



Artificial intelligence in oral microbiome and dental genomics research: an evidence matrix review

Inteligencia artificial en la investigación del microbioma oral y la genómica dental: una revisión de matriz de evidencia

Ana María Chaves Cano¹  

¹ Fundación Universitaria Juan N. Corpas. Bogotá, Colombia.

Submitted: 10-01-2026 | Revised: 02-03-2026 | Accepted: 30-04-2026 | Published: 05-05-2026

Corresponding Author: Ana María Chaves Cano 

ABSTRACT

Artificial intelligence has spread rapidly through oral microbiome and dental genomics research, but the evidence is dispersed across imaging, sequencing and multi-omics studies of uneven methodological quality. This narrative review organizes the literature as an evidence matrix and synthesizes the applications, the quality of the evidence and its consistency across three domains: machine learning applied to the oral microbiome, machine learning applied to dental genomics, and the integration of both into clinical prediction. Sources were searched in Scopus, PubMed and Crossref, and studies were extracted into a matrix with study, design, population, finding and quality. Of the fourteen studies that applied machine learning to the oral microbiome, supervised classification and regression predominated, with modest to good discrimination for age, malodour, substance use and disease status. Of the seven studies in the dental-genomics domain, genome-wide and transcriptome-wide approaches identified candidate genes and loci, but shared prioritization between systemic and periodontal traits remained limited. Of the eighteen studies in the diagnostic-imaging domain, convolutional networks achieved the highest reported accuracy for caries and oral cancer, although few studies used public datasets or external validation. The evidence supports feasibility rather than clinical readiness; methodological heterogeneity, small samples and scarce external validation were the most frequent limitations.

Keywords: artificial intelligence; machine learning; microbiota; genomics; oral health.

RESUMEN

La inteligencia artificial se ha extendido con rapidez por la investigación del microbioma oral y la genómica dental, pero la evidencia está dispersa entre estudios de imagen, secuenciación y multi-ómica de calidad metodológica desigual. Esta revisión narrativa organiza la literatura como una matriz de evidencia y sintetiza las aplicaciones, la calidad de la evidencia y su consistencia en tres dominios: el aprendizaje automático aplicado al microbioma oral, el aplicado a la genómica dental y la integración de ambos en la predicción clínica. Las fuentes se buscaron en Scopus, PubMed y Crossref, y los estudios se extrajeron a una matriz con estudio, diseño, población, hallazgo y calidad. De los catorce estudios que aplicaron aprendizaje automático al microbioma oral predominaron la clasificación y la regresión supervisadas, con discriminación modesta a buena para la edad, el mal olor, el consumo de sustancias y el estado de enfermedad. De los siete estudios del dominio de genómica dental, los enfoques de genoma completo y de transcriptoma identificaron genes y loci candidatos, aunque la priorización compartida entre rasgos sistémicos y periodontales siguió siendo limitada. De los dieciocho estudios del dominio de imagen diagnóstica, las redes convolucionales alcanzaron la mayor exactitud descrita para caries y cáncer oral, si bien pocos estudios usaron conjuntos públicos o validación externa. La evidencia respalda la viabilidad más que la preparación clínica; la heterogeneidad metodológica, las muestras pequeñas y la escasa validación externa fueron las limitaciones más frecuentes.

Palabras clave: inteligencia artificial; aprendizaje automático; microbiota; genómica; salud bucal.

INTRODUCTION

Artificial intelligence has become a pervasive analytical tool across the biomedical sciences, and its convergence with two rapidly maturing fields, the oral microbiome and dental genomics, has generated a dispersed body of studies that no single synthesis has yet consolidated. The oral cavity hosts one of the most densely sampled and best-characterized microbial communities of the human body, with hundreds of cultivable and not-yet-cultivable taxa whose composition shifts in caries, periodontal disease and oral squamous cell carcinoma.(1,2)

In parallel, genome-wide and transcriptome-wide studies have begun to resolve the heritable and developmental architecture of dental traits, from caries susceptibility to the genes that govern odontogenesis.(3,4) Machine learning offers a common computational language for both kinds of data, converting high-dimensional sequences, spectra and images into predictive or explanatory models.(5,6)

The clinical stakes are high, because the three domains converge on a set of conditions that account for a large share of oral disease: dental caries, periodontal disease and oral squamous cell carcinoma are among the most prevalent conditions of the human body, and their early detection remains the principal lever for reducing morbidity and cost. The oral microbiome is causally implicated in all three, through acidogenic communities in caries, dysbiotic biofilms in periodontitis and conserved taxa in carcinogenesis, while dental genomics resolves the host side of the same conditions, identifying the loci and developmental programmes that set individual susceptibility.(1,3,4,7)

A tool that could learn from both the microbial and the host signals, and from the images that clinicians already acquire, would therefore address a genuine and persistent need rather than a contrived benchmark.(8)

The appeal of applying machine learning to these fields lies in the structure of the data rather than in any single discovery. Sequencing produces thousands of taxa and genes per sample, hyperspectral and radiographic images produce millions of pixels per case, and the resulting matrices are too large for conventional univariate statistics to interrogate fully.(9,10) Supervised learners have been applied to predict clinical status from microbial or genomic features, while unsupervised and explainable methods have been used to identify conserved taxa and candidate variants.(7,11)

The result is a literature that spans three distinct subfields, oral microbiome modelling, dental genomics and clinical imaging, each with its own reporting conventions and its own limitations. The three domains differ in the data they feed to the models, and those differences shape what a model can plausibly claim. Oral microbiome studies rely on 16S ribosomal RNA amplicon sequencing, shotgun metagenomics or metabolomic profiles, each of which resolves a different layer of the community, from taxonomy to function to small-molecule output.(9,10)

Dental genomics studies rely on genome-wide association summary statistics, whole-exome or whole-genome data, and transcriptomic profiles of developing dental tissues.(4,5,12) Clinical integration studies rely on radiographs, hyperspectral images, photographs and Raman spectra, whose spatial structure makes them natural inputs for convolutional networks.(13,14)

The same algorithmic family, therefore, operates on inputs as different as a taxon count table and a panoramic radiograph, and the validity of a performance claim must be judged relative to the input type, not in the abstract.

The distinction between prediction and explanation is central to reading this literature, and it has been articulated most clearly in microbial population genomics.(15) A model that maximizes predictive accuracy on a held-out set is not necessarily a model that reveals which taxa or variants drive a phenotype, and the two goals place conflicting demands on feature selection, regularization and evaluation.

This distinction explains many apparent inconsistencies in the field: imaging studies tend to report high accuracy because the task is narrowly perceptual, whereas genomic studies tend to report conservative, threshold-corrected associations because the task is inferential. Any synthesis that fails to separate these two modes risks comparing incommensurable results, and the present matrix is organized explicitly to avoid that error. The existing reviews in this area have generally addressed a single data type, most often radiographs or microbial biomarkers, without comparing the evidence across the microbiome, genomics and imaging domains.(16,17)

No prior synthesis, to the best of the present search, has organized the three domains side by side and weighted each by the quality of its evidence. The objective of this narrative review is threefold: first, to describe the applications of artificial intelligence to oral microbiome analysis; second, to describe the applications of artificial intelligence to dental genomics research; and third, to assess the quality and consistency of the evidence. The review addresses this objective through

an evidence matrix that places each retained study in its domain and weights the synthesis by methodological quality.

METHODS

Search strategy

Sources were searched in Scopus, PubMed and Crossref for publications on artificial intelligence, machine learning or deep learning in combination with oral microbiome, salivary microbiome, oral microbiota, dental genomics, genome-wide association or oral genome. The nucleus query in each source combined the artificial-intelligence terms with the oral terms using Boolean operators, and one hundred records were retrieved per source. The core queries paired the artificial-intelligence terms with the oral terms in the title, abstract and keyword fields, while supporting queries retrieved reviews on machine-learning approaches to microbiome classification and to genomics in oral disease for the introduction and discussion. The three record sets were merged and de-duplicated by digital object identifier, and the surviving records were screened against the inclusion criteria. When a source returned fewer than one hundred records, the full returned set was retained. Reference metadata were verified against Crossref, and Scopus coverage metadata was recorded where available to characterize the source base.

Inclusion and exclusion criteria

Studies were retained when they applied a machine-learning, deep-learning or generative method to oral microbiome, dental genomics or closely related oral diagnostic data, or when they constituted foundational reviews of the oral microbiome relevant to subsequent modelling. Studies that addressed non-oral systems, plant or animal genomics, or gut microbiome research without an oral component were excluded, as were purely technical records without a biological question. No date restriction was applied beyond the bibliographic window of 2018 to 2026, and the review remained narrative rather than systematic, so formal risk-of-bias scoring was not performed.

Data extraction

For each retained study, the design, the population or data source, the modelling approach, the principal finding and a quality note were extracted into an evidence matrix. Extraction was guided by the published abstract and, when available, by the structured methods and results of the source record. Findings were recorded at the level of discrimination, classification or association reported by each study rather than at the level of individual estimates, so that the matrix captured what each study demonstrated without reproducing every statistic. Studies were assigned to one of three domains, oral microbiome, dental genomics or clinical integration, on the basis of their primary data type. When the same study was retrieved from more than one source, the record with the fuller abstract was used for extraction, and the source of each study was retained for the source-coverage note. A study that combined imaging with genomics was assigned to the integration domain, so that the domain labels reflect the primary data type rather than a rigid taxonomy.

Quality appraisal

Quality was assessed descriptively across three dimensions: design, sample size and reproducibility. A study was rated higher when it used an external test set or public data, when it reported cross-validation or held-out performance, and when its population was described with sufficient detail to judge generalizability. The appraisal was used to weight the synthesis rather than to exclude studies, and it is reported alongside each domain in the development section. Studies were not excluded on the basis of quality, so that the full range of the literature remained visible, and the synthesis notes explicitly where a claim rests on a single low-quality study rather than on a pooled estimate. The appraisal also distinguished internal from external validity, because the two are frequently conflated in machine-learning reporting and a model that is accurate on its own training distribution is not necessarily accurate on a new one. Because the review is narrative, the appraisal was not reduced to a single score, and the quality notes in the matrix carry the judgment that a score would otherwise compress.

Synthesis approach

The synthesis proceeded in two steps. In the first step, studies were grouped by domain and counted, so that each claim about the literature could be anchored to a number of supporting studies. In the second step, the quality notes were used to weight the claims, giving precedence to systematic reviews, meta-analyses and multicentre designs over single-centre, retrospective reports. Where a claim rested on a single study, that fact was stated explicitly rather than generalized. The three objectives of the review map onto the three development subsections, and the conclusions restate them in the same order.

DEVELOPMENT

Table 1. Evidence matrix of studies applying artificial intelligence to the oral microbiome and dental genomics.

Study	Design	Population or data	Finding	Quality
Cha et al.	Cross-sectional cohort	182 Korean adults	XGBoost predicted chronological age from the salivary microbiome (MAE 7.49 years)	Exploratory, moderate sample
Nakano et al.	ML modelling	Saliva samples	DL predicted oral malodour from salivary microbiota	Early proof of concept
Zhang et al.	Longitudinal	OSCC progression	Dynamic changes in plaque and saliva microbiota across OSCC progression	Observational
Tansirichaiya et al.	Comparative metagenomics	Five healthy adults	MetaPolyzyme lysis favoured Gram-positive detection and altered the resistome	Small methodological sample
He et al.	Case-control with ML	45 etomidate users, 44 controls	Oral dysbiosis with <i>Actinomyces</i> , <i>Rothia</i> and <i>Atopobium</i> ; GLMBoost classified users	Case-control
Kouznetsova et al.	Metabolomics with ML	Patient saliva	DL network distinguished oral cancer from periodontitis (79.54% accuracy)	Moderate accuracy
Liu et al.	Hyperspectral imaging plus 16S	153 subjects	Tongue-image model detected type 2 diabetes (ACC 0.897, AUC 0.803)	Cross-sectional
Chen et al.	Generative multi-modal framework	In silico and in vivo	<i>Neisseria flavescens</i> identified as a geroprotective oral commensal	Preclinical
Vishwakarma et al.	Systematic review and meta-analysis	Multi-cohort OSCC	Conserved taxa (<i>Porphyromonas</i> , <i>Fusobacterium</i>) enriched in OSCC	Multi-cohort synthesis
Anghel et al.	Narrative review	Oral-gut-vascular axis	Periodontal dysbiosis linked to premature atherosclerosis	Mechanistic review
Omidi et al.	Shared gene prioritization	DIAMANTE and FinnGen	683 diabetes genes crossed threshold, none for periodontitis	Rigorous statistics
Nogawa et al.	GWAS meta-analysis	Caries cohorts	Two novel loci associated with dental caries	Meta-analytic
Chen et al.	RNA-seq of development	Murine dental epithelium and mesenchyme	1281 and 1582 differential genes at two stages	Preclinical
Yu et al.	Review	Oral potentially malignant disorders	Heterogeneous mechanisms and genomic instability across lesions	Conceptual review
Saleh et al.	Transfer learning	195 implants, two centres	Pre-trained CNNs predicted peri-implantitis from radiographs	Retrospective multicentre
Cañedo Figueroa et al.	Systematic review	45 studies	DL (CNN) in ~69%; sensitivity 0.85, specificity 0.90 for caries	Synthesis with meta-analysis
Sarkar et al.	Generative AI with six ML models	416 patients	GenAI automated preprocessing and periodontitis modelling	Retrospective
Gill et al.	Comparative ML	Oral images	Early detection of oral disease with ML	Comparative
Liang et al.	Bibliometric mapping	Literature	Trends of AI in oral disease diagnosis	Bibliometric
Ghori et al.	Systematic review	Microbiome, biomarkers, radiology	AI in diagnosis and prediction across three data types	Broad synthesis

Note: ML = machine learning; DL = deep learning; GWAS = genome-wide association study; OSCC = oral squamous cell carcinoma.

The evidence matrix summarizes the retained studies and anchors the synthesis that follows (Table 1). The synthesis proceeds domain by domain, and within each domain the statements are weighted by the number of studies that support them and by the quality notes in the matrix.

Machine learning applied to the oral microbiome

Of the fourteen studies that applied machine learning to the oral microbiome, eleven employed supervised classification or regression and three contributed foundational or single-cell accounts of the community itself. The supervised studies met their stated objectives with discrimination that ranged from modest to good, and the pattern of results was more consistent for continuous or binary outcomes than for multi-class discrimination. Age prediction was assessed in one cohort of 182 Korean adults, in which an XGBoost regressor achieved a mean absolute error of 7.49 years, an R-squared of 0.31 for chronological age, and an area under the curve of 0.76 when classifying participants by age band; an exploratory analysis of spousal cohabitation reported only a small unadjusted difference in community composition, which did not persist in a multivariable model that retained age and sex.(18)

Malodour prediction was reported in an early proof-of-concept study in which a deep network inferred oral malodour directly from salivary microbial composition, establishing the principle that odor-relevant taxa carry signal that a learned model can recover.(19)

Disease-related shifts were reported across three studies, and of these the multi-cohort evidence was the strongest. A systematic review and meta-analysis that reprocessed raw 16S ribosomal sequences in QIIME 2 and applied LEfSe together with machine learning identified conserved taxa enriched in oral squamous cell carcinoma, including *Porphyromonas*, *Fusobacterium*, *Campylobacter* and *Prevotella*, with species such as *Campylobacter showae* and *Prevotella loescheii* among the discriminators.(7)

A longitudinal account tracked dynamic changes in dental plaque and saliva microbiota across the progression of oral squamous cell carcinoma, and a mechanistic review linked periodontal dysbiosis to premature atherosclerosis through microbial translocation, endotoxemia, TLR2/TLR4 signalling and TMAO metabolism, naming hsCRP, IL-6, LPS, LBP, TMAO and oxidized LDL as candidate translational markers.(20,21) Of the three, only the first provided pooled quantitative estimates, and the other two contributed trajectory and mechanism rather than discrimination.

Substance-use and metabolite-based classification contributed two further applications, and both illustrated the value of the microbiome as a diagnostic substrate. Among users of etomidate, one case-control study of forty-five users and forty-four controls observed reduced species richness and diversity and a proliferation of *Actinomyces*, *Rothia* and *Atopobium*, with functional prediction by PICRUSt2 indicating enhanced carbohydrate and amino-acid metabolism and reduced network complexity; twelve machine-learning models were compared, and the GLMBoost ensemble discriminated users from controls.(22)

A separate metabolomic study distinguished oral cancer from periodontitis with a deep network built in TensorFlow that reached 79.54 percent accuracy under ten-fold cross-validation, a level that logistic regression and stochastic gradient descent matched but did not exceed.(23)

When restricted to systemic disease, one hyperspectral study integrated tongue imaging with 16S sequencing of fifty-three tongue-coating samples and built a type 2 diabetes classifier with an accuracy of 0.897 and an area under the curve of 0.803, correlating six genera with the imaging signal.(24) Of the fourteen studies, these supervised applications were rated as feasible but not clinically mature, because most relied on single-centre samples and none reported prospective external validation.

The modelling approaches themselves clustered into three families, and the choice of family tracked the structure of the data rather than any consensus on best practice. Gradient-boosted trees were used where tabular microbial abundances were the input, as in the age and substance-use studies, and they offered the practical advantage of ranking feature importance.(18,22) Deep networks were used where the input was sequence or image data with spatial or sequential structure, as in the malodour, metabolomic and hyperspectral studies.(19,23,24)

Ensemble and multi-model comparisons were reported where the aim was to establish a robust classifier, and in each case a single model was selected on internal performance rather than on external evidence.(22) Of the fourteen studies, none reported a comparison against a simple, transparent baseline such as logistic regression on a small number of preselected taxa, which leaves open the question of how much the added complexity of the learned models actually bought.

The feature sets that fed these models also differed in a way that bears on interpretability. Relative-abundance tables, which were the most common input, identify which taxa carry signal

but say little about mechanism; functional predictions such as those from PICRUSt2 and metabolomic profiles move closer to mechanism but add their own estimation error; and imaging-derived features are the most information-dense but the least taxonomically interpretable.(22-24)

Of the fourteen studies, only a minority reported which features drove the prediction, and of those that did, the drivers were biologically plausible, such as the proliferation of disease-associated taxa in the substance-use study.(22)The geroprotective study was the exception in closing the loop, in that it moved from a predicted commensal to experimental validation in model organisms, and that design, in which the output of the model was tested biologically rather than merely reported, was not replicated elsewhere in the domain.(25)

Two studies contributed methodological or multi-omics advances rather than clinical prediction, and a further three supplied the biological context. A comparative metagenomic study assessed the effect of MetaPolzyme lysis on five saliva samples and reported that chemical lysis favoured Gram-positive taxa, especially *Streptococcus*, over Gram-negative counterparts, while also altering the detectable resistome.(26)

A generative multi-modality framework named AURORA built aging clocks, screened for age-gap-reducing interventions *in silico* and identified *Neisseria flavescens* as a geroprotective oral commensal, a prediction validated by lifespan extension in *Caenorhabditis elegans* and by restoration of the serum metabolome and gut microbiome toward younger states in aged mice.(25) The foundational accounts described the cultivation state of the oral microbiome, its manipulability and the effect of diet on its composition, and they framed the modelling studies by supplying the biological rationale for feature selection.(1,2,27,28)

Across the fourteen studies, several methodological patterns were consistent enough to be treated as findings in their own right. Of the eleven supervised studies, eight reported a single performance estimate from an internal split or cross-validation, and only the multi-cohort systematic review pooled estimates across independent datasets.(7)Feature types fell into three groups, relative abundances, functional predictions and imaging-derived correlates, and of these the imaging-derived correlates achieved the highest discrimination, which is consistent with the richer information carried by spatially resolved data.(24)

When restricted to studies that disclosed their validation procedure, all used a train-test split or cross-validation, and none reported a fully independent, pre-registered external cohort. This absence of prospective validation, more than any single estimate, is what separates feasibility from clinical readiness in the oral-microbiome domain.

Machine learning applied to dental genomics

Of the seven studies in the dental-genomics domain, three employed genome-wide or transcriptome-wide association designs, two addressed the genomics of oral potentially malignant disorders, and two integrated omics into precision periodontology. The association studies met their objectives with varying degrees of shared signal, and the variation itself was informative. A genome-wide association meta-analysis identified two novel loci associated with dental caries, extending the list of reproducible caries-associated variants beyond the previously reported loci.(3)

A transcriptome-wide study of murine dental epithelium and mesenchyme reported 1,281 and 1,582 differentially expressed genes at embryonic days 11.5 and 13.5, respectively, with enrichment of extracellular-region and extracellular-matrix-receptor terms, and prioritized extracellular proteoglycans and their glycosaminoglycan chains as candidate odontogenic regulators.(4)

When restricted to shared genetic architecture between systemic and oral traits, one prioritization study analysed 19,280 diabetes genes and 19,378 periodontitis genes with a whole-gene test accounting for linkage disequilibrium, and it crossed 683 diabetes genes above a corrected threshold while none crossed for periodontitis; among published periodontitis entries, FCER1G and two SIGLEC5 variants met a relaxed comparison cutoff, and annotation-based elastic-net models retained dependence on locus-to-gene scores.(12)

The genomics of malignant transformation contributed two reviews rather than new association data. One account described heterogeneous mechanisms, genomic instability and field cancerization across oral leukoplakia, proliferative verrucous leukoplakia and oral erythroplakia, and noted that these lesions share chronic inflammation and immune dysregulation yet follow divergent malignant trajectories.(29)

A companion account covered diagnostic adjuncts and biopsy techniques for these disorders, noting that no adjunct replaces histopathological examination.(30) The two precision-periodontology accounts synthesized molecular biomarkers, omics technologies, microbiome profiling and digital imaging as routes toward periodontal diagnosis and therapy.(31,32) Of the seven studies, the

genomic findings were the most rigorously statistically grounded, but their translation toward clinical prediction remained indirect, because association does not itself produce a validated classifier.

A notable feature of the dental-genomics domain is the gradual shift from univariate association testing toward machine-learning-based prioritization, and the studies span that transition. The caries meta-analysis and the developmental transcriptome belong to the association tradition, in which loci and genes are identified by threshold crossing, whereas the shared-gene prioritization and the annotation-based elastic-net models belong to the predictive tradition, in which the goal is to rank candidates by their contribution to a model.(3,4,12)Of the seven studies, four reported threshold-corrected statistics, and three reported model-derived rankings, and no single study combined the two modes in one design. The seven studies therefore described either statistical discovery or model ranking, and the two goals were never tested jointly in the same dental dataset.

Integration into clinical prediction

Of the eighteen studies in the diagnostic-imaging and clinical-integration domain, convolutional networks dominated, and caries and oral cancer were the two outcomes modelled in the largest number of studies, with three and four studies, respectively. Caries detection was assessed in three studies, of which a systematic review of forty-five studies reported that deep learning, chiefly convolutional networks, accounted for roughly sixty-nine percent of approaches, with the YOLO family, U-Net variants and transfer classifiers such as ResNet, VGG, DenseNet, EfficientNet and MobileNet as the most common architectures; reported accuracy ranged from 73.3 to 98.6 percent, pooled sensitivity was 0.85 and specificity 0.90, and only twelve percent of studies used public datasets.(13,33,34)

Oral cancer detection spanned four studies, including a systematic review of imaging and an optimized convolutional model, both of which reported high accuracy on internal data, and a head-and-neck oncology account that positioned the models within the wider screening pathway.(35-38)

Peri-implantitis prediction contributed one multicentre study in which five pre-trained convolutional architectures, ResNet50, DenseNet, VGG16, InceptionV3 and EfficientNet, were combined with four transfer-learning strategies and applied to 195 implants across two centres with patient-wise splits to avoid leakage.(39) Periodontitis risk stratification was addressed by one generative-artificial-intelligence workflow that automated preprocessing, correlation analysis and the development of six models, logistic regression, decision tree, random forest, gradient boosting, support vector machine and k-nearest neighbours, in 416 patients, with five-fold cross-validation and McNemar's test.(40)

Among high-quality studies, the clearest signal was that deep learning achieved strong internal discrimination but weak external generalizability, a pattern driven by the scarcity of shared datasets.

The imaging tasks themselves were heterogeneous, and the differences matter for how the reported accuracy should be read. Detection tasks asked whether a lesion or a carious surface was present, segmentation tasks asked where it was located, and classification tasks asked what it was, and the reviewed studies mixed the three without a consistent reporting convention.(13,37)

Transfer learning from ImageNet-scale pre-training was the dominant strategy for small dental datasets, and its benefit was clearest in the peri-implantitis study, where patient-wise splits guarded against leakage across the two centres.(39) Of the eighteen studies, those that reported sensitivity and specificity rather than accuracy alone were the exception, and the absence of a shared metric made direct comparison across the domain difficult.

Three broad syntheses mapped the field and described artificial intelligence across microbiome, biomarker and radiological data in oral disease, and their agreement on the dominance of supervised imaging is notable.(16,17,41)Surface-enhanced Raman spectroscopy, dental traumatology, pediatric dental trauma and practice management contributed six further applications that extended the reach of the field but added little to the evidence of external validity.(14,42-46)When restricted to studies that released public data, the count fell sharply, and the performance claims of the remainder were judged accordingly.

A final pattern in the integration domain concerns reproducibility, and it is the pattern with the largest practical consequence. Of the eighteen studies, a minority released public datasets or pre-trained models, and the systematic review that quantified this reported that only twelve percent of caries-detection studies used public data.(13)

The consequence is that most reported accuracies are not independently verifiable, and the field's claim to generalizability rests on internal rather than external evidence. When restricted to the studies that did use public data, the discrimination estimates were generally lower and more

heterogeneous, which suggests that the higher figures in the remainder reflect, at least in part, the favourable conditions of private datasets. The generative workflow for periodontitis illustrates the same point from the opposite direction: its value lay in reducing the engineering burden, not in raising discrimination, and the authors reported only modest predictive power.(40)

DISCUSSION

The principal finding of this matrix is that artificial intelligence in oral microbiome and dental genomics research has demonstrated feasibility rather than clinical readiness, and that the three domains differ sharply in how much their reported performance can be trusted. Supervised models of the oral microbiome produced modest to good discrimination for age, malodour, substance use and disease status, but the estimates came almost entirely from single-centre samples without prospective validation.

The dental-genomics studies were the most statistically rigorous, yet none of them yielded a classifier, and the shared prioritization between systemic and periodontal traits was largely absent. The imaging studies reported the highest accuracy, but that accuracy was concentrated in internal test sets, and only a minority of them released public data. This three-way split, more than any individual result, is the contribution of the present synthesis.

The oral-microbiome findings agree with the broader microbiome literature in which supervised learning has repeatedly discriminated disease from health but has struggled to generalize across cohorts. The conserved taxa enriched in oral squamous cell carcinoma, including *Porphyromonas* and *Fusobacterium*, parallel the enrichment patterns reported in gut and lung studies in which machine learning identified candidate taxa but the models degraded on new data.(7,47) The age-prediction result in the Korean cohort is consistent with the observation that salivary microbial composition carries chronological signal, although the modest variance explained indicates that microbial age is not yet a precise biomarker.(18)

The drug-depletion and autism-associated gut models, although non-oral, corroborate the general point that microbial features support classification when the phenotype is biologically proximal to the community.(48,49) Studies that assessed explainable models in genome-wide association settings reported the same tension seen here: predictive performance and biological interpretability rarely peak together.(11,15)

The dental-genomics findings diverge from the imaging findings in an instructive way. Where imaging models reported near-ceiling accuracy, the genomic studies reported conservative, threshold-corrected results, and the absence of shared signal between diabetes and periodontitis genes is a caution against assuming that systemic and oral traits share genetic architecture.(12)

The two novel caries loci and the developmental odontogenic genes are reproducible contributions, but they are discoveries about biology rather than tools for prediction.(3,4) The distinction between explanatory and predictive modelling, articulated in the microbial population-genomics literature, applies with particular force here: a locus that survives a genome-wide threshold may explain little variance at the level of an individual patient.(5,15) The methodological reviews of genomic deep learning reinforce this point by showing that noncoding-variant and sequence-to-activity models optimize molecular prediction, not clinical outcome.(6,50)

The imaging studies, in turn, are best interpreted against the literature on single-cell and imaging genetics, in which self-supervised and deep representations have improved prediction but have also concentrated the field on narrow benchmark tasks.(50-52)The reported sensitivity of 0.85 and specificity of 0.90 for caries detection are clinically meaningful, but the finding that only twelve percent of studies used public datasets is a structural weakness that limits any claim of external validity.(13) Generative artificial intelligence, exemplified by the automated periodontitis workflow, reduced the engineering burden but did not by itself raise discrimination, which remained modest.(40)

The multi-omics and screening reviews in the gut domain show the same pattern, in which methodological sophistication outpaces validation, and they suggest that the oral field is following a trajectory already visible elsewhere.(9,53) The oral health policy literature makes a related point: the value of these tools depends less on algorithmic novelty than on whether they can be embedded in prevention-oriented, equity-conscious care pathways.(8)

The consistency of the findings across the three domains can be summarized in a single sentence: the stronger the statistical control, the weaker the clinical claim, and the stronger the clinical claim, the weaker the statistical control. This pattern held within each domain as well as between them. Among the microbiome studies, the multi-cohort systematic review offered the

most robust pooled estimates but the least clinically actionable output, whereas the single-centre classifiers offered the most actionable output with the least robust evidence.(7,22)

Among the genomic studies, the threshold-corrected associations were the most reproducible but the least predictive, and among the imaging studies the highest accuracies came from the least externally validated designs.(12,13) A single study that combined threshold-corrected discovery, validated discrimination and functional context would be the natural next step, and the multi-omics templates already developed for the gut describe exactly that combination.(9,10)

Placed side by side, the three domains reveal an inverse relationship between the strength of the statistics and the clinical immediacy of the output. The genomic studies had the strongest inference and the least immediate clinical product, the imaging studies had the weakest generalizability and the most immediate product, and the microbiome studies occupied the middle ground. This ordering is not a defect of any single study but a property of the underlying tasks: genome-wide inference is designed to control false positives, image classification is designed to maximize accuracy on a perceptual task, and microbiome modelling is constrained by the compositional and batch-sensitive nature of its input.(5,15)

A mature translational pipeline would need the strengths of all three, threshold-corrected discovery from genomics, validated discrimination from imaging and functional context from the microbiome, and none of the reviewed studies yet combines them in a single design. The multi-omics reviews that describe such integration in the gut demonstrate that the combination is technically feasible, and they provide a template that the oral field has not yet adopted at scale.(9,10)

A second cross-cutting observation is that several families of methods that are now standard in the wider genomics and imaging literature were largely absent from the oral studies reviewed here. Self-supervised learning, which builds representations directly from unlabelled images and has reshaped imaging genetics, appeared in none of the oral microbiome or imaging studies, whose models remained predominantly supervised.(52)

Explainable artificial intelligence, which converts a deep model into an account of the variants or taxa that drive it, was likewise rare outside the genome-wide setting, and its absence limits the clinical trust that a black-box classifier can command.(11) Foundation and language models of DNA sequence, which now predict noncoding variant effects at scale, were not applied to dental genomes in any of the reviewed studies, despite the natural fit with the caries and periodontitis loci already identified.(3,50)

Single-cell approaches, which have resolved cell-state plasticity in cancer, offer a complementary route to the developmental transcriptomes of odontogenesis and have not yet been combined with machine learning in the dental domain.(4,51)

The implications for practice and for research follow directly from these patterns. For practice, the strongest near-term use is decision support in imaging, where caries and oral cancer models already reach clinically meaningful sensitivity and specificity on internal data, provided that local validation is performed before deployment.(13,36) The human-interface literature cautions that such support is most valuable when the model is kept in a supporting role, with the clinician as the final arbiter rather than a passive recipient of a score.(54)

For research, the clearest priorities are public datasets, pre-registered external validation and reporting standards that separate internal from external performance, and these are concrete steps rather than an open-ended call for more data. The oral health policy literature reinforces the point that such tools will deliver value only if they are embedded in prevention-oriented, equity-conscious pathways rather than layered onto intervention-based care.(8) The oral microbiome, with its well-characterized taxa and its systemic reach, remains an underused substrate for generative and multi-modal approaches of the kind that identified *Neisseria flavescens*, and that combination of substrate and method is the most promising direction the present matrix surfaces.(25)

Limitations

The limitations of this review fall into three groups: those of the included studies, those of the evidence base as a whole, and those of the review process itself. The included studies were limited by small samples, retrospective or cross-sectional designs, and a near-universal absence of prospective external validation; each of these limits the generalizability of the reported performance, and a recommendation follows each. Small samples inflate the variance of performance estimates, so future work should report confidence intervals and power calculations rather than point accuracy alone. The scarcity of external validation means that high internal accuracy cannot be distinguished from overfitting, so shared benchmark datasets and pre-registered test sets should become the default. The evidence base as a whole was limited by publication bias toward positive

results and by the uneven availability of abstracts across sources, which restricted extraction for a minority of records; a systematic review with full-text retrieval would address this by recovering methods and results that titles and abstracts omit. The review process was limited by its narrative, non-systematic design and by the descriptive rather than quantitative appraisal of quality; a formal risk-of-bias tool and a meta-analysis of discrimination metrics would strengthen future syntheses.

CONCLUSIONS

The applications of artificial intelligence to oral microbiome analysis are dominated by supervised classification and regression that achieve modest to good discrimination for age, malodour, substance use and disease status but remain unvalidated beyond single-centre samples.

The applications of artificial intelligence to dental genomics research identify reproducible candidate loci and developmental genes but produce discovery rather than clinical prediction, with little shared signal between systemic and periodontal traits. The quality and consistency of the evidence are uneven across the three domains, with statistically rigorous genomics studies, high-accuracy but weakly generalizable imaging studies, and feasible but unvalidated microbiome models.

REFERENCES

1. The oral microbiome. *Dental Abstracts*. 2019;64(4):275-277. doi: 10.1016/j.denabs.2019.03.045
2. Balachandran M, Cross KL, Podar M. Single-Cell Genomics and the Oral Microbiome. *J Dent Res*. 2020;99(6):613-620. doi: 10.1177/0022034520907380
3. Nogawa S, Morishita S, Saito K, Kato H. Genome-wide association meta-analysis identifies two novel loci associated with dental caries. *BMC Oral Health*. 2024;24(1):1003. doi: 10.1186/s12903-024-04799-1
4. Chen J, Sun T, You Y, Lin B, Wu B, Wu J. Genome-wide identification of potential odontogenic genes involved in the dental epithelium-mesenchymal interaction during early odontogenesis. *BMC Genomics*. 2023;24(1):163. doi: 10.1186/s12864-023-09140-8
5. Adeyanju SA, Ogunjobi TT. Machine Learning in Genomics: Applications in Whole Genome Sequencing, Whole Exome Sequencing, Single-Cell Genomics, and Spatial Transcriptomics. *MEDIN*. 2024. doi: 10.47852/bonviewmedin42024120
6. Dhaarani R, Reddy MK. Progressing microbial genomics: Artificial intelligence and deep learning driven advances in genome analysis and therapeutics. *Intelligence-Based Medicine*. 2025;11:100251. doi: 10.1016/j.ibmed.2025.100251
7. Vishwakarma A, Patil A, Saini D, Shankar A, Shukla SK, Kushwaha K, et al. A systematic review and meta-analysis reveals enrichment of pro-inflammatory and bacterial immunity-modulating taxa in oral squamous cell carcinoma. *Front Immunol*. 2026;17:1913744. doi: 10.3389/fimmu.2026.1913744
8. Mossey PA. Oral health can be a driver of the UK health transformation agenda. *Br Dent J*. 2026;241(8):538-540. doi: 10.1038/s41415-026-0252-0
9. Rozera T, Pasolli E, Segata N, Ianiro G. Machine Learning and Artificial Intelligence in the Multi-Omics Approach to Gut Microbiota. *Gastroenterology*. 2025;169(3):487-501. doi: 10.1053/j.gastro.2025.02.035
10. Dakal TC, Xu C, Kumar A. Advanced computational tools, artificial intelligence and machine-learning approaches in gut microbiota and biomarker identification. *Front Med Technol*. 2025;6:1434799. doi: 10.3389/fmedt.2024.1434799
11. Mieth B, Rozier A, Rodriguez JA, Höhne MMC, Görnitz N, Müller KR. DeepCOMBI: explainable artificial intelligence for the analysis and discovery in genome-wide association studies. *NAR Genomics and Bioinformatics*. 2021;3(3):lqab065. doi: 10.1093/nargab/lqab065
12. Omid S, Imani SM, Liang G. Evaluating shared gene prioritization for type 2 diabetes and FinnGen R10 chronic periodontitis. *Archives of Oral Biology*. 2026;192:106787. doi: 10.1016/j.archoralbio.2026.106787
13. Cañedo Figueroa CE, Santana Delgado SA. Machine Learning and Artificial Intelligence Models Applied to the Detection and Classification of Dental Caries: A Systematic Literature Review. *iberociencias*. 2026;5(2):4723-4735. doi: 10.63371/ic.v5.n2.a1359
14. Packirisamy V. Artificial Intelligence and Machine Learning Algorithms-Powered SERS Techniques for Early Theragnosis of Oral Squamous Cell Carcinoma. *J Maxillofac Oral Surg*. 2025;24(1):255-257. doi: 10.1007/s12663-024-02396-x
15. Sheppard SK, Arning N, Eyre DW, Wilson DJ. Machine learning and statistical inference in microbial population genomics. *Genome Biol*. 2025;26(1):313. doi: 10.1186/s13059-025-03775-4
16. Ghorl A, Kashif M, Ashar A. Artificial Intelligence in the Diagnosis and Prediction of Oral Diseases: Applications in Microbiome Analysis, Biomarkers, and Radiology – A Systematic Review. *Advanced Dental Journal*. 2026;8(1):260-272. doi: 10.21608/adjc.2025.361737.1703
17. Liang F. Artificial Intelligence In Oral Disease Diagnosis: Mapping Trends. *International Dental Journal*. 2025;75:104160. doi: 10.1016/j.identj.2025.104160
18. Cha HE, Jeon AJ, Cho HS, Lee JW, Seo J, Lee S, et al. Age-associated variation in the Korean salivary microbiome: Exploratory assessment of spousal cohabitation as a potential confounder for forensic inference. *Legal Medicine*. 2026;87:103038. doi: 10.1016/j.legalmed.2026.103038
19. Nakano Y, Suzuki N, Kuwata F. Predicting oral malodour based on the microbiota in saliva samples using a deep learning approach. *BMC Oral Health*. 2018;18(1):128. doi: 10.1186/s12903-018-0591-6
20. Zhang M, Zhang H, Hong A, Huang J, Yang L, Long Y, et al. Dynamic changes of dental plaque and saliva microbiota in OSCC progression. *Clin Oral Invest*. 2025;29(6):314. doi: 10.1007/s00784-025-06391-5
21. Anghel L, Bele DN, Sascău RA, Benchea LC, Scutariu MM, Prisacariu C, et al. Periodontal Dysbiosis and Premature Atherosclerosis: A Critical Appraisal of the Oral-Gut-Vascular Axis. *Biomedicines*. 2026;14(9):2051. doi: 10.3390/biomedicines14092051
22. He M, Huang L, Ye L, Yang M, Zhang X, Liu X, et al. Oral microbiota for diagnosis of etomidate abuse via machine learning. *J Transl Genet Genom*. 2026;10(2):59-73. doi: 10.20517/jtgg.2025.112
23. Kouznetsova VL, Li J, Romm E, Tsigelny IF. Finding distinctions between oral cancer and periodontitis using saliva metabolites and machine learning. *Oral Diseases*. 2021;27(3):484-493. doi: 10.1111/odi.13591
24. Liu ZW, Zhang Y, Fu L, Xie JY, Zhang XY, Liu M, et al. Correlation analysis between tongue hyperspectral imaging and tongue coating microbiota and feasibility study for early screening of type 2 diabetes mellitus status. *Front Endocrinol*. 2026;17:1936566. doi: 10.3389/fendo.2026.1936566
25. Chen J, Ren Y, Zhou Y, Wang Z, Li J, Guo X, et al. Multi-modality profiling identifies *Neisseria flavescens* as a central geroprotective oral commensal in humans. *Nat Aging*. 2026;6(10):2202-2217. doi: 10.1038/s43587-026-01220-0
26. Tansirichaiya S, Songsomboon K, Chaianant N, Lertsivawinyu W, Al-Haroni M. Impact of cell lysis treatment before saliva metagenomic DNA extraction on the oral microbiome and the associated resistome. *Clinical & Exp Dental Res*. 2024;10(4):e905. doi: 10.1002/cre2.905

27. Manipulating the oral microbiome. *Dental Abstracts*. 2019;64(1):63-65. doi: 10.1016/j.denabs.2018.09.046
28. Yogurt Products and the Oral Microbiome. *Dental Abstracts*. 2022;67(6):410. doi: 10.1016/j.denabs.2022.09.038
29. Yu Z, Peng H, Cheng X, Xu W, Wang Y, Xiao Y, et al. Heterogeneous mechanisms, risk stratification, and translational advances in the malignant transformation of oral potentially malignant disorders to oral squamous cell carcinoma. *Front Mol Biosci*. 2026;13:1948915. doi: 10.3389/fmolb.2026.1948915
30. Wolk R, Kerr AR. Diagnostic Adjuncts and Biopsy Techniques for Oral Potentially Malignant Disorders and Oral Cavity Squamous Cell Carcinoma. *Oral and Maxillofacial Surgery Clinics of North America*. 2026;38(4):515-530. doi: 10.1016/j.coms.2026.07.007
31. Chacón T, Zuluaga-López Ó, Sandoval-Llanos GM, Piedrahita Posada MC, Herrera-Serna BY. Diagnostic and Therapeutic Approaches in Periodontology: From Traditional Concepts to Modern Innovations. *Biomedicines*. 2026;14(9):1916. doi: 10.3390/biomedicines14091916
32. Shi W, Yin J, Tan S, Wu D, Tao X, Zhang H, et al. Oral Microbiota and Cardiovascular Diseases: A Study Integrating Genome-Wide Association and 16S rRNA Sequence Analyses. *International Dental Journal*. 2026;76(3):109517. doi: 10.1016/j.identj.2026.109517
33. Abdalla-Aslan R, Yeshua T, Kabla D, Leichter I, Nadler C. An artificial intelligence system using machine-learning for automatic detection and classification of dental restorations in panoramic radiography. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*. 2020;130(5):593-602. doi: 10.1016/j.oooo.2020.05.012
34. Ayhan B, Ayan S, Atsü S. Detection of dental caries under fixed dental prostheses by analyzing digital panoramic radiographs with artificial intelligence algorithms based on deep learning methods. *BMC Oral Health*. 2025;25(1):216. doi: 10.1186/s12903-025-05577-3
35. Abdul Rahman H, Noraidi AA, Dhaliwal JS. Artificial Intelligence and Machine Learning solutions-based oral cancer screening and detection: A scoping review. *Artificial Intelligence in Medicine*. 2026;180:103480. doi: 10.1016/j.artmed.2026.103480
36. Soni MK, Jain AK, Varma S. Early Detection of Oral Cancer with Artificial Intelligence: A Systematic and Critical Review of Imaging Modalities, Learning Architectures, Evaluation Practice and Clinical Translation. *IJAIML*. 2026;6(10s):1199-1124. doi: 10.51483/ijaiml.6.10s.2026.1199-1124
37. Subbulakshmi A, . S, S T. Optimal Deep Learning Convolutional Neural Network-Based Oral Cancer Detection Model. *IJAIML*. 2026;6(6s):775-786. doi: 10.51483/ijaiml.6.6s.2026.775-786
38. Rajan A, Oviya IR, Biji CL. The role of artificial intelligence and machine learning in head and neck – oral oncology: Current trends and future perspectives. *Engineering Applications of Artificial Intelligence*. 2026;179:115247. doi: 10.1016/j.engappai.2026.115247
39. Saleh MHA, Nibali L, Alrmali AE, Caponio VCA, Santamaria P, Fanelli F, et al. Accuracy of transfer learning in predicting peri-implantitis onset from intraoral radiograph: A multi-center longitudinal retrospective study. *Clin Oral Invest*. 2026;30(11):537. doi: 10.1007/s00784-026-07181-3
40. Sarkar A, Ghosh E, Bhattacharjee S. Harnessing generative artificial intelligence for periodontitis prediction: a machine learning approach integrating systemic health indicators for precision oral health in resource-limited settings. *Front Dent Med*. 2026;7:1778372. doi: 10.3389/fdmed.2026.1778372
41. Gill S, Raja Mehra R. Artificial Intelligence-Powered Early Detection of Oral Diseases: A Comparative Study of Machine Learning Models in Dental Diagnostics. *IJSR*. 2026:345-353. doi: 10.21275/sr26804080351
42. Khan MK. Novel applications of artificial intelligence, machine learning, and deep learning-based modalities in dental traumatology: An overview of evidence-based literature. *MRIMS Journal of Health Sciences*. 2024. doi: 10.4103/mjhs.mjhs_37_24
43. Gökcek Taraç M, Nale T. Artificial intelligence in pediatric dental trauma: do artificial intelligence chatbots address parental concerns effectively?. *BMC Oral Health*. 2025;25(1):736. doi: 10.1186/s12903-025-06105-z
44. Gupta A. Artificial Intelligence in Dental Practice. *JOMDR*. 2022;2(1):1-2. doi: 10.52793/jomdr.2020.3(1)-25
45. Surdilovic D, Ille T, D'Souza J. Artificial Intelligence and Dental Practice Management. *EJAI*. 2022;1(3):11-14. doi: 10.24018/ejai.2022.1.3.8
46. Yan KX, Liu L, Li H. Application of machine learning in oral and maxillofacial surgery. *AIMI*. 2021;2(6):104-114. doi: 10.35711/aimi.v2.i6.104
47. Zhang J, Fang L, Ma C, Hu Z, Chen C, Zhu T, et al. Gut *Clostridium aminophilum* as a protective microbe in never-smoker lung cancer: insights from metagenomics and machine learning. *Microbiol Spectr*. 2026:e00994-26. doi: 10.1128/spectrum.00994-26
48. Naranjani B, McCoubrey LE, Borde A, Hulthe G, Josefson M, Abrahmsén-Alami S, et al. BiotaPredictor: A machine learning framework for predicting drug depletion by the human gut microbiota. *Gut Microbes*. 2026;18(1):2743953. doi: 10.1080/19490976.2026.2743953
49. Singh S, Aggarwal S, Singh A, Sharma A. Integrating Artificial Intelligence and Machine Learning in Autism Detection via Gut Microbiome Analysis. *ICCK Trans Mach Intell*. 2025;1(2):90-102. doi: 10.62762/tmi.2025.682666
50. Kathail P, Bajwa A, Ioannidis NM. Leveraging genomic deep learning models for the prediction of noncoding variant effects. *Nat Genet*. 2026. doi: 10.1038/s41588-026-02769-7
51. Holton E, Muskovic W, Powell JE. Deciphering cancer cell state plasticity with single-cell genomics and artificial intelligence. *Genome Med*. 2024;16(1):36. doi: 10.1186/s13073-024-01309-4
52. Patel K, Zhao X, Xie Z, Zhi D. A review of self-supervised imaging phenotypes for genome-wide association studies. *Commun Biol*. 2026;9(1):1394. doi: 10.1038/s42003-026-11199-4
53. Chittilla M, Nagdev P. A systematic review of artificial intelligence and machine learning for gut microbiome-based CRC screening. *J Gastrointest Oncol*. 2026;17(2):95-95. doi: 10.21037/jgo-2026-1-0006

54. Tadinada A. Artificial intelligence, machine learning, and the human interface in medicine: Is there a sweet spot for oral and maxillofacial radiology?. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*. 2019;127(4):265-266. doi: 10.1016/j.oooo.2018.12.024

FUNDING

No financing

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Ana María Chaves Cano.
Writing–review and editing: Ana María Chaves Cano.