.6)	107 (18.0)	104 (17.5)	
Very High	1562 (65.2)	346 (61.6)	348 (60.7)	403 (67.7)	415 (69.9)	
Comorbidities						
Diabetes (%)	586 (24.5)	181 (31.4)	153 (26.6)	131 (22.6)	91 (15.8)	<0.001

Arthritis (%)	1185 (47.3)	303 (50.5)	268 (44.3)	276 (45.6)	297 (49.0)	0.109
Congestive Heart Failure (%)	154 (6.2)	47 (7.8)	50 (8.3)	28 (4.6)	19 (3.1)	<0.001
Coronary Heart Disease (%)	195 (7.8)	54 (8.9)	51 (8.5)	51 (8.4)	30 (5.0)	0.034
Angina (%)	121 (4.8)	30 (5.0)	36 (6.0)	30 (5.0)	20 (3.3)	0.182
Heart Attack (%)	206 (8.2)	63 (10.4)	52 (8.6)	51 (8.4)	29 (4.8)	0.003
Stroke (%)	167 (6.7)	64 (10.6)	46 (7.6)	21 (3.5)	23 (3.8)	<0.001
Liver (%)	136 (5.4)	28 (4.6)	40 (6.6)	34 (5.6)	31 (5.1)	0.48
Sleep (%)	286 (11.4)	57 (9.5)	76 (12.6)	66 (10.9)	76 (12.5)	0.265
Depression (PHQ9 Score), Mean (SD)	3.1 (4.2)	3.8 (5.1)	3.2 (4.3)	2.6 (3.8)	2.5 (3.4)	<0.001
Cognitive Function						
Memory, Mean (SD)	6.0 (1.8)	4.1 (1.3)	5.8 (1.1)	6.6 (1.1)	7.7 (1.0)	<0.001
Executive Function, Mean (SD)	3.6 (1.5)	2.3 (1.0)	3.2 (0.9)	4.0 (1.0)	5.20 (1.3)	<0.001
Processing Speed, Mean (SD)	4.4 (1.6)	2.7 (1.0)	3.8 (1.0)	4.89 (0.9)	6.2 (1.0)	<0.001
Global Cognitive Function, Mean (SD)	14.1 (3.9)	9.1 (1.8)	12.8 (0.8)	15.5 (0.8)	19.1 (1.8)	<0.001
Oral Health						
Missing Teeth, Mean (SD)	11.9 (10.2)	16.1 (9.9)	13.6 (10.0)	10.1 (9.5)	7.0 (8.6)	<0.001
Dentate (%)	2046 (81.6)	437 (72.2)	467 (77.2)	525 (86.8)	558 (92.1)	<0.001
DMFT, Mean (SD)	18.4 (7.3)	20.3 (7.5)	19.1 (7.25)	17.5 (7.2)	16.0 (6.6)	<0.001
Periodontitis Severity (%)						<0.001
None	1268 (50.6)	291 (48.1)	291 (48.1)	287 (47.4)	354 (58.4)	
Mild-Moderate	978 (39.0)	239 (39.5)	236 (39.0)	263 (43.5)	206 (34.0)	
Severe	262 (10.4)	75 (12.4)	78 (12.9)	55 (9.1)	46 (7.6)	
Oral Hygiene Behaviours						
Dental Visit, over 1 Year (%)	1060 (42.3)	356 (59.0)	277 (45.9)	226 (37.4)	153 (25.2)	<0.001
Dental Floss (%)						<0.001

No	1086 (43.4)	370 (61.4)	280 (46.5)	201 (33.2)	166 (27.4)
Not Every Day	585 (23.4)	102 (16.9)	122 (20.3)	160 (26.4)	193 (31.8)
Everyday	832 (33.2)	131 (21.7)	200 (33.2)	244 (40.3)	247 (40.8)

N, number of participants; SD, standard deviation.

Appendix Table a.2. Sample characteristics concerning the dentition and periodontal status: NHANES (n = 2,508), 2011-2014.

	Periodontal	**Status**		**Dentition status**		
	Healthy	Periodontal disease	P value	Edentulous	Dentate	P value
N	1268	1240		462	2046	
Demographic						
Gender, male (%)	536 (42.3)	721 (58.1)	<0.001	236 (51.1)	1021 (49.9)	0.684
Age, mean (SD)	69.76 (6.80)	68.91 (6.66)	0.001	71.83 (6.52)	68.78 (6.66)	<0.001
Race (%)			<0.001			<0.001
Mexican american	87 (6.9)	138 (11.1)		23 (5.0)	202 (9.9)	
Other hispanic	111 (8.8)	153 (12.3)		41 (8.9)	223 (10.9)	
Non-hispanic white	677 (53.4)	472 (38.1)		216 (46.8)	933 (45.6)	
Non-hispanic black	278 (21.9)	346 27.9)		146 (31.6)	478 (23.4)	
Other	115 (9.1)	131 (10.6)		36 (7.8)	210 (10.3)	
Education, college or above (%)	644 (50.8)	597 (48.2)	0.213	124 (26.8)	1117 (54.6)	<0.001
Marrital Status, widowed/not married (%)	550 (43.4)	520 (42.0)	0.491	258 (55.8)	812 (39.8)	<0.001
Income Ratio, mean (SD)	2.66 (1.63)	2.47 (1.56)	0.004	1.88 (1.34)	2.72 (1.61)	<0.001
Lifestyle						
At least 12 alcohol drink/ years, Yes (%)	837 (66.8)	853 (69.7)	0.133	298 (66.2)	1392 (68.7)	0.34
Number Of Cigarette Smoke in Last 30 Days, Mean (SD)	11.50 (8.10)	10.74 (8.81)	0.417	12.00 (8.42)	10.64 (8.54)	0.171
Smoking Status (%)			0.002			<0.001
Smoker	146 (11.5)	197 (15.9)		107 (23.2)	236 (11.5)	

Non-Smoker	466 (36.8)	469 (37.8)		199 (43.1)	736 (36.0)	
Unknown	656 (51.7)	574 (46.3)		156 (33.8)	1074 (52.5)	
Drug Use (%)			<0.001			<0.001
Yes	85 (6.7)	98 (7.9)		30 (6.5)	153 (7.5)	
No	543 (42.8)	614 (49.5)		126 (27.3)	1031 (50.4)	
Unknown	640 (50.5)	528 (42.6)		306 (66.2)	862 (42.1)	
Sugar Intake (gm/day), Mean (SD)	95.50 (51.51)	94.19 (52.69)	0.542	96.06 (60.67)	94.59 (50.00)	0.598
Carbohydrate Intake (gm/day), Mean (SD)	218.08 (86.09)	224.61 (91.68)	0.075	215.31 (96.21)	222.61 (87.20)	0.125
Kcal Intake (kcal/day), Mean (SD)	1779.71 (658.00)	1847.24 (723.80)	0.018	1731.11 (731.33)	1831.06 (681.64)	0.007
Physical Activity (%)	731 (57.6)	735 (59.3)	0.433	213 (46.1)	1253 (61.2)	<0.001
Anthropometric						
BMI (%)			0.088			0.585
Underweight	55 (4.4)	39 (3.2)		21 (4.6)	73 3.6)	
Normal	279 (22.2)	317 (25.9)		103 (22.8)	493 (24.3)	
Overweight	461 (36.7)	425 (34.7)		168 (37.2)	718 (35.4)	
Obese	461 (36.7)	445 (36.3)		160 (35.4)	746 (36.7)	
Systolic blood pressure, mean (SD)	132.67 (19.71)	133.63 (19.06)	0.217	134.59 (21.31)	132.82 (18.93)	0.081
Diastolic blood pressure, mean (SD)	70.31 (10.35)	70.68 (10.42)	0.406	69.06 (11.25)	70.80 (10.17)	0.002
Waist (%)			0.002			0.507
Low	160 (13.2)	215 (18.2)		63 (14.8)	312 (15.8)	

High	228 (18.8)	231 (19.5)		75 (17.6)	384 (19.5)	
Very high	824 (68.0)	738 (62.3)		288 (67.6)	1274 (64.7)	
Comorbidities						
Diabetes (%)	278 (23.0)	308 (26.1)	0.09	137 (30.8)	449 (23.1)	0.001
Arthritis (%)	621 (49.1)	564 (45.5)	0.077	227 (49.3)	958 (46.9)	0.367
Congestive heart failure (%)	78 (6.2)	76 (6.1)	1	48 (10.5)	106 (5.2)	<0.001
Coronary heart disease (%)	107 (8.5)	88 (7.1)	0.25	54 (11.7)	141 (6.9)	0.001
Angina (%)	66 (5.2)	55 (4.4)	0.426	30 (6.5)	91 (4.5)	0.086
Heart attack (%)	110 (8.7)	96 (7.7)	0.43	61 (13.3)	145 (7.1)	<0.001
Stroke (%)	101 (8.0)	66 (5.3)	0.01	56 (12.1)	111 (5.4)	<0.001
Liver (%)	72 (5.7)	64 (5.2)	0.639	27 (5.8)	109 (5.3)	0.751
Sleep (%)	128 (10.1)	158 (12.8)	0.041	42 (9.1)	244 (11.9)	0.097
Depression (PHQ9 score), mean (SD)	3.28 (4.42)	2.82 (4.03)	0.006	3.53 (4.58)	2.94 (4.15)	0.007
Cognitive function						
Memory, mean (SD)	6.06 (1.81)	5.86 (1.75)	0.004	5.39 (1.80)	6.09 (1.76)	<0.001
Executive function, mean (SD)	3.69 (1.53)	3.59 (1.44)	0.069	3.19 (1.36)	3.74 (1.49)	<0.001
Processing speed, mean (SD)	4.50 (1.71)	4.24 (1.56)	<0.001	3.52 (1.43)	4.56 (1.63)	<0.001
Global cognitive function, mean (SD)	14.40 (4.05)	13.81 (3.73)	<0.001	12.30 (3.60)	14.50 (3.86)	<0.001
Oral health						
Missing Tooth, mean (SD)	14.56 (11.81)	9.23 (7.16)	<0.001	28.00 (0.00)	8.30 (7.40)	<0.001
Dentate (%)	806 (63.6)	1240 (100.0)	<0.001	0 (0.0)	2046 (100.0)	<0.001
DMFT, mean (SD)	20.21 (7.68)	16.45 (6.41)	<0.001	27.52 (3.61)	16.28 (6.29)	<0.001
Periodontitis severity (%)			<0.001			<0.001
None	1268 (100.0)	0 (0.0)		462 (100.0)	806 (39.4)	

Mild-Moderate	0 (0.0)	978 (78.9)		0 (0.0)	978 (47.8)	
Severe	0 (0.0)	262 (21.1)		0 (0.0)	262 (12.8)	
Oral hygiene behaviors						
Dental Visit, over 1 year (%)	555 (43.8)	505 (40.8)	0.129	374 (81.3)	686 (33.5)	<0.001
Use of dental Floss (%)			<0.001			<0.001
No	608 (48.1)	478 (38.6)		427 (93.0)	659 (32.2)	
Not every day	287 (22.7)	298 (24.1)		22 (4.8)	563 (27.5)	
Everyday	369 (29.2)	463 (37.4)		10 (2.2)	822 (40.2)	

N, number of participants; SD, standard deviation; NA, data not applicable due to insufficient data.

Appendix Table a.3. Sample characteristics concerning the cognitive function pre-tests results: NHANES (n = 2,508), 2011-2014.

		Animal fluency pretest		Digit symbol substitution pretest		
	Passed	Failed	P value	Passed	Failed	P value
N	2498	10		2430	78	
Demographic						
Gender, male (%)	1252 (50.1)	5 (50.0)	1	1220 (50.2)	37 (47.4)	0.714
Age, mean (SD)	69.33 (6.74)	71.60 (6.54)	0.289	69.21 (6.71)	73.35 (6.65)	<0.001
Race (%)			0.02			<0.001
Mexican American	225 (9.0)	0 (0.0)		209 (8.6)	16 (20.5)	
Other hispanic	264 (10.6)	0 (0.0)		254 (10.5)	10 (12.8)	
Non-hispanic white	1147 (45.9)	2 (20.0)		1133 (46.6)	16 (20.5)	
Non-hispanic black	617 (24.7)	7 (70.0)		597 (24.6)	27 (34.6)	
Other	245 (9.8)	1 (10.0)		237 (9.8)	9 (11.5)	
Education, college or above (%)	1239 (49.6)	2 (20.0)	0.12	1226 (50.5)	15 (19.2)	<0.001
Marital Status, widowed/not married (%)	1065 (42.7)	5 (50.0)	0.884	1026 (42.3)	44 (57.1)	0.013
Income Ratio, mean (SD)	2.57 (1.60)	2.15 (1.24)	0.433	2.59 (1.60)	1.76 (1.34)	<0.001
Lifestyle						
At least 12 alcohol drink/ years, Yes (%)	1684 (68.2)	6 (75.0)	0.975	1661 (69.0)	29 (40.8)	<0.001
Number Of Cigarette Smoke in Last 30 Days, Mean (SD)	11.06 (8.52)	15.00 (NA)	NA	11.06 (8.53)	11.25 (8.14)	0.94
Smoking Status (%)			0.704			0.096
Smoker	342 (13.7)	1 (10.0)		330 (13.6)	13 (16.7)	

Non-Smoker	930 (37.2)	5 (50.0)		915 (37.7)	20 (25.6)	
Unknown	1226 (49.1)	4 (40.0)		1185 (48.8)	45 (57.7)	
Drug Use (%)			0.25			<0.001
Yes	182 (7.3)	1 (10.0)		182 (7.5)	1 (1.3)	
No	1155 (46.2)	2 (20.0)		1142 (47.0)	15 (19.2)	
Unknown	1161 (46.5)	7 (70.0)		1106 (45.5)	62 (79.5)	
Sugar Intake (gm/day), Mean (SD)	94.94 (52.11)	66.64 (35.38)	0.151	95.05 (51.65)	86.69 (67.83)	0.239
Carbohydrate Intake (gm/day), Mean (SD)	221.44 (88.93)	170.06 (77.19)	0.127	221.70 (88.16)	203.97 (116.05)	0.144
Kcal Intake (kcal/day), Mean (SD)	1814.19 (691.89)	1392.00 (562.80)	0.107	1818.23 (689.57)	1591.24 (754.29)	0.016
Physical Activity (%)	1464 (58.6)	2 (20.0)	0.031	1437 (59.1)	29 (37.2)	<0.001
Anthropometric						
BMI (%)			0.884			0.716
Underweight	94 (3.8)	0 (0.0)		93 (3.9)	1 (1.3)	
Normal	593 (24.0)	3 (30.0)		577 (24.0)	19 (25.0)	
Overweight	883 (35.7)	3 (30.0)		859 (35.7)	27 (35.5)	
Obese	902 (36.5)	4 (40.0)		877 (36.5)	29 (38.2)	
Systolic blood pressure, mean (SD)	133.13 (19.38)	137.58 (23.77)	0.517	132.90 (19.11)	140.96 (25.81)	<0.001
Diastolic blood pressure, mean (SD)	70.50 (10.40)	67.24 (4.08)	0.406	70.51 (10.33)	69.99 (12.11)	0.686
Waist (%)			0.725			0.759
Low	374 (15.7)	1 (11.1)		367 (15.7)	8 (12.5)	
High	458 (19.2)	1 (11.1)		447 (19.2)	12 (18.8)	
Very high	1555 (65.1)	7 (77.8)		1518 (65.1)	44 (68.8)	
Comorbidities						

Diabetes (%)	581 (24.4)	5 (50.0)	0.131	560 (24.2)	26 (35.1)	0.043
Arthritis (%)	1181 (47.4)	4 (40.0)	0.882	1147 (47.3)	38 (48.7)	0.895
Congestive heart failure (%)	152 (6.1)	2 (20.0)	0.244	145 (6.0)	9 (11.7)	0.071
Coronary heart disease (%)	194 (7.8)	1 (10.0)	1	187 (7.7)	8 (10.5)	0.497
Angina (%)	121 (4.9)	0 (0.0)	1	116 (4.8)	5 (6.5)	0.675
Heart attack (%)	203 (8.1)	3 (30.0)	0.053	197 (8.1)	9 (11.5)	0.383
Stroke (%)	165 (6.6)	2 (20.0)	0.29	156 (6.4)	11 (14.3)	0.013
Liver (%)	136 (5.5)	0 (0.0)	0.952	133 (5.5)	3 (3.9)	0.727
Sleep (%)	286 (11.5)	0 (0.0)	0.523	275 (11.3)	11 (14.1)	0.564
Depression (PHQ9 score), mean (SD)	3.04 (4.23)	4.20 (4.73)	0.389	3.03 (4.23)	3.74 (4.34)	0.141
Cognitive function						
Memory, mean (SD)	5.97 (1.78)	3.53 (1.87)	<0.001	6.04 (1.73)	3.62 (1.89)	<0.001
Executive function, mean (SD)	3.64 (1.49)	NA	NA	3.68 (1.48)	2.38 (1.19)	<0.001
Processing speed, mean (SD)	4.38 (1.63)	1.82 (1.75)	<0.001	4.37 (1.64)	NA	NA
Global cognitive function, mean (SD)	14.11 (3.90)	NA	NA	14.11 (3.90)	NA	NA
Oral health						
Missing Teeth, mean (SD)	11.90 (10.15)	19.20 (8.36)	0.023	11.73 (10.10)	18.04 (9.67)	<0.001
Tooth present (at least 1) (%)	2040 (81.7)	6 (60.0)	0.175	1992 (82.0)	54 (69.2)	0.007
DMFT, mean (SD)	18.34 (7.33)	22.10 (5.92)	0.105	18.26 (7.30)	21.21 (7.68)	<0.001
Periodontitis severity (%)			0.997			0.911
None	1263 (50.6)	5 (50.0)		1228 (50.5)	40 (51.3)	
Mild-Moderate	974 (39.0)	4 (40.0)		947 (39.0)	31 (39.7)	
Severe	261 (10.4)	1 (10.0)		255 (10.5)	7 (9.0)	

Oral hygiene behaviors						
Dental Visit = over 1 year (%)	1054 (42.2)	6 (60.0)	0.416	1017 (41.9)	43 (55.1)	0.027
Dental Floss (%)			0.01			<0.001
No	1077 (43.2)	9 (90.0)		1025 (42.3)	61 (78.2)	
Not every day	584 (23.4)	1 (10.0)		578 (23.8)	7 (9.0)	
Everyday	832 (33.4)	0 (0.0)		822 (33.9)	10 (12.8)	

N, number of participants; SD, standard deviation; NA, data not applicable due to insufficient data.

Appendix Table a.4. Oral health's association with cognitive function outcomes: NHANES (n = 1,987), 2011-2014.

		Cognitive outcomes (beta, 95% CI)			
		Memory	Executive function	Processing speed	Global cognitive score
N		1987	1987	1987	1987
	Models	Linear	Linear	Linear	Linear
Missing	Unadjusted	-0.04	-0.05	-0.08	-0.17
Teeth		[-0.05, -0.03]***	[-0.06,-0.04]***	[-0.09,-0.07]***	[-0.19,-0.15]***
	Adjusted For				
	+Demographics	-0.01	-0.02	-0.03	-0.06
		[-0.02,0.00]	[-0.03,-0.01]***	[-0.04,-0.02]***	[-0.08,-0.04]***
	+Lifestyles	-0.01	-0.02	-0.03	-0.05
		[-0.02,0.00]	[-0.03,-0.01]***	[-0.04,-0.02]***	[-0.07,-0.03]***
	+Anthropometric	-0.01	-0.02	-0.03	-0.05
		[-0.02,0.00]	[-0.03,-0.01]***	[-0.03,-0.02]***	[-0.07,-0.03]***
	+Comorbidities	-0.01	-0.01	-0.02	-0.04
		[-0.02,0.00]	[-0.02,-0.01]**	[-0.03,-0.02]***	[-0.07,-0.02]***
	+Oral Hygiene	-0.00	-0.01	-0.02	-0.04
		[-0.01,0.01]	[-0.02,-0.00]**	[-0.03,-0.01]***	[-0.06,-0.02]***
DMFT	Unadjusted	-0.03	-0.04	-0.04	-0.09
		[-0.04,-0.02]***	[-0.03,-0.01]***	[-0.05,-0.02]***	[-0.11,-0.06]***
	Adjusted For				
	+Demographics	-0.01	-0.01	-0.02	-0.04
		[-0.02,-0.00]*	[-0.02,0.00]	[-0.03,-0.01]***	[-0.06,-0.02]***

	+Lifestyles	-0.01 [-0.02,-0.00]*	-0.01 [-0.02,-0.0]**	-0.02 [-0.03,-0.01]***	-0.04 [-0.06,-0.02]***
	+Anthropometric	-0.01 [-0.02,0.00]*	-0.01 [-0.02,0.00]	-0.02 [-0.03,-0.01]***	-0.04 [-0.06,-0.02]***
	+Comorbidities	-0.01 [-0.02,0.00]	-0.01 [-0.02,0.00]	-0.02 [-0.03,-0.01]***	-0.04 [-0.06,-0.02]***
	+Oral Hygiene	-0.01 [-0.02,0.00]	-0.01 [-0.02,0.00]	-0.01 [-0.02,-0.01]**	-0.03 [-0.06,-0.01]**
Periodontal Disease (Yes)	Unadjusted	-0.57 [-0.72,-0.41]***	-0.79 [-0.53,-0.26]***	-0.79 [-0.94,-0.65]***	-1.75 [-2.10,-1.41]***
	Adjusted For				
	+Demographics	-0.18 [-0.32,-0.03]*	-0.11 [-0.24,0.02]	-0.21 [-0.33,-0.09]***	-0.49 [-0.79,-0.19]***
	+Lifestyles	-0.16 [-0.31,-0.01]*	-0.10 [-0.23,0.03]	-0.18 [-0.30,-0.06]**	-0.43 [-0.73,-0.14]**
	+Anthropometric	-0.01 [-0.16,0.14]*	-0.09 [-0.22,0.04]	-0.17 [-0.29,-0.05]**	-0.41 [-0.71,-0.12]**
	+Comorbidities	-0.16 [-0.32,-0.01]*	-0.10 [-0.23,0.03]	-0.19 [-0.31,-0.07]**	-0.46 [-0.75,-0.16]**
	+Oral Hygiene	-0.15 [-0.30,0.00]*	-0.08 [-0.22,0.05]	-0.16 [-0.28,-0.04]**	-0.39 [-0.69,-0.10]**

N, number of participants. *Note:* ***p<.001, **p<.01, *p<.05

Appendix Table a.5. Cognitive function's association with oral health outcomes. NHANES (n = 1,987), 2011-2014.

		Periodontal disease	DMFT	Number of Missing teeth
		OR 95% CI	Beta, 95%CI	Beta, 95%CI
	Models	Binomial logistic regression	Linear	Negative binomial
N		1987	1987	1987
Memory	Unadjusted	0.82 [0.77,0.86]***	-0.42 [-0.58,-0.25]***	-0.10 [-0.12,-0.07]***
	Adjusted For			
	+Demographics	0.93 [0.87,0.99]*	-0.19 [-0.37,-0.01]*	-0.03 [-0.05,0.00]*
	+Lifestyles	0.94 [0.88,1.00]*	-0.18 [-0.36,-0.01]*	-0.02 [-0.05,0.01]
	+Anthropometric	0.94 [0.88,1.00]	-0.17 [-0.35,0.00]*	-0.02 [-0.05,0.01]
	+Comorbidities	0.93 [0.87,0.99]*	-0.17 [-0.35,0.01]	-0.02 [-0.05,0.01]
	+Oral Health	0.93 [0.87,1.00]*	-0.15 [-0.33,0.02]	-0.01 [-0.04,0.02]
Executive Function	Unadjusted	0.94 [0.89,1.00]*	-0.35 [-0.54,-0.16]***	-0.15 [-0.18,-0.12]***
	Adjusted For			
	+Demographics	0.94 [0.88,1.01]	-0.19 [-0.40,0.01]	-0.08 [-0.11,-0.04]***
	+Lifestyles	0.88 [0.88,1.02]	-0.21 [-0.41,-0.00]*	-0.07 [-0.10,-0.04]***
	+Anthropometric	0.95 [0.88,1.03]	-0.20 [-0.40,0.00]	-0.07 [-0.10,-0.04]***
	+Comorbidities	0.94 [0.87,1.02]	-0.18 [-0.39,0.02]	-0.07 [-0.10,-0.04]***
	+Oral Health	0.95 [0.88,1.03]	-0.17 [-0.38,0.03]	-0.06 [-0.09,-0.03]***
Processing Speed	Unadjusted	0.91 [0.86,0.95]***	-0.54 [-0.71,-0.36]***	-0.22 [-0.25,-0.19]***
	Adjusted For			
	+Demographics	0.87 [0.81,0.94]***	-0.47 [-0.69,-0.26]***	-0.12 [-0.15,-0.08]***

	+Lifestyles	0.88 [0.81,0.96]**	-0.47 [-0.69,-0.25]***	-0.10 [-0.14,-0.07]***
	+Anthropometric	0.89 [0.82,0.97]**	-0.45 [-0.67,-0.23]***	-0.10 [-0.13,-0.07]***
	+Comorbidities	0.87 [0.80,0.94]***	-0.43 [-0.65,-0.21]***	-0.09 [-0.13,-0.06]***
	+Oral Health	0.88 [0.81,0.96]**	-0.36 [-0.58,-0.14]**	-0.07 [-0.11,-0.04]***
Global Cognitive Function	Unadjusted	0.88 [0.86,0.91]***	-0.23 [-0.30,-0.16]***	-0.08 [-0.09,-0.07]***
	Adjusted For			
	+Demographics	0.95 [0.92,0.98]**	-0.16 [-0.25,-0.07]***	-0.04 [-0.05,-0.03]***
	+Lifestyles	0.95 [0.92,0.99]**	-0.17 [-0.26,-0.08]***	-0.04 [-0.05,-0.02]***
	+Anthropometric	0.96 [0.92,0.99]**	-0.16 [-0.25,-0.07]***	-0.03 [-0.05,-0.02]***
	+Comorbidities	0.95 [0.92,0.98]**	-0.15 [-0.24,-0.06]***	-0.03 [-0.05,-0.02]***
	+Oral Health	0.95 [0.92,0.99]**	-0.13 [-0.23,-0.04]**	-0.03 [-0.04,-0.01]***

N, number of participants

Note: ***p<.001, **p<.01, *p<.05

Appendix Table a.6. Sensitivity analysis on non-Imputed data: Cognitive function's association with oral health outcomes. NHANES (n = 1,987), 2011-2014.

		Periodontal disease OR, 95% CI binomial logistic	DMFT	Number of Missing teeth
	models	regression	linear	negative binomial
Sample Size N		1987	1987	1987
memory	unadjusted	0.82 (0.76,0.87)***	-0.42 (-0.58,-0.26)***	-0.10 (-0.12,-0.07)***
	adjusted for			
	+demographics	0.93 (0.87,1.00)*	-0.22 (-0.40,-0.04)*	-0.03 (-0.06,-0.00)*
	+lifestyles	0.94 (0.87,1.01)	-0.18 (-0.37,0.01)	-0.02 (-0.05,0.01)
	+anthropometric	0.95 (0.88,1.03)	-0.22 (-0.42,-0.02)*	-0.02 (-0.05,0.01)
	+comorbidities	0.95 (0.87,1.03)	-0.21 (-0.42,0.00)	-0.02 (-0.05,0.02)
	+oral health	0.95 (0.88,1.03)	-0.19 (-0.40,0.01)	-0.01 (-0.04,0.02)
executive function	unadjusted	0.83 (0.77,0.90)***	-0.35 (-0.53,-0.16)***	-0.15 (-0.18,-0.12)***
	adjusted for			
	+demographics	0.96 (0.89,1.03)	-0.22 (-0.43,-0.01)*	-0.08 (-0.11,-0.05)***
	+lifestyles	0.96 (0.88,1.04)	-0.20 (-0.41,0.01)	-0.07 (-0.10,-0.04)***
	+anthropometric	0.99 (0.91,1.08)	-0.21 (-0.44,0.02)	-0.07 (-0.10,-0.03)***
	+comorbidities	0.98 (0.90,1.07)	-0.15 (-0.38,0.09)	-0.06 (-0.10,-0.03)***
	+oral health	0.99 (0.90,1.08)	-0.14 (-0.38,0.09)	-0.06 (-0.09,-0.02)**
speed	unadjusted	0.73 (0.67,0.79)***	-0.54 (-0.70,-0.37)***	-0.22 (-0.24,-0.19)***

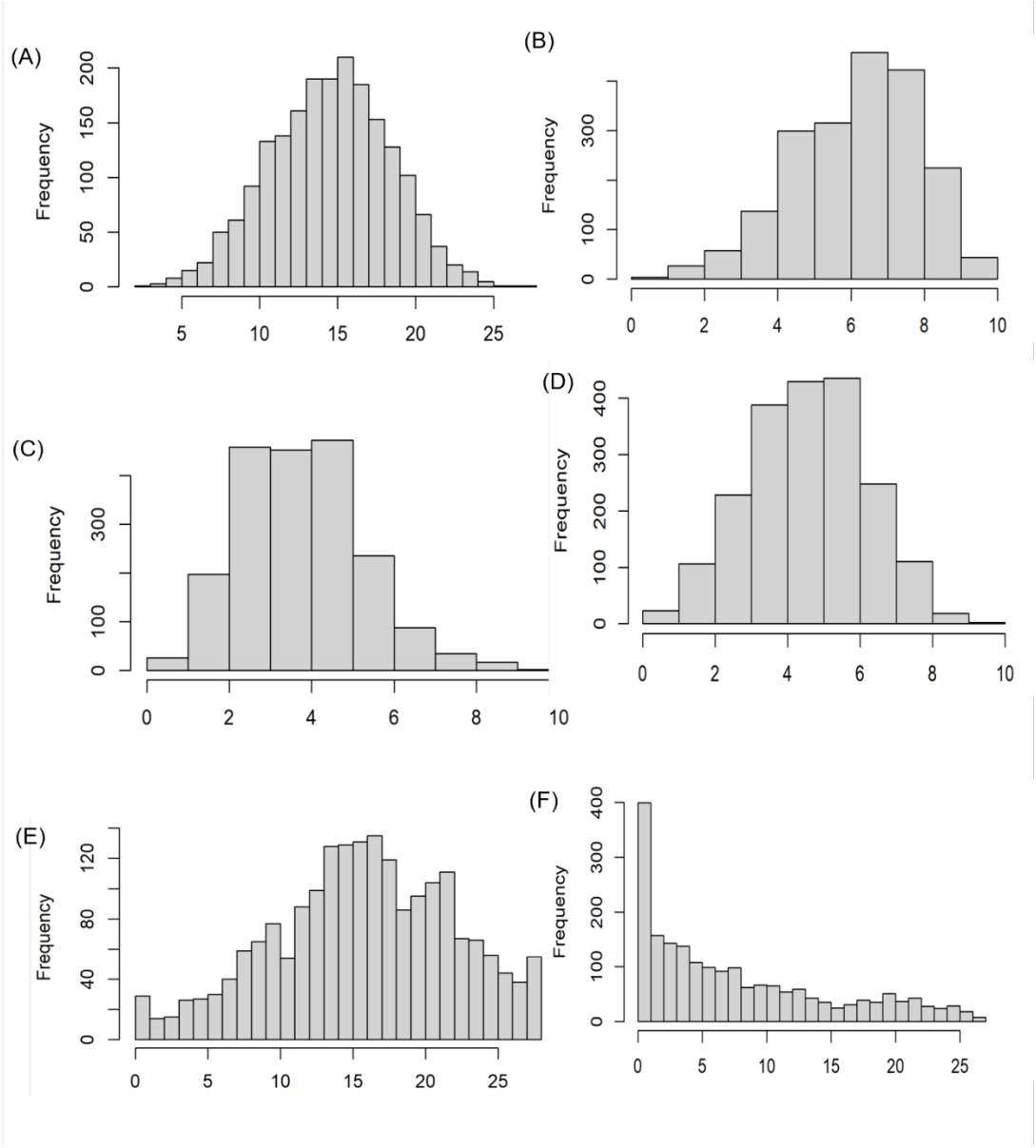
		adjusted for			
		+demographics	0.89 (0.81,0.97)**	-0.44 (-0.66,-0.22)***	-0.11 (-0.14,-0.08)***
		+lifestyles	0.91 (0.82,0.99)*	-0.41 (-0.64,-0.18)***	-0.10 (-0.13,-0.06)***
		+anthropometric	0.91 (0.82,1.00)	-0.34 (-0.59,-0.09)**	-0.09 (-0.12,-0.05)***
		+comorbidities	0.92 (0.83,1.02)	-0.27 (-0.53,-0.01)*	-0.07 (-0.11,-0.03)***
		+oral health	0.93 (0.83,1.03)	-0.21 (-0.47,0.05)	-0.05 (-0.09,-0.01)**
Global function	cognitive	unadjusted	0.88 (0.86,0.91)***	-0.23 (-0.30,-0.16)***	-0.08 (-0.09,-0.07)***
		adjusted for			
		+demographics	0.96 (0.92,0.99)**	-0.17 (-0.26,-0.08)***	-0.04 (-0.05,-0.03)***
		+lifestyles	0.96 (0.93,1.00)*	-0.16 (-0.25,-0.06)**	-0.03 (-0.05,-0.02)***
		+anthropometric	0.97 (0.93,1.01)	-0.16 (-0.26,-0.06)**	-0.03 (-0.05,-0.02)***
		+comorbidities	0.97 (0.93,1.01)	-0.13 (-0.24,-0.02)*	-0.03 (-0.04,-0.01)***
		+oral health	0.97 (0.93,1.01)	-0.12 (-0.22,-0.01)*	-0.02 (-0.04,-0.01)**

Appendix Table a.7. Sensitivity analysis on non-Imputed data: Oral health's association with cognitive function outcomes. NHANES (n = 1,987), 2011-2014.

		cognitive outcomes (beta, 95% CI)			
		memory	executive function	processing speed	global cognitive score
sample size					
N		1987	1987	1987	1987
	models				
NmissingTeeth	unadjusted	-0.04 (-0.05,-0.03)***	-0.05 (-0.06,-0.04)***	-0.08 (-0.09,-0.07)***	-0.17 (-0.19,-0.15)***
	adjusted for				
	+demographics	-0.01 (-0.02,-0.00)*	-0.02 (-0.03,-0.01)***	-0.03 (-0.04,-0.02)***	-0.06 (-0.08,-0.04)***
	+lifestyles	-0.01 (-0.02,0.00)	-0.02 (-0.03,-0.01)**	-0.02 (-0.03,-0.02)***	-0.05 (-0.07,-0.03)***
	+anthropometric	-0.01 (-0.02,0.00)	-0.02 (-0.03,-0.01)**	-0.02 (-0.03,-0.01)***	-0.05 (-0.07,-0.02)***
	+comorbidities	-0.01 (-0.02,0.01)	-0.01 (-0.02,-0.00)*	-0.02 (-0.03,-0.01)***	-0.04 (-0.06,-0.01)**
	+oral hygeine	-0.00 (-0.02,0.01)	-0.01 (-0.02,-0.00)*	-0.01 (-0.02,-0.00)**	-0.03 (-0.05,-0.01)*
DMFT	unadjusted	-0.03 (-0.04,-0.02)***	-0.02 (-0.03,-0.01)***	-0.04 (-0.05,-0.02)***	-0.09 (-0.11,-0.06)***
	adjusted for				
	+demographics	-0.01 (-0.03,-0.00)*	-0.01 (-0.02,-0.00)*	-0.02 (-0.03,-0.01)***	-0.04 (-0.07,-0.02)***
	+lifestyles	-0.01 (-0.02,0.00)	-0.01 (-0.02,0.00)	-0.02 (-0.03,-0.01)***	-0.04 (-0.06,-0.01)**
	+anthropometric	-0.01 (-0.03,-0.00)*	-0.01 (-0.02,0.00)	-0.01 (-0.02,-0.00)**	-0.04 (-0.06,-0.01)**

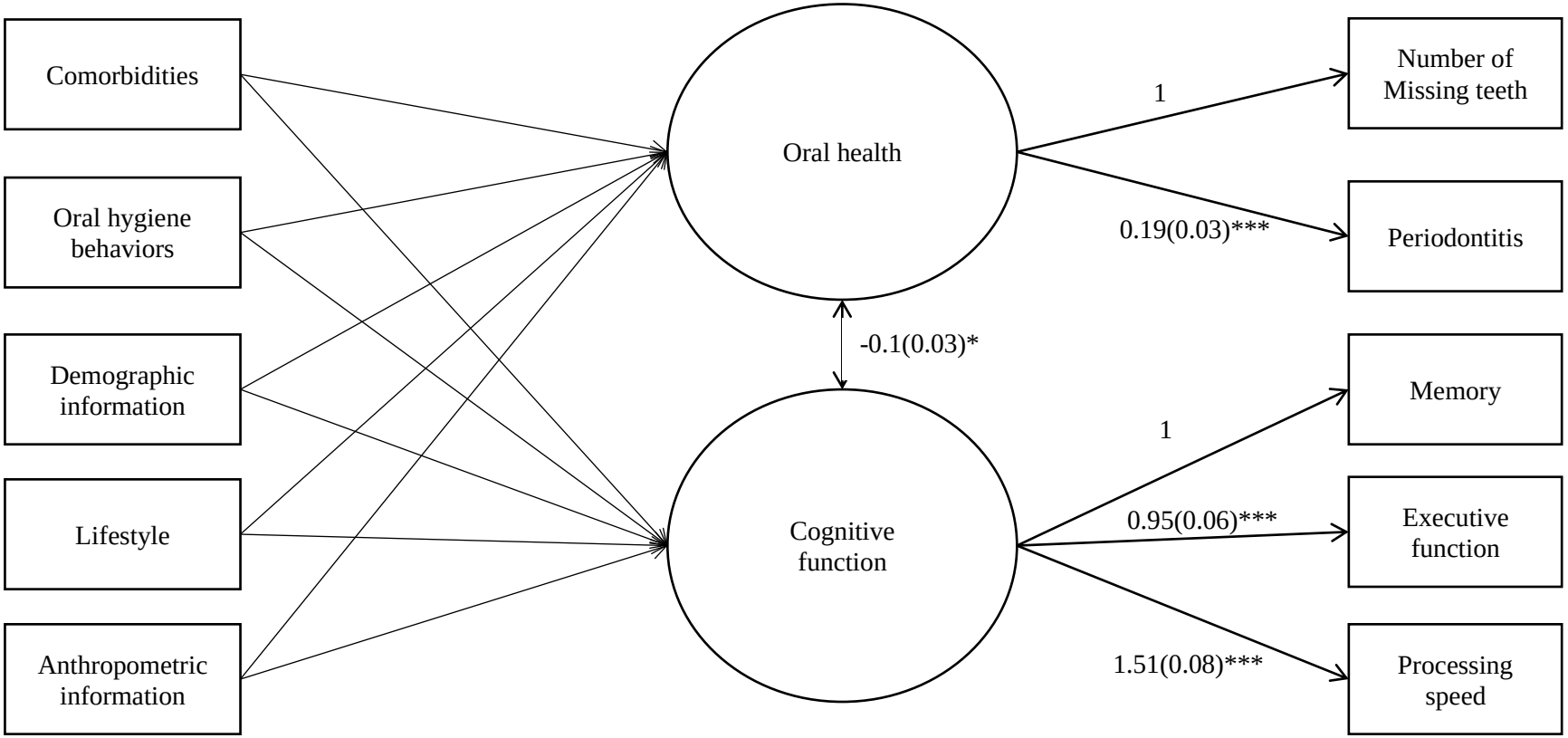
	+comorbidities	-0.01 (-0.03,0.00)	-0.01 (-0.02,0.00)	-0.01 (-0.02,-0.00)*	-0.03 (-0.06,-0.01)*
	+oral hygeine	-0.01 (-0.02,0.00)	-0.01 (-0.02,0.00)	-0.01 (-0.02,0.00)	-0.03 (-0.05,-0.00)*
Periodontal disease (have Periodontal disease)	unadjusted	-0.57 (-0.72,-0.41)***	-0.40 (-0.53,-0.26)***	-0.79 (-0.93,-0.65)***	-1.75 (-2.09,-1.42)***
	adjusted for				
	+demographics	-0.16 (-0.32,-0.01)*	-0.08 (-0.21,0.06)	-0.17 (-0.29,-0.05)**	-0.41 (-0.71,-0.11)**
	+lifestyles	-0.14 (-0.30,0.01)	-0.07 (-0.21,0.07)	-0.14 (-0.26,-0.01)*	-0.35 (-0.65,-0.04)*
	+anthropometric	-0.11 (-0.27,0.05)	-0.01 (-0.15,0.13)	-0.12 (-0.25,0.00)	-0.25 (-0.56,0.07)
	+comorbidities	-0.11 (-0.28,0.06)	-0.04 (-0.19,0.11)	-0.11 (-0.24,0.02)	-0.26 (-0.58,0.07)
	+oral hygeine	-0.11 (-0.27,0.06)	-0.02 (-0.17,0.13)	-0.09 (-0.23,0.04)	-0.22 (-0.55,0.10)

Appendix Figure a.1. Histogram of all cognitive and oral health outcomes



(A) global cognitive function, (B) memory, (C) executive function, (D) processing speed, (E) the number of DMFT and (F) missing tooth. X axis display the frequency of each value in the sample (n=1,987), y axis display the scores/value in the cognitive function and oral health outcomes.

Appendix Figure a.2. Structure equation model result



The cycle represented the latent variable; the rectangle represented observed values. The coefficient estimate and standardised error were marked next to the solid line, which indicated the association. Covariates were summarised and represented by its category in this diagram: demographic variables (age, sex, ethnicity, education qualification, marital status, and ratio of family income to poverty), anthropometric measures (BMI, waist measurement), lifestyle factors (smoking status, alcohol intake, substance misuse, physical activity, intake of sugar, carbohydrate, and energy); comorbidities (cardiovascular diseases, diabetes, liver disease, arthritis, depression and sleep disorder); dental hygiene behaviours (time since last dental visit, use dental floss). The results showed that the model is well fitted: (CFI=0.99, RMSEA =0.05, SRMR=0.02).

Note. ***p<.001, **p<.01, *p<.05

Appendix b. Chapter 3.1 supplementary file

Appendix Table b.1. The search strategies in the GWAS catalog

Search terms & search strategy	Search results
Periodontal disease	3
periodontitis	13
periodon	0

Appendix Table b.2. The search strategies in PubMed

category	Search strategy	results
1. periodontitis	((periodontal disease[Title/Abstract]) OR (periodontitis[Title/Abstract])) OR (periodon*[Title/Abstract])	
2. GWAS	((((((GWA[Title/Abstract]) OR (genome wide[Title/Abstract])) OR (GWAS[Title/Abstract])) OR (genome wide association[Title/Abstract])) OR (whole genome association[Title/Abstract])) OR (WGA[Title/Abstract])) OR (WGAS[Title/Abstract])	
1 and 2	(((periodontal disease[Title/Abstract]) OR (periodontitis[Title/Abstract])) OR (periodon*[Title/Abstract])) AND (((((((GWA[Title/Abstract]) OR (genome wide[Title/Abstract])) OR (GWAS[Title/Abstract])) OR (genome wide association[Title/Abstract])) OR (whole genome association[Title/Abstract])) OR (WGA[Title/Abstract])) OR (WGAS[Title/Abstract]))	202

Appendix Table b.3. The search strategies in ScienceDirect

Search terms/search strategies	search field	No. of Search Results
WGA periodontal disease	Title, abstract or author-specified keywords	1
WGA periodontitis		0
whole genome association periodontal disease		1
whole genome association periodontitis		0
genome wide association periodontal disease		5
genome wide association periodontitis		10
GWAS periodontitis		6
GWAS periodontal disease		4
GWA periodontal disease		4
GWA periodontitis		6
WGAS periodontitis		0
WGAS periodontal disease		1

Appendix Table b.4. The search strategies in EMBASE, GLOBAL HEALTH, MEDLINE (via OVID)

	search terms	results
1	genome-wide association.ab. or genome-wide association.ti. or genome-wide association.kf.	102642
2	periodontitis.ab. or periodontitis.ti. or periodontitis.kf.	75165
3	periodontal disease.ab. or periodontal disease.ti. or periodontal disease.kf.	44806
4	periodon*.ab. or periodon*.ti. or periodon*.kf.	176245
5	GWA.ab. or GWA.ti. or GWA.kf.	3544
6	genome wide.ab. or genome wide.ti. or genome wide.kf.	264521
7	whole genome association.ab. or whole genome association.ti. or whole genome association.kf.	1272
8	WGA.ab. or WGA.ti. or WGA.kf.	10212
9	WGAS.ab. or WGAS.ti. or WGAS.kf.	132
10	GWAS.ab. or GWAS.ti. or GWAS.kf.	60911
11	2 or 3 or 4	176245
12	1 or 5 or 6 or 7 or 8 or 9 or 10	287085
13	11 and 12	491

Appendix Table b.5. PRISMA 2020 abstract checklist

Section and Topic	Item #	Checklist item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesise results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			

Section and Topic	Item #	Checklist item	Reported (Yes/No)
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	No (mentioned in acknowledgement section not abstract)
Registration	12	Provide the register name and registration number.	No (Mentioned in method section not abstract)

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Appendix Table b.6. PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	See abstract check list in supplements
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 4
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 4-5
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 5
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 5
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 5 & supplements
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 5 -6

Section and Topic	Item #	Checklist item	Location where item is reported
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 5-6
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 6-7
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 6-7
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Page 7
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 7
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Page 6- 7
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 6-7
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 6-7

Section and Topic	Item #	Checklist item	Location where item is reported
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page 6-7
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Page 6-7
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Not performed as irrelevant.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Page 7
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page 7
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 7 (figure 1)
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	

Section and Topic	Item #	Checklist item	Location where item is reported
Study characteristics	17	Cite each included study and present its characteristics.	Page 7 -9 (table 2-4)
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 8 (table 1)
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Page 7-9 (table 2-4)
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page 8
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Page 8-9
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Table 2.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Due to missing data meta-analysis was not performed

Section and Topic	Item #	Checklist item	Location where item is reported
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 9-10
	23b	Discuss any limitations of the evidence included in the review.	Page 10
	23c	Discuss any limitations of the review processes used.	Page 11
	23d	Discuss implications of the results for practice, policy, and future research.	Page 11
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page 5
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 5
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Page 5
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 12
Competing interests	26	Declare any competing interests of review authors.	Page 12
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Page 12

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Appendix c. Chapter 3.2 supplementary file

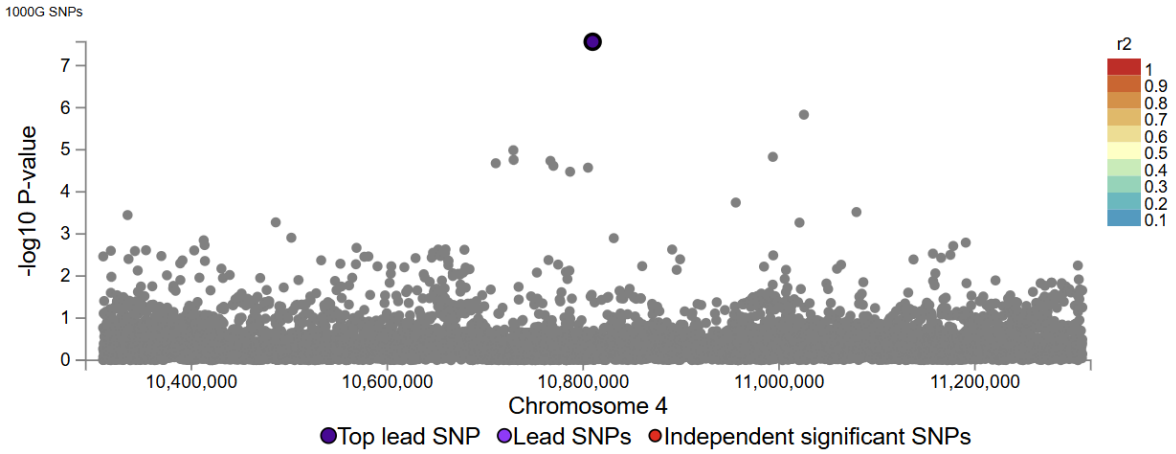
Appendix c. Supplementary information

Ethical approval and UKB approval reference

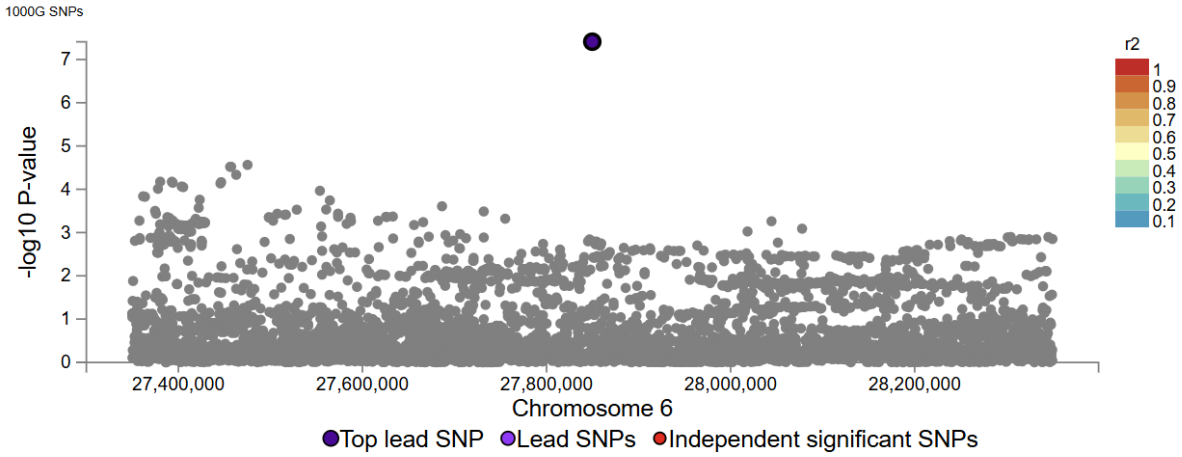
UK Biobank study is ethnically approved by the North West Multicentre Research Ethnic Committee, UK (REF: 16/NW/0274). Participants were recruited prior to the data collect, and participants are free to withdraw their data at any time point of the study. This current study was provided access to the relevant data (UKB approval reference: 54633).

Appendix Figure c.1. The regional plot of two significant SNVs

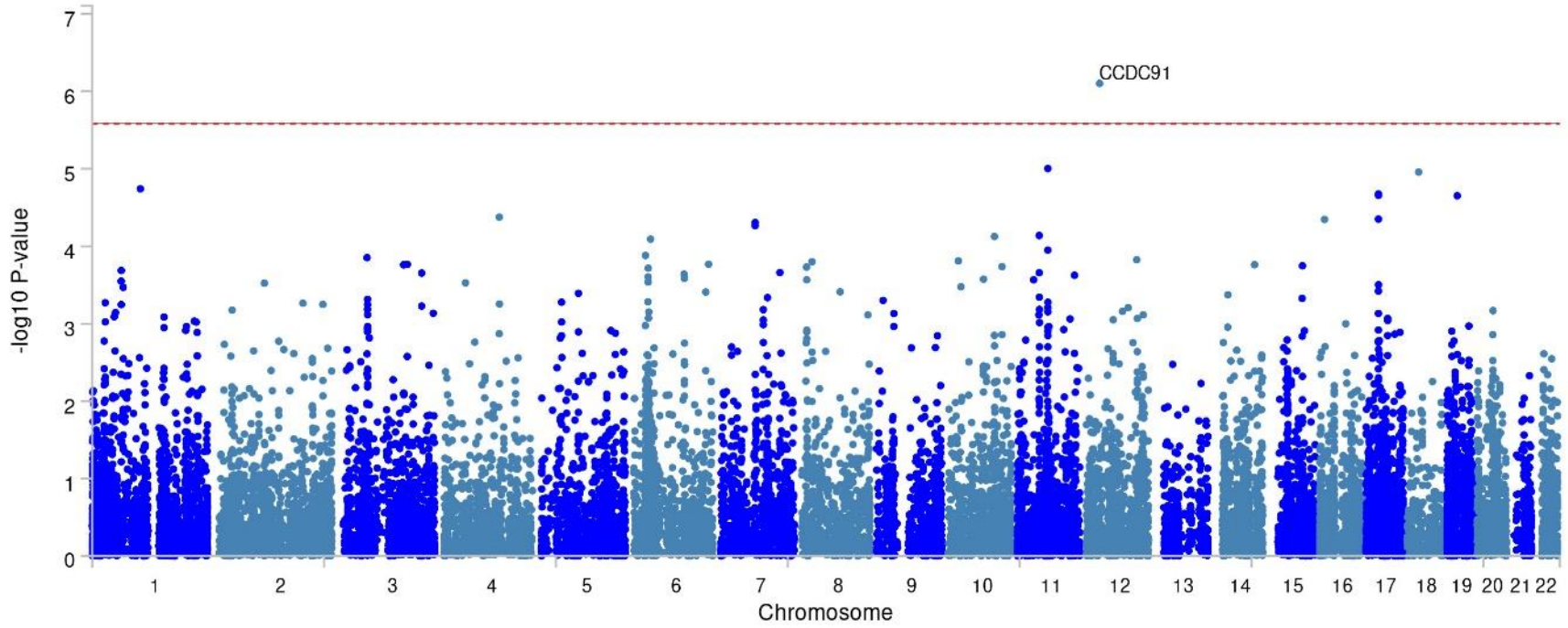
(a) rs149922301



(b) rs368467810

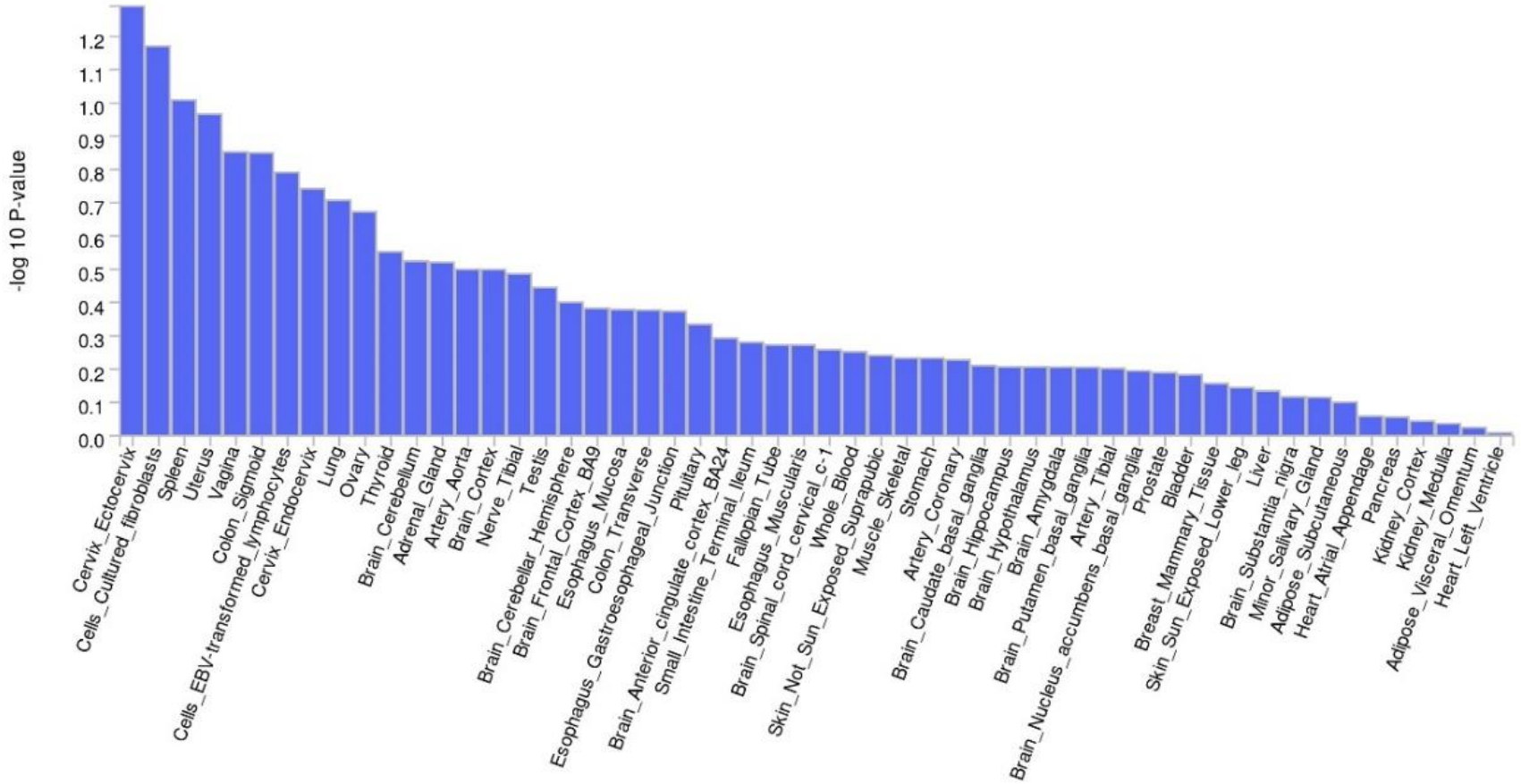


Appendix Figure c.2. The gene-based Manhattan plot



The name of the top significant gene was marked in the Manhattan plot. The red line represents the significance threshold ($p <= 2.608x10^{-6}$

Appendix Figure c.3. MAGMA tissue expression analysis.



Note: The blue bars in the figure indicates no significant expression of corresponding tissues. Tissue names were noted at x-axis

Appendix Table c.1. Previous reported results and current GWAS results of previously reported significant SNPs (n=52).

	Previously reported significant SNPs				association results in current study		
Study	SNP [EA]	Nearest Gene	OR (95%CI)	P	EA	OR (95% CI)	p
Bevilacqua et al. (2017)	rs242016 [A]	*CRACR2A*	3.7 (2.32, 6.29)	1.50E-08	A	1.00 (0.99, 1.02)	0.97
	rs242014 [T]	*CRACR2A*	3.7 (2.32, 6.29)	1.60E-08	T	1.00 (0.98, 1.02)	0.99
	rs10491972 [G]	*CRACR2A*	3.7 (2.32, 6.29)	1.70E-08	G	1.00 (0.98, 1.02)	1
	rs242002 [T]	*CRACR2A*	3.6 (2.28, 6.13)	2.73E-08	T	1.00 (0.98, 1.01)	0.95
deCoo et al. (2021)	rs35709256 [A]	*FAT3*	2.07 (1.55, 2.77)	9.48E-07	A	1.00 (0.98, 1.02)	0.87
	rs4807188 [A]	*CSNK1G2*	2.8 (1.83, 4.27)	1.81E-06	A	0.95 (0.90, 1.00)	0.047
	rs2074872 [A]	*MYH13, LOC107985004*	2.12 (1.55, 2.91)	2.84E-06	A	1.00 (0.99, 1.01)	0.92
	rs116611488 [T]	*none*	4.22 (2.31, 7.72)	2.90E-06	T	0.98 (0.93, 1.04)	0.58
	rs4854545 [G]	*ANTXR1*	4.34 (2.34, 8.02)	2.97E-06	T	0.99 (0.95, 1.03)	0.47
	rs78672540 [C]	***none***	2.32 (1.62, 3.31)	3.78E-06	C	0.99 (0.96, 1.01)	0.16
	rs13439823 [A]	*ANGPT1*	2.01 (1.49, 2.7)	4.22E-06	A	1.00 (0.99, 1.02)	0.5
	rs11993287 [A]	*none*	0.58 (0.46, 0.73)	4.31E-06	A	1.00 (0.99, 1.01)	0.98
Divaris et al. (2013)	rs2521634 [G]	*LOC107986777*	1.49 (1.28, 1.73)	3.50E-07	A	1.00 (0.99, 1.01)	0.95
	rs7762544 [G]	*none*	1.4 (1.24, 1.59)	7.50E-08	A	0.99 (0.98, 1.01)	0.26
	rs3826782 [A]	*ADGRE1*	2.01 (1.52, 2.65)	8.20E-07	A	1.01 (0.99, 1.04)	0.27
Hong et al. (2015)	rs4242220 [MA: C]	*TENM2*	0.53 (0.41, 0.69)	2.84E-06	G	1.00 (0.99, 1.01)	0.95
	rs12969041 [MA: A]	*none*	2.86 (1.92, 4.27)	2.79E-07	T	1.00 (0.98, 1.01)	0.56
	rs2027756 [MA: A]	*none*	2.86 (1.92, 4.27)	2.79E-07	T	1.00 (0.98, 1.01)	0.55
	rs2978951 [A]	*none*	1.25 (1.16, 1.35)	2.06E-08	G	1.01 (1.00, 1.02)	0.21

Munz et al. (2017)	rs2738058 [T]	*none*	1.28 (1.18, 1.38)	6.78E-10	C	1.02 (1.00, 1.03)	0.01
	rs4284742 [G]	*SIGLEC5*	1.34 (1.21, 1.48)	1.34E-08	G	0.99 (0.97, 1.00)	0.08
	rs4970469 [G]	*none*	1.52 (1.29, 1.81)	1.20E-06	A	1.01 (0.99, 1.03)	0.42
	rs1122900 [A]	*none*	1.27 (1.16, 1.4)	8.00E-07	C	0.99 (0.98, 1.01)	0.3
	rs2070901 [T]	*FCER1G*	1.29 (1.16, 1.44)	4.36E-06	T	1.01 (1.00, 1.02)	0.1
Munz et al. (2019)	rs729876 [T]	*LOC107984137*	1.23 (1.15, 1.32)	1.21E-08	C	1.01 (0.99, 1.02)	0.44
	rs11084095 [A]	*SIGLEC5 - AC018755.18*	1.17 (1.11, 1.24)	5.09E-08	A	1.02 (1.01, 1.03)	4.58E-04
	rs9982623 [C]	*MCM3AP*	1.23 (1.13, 1.33)	8.65E-07	T	1.02 (1.00, 1.04)	0.04
Sanders et al. (2017)	rs149133391 [T]	*TSNAX-DISC1*	beta: -0.139 (-0.09, -0.19)	7.90E-09	C	1.08 (0.92, 1.27)	0.33
	rs75715012 [G]	*none*	beta: 0.045 (0.03, 0.06)	1.10E-07	A	0.99 (0.97, 1.01)	0.38
	rs186066047 [G]	*none*	beta: 0.23 (0.14, 0.31)	1.70E-07	A	0.34 (0.08, 1.47)	0.15
	rs10456847 [C]	*none*	beta: -0.03(-0.04, -0.02)	2.60E-07	G	0.99 (0.98, 1.00)	0.06
	rs79308117 [A]	*ASH1L*	beta: 0.18 (0.11, 0.25)	2.80E-07	C	0.55 (0.23, 1.33)	0.19
Schaefer et al. (2010)	rs1537415 [G]	*GLT6D1*	1.59 (1.36, 1.86)	5.51E-09	C	1.02 (1.00,1.03)	8.14E-03
Shaffer et al. (2014)	rs733048 [not found]	*LOC105374494*	2.4 ()	1.00E-06	A	1.01 (1.00, 1.02)	0.16
	rs10457525 [not found]	*LOC102723409*	2.33 ()	3.50E-06	T	1.00 (0.98, 1.01)	0.66

	rs7749983 [not found]	*LOC102723409*	2.39 ()	2.40E-06	A	0.99 (0.98, 1.01)	0.35
Shimizu et al. (2015)	rs9446777 [A]	*KCNQ5*	0.86 (0.80, 0.93)	4.83E-06	G	0.99 (0.96, 1.03)	0.62
	rs2392510 [C]	*GPR141*	0.87 (0.82, 0.92)	4.17E-06	T	1.00 (0.98, 1.01)	0.52
Shungin et al. (2019)	rs12461706 [T]	*SIGLEC5*	1.05 ()	3.90E-09	T	1.02 (1.01, 1.03)	5.34E-04
Tegelberg et al. (2021)	rs200392355 [CT]	*LOC102724234*	beta: 0.16 (0.09, 0.22)	1.22E-06	NA		
	rs2409703 [C]	*XKR6*	beta: 0.28 (0.16, 0.39)	1.61E-06	C	1.02 (1.00, 1.04)	0.1
	rs11630851 [T]	*DNM1P35*	beta: 0.30 (0.18, 0.42)	9.39E-07	T	1.00 (0.97, 1.02)	0.8
	rs4444613 [A]	*TASP1*	beta: -0.28 (-0.38, -0.18)	1.35E-07	A	0.99 (0.97, 1.01)	0.26
	rs2003705 [T]	*LOC107985448*	beta: -0.16 (-0.23, -0.10)	1.68E-06	T	1.01 (1.00,1.02)	0.21
Petty et al., (2023)	rs12036106 [T]	*RAP1GAP,USP48*		5.07E-07	T	1.02 (1.00, 1.04)	0.03
	rs13031512 [A]	*TFPI,LINC01090*		2.60E-06	A	0.99 (0.97, 1.00)	0.15
	rs72870126 [G]	*MEPE,SPP1*		4.80E-06	G	0.99 (0.97, 1.02)	0.55
	rs369717575 [TA]	*PALLD*		1.96E-06	NA		
	rs36793 [T]	*LINC01848,TMEM232*		2.07E-06	T	0.98 (0.93, 1.04)	0.5

rs148550758 [C]	*LINC02487,LINC01558*	2.04E-06	C	0.95 (0.89, 1.02)	0.14
rs7835237 [G]	*STC1,ADAM28*	3.62E-06	G	1.01 (0.97, 1.05)	0.57
rs12800372 [C]	*TPCN2,LOC338694*	4.34E-06	C	1.01 (1.00, 1.02)	0.2

EA, effect allele; MA, minor allele; OR, Odds Ratio; CI, Confidence interval. Note. The beta marked in OR column means the results are from linear model and the results are corresponding coefficient beta and confidence interval.

Appendix Table c.2. The top 10 associated gene sets

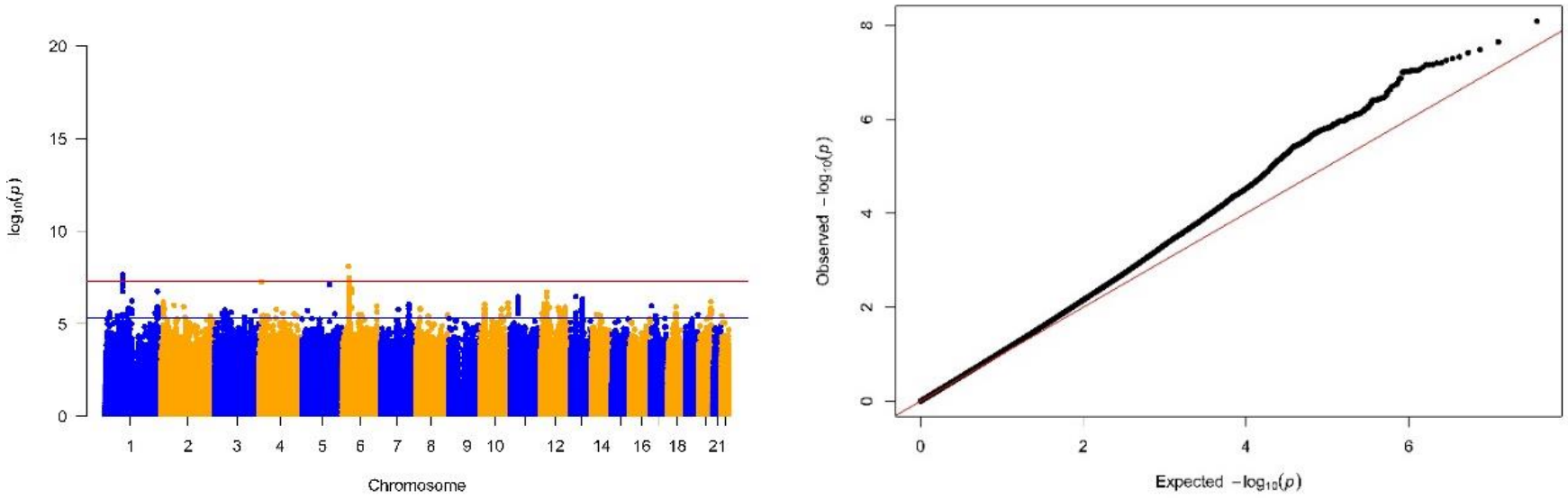
Gene Set	N genes	Beta	Beta STD	SE	P
GOBP_REGULATION_OF_EPITHELIAL_CELL_DIFFERENTIATION	148	0.26018	0.02277	0.075863	0.00030296
HOWLIN_CITED1_TARGETS_1_UP	33	0.54028	0.022395	0.15891	0.00033767
WP_MELANOMA	66	0.37145	0.021756	0.10986	0.00036184
GOMF_CALCIUM_ACTIVATED_CATION_CHANNEL_ACTIVITY	25	0.56123	0.020253	0.16671	0.00038154
DAZARD_RESPONSE_TO_UV_NHEK_DN	294	0.18309	0.022498	0.054564	0.00039681
GOBP_CELLULAR_RESPONSE_TO_UV_B	12	0.83387	0.020855	0.24917	0.00040985
GOCC_DENDRITIC_TREE	575	0.12584	0.021463	0.037756	0.0004305
PHONG_TNF_RESPONSE_VIA_P38_COMPLETE	216	0.19809	0.020907	0.059507	0.00043681
GOBP_CELL_CYCLE_PROCESS	1211	0.0818	0.019898	0.024652	0.00045382
BOYAULT_LIVER_CANCER_SUBCLASS_G1_UP	113	0.26765	0.020487	0.08073	0.00045846

$P_{bon,}$ p value after Bonferroni correction; STD, standard deviation; SE, standard error.

Appendix d. Chapter 4.1 supplementary file

Appendix Figure d.1. Manhattan plot and Q-Q plot of periodontitis group GWAS

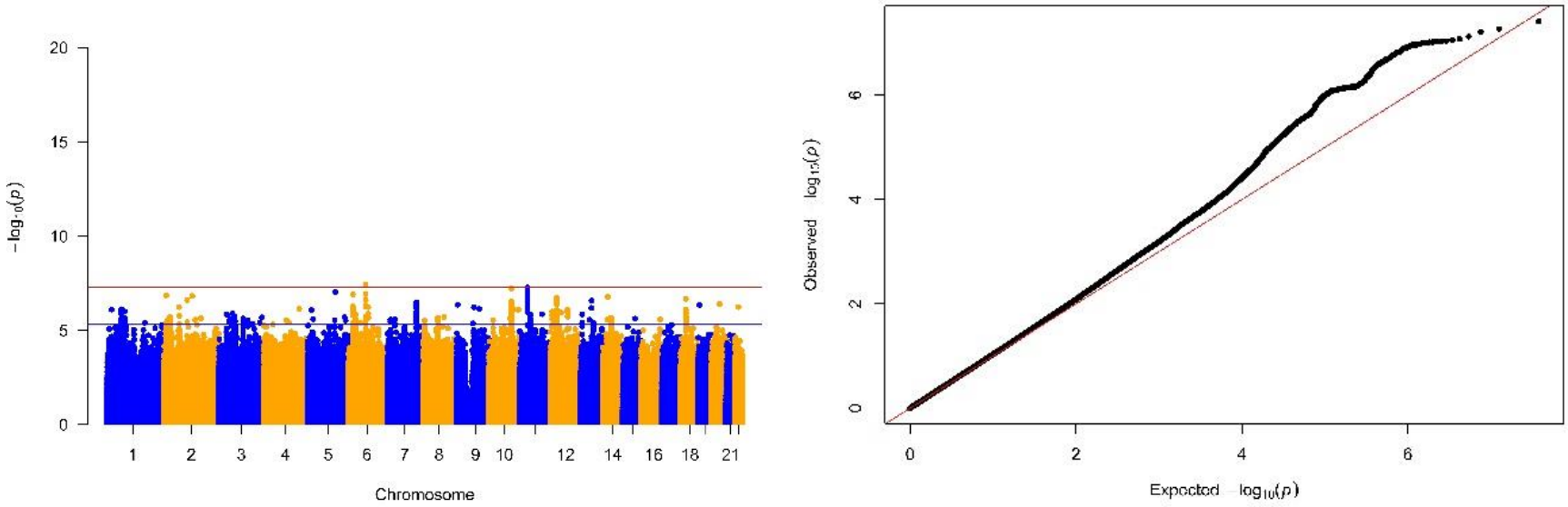
(a) (b)



(a) Manhattan Plot of periodontitis GWAS in periodontitis group. The red line indicated the conventional GWAS significant level (p<5e-8). The blue line indicates the suggestive significant level (p<1e-5). (b) Q-Q plot of periodontitis GWAS in periodontitis group (genomic control inflation factor lambda = 1.07). The red line represents the reference line.

Appendix Figure d.2. Manhattan plot and Q-Q plot of periodontitis train group GWAS

(a) (b)



(a) Manhattan Plot of periodontitis GWAS in periodontitis train group. The red line indicated the conventional GWAS significant level (p<5e-8). The blue line indicates the suggestive significant level (p<5e-6). (b) Q-Q plot of periodontitis GWAS in periodontitis train group (genomic control inflation factor lambda = 1.05). The red line represents the reference line.

Appendix Table d.1. Sample characteristics for periodontitis and periodontitis train group participants.

	Periodontitis Group		**Periodontitis Train Group**	
	Case	**Control**	**Case**	**Control**
N	271619	67591	181079	45061
Age (Mean (SD))	69.95 (7.85)	69.11 (7.75)	69.68 (7.84)	69.08 (7.75)
Sex = Female (%)	144554 (53.2)	40382 (59.7)	99714 (55.1)	26934 (59.8)
Periodontitis = Yes (%)	0 (0.0)	67591 (100.0)	0 (0.0)	45061 (100.0)
Dementia = Yes (%)	263906 (100.0)	66062 (100.0)	175797 (100.0)	44052 (100.0)
Household Income (%)				
Less Than 18,000	48318 (20.6)	13326 (22.5)	31891 (20.4)	8852 (22.5)
18,000 To 30,999	59151 (25.2)	15001 (25.4)	39196 (25.0)	10037 (25.5)
31,000 To 51,999	63129 (26.9)	15952 (27.0)	42434 (27.1)	10567 (26.8)
52,000 To 100,000	50561 (21.5)	12152 (20.6)	33955 (21.7)	8143 (20.7)
Greater Than 100,000	13532 (5.8)	2678 (4.5)	9156 (5.8)	1806 (4.6)
Smoking Status				
Never	151466 (56.0)	34825 (51.7)	101509 (56.2)	23240 (51.7)
Ex-Smoker	91685 (33.9)	25602 (38.0)	60534 (33.5)	17034 (37.9)
Current Smoker	27551 (10.2)	6960 (10.3)	18440 (10.2)	4656 (10.4)
Alcohol Status				
Never	8306 (3.1)	1764 (2.6)	5582 (3.1)	1184 (2.6)
Former Drinker	8929 (3.3)	2541 (3.8)	5958 (3.3)	1671 (3.7)
Current Drinker	254165 (93.6)	63229 (93.6)	169392 (93.6)	42166 (93.7)
BMI (Mean(SD))	27.30 (4.70)	27.82 (5.07)	27.26 (4.72)	27.82 (5.10)

C-Reactive Protein Level (Mean(SD))	2.09 (2.37)	2.29 (2.52)	2.09 (2.38)	2.29 (2.51)

SD, Standard Deviation; BMI, Body Mass Index