OPEN ACCESS

Public Health ARTICLE

Development and characterization of OralAge: An integrated microbiome-clinical phenotype of oral aging

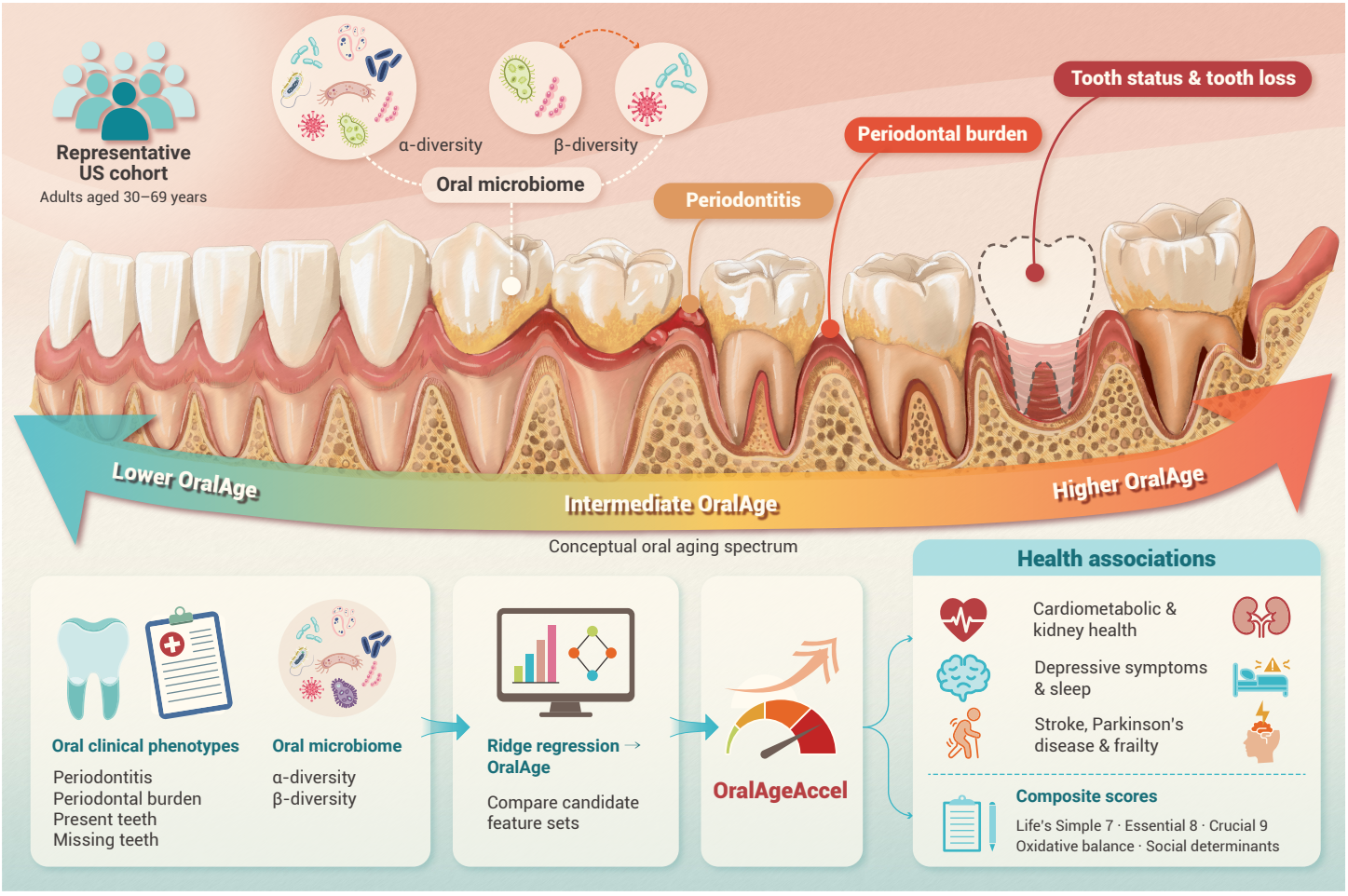
Yanwei You[1,2,*]

*Correspondence: Yanwei You (yyw22@tsinghua.org.cn)

Received: July 14, 2026; Accepted: November 10, 2026; Published Online: November 10, 2026; https://doi.org/10.59717/j.tiis.2026.100019

© 2026 The Author(s). This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Oral health age (OralAge) combines oral microbiome, gum, and tooth data.
- Older-than-expected oral profiles were linked to heart disease, diabetes, chronic kidney disease, and stroke.
- Better cardiovascular health was linked to younger oral profiles.
- Greater social disadvantage was linked to older oral profiles.

ARTICLE Public Health

OPEN ACCESS

Development and characterization of OralAge: An integrated microbiome-clinical phenotype of oral aging

Yanwei You[1,2,*]

[1]Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117549, Singapore

[2]Division of Sports Science & Physical Education, Tsinghua University, Beijing 100084, China

*Correspondence: Yanwei You (yyw22@tsinghua.org.cn)

Received: July 14, 2026; Accepted: November 10, 2026; Published Online: November 10, 2026; https://doi.org/10.59717/j.tiis.2026.100019

© 2026 The Author(s). This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Citation: You Y. (2026). Development and characterization of OralAge: An integrated microbiome-clinical phenotype of oral aging. *The Innovation Insights* **1**:100019.

This study aims to develop an oral health age phenotype, OralAge, and examine whether OralAge acceleration is associated with oral deterioration, chronic disease burden, and composite health profiles. It included 5,323 US adults aged 30–69 years with complete oral microbiome, periodontal, and tooth-status information. Ridge regression was used to estimate OralAge from oral alpha-diversity and beta-diversity measures, periodontal indicators, and tooth-level status variables. OralAge acceleration was defined as the residual of predicted OralAge after adjustment for chronological age. Survey-weighted regression models evaluated associations with chronic diseases, disease-domain burden, disease counts, and composite health scores. OralAge acceleration was higher among participants with periodontitis and was greater among those with more missing teeth. Each 1-SD higher OralAge acceleration was associated with higher odds of cardiovascular disease, diabetes, chronic kidney disease, and stroke, as well as greater cardiometabolic, renal, core chronic, and extended disease burden. Higher Life's Simple 7, Life's Essential 8, and Life's Crucial 9 scores were associated with lower OralAge acceleration, whereas greater social disadvantage was associated with higher acceleration. OralAge acceleration may capture age-independent oral deterioration and health burden. OralAge is an internally evaluated research phenotype for characterizing heterogeneity in oral ageing and oral-systemic health profiles. Independent and longitudinal validation is required before clinical application.

INTRODUCTION

Oral health is an important component of healthy ageing. Oral diseases are highly prevalent worldwide and contribute to pain, tooth loss, impaired chewing, poor nutrition, reduced quality of life, and health-care burden.[1,2] With age, the oral cavity accumulates multiple forms of biological and clinical change, including shifts in microbial community structure, periodontal tissue destruction, tooth loss, dental restoration patterns, and root fragments. These changes rarely occur in isolation. They jointly reflect long-term exposure to microbial, inflammatory, behavioral, and social conditions. A composite oral ageing phenotype may therefore provide information beyond a single oral disease label or a single microbiome measure.

The oral microbiome is one of the most complex microbial ecosystems in the human body. It varies across oral niches and is closely related to local immune, metabolic, and environmental conditions.[3] Oral microbial communities also contribute to oral disease through ecological shifts, polymicrobial interactions, and host inflammatory responses.[4,5] These observations support the value of combining microbial community information with clinical oral indicators when studying oral ageing.

Oral deterioration has systemic health relevance. Periodontitis has been linked to inflammatory comorbidities through mechanisms involving systemic inflammation, microbial translocation, immune activation, and vascular dysfunction.[6] Clinical and epidemiological evidence also supports associations of periodontitis with cardiovascular disease and diabetes.[7,8] Periodontal disease has been associated with chronic kidney disease in meta-analytic evidence.[9] A large UK cohort study further reported associations between periodontal diseases and a broad chronic disease burden across cardiovascular, cardiometabolic, autoimmune, and mental health outcomes.[10] Tooth loss represents another cumulative oral health marker and has been associated with higher risks of atherosclerotic cardiovascular disease and death in meta-analyses of survival data.[11] These findings suggest that oral ageing may reflect systemic vulnerability, especially cardiometabolic, renal, cerebrovascular, and multimorbidity-related burden.

Recent work has begun to examine the oral microbiome as an ageing-related biomarker. Human microbiome profiles from several body sites can predict chronological age, and the oral microbiome has shown age-associated signals in this context.[12] Oral microbial signatures have also been associated with age and frailty in adults, suggesting that oral microbial composition may capture dimensions of biological and functional ageing.[13] This line of evidence indicates that the oral microbiome contains ageing-related information. It also leaves an important question: whether an oral ageing phenotype that integrates microbiome features with periodontal and tooth-status information can better represent accumulated oral deterioration and its systemic correlates.

Biological ageing clocks provide a useful conceptual framework for this question. Modern biological age measures use high-dimensional molecular, clinical, or physiological data to estimate ageing-related variation that chronological age alone cannot capture.[14] A similar strategy can be applied to the oral cavity. Predicted oral age can be estimated from oral microbial and clinical features, and the residual difference between predicted oral age and chronological age can be used to define oral health age acceleration. This residualized index may capture whether an individual's oral profile appears older or younger than expected for their age. Understanding determinants of accelerated oral ageing is also important. Composite cardiovascular, multidimensional health, oxidative balance, and social disadvantage may help identify behavioral and social-environmental profiles linked to oral ageing. Examining these factors alongside disease burden can provide a broader epidemiological characterization of oral health age acceleration.

The overall study design and analytical workflow are shown in Figure 1. In this study, we developed an oral health age phenotype, OralAge, using oral microbiome features, periodontal status, periodontal burden, and tooth-status indicators in a representative US adult cohort. We then derived OralAge acceleration as the residual of OralAge after accounting for chronological age. We examined the clinical characterization of OralAge acceleration using periodontitis and tooth-loss gradients, evaluated its associations with individual chronic diseases and disease-burden measures, and explored its relationships with composite health scores.

METHODS

Study design and data source

This cross-sectional study used data from the National Health and Nutrition Examination Survey (NHANES) 2009–2010 and 2011–2012 cycles. NHANES is a nationally representative survey of the civilian, non-institutionalized US population and uses a complex, multistage probability sampling design that combines household interviews, physical examinations, and laboratory assessments.[15,16] The present analysis integrated demographic information, oral microbiome data, periodontal examination data, tooth-status records, chronic disease variables, and composite health scores using the unique participant identifier. All NHANES protocols were approved by the National Center for Health Statistics Research Ethics Review Board, and all participants provided written informed consent. This study used publicly available de-identified data and did not require additional institutional review board approval.

Study population

The initial NHANES 2009–2012 sample included 19,591 individuals. We

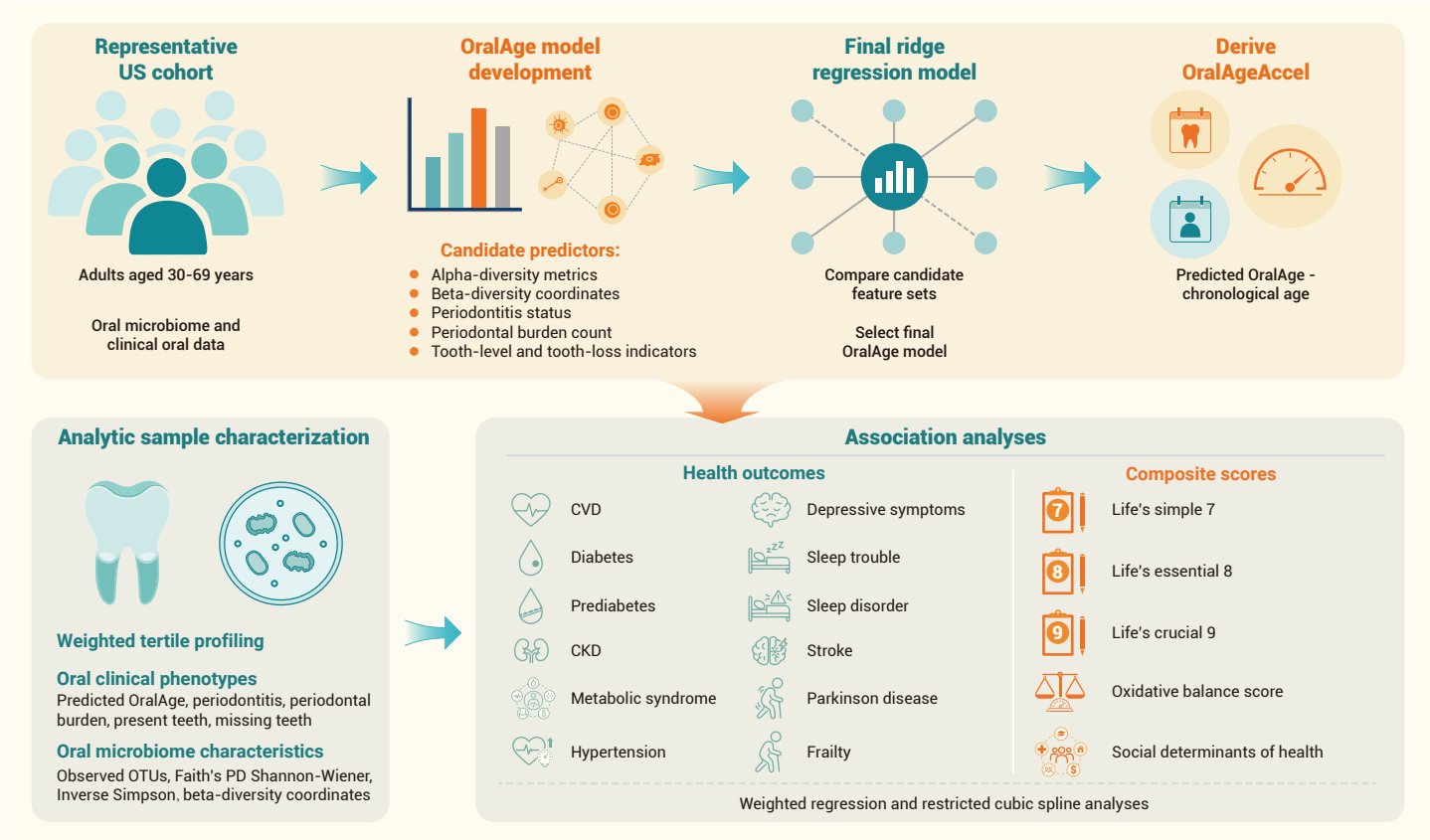


Figure 1. Study design and analytical workflow of the OralAge study Notes: The workflow shows the development of OralAge from oral microbiome, periodontal, and tooth-status features in a representative US adult cohort, derivation of OralAgeAccel, tertile-based sample characterization, and association analyses with chronic diseases and composite health scores. The age-difference notation in the schematic illustrates the concept of accelerated oral ageing. In the analyses, OralAgeAccel was defined as the residual from a linear regression of out-of-fold OralAge on chronological age.

first restricted the sample to participants aged ≥ 20 and < 70 years, leaving 9,501 individuals. Participants with missing oral microbiome alpha-diversity measures were excluded, leaving 8,002 individuals. We further excluded participants without beta-diversity measures, resulting in 7,736 participants with complete microbiome-derived features. For OralAge construction, the analytic sample was restricted to adults aged 30–69 years. The lower bound reflected the availability of full-mouth periodontal examinations among adults aged ≥ 30 years, and the upper bound reflected the concurrent age range of comparable oral microbiome and oral examination data in these cycles. This restriction left 6,073 participants. After excluding participants with missing periodontal diagnosis, periodontal burden count, or tooth-status information, 5,323 participants were included in the final OralAge analysis.

Oral microbiome features

Oral microbiome data were obtained from oral rinse samples collected during the NHANES 2009–2010 and 2011–2012 cycles. These data provide population-scale characterization of the adult oral microbiome in the United States and include amplicon sequence variant-based diversity metrics derived from 16S rRNA gene sequencing.[17] Sequencing was performed for the V4 region of the 16S rRNA gene using the 515F/806R primer pair on the Illumina platform, a protocol widely used for high-throughput microbial community profiling.[18] Sequence processing and amplicon sequence variant inference were based on the DADA2 pipeline, which models and corrects Illumina amplicon sequencing errors to infer high-resolution sequence variants.[19]

The microbiome feature set included four alpha-diversity metrics: observed ASVs, Faith's phylogenetic diversity, Shannon–Wiener diversity, and inverse Simpson diversity. Beta-diversity features included the first two principal coordinate axes and cluster assignments derived from Bray–Curtis dissimilarity, unweighted UniFrac distance, and weighted UniFrac distance. UniFrac-based measures incorporate phylogenetic information when comparing microbial communities across samples.[20] Referring to previous literature,[21,22] these microbiome-derived features were used as candidate predictors in OralAge model development and were also summarized across OralAgeAccel tertiles to describe microbiome-related differences in the analytic sample. Each cluster assignment was represented as a four-level categorical variable (clusters 1–4) and entered as a nominal predictor; the available cluster labels were not interpreted as ordered categories.

Oral clinical and tooth-status features

Oral clinical phenotypes included periodontitis status, periodontal burden count, dental implant status, number of present permanent teeth, number of missing teeth, number of root fragments, number of implant teeth, number of natural or restored teeth, and number of non-natural tooth-status findings. The periodontal burden count represented the number of examined non-third-molar tooth positions classified as periodontally affected (range, 0–28). Tooth-level status was recorded at 32 positions using five observed categories: permanent tooth present, tooth not present, permanent dental root fragment present, dental implant, and primary (deciduous) tooth present. Position-specific categories were entered as categorical predictors, and aggregate counts were obtained by summing the relevant status across positions. A derived tooth-loss group variable was created according to the number of missing teeth: 0, 1–4, 5–8, and ≥ 9 missing teeth.

Periodontal information in NHANES 2009–2012 was based on full-mouth periodontal examination protocols among adults aged ≥ 30 years, with probing measurements obtained at six sites per tooth, excluding third molars.[23] Periodontitis was classified using the CDC/AAP surveillance case definition, which defines mild, moderate, and severe periodontitis using combinations of clinical attachment loss and probing depth at interproximal sites.[24] In the present analysis, periodontitis was analyzed as a binary variable indicating any periodontitis versus no periodontitis.

Development of OralAge

Chronological age was used as the target variable for OralAge model training.

Candidate predictors were grouped into four feature sets: microbiome features only, oral clinical features only, combined derived oral features, and combined features further incorporating tooth-level status. Chronological age itself was not included as a predictor. Categorical variables were dummy-coded, continuous variables were standardized within the modeling process, and predictors with only one observed level or zero variance were removed before model fitting.

Ridge regression was used to construct OralAge. This approach was selected to handle correlated predictors and stabilize estimation in the presence of multiple microbiome, clinical, and tooth-level variables. Ridge models were fitted using penalized regression with the mixing parameter fixed at $\alpha = 0$, and the penalty parameter λ was selected by cross-validation using the glmnet framework.[25] Model performance was evaluated using 5-fold cross-validation. Out-of-fold predicted age values from the final model were used as OralAge to reduce optimism from in-sample prediction. Model performance was assessed using mean absolute error, root mean square error, coefficient of determination, and Pearson correlation between OralAge and chronological age. The four candidate feature sets were specified before performance comparison and evaluated using the same fold assignments. Reported metrics were calculated from held-out predictions, so the current evaluation represents internal cross-validation rather than independent external validation.

Derivation of OralAgeAccel

OralAgeAccel was calculated to quantify whether an individual's oral profile appeared older or younger than expected for their chronological age. Specifically, out-of-fold OralAge was regressed on chronological age in the analytic sample, and the residual from this model was defined as OralAgeAccel. This residual-based approach follows the logic used in biological ageing clocks, where age acceleration is often defined as residual variation in predicted biological age after accounting for chronological age.[26] A positive OralAgeAccel value indicates an older-than-expected oral ageing profile, whereas a negative value indicates a younger-than-expected oral ageing profile. By construction, residualization makes the sample linear correlation between OralAgeAccel and chronological age approximately zero; this property is descriptive and cannot serve as independent validation.

Model interpretation and clinical characterization

To describe the relative contribution of predictors in the final OralAge model, the selected ridge model was refitted in the full analytic sample using standardized predictors. Predictors were ranked according to the absolute value of their standardized ridge coefficients. Clinical characterization of OralAgeAccel was performed by comparing its distribution across periodontitis status and tooth-loss groups. These comparisons quantify expected concordance with components of the phenotype and do not provide independent validation.

Assessment of disease outcomes and disease burden

Thirteen individual disease or health-condition outcomes were evaluated in the disease-association analyses: cardiovascular disease (CVD), diabetes, prediabetes, diabetes or prediabetes, chronic kidney disease (CKD), metabolic syndrome, hypertension, depressive symptoms, sleep trouble, sleep disorder, stroke, Parkinson disease, and frailty.

CVD was defined using self-reported physician-diagnosed cardiovascular conditions available in NHANES, consistent with prior NHANES-based cardiovascular analyses. Diabetes and prediabetes were defined using a combination of self-reported diagnosis, medication use, and laboratory-based glycemic criteria, following American Diabetes Association diagnostic thresholds.[27] A combined dysglycemia outcome was also created to identify participants with either diabetes or prediabetes. Hypertension was defined using measured blood pressure, self-reported physician diagnosis, or antihypertensive medication use, based on conventional JNC 7 criteria.[28] CKD was defined using estimated glomerular filtration rate and albuminuria, with CKD classified as estimated glomerular filtration rate < 60 mL/min/1.73 m^2 or urinary albumin-to-creatinine ratio ≥ 30 mg/g.[29] Metabolic syndrome was defined according to the harmonized criteria, requiring at least three of the following components: elevated waist circumference, elevated triglycerides, reduced high-density lipoprotein cholesterol, elevated blood pressure or antihypertensive medication use, and elevated fasting glucose or glucose-lowering medication use.[30]

Depressive symptoms were assessed using the Patient Health Questionnaire-9 (PHQ-9). Participants with a PHQ-9 score ≥ 10 were classified as having clinically relevant depressive symptoms.[31] Sleep trouble and sleep disorder were derived from the NHANES Sleep Disorders Questionnaire, including self-reported history of being told by a doctor or other health professional about trouble sleeping or a sleep disorder. Stroke and Parkinson disease were identified using the corresponding self-reported physician-diagnosed disease variables. Frailty was defined using a deficit-accumulation frailty index; participants meeting the prespecified frailty threshold were classified as frail.[32]

To summarize disease burden across systems, individual outcomes were grouped into disease domains. Cardiometabolic conditions included CVD, diabetes, metabolic syndrome, and hypertension. Metabolic abnormality included diabetes, prediabetes, metabolic syndrome, and hypertension. Renal disease included CKD. Neurological conditions included stroke and Parkinson disease. Mental and sleep-related conditions included depressive symptoms, sleep trouble, and sleep disorder. Geriatric condition included frailty. For each domain, a binary indicator was created to indicate the presence of at least one condition in that domain. Count variables were also calculated by summing the number of conditions within each domain.

Two overall disease-burden indicators were constructed. Core chronic condition burden included CVD, diabetes, CKD, metabolic syndrome, hypertension, stroke, Parkinson disease, and frailty. Extended condition burden further incorporated prediabetes, depressive symptoms, sleep trouble, and sleep disorder. For each burden definition, both a binary indicator of any condition and a count variable were generated.

Composite health scores

To examine potential determinants of OralAgeAccel, this study assessed five composite health or social-environmental scores: Life's Simple 7 score, Life's Essential 8 score, Life's Crucial 9 score, oxidative balance score, and social determinants of health score. Higher Life's Simple 7, Life's Essential 8, Life's Crucial 9, and oxidative balance scores represented more favorable health profiles, whereas a higher social determinants of health score represented greater social disadvantage.

Life's Simple 7 score was constructed to reflect seven cardiovascular health components: diet, physical activity, smoking status, body mass index, blood pressure, total cholesterol, and fasting glucose.[33] Life's Essential 8 score extended the cardiovascular health framework by incorporating diet, physical activity, nicotine exposure, sleep health, body mass index, blood lipids, blood glucose, and blood pressure.[34,35] Life's Crucial 9 score was derived as an extension of Life's Essential 8 by incorporating psychological health, operationalized from PHQ-9 depressive symptom categories (scores 0–9 and 10–27), as an additional domain.[36] The oxidative balance score summarized dietary and lifestyle pro-oxidant and antioxidant exposures; higher values indicated a more antioxidant-favorable profile.[37] The social determinants of health score summarized cumulative social disadvantage. In the present analysis, this score included socioeconomic and social factors available in NHANES, including educational attainment, marital status, and poverty-to-income ratio; therefore, these factors were not included as covariates in the fully adjusted model for the social determinants of health score.[38]

Covariates

Covariates were selected according to some prior NHANES-based population analyses.[39,40] Demographic covariates included chronological age, sex, and race/ethnicity. Socioeconomic covariates included educational attainment, marital status, and poverty-to-income ratio. Behavioral covariates included smoking status and alcohol use. These covariates were used to define the sequential adjustment models described below. When a composite score already included specific socioeconomic or behavioral variables, the corresponding covariates were not simultaneously included in the fully adjusted model for that score to avoid adjustment for score components.

Statistical analysis

Continuous variables were summarized using means and standard deviations or medians and interquartile ranges, as appropriate. Categorical variables were summarized as counts and percentages. Weighted characteristics of the analytic sample were described across tertiles of OralAgeAccel. The primary OralAge variable for analysis was the out-of-fold prediction from the selected ridge model, and the primary oral ageing acceleration index was the residualized OralAgeAccel.

All main association analyses used NHANES analytic weights, strata, and primary sampling units to account for unequal selection probability, stratified multistage sampling, and clustering. OralAgeAccel was standardized before regression analysis; estimates therefore correspond to the association per 1-SD higher OralAgeAccel.

For individual diseases and binary disease-domain outcomes, survey-weighted logistic regression models were fitted, and results were reported as odds ratios with 95% confidence intervals. For disease-count outcomes, survey-weighted linear regression models were fitted, and results were reported as beta coefficients with 95% confidence intervals. Three models were fitted sequentially. The crude model included OralAgeAccel only. Model 1 adjusted for chronological age and sex. Model 2 further adjusted for race/ethnicity, educational attainment, marital status, poverty-to-income ratio, smoking status, and alcohol use. Model 2 was treated as the primary model for interpretation. Linear models were selected to estimate adjusted absolute differences in the mean number of conditions per 1-SD higher OralAgeAccel.

Composite scores were standardized before regression analysis. Associations between composite scores and OralAgeAccel were examined using survey-weighted linear regression models. The crude model included the score only. Model 1 adjusted for age and sex. For Life's Simple 7 score, Life's Essential 8 score, Life's Crucial 9 score, and oxidative balance score, Model 2 further adjusted for race/ethnicity, educational attainment, marital status, and poverty-to-income ratio. For the social determinants of health score, Model 2 adjusted for age, sex, race/ethnicity, smoking status, and alcohol use, given that educational attainment, marital status, and poverty-to-income ratio were components of the score. Results were reported as beta coefficients and 95% confidence intervals per 1-SD higher score.

Restricted cubic spline analyses were performed to visualize potential non-linear exposure–response patterns. For disease outcomes, spline analyses were performed for 12 non-overlapping individual outcomes after excluding the combined diabetes-or-prediabetes outcome from the spline figure to avoid redundancy. For composite-score analyses, spline models examined each standardized composite score in relation to OralAgeAccel. Spline analyses were conducted without survey weights and used the same covariate structure as the corresponding fully adjusted model. Four knots were placed at the 5th, 35th, 65th, and 95th percentiles of the relevant exposure distribution. For OralAgeAccel spline models, the reference value was set at OralAgeAccel = 0 SD. For composite-score spline models, the reference value was set at the standardized score value of 0 SD. Overall associations were assessed by comparing the spline model with a covariate-only model, and non-linearity was assessed by comparing the spline model with a model containing only the linear exposure term.[41] These unweighted spline models were used as secondary descriptive shape analyses; population-level inference was based on the survey-weighted regression models.

All analyses were conducted using R version 4.5.2. Statistical tests were two-sided, and P values < 0.05 were treated as nominal evidence of association. Given the clinically distinct and correlated outcomes, multi-outcome association results were interpreted as exploratory, with emphasis on effect estimates and 95% confidence intervals.

RESULTS

Development and performance of OralAge

The OralAge model was developed in 5,323 adults aged 30–69 years with complete oral microbiome, periodontal, and tooth-status information. Four candidate feature sets were evaluated, including microbiome-derived features alone, oral clinical features alone, combined derived oral features, and combined features further incorporating tooth-level status. The microbiome-only model showed limited age-prediction performance, whereas the inclusion of oral clinical features improved model accuracy. The combined tooth-level model showed the best overall performance across candidate models and was therefore selected as the final OralAge model (Figure 2A, Table S1). Across Table S1, clinical features accounted for most of the predictive performance, microbiome features provided incremental information in the combined model, and tooth-position coding produced a small additional gain.

Using out-of-fold predictions from the final ridge regression model, OralAge was correlated with chronological age in the analytic sample (Pearson r = 0.538). The model achieved a mean absolute error of 7.740 years, a root mean square error of 9.390 years, and an R^2 of 0.290 (Figure 2B). OralAgeAccel was then derived as the residual of OralAge after adjustment for chronological age. As expected from residualization in the same analytic sample, the linear correlation between OralAgeAccel and chronological age was close to zero (Figure 2C). This mathematical property does not constitute independent validation of age independence in a separate population.

The leading predictors in the final model included beta-diversity coordinates, periodontal burden count, periodontitis status, and selected tooth-level status indicators (Figure 2D, Table S2). Among microbiome-derived features, unweighted UniFrac PCoA1 showed the largest positive coefficient, followed by Bray–Curtis PCoA2 and other beta-diversity features. Clinical oral indicators also contributed prominently, particularly periodontal burden count and periodontitis status. Several tooth-level features related to tooth absence, tooth presence, and root fragments were also retained among the top predictors, supporting the contribution of both microbial community structure and clinical oral deterioration to OralAge estimation.

OralAgeAccel showed clear clinical gradients across established oral health indicators. Participants with periodontitis had higher OralAgeAccel than those without periodontitis, with mean values of 2.36 and −2.60 years, respectively (Figure 2E). A similar dose–response pattern was observed across tooth-loss burden. Mean OralAgeAccel increased from −3.44 years among participants with no missing teeth to −1.98 years among those with 1–4 missing teeth, −0.21 years among those with 5–8 missing teeth, and 4.53 years among those with nine or more missing teeth (Figure 2F). These findings show expected clinical concordance with model components and characterize the phenotype; they do not provide independent validation.

Associations of OralAgeAccel with chronic disease burden

Weighted characteristics of the analytic sample according to OralAgeAccel tertiles are shown in Table 1. The final disease-analysis sample included 5,323 adults aged 30–69 years, representing approximately 109.2 million US adults. OralAgeAccel tertiles showed clear gradients in oral clinical phenotypes. Across the lowest, middle, and highest tertiles, weighted mean predicted OralAge increased from 43.261 to 46.888 and 53.199 years, while weighted mean OralAgeAccel increased from −5.247 to −0.372 and 5.401 years. Periodontal burden count increased from 1.259 to 3.278 and 6.834, whereas the number of present permanent teeth decreased from 27.891 to 26.620 and 21.814. The weighted mean number of missing teeth increased from 3.887 in the lowest tertile to 9.698 in the highest tertile. Periodontitis was also more frequent across increasing OralAgeAccel tertiles, with weighted prevalences of 18.434%, 42.922%, and 77.306%. Participants in the highest OralAgeAccel tertile also showed lower alpha-diversity metrics, including observed ASVs and Faith's phylogenetic diversity, together with differences in beta-diversity coordinates. These patterns suggest that higher OralAgeAccel was accompanied by both reduced within-sample microbial diversity and altered between-person microbial community structure.

In weighted logistic regression models, higher OralAgeAccel was associated with some chronic diseases after full adjustment for demographic, socioeconomic, and behavioral covariates (Figure 3A, Table S3). Per 1-SD increase in OralAgeAccel, the odds were higher for CVD (OR = 1.134, 95% CI: 1.004–1.280; P = 0.043), diabetes (OR = 1.134, 95% CI: 1.038–1.238; P = 0.005), CKD (OR = 1.139, 95% CI: 1.033–1.256; P = 0.009), and stroke (OR = 1.216, 95% CI: 1.003–1.475; P = 0.047). Association with frailty (OR = 1.076, 95% CI: 0.991–1.170; P = 0.083) was positive but did not reach conventional statistical significance. No clear independent associations were observed for prediabetes, metabolic syndrome, hypertension, depressive symptoms, sleep trouble, sleep disorder, or Parkinson disease in

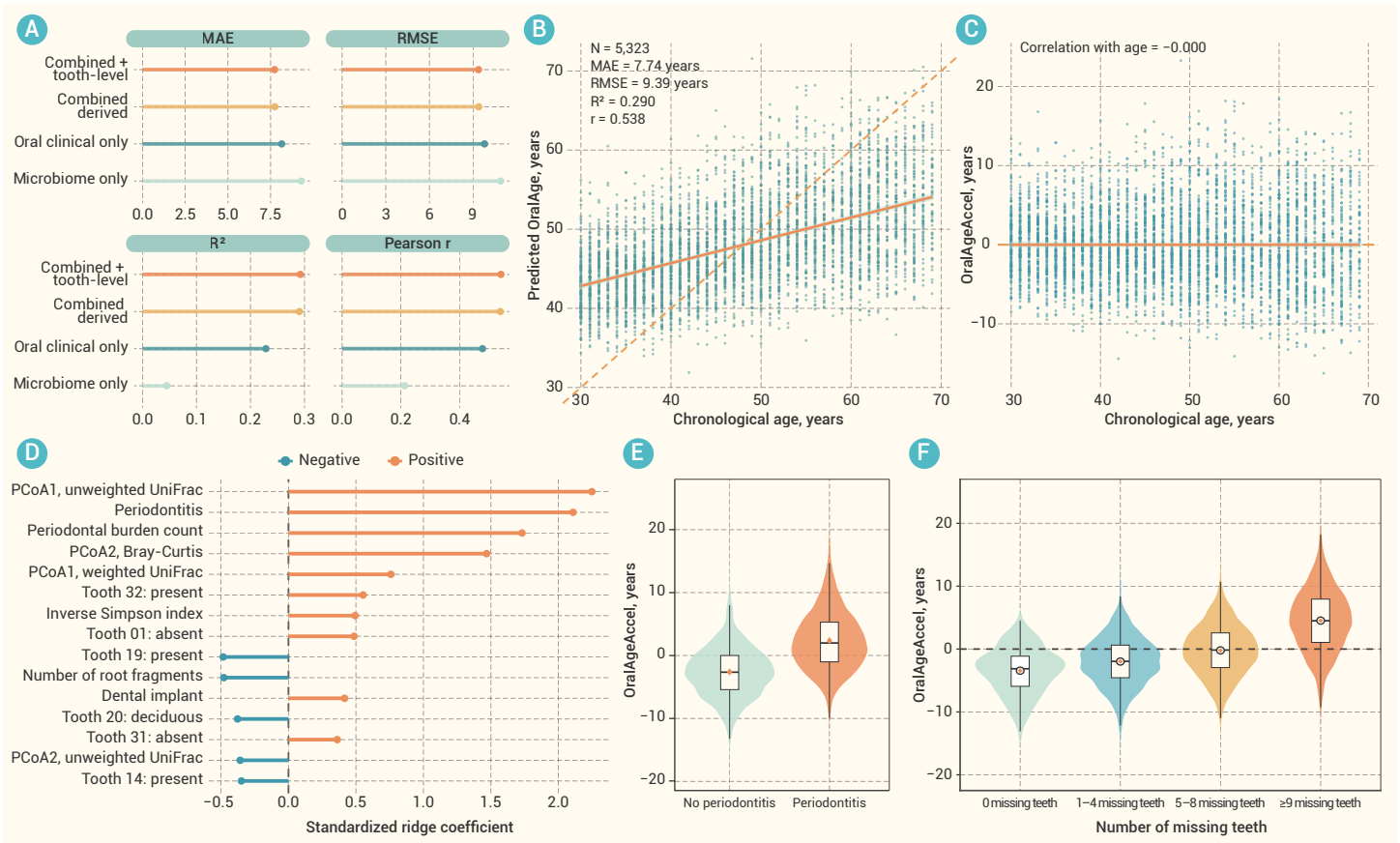


Figure 2. Development, performance, and clinical characterization of OralAge Notes: (A) Cross-validated performance of candidate OralAge models based on microbiome-only, oral clinical-only, combined derived, and combined tooth-level feature sets. (B) Association between predicted OralAge and chronological age in the final ridge regression model. (C) Association between OralAgeAccel and chronological age after residualization; the near-zero linear correlation is expected by construction and is descriptive. (D) Top predictors ranked by standardized ridge coefficients in the final OralAge model. (E) Distribution of OralAgeAccel according to periodontitis status. (F) Distribution of OralAgeAccel according to tooth-loss burden.

the fully adjusted model. Figure 3A presents the fully adjusted odds ratios and 95% confidence intervals for 13 outcomes, including the combined diabetes-or-prediabetes outcome. Table S3 provides the corresponding numerical estimates and P values across all three models.

At the disease-domain level, higher OralAgeAccel was associated with greater cardiometabolic, renal, and overall chronic disease burden (Figure 3B, Table S4). Each 1-SD higher OralAgeAccel was associated with higher odds of having any cardiometabolic condition (OR = 1.088, 95% CI: 1.018–1.162; P = 0.012), renal disease (OR = 1.139, 95% CI: 1.033–1.256; P = 0.009), any core chronic condition (OR = 1.119, 95% CI: 1.045–1.198; P = 0.001), and any extended condition (OR = 1.076, 95% CI: 1.002–1.155; P = 0.045).

Disease-count models provided an assessment of multimorbidity burden (Figure 3C, Table S5). In the fully adjusted weighted linear models, each 1-SD increase in OralAgeAccel was associated with a higher cardiometabolic condition count (β = 0.048, 95% CI: 0.012–0.085; P = 0.009), renal count (β = 0.014, 95% CI: 0.003–0.025; P = 0.015), geriatric count (β = 0.012, 95% CI: 0.000–0.024; P = 0.044), core chronic condition count (β = 0.083, 95% CI: 0.037–0.129; P < 0.001), and extended condition count (β = 0.074, 95% CI: 0.014–0.134; P = 0.016).

Restricted cubic spline analyses were used to visualize the adjusted exposure–response patterns between OralAgeAccel and individual diseases (Figure 4). As secondary unweighted descriptive analyses, the curves generally aligned with the directions of the weighted regression estimates. Increasing patterns were apparent for CVD, diabetes, CKD, stroke, and frailty. Curves for metabolic syndrome, hypertension, depressive symptoms, sleep-related outcomes, and Parkinson disease did not show stable positive gradients. Overall, these results indicate that accelerated oral ageing is most consistently related to cardiometabolic, renal, cerebrovascular, and multimorbidity-related health burden rather than to mental or sleep-related conditions.

Associations of Composite health scores with OralAgeAccel

Associations between composite health scores and OralAgeAccel are shown in Figure 5 and Table S6. In weighted linear regression models, higher Life's Simple 7, Life's Essential 8, and Life's Crucial 9 scores were consistently associated with lower OralAgeAccel. In the fully adjusted model, each 1-SD higher Life's Simple 7 score was associated with a 0.300-year lower OralAgeAccel (β = −0.300, 95% CI: −0.495 to −0.105; P = 0.003). Similar inverse associations were observed for Life's Essential 8 score (β = −0.317, 95% CI: −0.530 to −0.103; P = 0.004) and Life's Crucial 9 score (β = −0.284, 95% CI: −0.502 to −0.067; P = 0.010). The oxidative balance score showed inverse associations in the crude and age- and sex-adjusted models, but the association was attenuated after full adjustment (β = −0.132, 95% CI: −0.315 to 0.050; P = 0.156). By contrast, higher social determinants of health score, indicating greater social disadvantage, was associated with higher OralAgeAccel across all models; in the fully adjusted model, each 1-SD higher score was associated with a 0.393-year higher OralAgeAccel (95% CI: 0.182 to 0.605; P < 0.001).

Tertile-based analyses provided a visual summary of the distribution of OralAgeAccel across composite-score groups (Figure 6A). OralAgeAccel was generally lower among participants in higher tertiles of Life's Simple 7, Life's Essential 8, and Life's Crucial 9 scores, whereas the opposite pattern was observed for the social determinants of health score. The oxidative balance score showed a weaker and less consistent tertile gradient. Restricted cubic spline analyses further characterized the exposure–response patterns between composite scores and OralAgeAccel (Figure 6B, Table S7). Life's Simple 7, Life's Essential 8, and Life's Crucial 9 scores showed inverse and non-linear associations with OralAgeAccel, with lower OralAgeAccel across higher score ranges and a partial flattening or rebound at the upper tail of the distributions. By contrast, the social determinants of health score showed a positive overall association with OralAgeAccel, with no strong evidence of

Table 1. Weighted characteristics of the analytic sample according to OralAgeAccel tertiles.

Characteristic	Overall	T1 (lowest)	T2	T3 (highest)	P value
Demographic and lifestyle characteristics					
Age, years	47.507 (10.719)	49.703 (10.675)	45.375 (10.325)	47.241 (10.694)	< 0.001
Poverty income ratio	3.177 (1.648)	3.388 (1.597)	3.220 (1.646)	2.874 (1.666)	< 0.001
Body mass index, kg/m^2	29.129 (6.451)	28.910 (6.245)	29.270 (6.356)	29.234 (6.781)	0.175
Female	2,666 (50.016)	973 (54.337)	886 (50.027)	807 (44.856)	< 0.001
Male	2,657 (49.984)	802 (45.663)	887 (49.973)	968 (55.144)	
Mexican American	899 (9.024)	290 (7.623)	318 (9.968)	291 (9.653)	< 0.001
Other Hispanic	567 (5.648)	204 (5.380)	190 (6.097)	173 (5.472)	
Non-Hispanic White	2,013 (66.002)	702 (70.470)	657 (64.041)	654 (62.841)	
Non-Hispanic Black	1,223 (11.723)	391 (10.498)	395 (11.867)	437 (13.026)	
Other Race - Including Multi-Racial	621 (7.603)	188 (6.030)	213 (8.027)	220 (9.009)	
Less than 9th grade	568 (5.590)	154 (3.807)	182 (5.690)	232 (7.608)	< 0.001
9-11th grade	751 (10.434)	216 (8.277)	212 (9.043)	323 (14.546)	
High school graduate/GED or equivalent	1,128 (20.364)	355 (18.976)	344 (17.531)	429 (25.152)	
Some college or AA degree	1,469 (29.366)	512 (28.420)	524 (32.874)	433 (26.618)	
College graduate or above	1,401 (34.246)	537 (40.521)	509 (34.863)	355 (26.075)	
Married	3,070 (63.002)	1,041 (66.268)	1,055 (62.701)	974 (59.435)	< 0.001
Living with partner	423 (7.380)	125 (5.901)	145 (8.047)	153 (8.408)	
Widowed	196 (2.461)	69 (2.391)	59 (2.664)	68 (2.318)	
Divorced	703 (12.876)	247 (13.141)	204 (11.434)	252 (14.154)	
Separated	250 (3.042)	77 (2.565)	85 (3.061)	88 (3.592)	
Never married	677 (11.239)	215 (9.733)	224 (12.092)	238 (12.093)	
Never smoker	2,999 (56.218)	1,128 (63.445)	1,038 (56.939)	833 (46.813)	< 0.001
Former smoker	1,183 (24.356)	380 (24.373)	375 (23.975)	428 (24.757)	
Current smoker	1,139 (19.426)	266 (12.183)	359 (19.086)	514 (28.430)	
Never alcohol user	578 (8.925)	191 (9.062)	188 (8.315)	199 (9.435)	< 0.001
Former alcohol user	828 (14.633)	222 (10.857)	252 (13.819)	354 (20.102)	
Mild alcohol user	1,643 (37.815)	622 (44.743)	546 (35.991)	475 (31.448)	
Moderate alcohol user	779 (18.400)	288 (18.995)	278 (20.385)	213 (15.487)	
Heavy alcohol user	1,026 (20.226)	304 (16.343)	336 (21.489)	386 (23.527)	
Oral ageing and oral health characteristics					
OralAge, years	47.492 (5.532)	43.261 (3.291)	46.888 (3.135)	53.199 (4.774)	< 0.001
OralAgeAccel, years	-0.383 (4.903)	-5.247 (2.312)	-0.372 (1.172)	5.401 (3.113)	< 0.001
Periodontal burden count	3.627 (5.404)	1.259 (2.780)	3.278 (5.012)	6.834 (6.496)	< 0.001
Present permanent teeth, n	25.621 (5.299)	27.891 (2.671)	26.620 (3.551)	21.814 (6.956)	< 0.001
Missing teeth, n	6.047 (5.047)	3.887 (2.475)	5.069 (3.202)	9.698 (6.735)	< 0.001
No periodontitis	2,530 (55.497)	1,348 (81.566)	851 (57.078)	331 (22.694)	< 0.001
Periodontitis	2,793 (44.503)	427 (18.434)	922 (42.922)	1,444 (77.306)	
No dental implant	5,213 (97.993)	1,755 (98.552)	1,739 (98.051)	1,719 (97.262)	< 0.001
Dental implant	110 (2.007)	20 (1.448)	34 (1.949)	56 (2.738)	
0 missing teeth	369 (6.665)	217 (11.722)	122 (5.866)	30 (1.522)	< 0.001
1-4 missing teeth	2,167 (46.100)	1,002 (60.025)	809 (50.767)	356 (24.359)	

Table 1. *(Continued)*

Characteristic	Overall	T1 (lowest)	T2	T3 (highest)	P value
5-8 missing teeth	1,488 (29.192)	442 (23.791)	573 (32.934)	473 (31.498)	
≥ 9 missing teeth	1,299 (18.043)	114 (4.462)	269 (10.433)	916 (42.622)	
Oral microbiome characteristics					
Observed OTUs	116.119 (36.322)	124.327 (35.565)	115.668 (36.193)	106.836 (35.042)	< 0.001
Faith's phylogenetic diversity	13.439 (3.099)	14.144 (2.981)	13.402 (3.177)	12.640 (2.947)	< 0.001
Shannon-Wiener diversity	4.579 (0.650)	4.616 (0.672)	4.589 (0.650)	4.523 (0.619)	< 0.001
Inverse Simpson diversity	0.902 (0.061)	0.897 (0.069)	0.903 (0.060)	0.906 (0.051)	< 0.001
Bray-Curtis PCoA1	-0.011 (0.171)	0.013 (0.180)	-0.005 (0.167)	-0.045 (0.160)	< 0.001
Unweighted UniFrac PCoA1	0.018 (0.132)	-0.015 (0.123)	0.025 (0.130)	0.050 (0.135)	< 0.001
Weighted UniFrac PCoA1	0.004 (0.120)	0.001 (0.114)	0.001 (0.117)	0.011 (0.129)	0.015
Chronic disease and health conditions					
CVD	333 (5.408)	77 (3.637)	90 (4.704)	166 (8.295)	< 0.001
Diabetes	912 (12.439)	273 (10.812)	274 (12.897)	365 (13.866)	< 0.001
Prediabetes	412 (8.033)	130 (8.931)	141 (7.762)	141 (7.266)	< 0.001
CKD	601 (9.346)	166 (7.974)	195 (9.897)	240 (10.375)	< 0.001
Metabolic syndrome	1,288 (23.105)	415 (22.907)	408 (23.320)	465 (23.105)	< 0.001
Hypertension	2,054 (35.173)	686 (35.101)	630 (33.644)	738 (36.945)	< 0.001
Depressive symptoms	457 (7.487)	137 (6.178)	138 (7.575)	182 (8.977)	< 0.001
Sleep trouble	1,301 (26.833)	426 (26.264)	412 (26.478)	463 (27.902)	< 0.001
Sleep disorder	458 (9.073)	156 (8.635)	142 (9.057)	160 (9.614)	< 0.001
Stroke	131 (1.795)	33 (1.294)	32 (1.334)	66 (2.901)	< 0.001
Parkinson disease	36 (0.939)	13 (1.155)	7 (0.284)	16 (1.405)	< 0.001
Frailty	965 (14.106)	273 (11.478)	281 (14.372)	411 (16.945)	< 0.001
Disease burden counts					
Cardiometabolic count	0.760 (0.985)	0.724 (0.952)	0.744 (0.990)	0.821 (1.016)	0.009
Renal count	0.093 (0.291)	0.080 (0.271)	0.099 (0.299)	0.104 (0.305)	0.037
Neurological count	0.027 (0.163)	0.024 (0.155)	0.016 (0.126)	0.043 (0.203)	< 0.001
Mental/sleep count	0.428 (0.720)	0.407 (0.704)	0.425 (0.716)	0.458 (0.741)	0.107
Geriatric count	0.141 (0.348)	0.115 (0.319)	0.144 (0.351)	0.169 (0.375)	< 0.001
Core chronic condition count	1.178 (1.471)	1.084 (1.383)	1.159 (1.462)	1.312 (1.568)	< 0.001
Extended condition count	1.526 (1.684)	1.435 (1.590)	1.501 (1.677)	1.664 (1.790)	< 0.001

Note. Values are weighted mean (SD) for continuous variables and n (weighted %) for categorical variables. NHANES oral microbiome sample weights were used for descriptive estimates. OralAgeAccel tertiles were defined from the analytic sample distribution: T1, lowest; T3, highest. P values were estimated using weighted linear regression for continuous variables and weighted chi-square tests for categorical variables. Beta-diversity indicators refer to the first principal coordinate from Bray-Curtis, unweighted UniFrac, and weighted UniFrac distance matrices.

non-linearity. Additional exploratory associations between dietary food groups and OralAgeAccel across the crude, age- and sex-adjusted, and fully adjusted models are presented in Table S8.

DISCUSSIONS

In this study, we developed OralAge as a microbiome-clinical oral ageing phenotype using oral microbiome features, periodontal indicators, and tooth-status information. OralAgeAccel, defined as the residual of predicted OralAge after accounting for chronological age, was higher in participants with periodontitis and was greater among those with more missing teeth. Higher OralAgeAccel was also associated with chronic diseases and broader disease-burden measures, particularly CVD, diabetes, CKD, stroke, cardiometabolic burden, renal burden, and core chronic condition count. In determinant analyses, more favorable cardiovascular and multidimensional health scores were associated with lower OralAgeAccel, whereas greater social disadvantage was associated with higher OralAgeAccel.

The construction of OralAge extends the biological age framework to the oral cavity. Biological ageing measures are commonly designed to summarize age-related variation that chronological age cannot fully represent.[42] In this study, the oral cavity was treated as a distinct ageing-relevant system, with microbial community structure, periodontal status, periodontal burden, and tooth-level information integrated into a single predicted-age phenotype. This

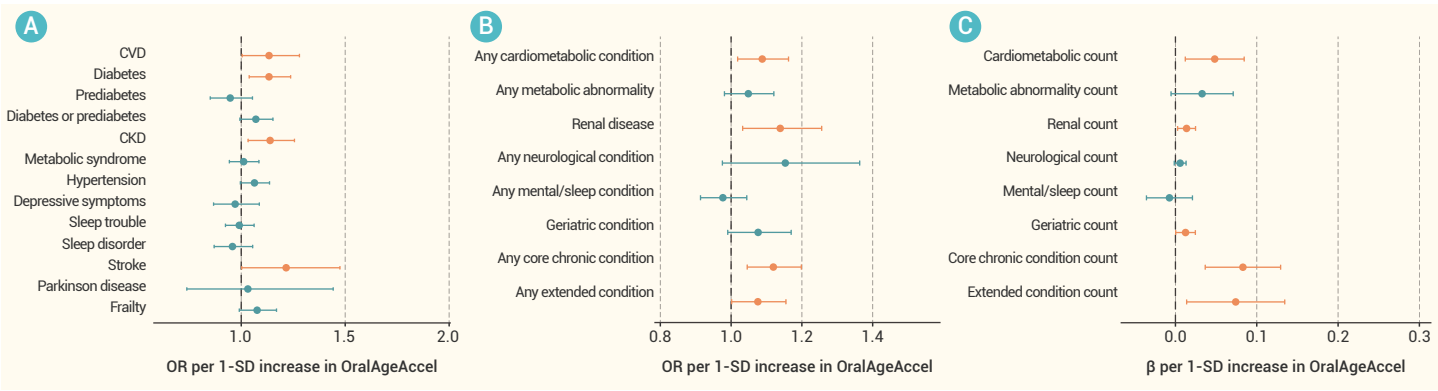


Figure 3. Associations of OralAgeAccel with chronic disease burden Notes: (A) Fully adjusted weighted associations of OralAgeAccel with individual chronic diseases. (B) Fully adjusted weighted associations of OralAgeAccel with binary disease-domain and overall disease-burden indicators. (C) Fully adjusted weighted associations of OralAgeAccel with disease-count outcomes. Estimates in panels A and B are odds ratios per 1-SD higher OralAgeAccel; estimates in panel C are beta coefficients per 1-SD higher OralAgeAccel.

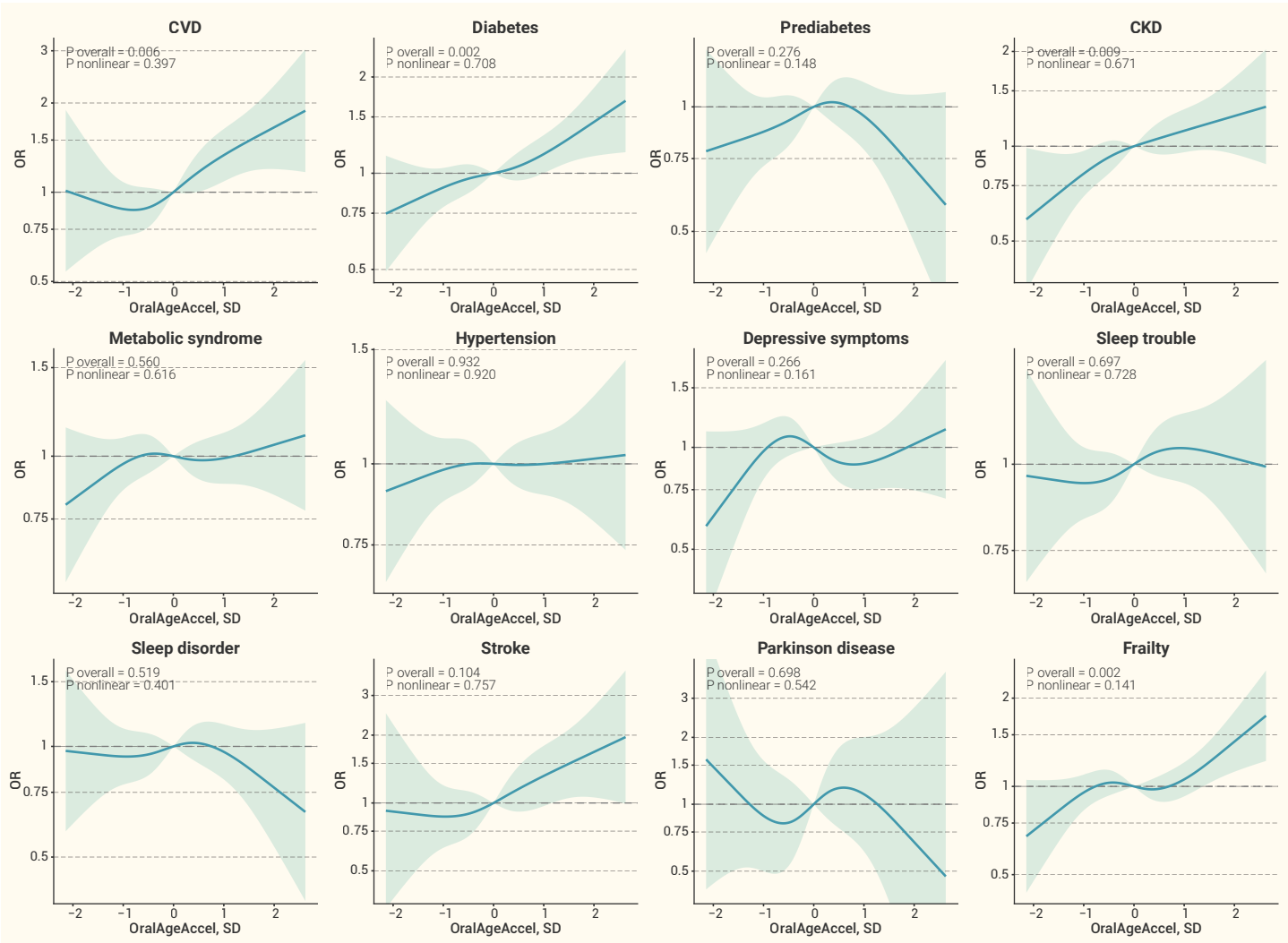


Figure 4. Restricted cubic spline associations between OralAgeAccel and individual diseases Notes: Restricted cubic spline curves showing adjusted exposure–response patterns between OralAgeAccel and 12 non-overlapping individual disease outcomes. Shaded areas indicate 95% confidence intervals.

approach is consistent with the concept that ageing phenotypes can be derived from tissue- or system-specific data when the selected features reflect cumulative biological and clinical change.[43,44] The OralAge model considered both microbiome-derived and clinical oral features, supporting the need to combine ecological and clinical information when characterizing oral ageing. The strongest predictors included beta-diversity coordinates, periodontitis status, periodontal burden count, and selected tooth-level indicators. Periodontitis is a chronic inflammatory disease involving dysbiotic microbial communities and host-mediated tissue destruction.[45] Modern periodontal classification also emphasizes severity, complexity, and risk of progression, which aligns with the broader idea that periodontal deterioration accumulates across time and reflects more than a binary diagnosis.[46] Tooth loss and root fragments represent later-stage or cumulative oral damage and can also indicate long-term functional decline. The present model therefore has potential to capture a gradient from microbial community differences to clinical periodontal damage and tooth-level deterioration. The feature-set

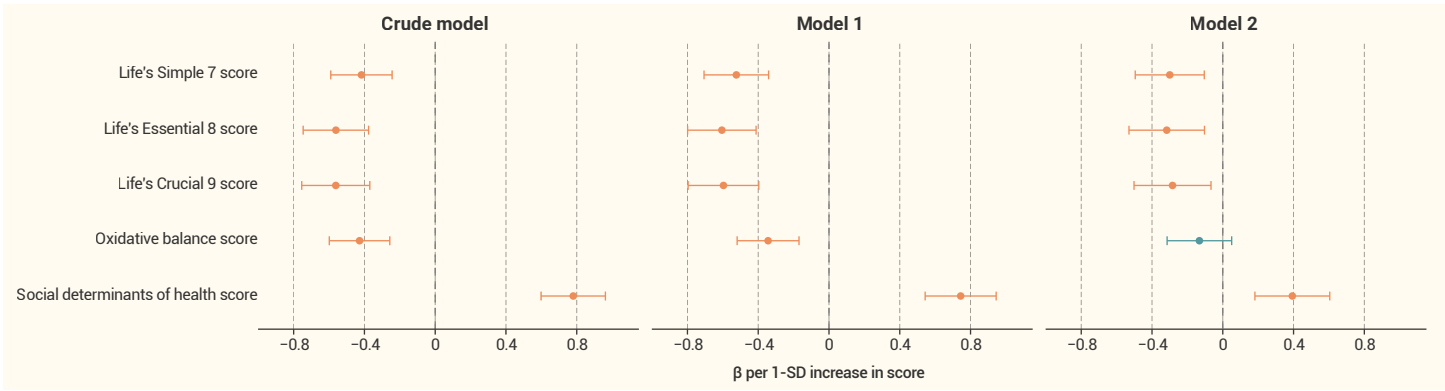


Figure 5. Associations of composite health scores with OralAgeAccel Notes: Survey-weighted linear regression associations between standardized composite health scores and OralAgeAccel across crude, age- and sex-adjusted, and fully adjusted models. Estimates are beta coefficients per 1-SD higher score.

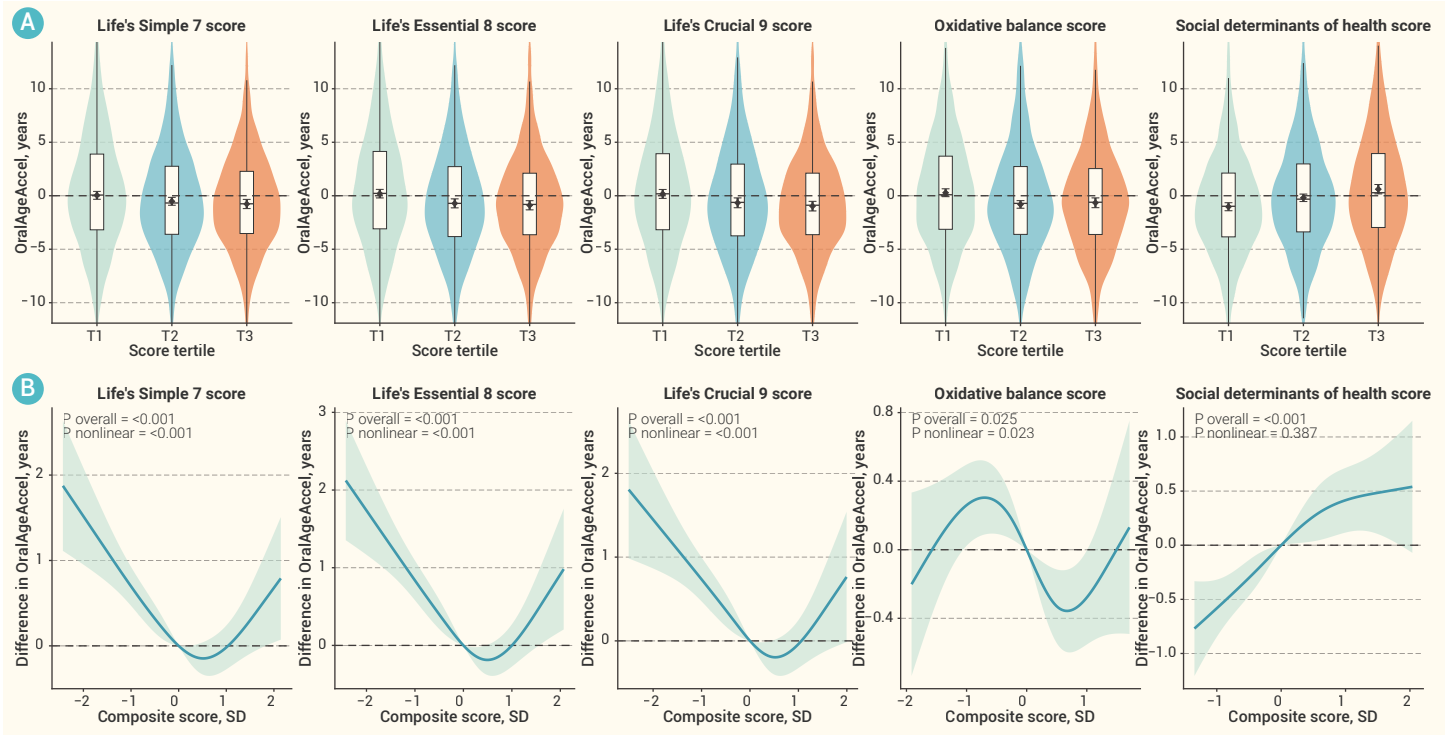


Figure 6. Distributions and dose–response associations of OralAgeAccel across composite health scores Notes: (A) Distribution of OralAgeAccel across tertiles of each composite score. (B) Restricted cubic spline curves showing adjusted exposure–response associations between standardized composite scores and OralAgeAccel. Shaded areas indicate 95% confidence intervals.

comparison further indicates that oral clinical measures contributed most of the age-prediction signal, microbiome features added incremental information, and tooth-position variables yielded only a small additional improvement.

The clinical-characterization findings further support the interpretation of OralAgeAccel as an oral ageing burden indicator. Participants with periodontitis had substantially higher OralAgeAccel than those without periodontitis, and OralAgeAccel increased stepwise across tooth-loss categories. This pattern is compatible with consensus statements emphasizing dental caries and periodontal diseases as major threats to oral health and well-being in ageing populations.[47] In the descriptive tertile analysis, higher OralAgeAccel was accompanied by lower observed ASVs, lower Faith's phylogenetic diversity, lower Shannon–Wiener diversity, and differences in beta-diversity coordinates. This pattern suggests that accelerated oral ageing involves both clinical deterioration and microbial community reorganization. Since periodontitis and tooth loss contributed to model construction, these gradients represent expected clinical concordance rather than independent validation.

This study found the associations between OralAgeAccel and cardiometabolic, renal, cerebrovascular, and multimorbidity-related outcomes. Oral dysbiosis and periodontal inflammation may contribute to systemic disease through recurrent microbial exposure, inflammatory mediators, endothelial activation, and oral-gut microbial interactions. A recent review summarized evidence linking oral dysbiosis with cardiometabolic health, including pathways involving inflammation, nitric oxide metabolism, and vascular function.[48] Oral bacteria may also translocate to the gut through enteral and hematogenous routes, providing a potential pathway by which oral microbial disturbances may influence distal mucosal and systemic immune environments.[49] The diabetes finding is consistent with clinical evidence showing an interaction between periodontal disease and glycemic control. Periodontal inflammation can increase systemic inflammatory burden, and diabetes can impair periodontal tissue response and host defense. Intervention evidence also supports clinical relevance: an updated Cochrane review reported that periodontal treatment using subgingival instrumentation improved glycemic control in people with both periodontitis and diabetes.[50] The association with CKD also deserves attention. CKD is characterized by systemic inflammation, immune dysregulation, altered mineral metabolism, vascular dysfunction, and high multimorbidity burden. Periodontitis may add to this inflammatory burden. In an NHANES III linked mortality study of adults with stages 3–5 CKD, periodontitis was associated with higher mortality, suggesting that periodontal health may be relevant in

medically vulnerable renal populations.[51] The present findings extend this line of evidence by showing that an age-independent oral ageing phenotype was associated with CKD and renal burden in a representative adult sample. The stroke association was also positive. Systematic review evidence has reported an association between periodontitis and stroke risk.[52] Mechanistically, oral inflammation may contribute to cerebrovascular risk through systemic inflammatory activation, endothelial dysfunction, platelet activation, and atherosclerotic processes. Sustained inflammatory exposure could also affect endothelial function and mitochondrial homeostasis, while disturbances in mitochondrial quality control, organelle dynamics, and endoplasmic reticulum–mitochondria communication are increasingly implicated in neuronal vulnerability and neurodegenerative processes.[53]

The association with frailty was positive but did not reach conventional statistical significance in the fully adjusted model. The disease-count model and spline pattern nevertheless suggested a geriatric burden signal. Oral frailty research provides a useful context for this finding. In community-dwelling older adults, accumulated poor oral status has been shown to predict physical frailty, disability, and mortality.[54] A recent systematic review and meta-analysis also reported that oral frailty is prevalent among older adults and may be related to unfavorable outcomes.[55] In the present study, the sample was limited to adults aged 30–69 years, which may reduce the ability to detect strong frailty-related associations. Weak associations were identified for depressive symptoms and sleep-related outcomes. These outcomes were based on symptom or questionnaire-derived information and may reflect more heterogeneous pathways than cardiometabolic or renal outcomes. Oral health can influence social participation, pain, sleep quality, and psychological well-being, yet these relationships are often shaped by socioeconomic conditions, health behavior, comorbidities, medication use, and access to care. The weaker neurological associations may reflect the blood-brain barrier, low event counts for stroke and Parkinson disease, self-reported ascertainment, and limited representation of older adults; these constraints reduce power and may attenuate observable gradients relative to renal and vascular outcomes.

The composite-score analyses provide additional insight into the determinants of accelerated oral ageing. Higher Life's Simple 7, Life's Essential 8, and Life's Crucial 9 scores were associated with lower OralAgeAccel, whereas higher social determinants of health score were associated with higher OralAgeAccel. These findings support a common-risk-factor interpretation: the same behavioral, metabolic, and social conditions that shape systemic health also shape oral ageing. Consistent with this broader framework, higher physical activity has been associated with younger age estimates across several DNA methylation clocks and with lower PhenoAge, with telomere length accounting for a small proportion of the physical activity–PhenoAge association.[56,57] Recent studies have reported inverse associations between Life's Essential 8 and periodontitis, supporting the relevance of cardiovascular health profiles to periodontal health.[58,59] Social disadvantage is also strongly related to oral health outcomes in representative US adults, including untreated caries, tooth loss, and limited access-related outcomes.[60]

This study has several strengths. First, it used a representative US adult sample with oral microbiome, periodontal examination, and tooth-status data. Second, the OralAge model incorporated both microbial and clinical oral features, allowing the phenotype to capture multiple layers of oral ageing. Third, out-of-fold prediction and residualization were used to reduce in-sample optimism and define an age-independent OralAgeAccel index. Fourth, the study examined clinical characterization, individual diseases, disease-domain burden, disease-count outcomes, and composite health scores within one analytic framework. This design provides a broad epidemiological characterization of oral health age acceleration.

Several limitations should also be noted. The cross-sectional design prevents causal inference and temporal ordering. OralAgeAccel may precede systemic disease, follow systemic disease, or reflect shared upstream determinants. The oral microbiome was measured from oral rinse samples using 16S rRNA sequencing, which provides useful community-level information but does not resolve all species-, strain-, gene-, or functional-level pathways. Human microbiome-disease associations are vulnerable to host-variable confounding, and factors such as medication use, diet, smoking, socioeconomic conditions, oral hygiene, and comorbidities may influence microbial profiles.[61] The near-zero age correlation is a direct consequence of within-sample residualization, and model performance may remain optimistic after feature-set comparison in the same dataset. Clinical gradients across periodontitis and tooth loss represent concordance with model components rather than independent validation. Findings may not generalize to adults aged ≥ 70 years or fully edentulous populations. The unweighted spline analyses are descriptive and may be less stable at distribution tails with sparse events. Residual confounding from adiposity, detailed dietary patterns, daily oral hygiene, medication use, and access to dental care remains possible. Longitudinal cohorts with repeated microbiome and oral examinations, incident disease follow-up, and external populations are needed to establish temporal ordering and transportability. Metagenomics and metabolomics could clarify strain-level functions and short-chain fatty acid pathways, while host genomic and cell-specific expression analyses may help determine whether oral–systemic multimorbidity reflects shared susceptibility or outcome-specific biological processes, following approaches used in cross-disorder comorbidity research.[62] Candidate host or microbial pathways emerging from these analyses could subsequently be examined with precision genome-engineering tools, ranging from targeted editing to emerging genome-writing systems, to support functional validation in experimental models.[63]

CONCLUSION

This study developed OralAge as an integrated oral ageing phenotype based on microbiome-derived diversity features, periodontal measures, and tooth-status information. OralAgeAccel, derived as the age-independent residual of predicted OralAge, was higher among participants with periodontitis and greater tooth-loss burden, supporting its relevance as a marker of accumulated oral deterioration. Higher OralAgeAccel was associated with CVD, diabetes, CKD, stroke, and greater multisystem disease burden. More favorable cardiovascular and multidimensional health scores were associated with lower OralAgeAccel, whereas greater social disadvantage was associated with higher OralAgeAccel. These findings suggest that accelerated oral ageing may reflect both local oral deterioration and broader health vulnerability. Further longitudinal studies are needed to validate OralAgeAccel, clarify temporal relationships with chronic disease, and evaluate whether oral health, behavioral, and social interventions can modify oral ageing trajectories. At present, OralAge should be viewed as a research phenotype requiring independent external and longitudinal validation before clinical application.

Data availability

The data used in this study were derived from the NHANES and were handled in accordance with the relevant NHANES data-use policies. The analytic dataset and other related materials are available from the corresponding author upon reasonable request.

Abbreviations

AA, Associate of Arts; ADA, American Diabetes Association; AMPM, Automated Multiple-Pass Method; ASV, amplicon sequence variant; BMI, body mass index; CDC/AAP, Centers for Disease Control and Prevention/American Academy of Periodontology; CI, confidence interval; CKD, chronic kidney disease; CVD, cardiovascular disease; CV, cross-validation; DADA2, Divisive Amplicon Denoising Algorithm 2; eGFR, estimated glomerular filtration rate; GED, General Educational Development; HDL-C, high-density lipoprotein cholesterol; JNC 7, Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure; LC9, Life's Crucial 9; LE8, Life's Essential 8; LS7, Life's Simple 7; MAE, mean absolute error; MetS, metabolic syndrome; NHANES, National Health and Nutrition Examination Survey; OBS, oxidative balance score; OralAge, oral health age; OralAgeAccel, oral health age acceleration; OR, odds ratio; OTU, operational taxonomic unit; PCoA, principal coordinate analysis; PD, phylogenetic diversity; PHQ-9, Patient Health Questionnaire-9; PIR, poverty-to-income ratio; RCS, restricted cubic spline; RMSE, root mean square error; SD, standard deviation; SDoH, social determinants of health; SEQN, respondent sequence number; T1–T3, tertile 1 to tertile 3; UACR, urinary albumin-to-creatinine ratio.

REFERENCES

1. Peres M. A., Macpherson L. M. D., Weyant R. J., et al. (2019). Oral diseases: A global public health challenge. *Lancet* **394**:249–260. DOI:10.1016/S0140-6736(19)31146-8
2. GBD 2021 Oral Disorders Collaborators. (2025). Trends in the global, regional, and national burden of oral conditions from 1990 to 2021: A systematic analysis for the Global Burden of Disease Study 2021. *Lancet* **405**:897–910. DOI:10.1016/S0140-6736(24)02811-3
3. Baker J. L., Mark Welch J. L., Kauffman K. M., et al. (2024). The oral microbiome: Diversity, biogeography and human health. *Nat. Rev. Microbiol.* **22**:89–104. DOI:10.1038/s41579-023-00963-6
4. Lamont R. J., Koo H. and Hajishengallis G. (2018). The oral microbiota: Dynamic communities and host interactions. *Nat. Rev. Microbiol.* **16**:745–759. DOI:10.1038/s41579-018-0089-x
5. Sedghi L., DiMassa V., Harrington A., et al. (2021). The oral microbiome: Role of key organisms and complex networks in oral health and disease. *Periodontol 2000* 87:107-131. DOI:10.1111/prd.12393
6. Hajishengallis G. and Chavakis T. (2021). Local and systemic mechanisms linking periodontal disease and inflammatory comorbidities. *Nat. Rev. Immunol.* **21**:426–440. DOI:10.1038/s41577-020-00488-6
7. Sanz M., Marco Del Castillo A., Jepsen S., et al. (2020). Periodontitis and cardiovascular diseases: Consensus report. *J. Clin. Periodontol.* **47**:268–288. DOI:10.1111/jcpe.13189
8. Sanz M., Ceriello A., Buysschaert M., et al. (2018). Scientific evidence on the links between periodontal diseases and diabetes: Consensus report and guidelines of the joint workshop on periodontal diseases and diabetes by the International diabetes Federation and the European Federation of Periodontology. *Diabetes Res. Clin. Pract.* **137**:231–241. DOI:10.1016/j.diabres.2017.12.001
9. Kapellas K., Singh A., Bertotti M., et al. (2019). Periodontal and chronic kidney disease association: A systematic review and meta-analysis. *Nephrology (Carlton)* **24**:202–212. DOI:10.1111/nep.13225
10. Zemedikun D. T., Chandan J. S., Raindi D., et al. (2021). Burden of chronic diseases associated with periodontal diseases: A retrospective cohort study using UK primary care data. *BMJ Open* **11**:e048296. DOI:10.1136/bmjopen-2020-048296
11. Beukers N., Su N., Loos B. G., et al. (2021). Lower number of teeth is related to higher risks for ACVD and death-systematic review and meta-analyses of survival data. *Front. Cardiovasc Med.* **8**:621626. DOI:10.3389/fcvm.2021.621626
12. Huang S., Haiminen N., Carrieri A. P., et al. (2020). Human skin, oral, and gut microbiomes predict chronological age. *mSystems* 5. DOI:10.1128/mSystems.00630-19
13. DeClercq V., Wright R. J., Nearing J. T., et al. (2024). Oral microbial signatures associated with age and frailty in Canadian adults. *Sci. Rep.* **14**:9685. DOI:10.1038/s41598-024-60409-8
14. Rutledge J., Oh H. and Wyss-Coray T. (2022). Measuring biological age using omics data. *Nat. Rev. Genet.* **23**:715–727. DOI:10.1038/s41576-022-00511-7
15. Curtin L. R., Mohadjer L. K., Dohrmann S. M., et al. (2013). National health and nutrition examination survey: Sample design, 2007-2010. *Vital. Health Stat.* **2**:1–23
16. Johnson C. L., Dohrmann S. M., Burt V. L., et al. (2014). National health and nutrition examination survey: Sample design, 2011-2014. *Vital Health Stat.* **2**:1–33
17. Chaturvedi A. K., Vogtmann E., Shi J., et al. (2025). Oral Microbiome Profile of the US Population. *JAMA Netw Open* **8**:e258283. DOI:10.1001/jamanetworkopen.2025.8283
18. Caporaso J. G., Lauber C. L., Walters W. A., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**:1621–1624. DOI:10.1038/ismej.2012.8
19. Callahan B. J., McMurdie P. J., Rosen M. J., et al. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**:581–583. DOI:10.1038/nmeth.3869
20. Lozupone C. and Knight R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**:8228–8235. DOI:10.1128/AEM.71.12.8228-8235.2005
21. You Y., Cui Y., Zheng K., et al. (2026). Oral microbiome diversity shapes the association between leisure-time physical activity and cognitive function among older adults. *iScience* **29**:115086. DOI:10.1016/j.isci.2026.115086
22. Zhao J. J., Hu M., Li S., et al. (2026). Oral microbiome signatures predict biological age and host health. *Nat. Commun.* 17. DOI:10.1038/s41467-026-72096-2
23. Eke P. I., Dye B. A., Wei L., et al. (2012). Prevalence of periodontitis in adults in the United States: 2009 and 2010. *J. Dent. Res.* **91**:914–920. DOI:10.1177/0022034512457373
24. Eke P. I., Page R. C., Wei L., et al. (2012). Update of the case definitions for population-based surveillance of periodontitis. *J. Periodontol.* **83**:1449–1454. DOI:10.1902/jop.2012.110664
25. Friedman J., Hastie T. and Tibshirani R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**:1–22
26. Horvath S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol.* **14**:R115. DOI:10.1186/gb-2013-14-10-r115
27. American Diabetes A. (2014). Standards of medical care in diabetes--2014. *Diabetes Care* 37 Suppl 1:S14-80. DOI:10.2337/dc14-S014
28. Chobanian A. V., Bakris G. L., Black H. R., et al. (2003). The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: The JNC 7 report. *JAMA* **289**:2560–2572. DOI:10.1001/jama.289.19.2560
29. Levey A. S., Eckardt K. U., Tsukamoto Y., et al. (2005). Definition and classification of chronic kidney disease: A position statement from Kidney Disease: Improving Global Outcomes (KDIGO). *Kidney Int* **67**:2089–2100. DOI:10.1111/j.1523-1755.2005.00365.x
30. Alberti K. G., Eckel R. H., Grundy S. M., et al. (2009). Harmonizing the metabolic syndrome: A joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* **120**:1640–1645. DOI:10.1161/CIRCULATIONAHA.109.192644
31. Kroenke K., Spitzer R. L. and Williams J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *J. Gen. Intern. Med.* **16**:606–613. DOI:10.1046/j.1525-1497.2001.016009606.x
32. Searle S. D., Mitnitski A., Gahbauer E. A., et al. (2008). A standard procedure for creating a frailty index. *BMC Geriatr* **8**:24. DOI:10.1186/1471-2318-8-24
33. Lloyd-Jones D. M., Hong Y., Labarthe D., et al. (2010). Defining and setting national goals for cardiovascular health promotion and disease reduction: The American Heart Association's strategic Impact Goal through 2020 and beyond. *Circulation* **121**:586–613. DOI:10.1161/CIRCULATIONAHA.109.192703
34. Lloyd-Jones D. M., Allen N. B., Anderson C. A. M., et al. (2022). Life's Essential 8: Updating and Enhancing the American Heart Association's Construct of Cardiovascular Health: A Presidential Advisory From the American Heart Association. *Circulation* **146**:e18–e43. DOI:10.1161/CIR.0000000000001078
35. You Y., Zheng K., Ablitip A., et al. (2025). Life's Essential 8 and Depression: A National Cross-Sectional Analysis in US Emerging Adults. *J. Adolesc Health* **77**:84–93. DOI:10.1016/j.jadohealth.2025.03.023
36. Ge J., Peng W. and Lu J. (2024). Predictive Value of Life's Crucial 9 for Cardiovascular and All-Cause Mortality: A Prospective Cohort Study From the NHANES 2007 to 2018. *J. Am. Heart Assoc.* **13**:e036669. DOI:10.1161/JAHA.124.036669
37. Liu X., Liu X., Wang Y., et al. (2023). Association between depression and oxidative balance score: National Health and Nutrition Examination Survey (NHANES) 2005-2018. *J. Affect Disord* **337**:57–65. DOI:10.1016/j.jad.2023.05.071
38. Palacio A., Mansi R., Seo D., et al. (2020). Social determinants of health score: Does it help identify those at higher cardiovascular risk. *Am J. Manag. Care* **26**:e312–e318. DOI:10.37765/ajmc.2020.88504
39. You Y., Li J., Zhang Y., et al. (2026). Exploring the role of beta2-microglobulin in the relationship between physical activity and DNAm-predicted PhenoAge: Evidence from a population-based and mice single-cell RNA-sequencing study. *J. Adv. Res.* **86**:743–755. DOI:10.1016/j.jare.2025.11.047
40. You Y. (2024). Accelerometer-measured physical activity and sedentary behaviour are associated with C-reactive protein in US adults who get insufficient sleep: A threshold and isotemporal substitution effect analysis. *J. Sports Sci.* **42**:527–536. DOI:10.1080/02640414.2024.2348906
41. Desquilbet L. and Mariotti F. (2010). Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med.* **29**:1037–1057. DOI:10.1002/sim.3841
42. You Y., Cui Y., Chen Y., et al. (2026). Biological age in the meta omics era: Integrating physical activity, sleep, and diet to reshape paths of healthy ageing. *iMetaOmics* 3. DOI:10.1002/imo2.70127
43. Levine M. E., Lu A. T., Quach A., et al. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* **10**:573–591. DOI:10.18632/aging.101414
44. Belsky D. W., Caspi A., Corcoran D. L., et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* 11. DOI:10.7554/eLife.73420
45. Kinane D. F., Stathopoulou P. G. and Papapanou P. N. (2017). Periodontal diseases. *Nat. Rev. Dis. Primers* **3**:17038. DOI:10.1038/nrdp.2017.38
46. Papapanou P. N., Sanz M., Buduneli N., 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. Periodontol* 89 Suppl 1:S173-S182. DOI:10.1002/JPER.17-0721
47. Tonetti M. S., Bottenberg P., Conrads G., et al. (2017). Dental caries and periodontal diseases in the ageing population: call to action to protect and enhance oral health and well-being as an essential component of healthy ageing - Consensus report of group 4 of the joint EFP/ORCA workshop on the boundaries between caries and periodontal diseases. *J. Clin. Periodontol* 44 Suppl 18:S135-S144. DOI:10.1111/jcpe.12681
48. Li Y., Zhu M., Liu Y., et al. (2022). The oral microbiota and cardiometabolic health: A comprehensive review and emerging insights. *Front. Immunol* **13**:1010368. DOI:10.3389/fimmu.2022.1010368
49. Kitamoto S., Nagao-Kitamoto H., Hein R., et al. (2020). The Bacterial Connection between the Oral Cavity and the Gut Diseases. *J. Dent Res.* **99**:1021–1029. DOI:10.1177/0022034520924633
50. Simpson T. C., Clarkson J. E., Worthington H. V., et al. (2022). Treatment of periodontitis for glycaemic control in people with diabetes mellitus. *Cochrane Database Syst. Rev.* **4**:CD004714. DOI:10.1002/14651858.CD004714.pub4
51. Sharma P., Dietrich T., Ferro C. J., et al. (2016). Association between periodontitis and mortality in stages 3-5 chronic kidney disease: NHANES III and linked mortality study. *J. Clin. Periodontol* **43**:104–113. DOI:10.1111/jcpe.12502
52. Fagundes N. C. F., Almeida A., Vilhena K. F. B., et al. (2019). Periodontitis as a risk

factor for stroke: A systematic review and meta-analysis. *Vasc Health Risk Manag* **15**:519–532. DOI:10.2147/VHRM.S204097
53. Huang W., Huang J., Kuang N., et al. (2026). Mitochondrial homeostasis dysregulation: Potential mechanisms of Alzheimer's disease mediated by TDP-43. *Ageing Res. Rev.* **121**:103290. DOI:10.1016/j.arr.2026.103290
54. Tanaka T., Takahashi K., Hirano H., et al. (2018). Oral frailty as a risk factor for physical frailty and mortality in community-dwelling elderly. *J. Gerontol A Biol. Sci. Med. Sci.* **73**:1661–1667. DOI:10.1093/gerona/glx225
55. Zhu S. R., Wei L. Y., Jia K., et al. (2024). Prevalence and unfavourable outcome of oral frailty in older adult: A systematic review and meta-analysis. *Front. Public Health* **12**:1501793. DOI:10.3389/fpubh.2024.1501793
56. You Y., Chen Y., Ding H., et al. (2025). Relationship between physical activity and DNA methylation-predicted epigenetic clocks. *NPJ Aging* **11**:27. DOI:10.1038/s41514-025-00217-0
57. You Y., Wang D., Ding H., et al. (2025). Mediation role of telomere length in the relationship between physical activity and PhenoAge: A population-based study. *J. Exerc Sci. Fit.* **23**:149–156. DOI:10.1016/j.jesf.2025.03.004
58. Chen X., Sun J., Zeng C., et al. (2024). Association between life's essential 8 and periodontitis: A population-based study. *BMC Oral Health* **24**:19. DOI:10.1186/s12903-023-03816-z
59. Hou K., Zhang H., Song W., et al. (2024). Association between life's essential 8 and periodontitis: A study based on NHANES 2009-2014. *Front. Med. (Lausanne)* **11**:1342792. DOI:10.3389/fmed.2024.1342792
60. Obeidat R., Heaton L. J., Tranby E. P., et al. (2024). Social determinants of health linked with oral health in a representative sample of U. *S. adults. BMC Oral Health* **24**:1518. DOI:10.1186/s12903-024-05257-8
61. Vujkovic-Cvijin I., Sklar J., Jiang L., et al. (2020). Host variables confound gut microbiota studies of human disease. *Nature* **587**:448–454. DOI:10.1038/s41586-020-2881-9
62. Kuang N., Liu Z., Yu G., et al. (2023). Neurodevelopmental risk and adaptation as a model for comorbidity among internalizing and externalizing disorders: Genomics and cell-specific expression enriched morphometric study. *BMC Med.* **21**:291. DOI:10.1186/s12916-023-02920-9
63. Huang K., Tian J., Zhao W., et al. (2026). Strategies and mechanisms of precision genome engineering: From gene editing to genome writing. *iMetaOmics* 3. DOI:10.1002/imo2.70115.

FUNDING AND ACKNOWLEDGMENTS

This study is supported by Guangzhou Concord Medical Humanities Research and Education Fund (23000-3050070) and Tsinghua University Computational Social Science Research Program (202501020008). The author would like to thank Yicong Cui, Lecturer at the School of Sports Science and Engineering, East China University of Science and Technology, for assistance in refining Figure 1. The author also thanks the editorial office of The Innovation Insights for guidance on the design of the graphical abstract and assistance with image formatting. The funders had no role instudy design, data collection and analysis, decision to publish, or preparation of themanuscript.

AUTHOR CONTRIBUTIONS

Y.Y.: Conceptualization, Data curation, Formal analysis, Visualization, Validation, Writing - original draft, Writing - review and editing.

DECLARATION OF INTERESTS

The author declares no competing interests.

ETHICAL STATEMENT AND PATIENT CONSENT

All NHANES protocols were approved by the National Center for Health Statistics Research Ethics Review Board, and all participants provided written informed consent. This study used publicly available de-identified data and did not require additional institutional review board approval.

DATA AND CODE AVAILABILITY

The data used in this study were derived from the NHANES and were handled in accordance with the relevant NHANES data-use policies. The analytic dataset and other related materials are available from the corresponding author upon reasonable request.

SUPPLEMENTAL INFORMATION

It can be found online at https://doi.org/10.59717/j.tiis.2026.100019.