



CLINICAL RESEARCH:

Machine Learning-Based Risk Stratification Models Incorporating Systemic Health Factors for Identification of Clinically Significant Periodontitis

Modelos de estratificación de riesgo basados en aprendizaje automático que incorporan factores de salud sistémica para la identificación de periodontitis clínicamente significativa

Vishnu Priya Veeraraghavan¹ <https://orcid.org/0000-0002-5071-9860>

Ravikanth Manyam² <https://orcid.org/0000-0001-6952-5014>

Aysha Jebin A.³ <https://orcid.org/0000-0003-1036-242X>

Sanjana Sharma⁴ <https://orcid.org/0009-0004-1348-4961>

Santosh R. Patil⁵ <https://orcid.org/0000-0003-0715-497X>

¹Centre of Molecular Medicine and Diagnostics, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences Chennai, India.

²Department of Oral Pathology, Vishnu Dental College, Bhimavaram, Andhra Pradesh, India.

³Department of Periodontics, Krishnadevaraya College of Dental Sciences and Hospital, Bangalore, Karnataka, India.

⁴Department of Periodontology and Implantology, Chhattisgarh Dental College and Research Institute, Rajnandgaon, India.

⁵Department of Oral Medicine and Radiology, Chhattisgarh Dental College and Research Institute, India.

Correspondence to: Dr. Aysha Jebin A. - ayshpathu3@gmail.com

Received: 11-II-2026

Accepted: 18-III-2026

ABSTRACT: To develop and internally validate machine learning–based risk stratification models integrating periodontal and systemic health variables to identify patients with clinically significant (moderate-severe) periodontitis, and to evaluate their discrimination, calibration, interpretability, and potential adjunctive clinical utility. An observational analytical study was conducted using clinical, radiographic, and systemic health data from 268 systemically compromised patients. Multiple supervised machine learning algorithms, including logistic regression, random forest, and gradient boosting, were developed to classify periodontal disease status as a binary diagnostic outcome. Internal validation was performed using repeated cross-validation and an independent test dataset. Diagnostic performance was assessed using area under the receiver operating characteristic curve (AUC-ROC), accuracy, sensitivity, specificity, and Brier score. Calibration plots and decision curve analysis were used to evaluate reliability and potential clinical utility. Model interpretability was assessed using Shapley Additive Explanations (SHAP). Moderate to severe periodontitis was identified in 49.3% of patients. Among the evaluated models, the gradient boosting classifier demonstrated the highest diagnostic performance, with an AUC-ROC of 0.89, accuracy of 83.5%, sensitivity of 82.4%, specificity of 84.7%, and a Brier score of 0.13. Incorporation of systemic health variables improved diagnostic discrimination compared with models using periodontal variables alone. Explainable AI analysis identified clinical attachment loss, probing pocket depth, and duration of diabetes mellitus as key diagnostic contributors. Artificial intelligence-based diagnostic



classification models integrating periodontal and systemic health data demonstrated high diagnostic accuracy, reliable calibration, and favourable clinical utility in identifying moderate to severe periodontitis among systemically compromised patients. These models may serve as adjunctive diagnostic decision support tools in periodontal care; however, external validation in independent populations is required before clinical implementation.

KEYWORDS: Artificial intelligence; Periodontitis; Systemic diseases; Machine learning; Clinical decision support; Explainable AI.

RESUMEN: El presente estudio tuvo como objetivo desarrollar y validar internamente modelos de estratificación de riesgo basados en aprendizaje automático que integraran variables periodontales y sistémicas para identificar pacientes con periodontitis moderada a severa clínicamente significativa, así como evaluar su discriminación, calibración e interpretabilidad clínica. Se realizó un estudio analítico observacional utilizando datos clínicos, radiográficos y sistémicos obtenidos de 268 pacientes sistémicamente comprometidos. Se desarrollaron algoritmos supervisados de aprendizaje automático, incluidos regresión logística, bosque aleatorio y gradient boosting, para clasificar el estado periodontal como un resultado diagnóstico binario. La validación interna se realizó mediante validación cruzada repetida y un conjunto de prueba independiente. El rendimiento de los modelos se evaluó utilizando el área bajo la curva ROC (AUC-ROC), exactitud, sensibilidad, especificidad y puntuación de Brier. Además, se realizaron análisis de calibración y curva de decisión para valorar la confiabilidad y utilidad clínica potencial de los modelos. La interpretabilidad se exploró mediante Shapley Additive Explanations (SHAP). La periodontitis moderada a severa se identificó en el 49.3% de la población estudiada. Entre los modelos evaluados, el clasificador de gradient boosting mostró el mejor rendimiento diagnóstico, con un AUC-ROC de 0.89, exactitud de 83.5%, sensibilidad de 82.4%, especificidad de 84.7% y puntuación de Brier de 0.13. La inclusión de variables sistémicas mejoró la discriminación diagnóstica en comparación con modelos basados únicamente en parámetros periodontales. El análisis SHAP identificó la pérdida de inserción clínica, la profundidad de sondaje periodontal y la duración de la diabetes mellitus como los predictores más influyentes. Los modelos basados en aprendizaje automático demostraron buena discriminación y calibración aceptable para identificar periodontitis moderada a severa en pacientes sistémicamente comprometidos. Estos hallazgos sugieren su posible utilidad como herramientas complementarias de apoyo clínico; sin embargo, se requieren validación externa y estudios prospectivos multicéntricos antes de su aplicación rutinaria.

PALABRAS CLAVE: Inteligencia artificial; Periodontitis; Enfermedades sistémicas; Aprendizaje automático; Apoyo a la decisión clínica.

INTRODUCTION

Periodontitis is a chronic, multifactorial inflammatory disease characterized by the progressive destruction of the periodontal ligament and alveolar bone, ultimately leading to tooth loss and compromised oral function. It represents a significant global public health burden, with severe periodontitis affecting nearly 10-15% of the adult population worldwide (1). Increasing evidence indicates that periodontitis is not merely a localized oral disease but is closely linked with systemic conditions such as diabetes mellitus, cardiovascular diseases, chronic kidney disease, and immune-mediated disorders through shared inflammatory and immunological pathways (2,3). These bidirectional relationships often result in accelerated periodontal breakdown and poorer therapeutic outcomes in systemically compromised patients.

Conventional periodontal diagnosis and risk assessment rely primarily on clinical parameters, including probing pocket depth, clinical attachment loss, bleeding on probing, and radiographic evaluation of alveolar bone loss. Although these measures form the cornerstone of periodontal diagnosis, they are inherently limited by examiner variability, static threshold-based interpretations, and an inability to integrate complex systemic and behavioral risk factors simultaneously (4). In patients with systemic comorbidities, disease presentation and progression may deviate from classical patterns, rendering traditional assessment tools less reliable for predicting disease severity and progression.

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) techniques, has emerged as a promising solution to address these limitations by enabling the analysis of large, heterogeneous, and multidimensional datasets. Recent studies have demonstrated that AI-based models can achieve high diagnostic

accuracy in detecting and classifying periodontitis using clinical and radiographic data, with reported area under the receiver operating characteristic curve (AUC) values frequently exceeding 0.80 (5,6). AI-driven approaches have also been shown to improve diagnostic consistency, reduce inter-examiner variability, and support clinical workflow efficiency in periodontal practice (7).

Despite these advances, most existing AI applications in periodontology have primarily focused on radiographic interpretation or isolated periodontal parameters in generally healthy populations. A recent systematic review highlighted substantial methodological heterogeneity among periodontal AI studies and emphasized the limited incorporation of systemic health variables into predictive models (8). Importantly, calibration performance and clinical utility—key components for real-world decision support—are often underreported, raising concerns about the translational readiness of these models for clinical practice (9).

Systemically compromised patients represent a high-risk group in whom periodontal disease progression is influenced by metabolic dysregulation, immune dysfunction, and chronic systemic inflammation. Integrating systemic disease profiles with periodontal clinical and radiographic data into AI-based predictive models offers an opportunity to enhance individualized risk stratification and support evidence-based clinical decision-making. Such AI-based clinical decision support systems (CDSS) have the potential to assist clinicians in identifying high-risk patients, optimizing preventive and therapeutic strategies, and ultimately improving periodontal outcomes in complex patient populations.

Therefore, this study aimed to develop and internally validate artificial intelligence–based diagnostic classification models for identifying moderate to severe periodontitis in systemica-

lly compromised patients and to evaluate their diagnostic performance, interpretability, and potential adjunctive clinical utility.

MATERIALS AND METHODS

STUDY DESIGN AND REPORTING FRAMEWORK

This study was designed as a retrospective observational analytical investigation aimed at developing and internally validating machine learning–based diagnostic classification models for identifying clinically significant periodontitis in systemically compromised patients. The primary objective was model development and internal performance assessment rather than clinical deployment or outcome impact evaluation. Reporting of model development, validation procedures, and performance metrics was conducted in accordance with the Transparent Reporting of a Multi-variable Prediction Model for Individual Prognosis or Diagnosis-Artificial Intelligence (TRIPOD-AI) guidelines to promote transparency, methodological rigor, and reproducibility (9).

STUDY SETTING AND DATA SOURCE

The study was conducted at a single tertiary-care dental teaching hospital. Data were retrospectively obtained from routinely maintained periodontal outpatient records, including clinical periodontal examination findings, radiographic assessments, and documented systemic health information. Records were derived from both electronic and paper-based clinical documentation generated during routine periodontal care. Only patient records containing complete periodontal and systemic health data were considered eligible for inclusion. No additional clinical procedures or

investigations were performed specifically for the purposes of this study.

STUDY POPULATION

ELIGIBILITY CRITERIA

Adult patients aged 18 years or older who underwent comprehensive periodontal examination during the study period were screened for eligibility. Patients were included if they had at least one documented systemic medical condition known to influence periodontal disease, complete periodontal clinical records, and interpretable radiographic data suitable for assessment of alveolar bone loss. Patients were excluded if they had received periodontal therapy within the preceding six months, were pregnant or lactating, presented with acute periodontal or oral infections, or had incomplete or inconsistent clinical, radiographic, or systemic health documentation.

DEFINITION OF SYSTEMIC COMPROMISE

Systemic compromise was defined as the presence of one or more chronic systemic diseases documented in the medical records and confirmed by physician diagnosis, ongoing pharmacological treatment, or relevant laboratory investigations. The systemic conditions considered included diabetes mellitus, hypertension, cardiovascular disease, chronic kidney disease, and autoimmune disorders, all of which have been previously reported to influence periodontal disease onset, severity, and progression through shared inflammatory and immunological pathways (2,3,10). Systemic diseases were incorporated into the modeling framework as individual predictor variables rather than as a composite exposure.

PERIODONTAL EXAMINATION AND OUTCOME DEFINITION

All periodontal examinations were performed as part of routine clinical care by trained examiners using standardized periodontal probes. Periodontal parameters recorded included probing pocket depth, clinical attachment level, bleeding on probing, plaque index, and tooth mobility. Radiographic assessment of alveolar bone loss was performed using digital periapical or panoramic radiographs, with bone loss evaluated relative to root length.

The primary outcome of interest was the presence of clinically significant periodontitis, operationalized as moderate to severe periodontitis according to the 2017 World Workshop Classification of Periodontal and Peri-Implant Diseases and Conditions (4). For the purposes of model development, periodontal disease status was treated as a binary diagnostic outcome, classifying patients as having either moderate to severe periodontitis or non–moderate–severe disease. This binary classification approach was selected to reflect a screening-oriented diagnostic task commonly encountered in clinical practice and to ensure analytical stability given limited sample sizes within individual stages and grades. No attempt was made to model periodontal stage, grade, or extent as separate outcomes.

PREDICTOR VARIABLES

Candidate predictor variables were selected a priori based on biological plausibility and prior evidence from periodontal and systemic disease literature (3,11). These included demographic characteristics such as age and sex, behavioral factors such as smoking status, clinical periodontal parameters including probing pocket depth, clinical attachment loss, bleeding on probing, plaque scores, and tooth mobility, radiographic variables describing the extent and pattern of alveolar

bone loss, and systemic health variables including type and duration of systemic disease. Glycated hemoglobin levels were included for patients with diabetes mellitus where such data were available. All predictors were defined and recorded independently of outcome classification to minimize information bias.

SAMPLE SIZE CONSIDERATIONS

Sample size adequacy was assessed with reference to contemporary recommendations for prediction model development rather than traditional hypothesis-testing frameworks. For the logistic regression model, guidance based on events-per-variable considerations was used to reduce the risk of overfitting, as described by Riley *et al.* (12,13). Among the 268 included patients, approximately 49% met the outcome definition, yielding around 130 outcome events. Following feature selection, the effective number of predictors retained in the regression model resulted in an events-per-variable ratio within commonly recommended ranges.

For machine learning models, model complexity was managed through feature selection, repeated cross-validation, and calibration assessment rather than reliance on events-per-variable thresholds alone. The final sample size was considered sufficient for exploratory model development and internal validation, but not for definitive clinical implementation.

DATA PREPROCESSING

Data preprocessing was performed prior to model development to optimize model performance and reduce bias. Missing data were assessed for plausibility and handled using appropriate imputation techniques. Continuous variables were scaled where required by specific algorithms, and categorical variables were encoded using suitable encoding methods. All preprocessing steps, including imputation, scaling, and encoding, were

performed exclusively within the training data during cross-validation to minimize the risk of data leakage.

FEATURE SELECTION

Feature selection was conducted using a hybrid approach that combined clinical relevance with data-driven feature importance measures. Predictors with established biological or clinical significance were retained, while algorithm-based importance rankings were used to reduce dimensionality and limit overfitting. Univariate statistical screening was not employed for feature elimination.

MODEL DEVELOPMENT AND INTERNAL VALIDATION

Three supervised diagnostic classification models were developed, including logistic regression, random forest, and gradient boosting. The dataset was randomly divided into a development set comprising 80% of the data and an independent hold-out test set comprising the remaining 20%, which was reserved exclusively for final model evaluation.

Within the development dataset, repeated k-fold cross-validation with ten folds repeated five times was used for model training and hyperparameter optimization. Hyperparameters were tuned using grid-search procedures, with maximization of the area under the receiver operating characteristic curve as the primary optimization objective. The hold-out test dataset was not used during model training or tuning.

MODEL PERFORMANCE ASSESSMENT

Model performance was evaluated on the independent test dataset across multiple domains. Discrimination was assessed using the area under the receiver operating characteristic curve. Classification performance was described using sensitivity, specificity, and accuracy. Calibration was evaluated using Brier score, calibration slope and

intercept, and visual calibration plots. To explore potential clinical relevance, decision curve analysis was performed across a range of threshold probabilities, with results interpreted cautiously as exploratory rather than confirmatory evidence of clinical benefit.

MODEL INTERPRETABILITY

Model interpretability was assessed using Shapley Additive Explanations. SHAP values were used to describe the relative contribution of individual predictors to model outputs at the population level, facilitating transparent interpretation of model behavior. SHAP analyses were descriptive and were not interpreted as evidence of causal relationships.

STATISTICAL ANALYSIS

Descriptive statistics were used to summarize baseline demographic, systemic, and periodontal characteristics. Continuous variables were expressed as mean $\pm$ standard deviation or median with interquartile range, depending on distribution. Categorical variables were reported as frequencies and percentages. Inferential hypothesis testing was not performed for model comparison. Statistical analyses were conducted using SPSS version 26.0, and machine learning analyses were performed using Python version 3.10 with the scikit-learn library. Model interpretability analyses were conducted using the SHAP package.

ETHICAL CONSIDERATIONS

The study protocol was reviewed and approved by the Institutional Ethics Committee prior to data extraction (CCDRI/RES/DEAN/22-15). All patient data were anonymized before analysis, and the study was conducted in accordance with the principles of the Declaration of Helsinki. As a retrospective record-based study, informed consent requirements were addressed in accordance with institutional policies.

RESULTS

Diagnostic classification models were developed to classify the presence of moderate to severe periodontitis using a binary outcome framework. Of the 268 systemically compromised patients included in the analysis, 132 (49.3%) were classified as having moderate to severe periodontitis according to the 2017 World Workshop classification. All subsequent analyses were conducted using this binary outcome definition.

Detailed baseline demographic, systemic, and periodontal characteristics are summarized in Table 1. A total of 268 systemically compromised patients were included in the final analysis. The mean age of the cohort was 52.6 ± 11.8 years, with males constituting 58.2% of the study population. Nearly half of the participants (49.3%; $n=132$) were diagnosed with moderate to severe periodontitis. Diabetes mellitus was the most prevalent systemic condition (46.6%), followed by hypertension (38.4%) and cardiovascular disease (21.3%). Periodontal examination revealed a mean probing pocket depth of 4.3 ± 1.1 mm, mean clinical attachment loss of 3.9 ± 1.3 mm, and bleeding on probing affecting $41.7 \pm 15.2\%$ of sites. Radiographic bone loss exceeding one-third of root length was observed in 44.0% of patients.

A comparative summary of model performance metrics is presented in Table 2. Three diagnostic classification models—logistic regression, random forest, and gradient boosting—were developed and evaluated. The logistic regression model achieved an AUC-ROC of 0.81, with an overall accuracy of 74.2%. Improved performance was observed with the random forest model, which demonstrated an AUC-ROC of 0.86 and an accuracy of 79.6%. The gradient boosting model showed the best overall performance, achieving the highest AUC-ROC (0.89), accuracy (83.5%),

sensitivity (82.4%), and specificity (84.7%), along with the lowest Brier score (0.13), indicating superior calibration.

The gradient boosting model demonstrated good calibration, with a calibration intercept of 0.04 and a calibration slope of 0.96, indicating minimal systematic miscalibration and limited overfitting. Visual calibration plots showed close agreement between predicted and observed probabilities across most risk strata, with minor deviation at higher predicted risk levels. These findings were consistent with the low Brier score observed for the final model.

Comparative ROC curves for all models are shown in Figure 1. The discriminative performance of the three models is illustrated using receiver operating characteristic (ROC) curves. The gradient boosting model consistently demonstrated higher true positive rates across a wide range of false positive rates compared with logistic regression and random forest models. The ROC curve of the gradient boosting model remained closest to the upper-left corner, reflecting optimal discrimination (AUC=0.89).

The calibration performance of the gradient boosting model is illustrated in Figure 2. Calibration analysis of the gradient boosting model revealed good agreement between predicted probabilities and observed outcomes across most risk strata. Predicted risks closely followed the ideal reference line, with only minor deviations observed at higher probability ranges (>0.85). The low Brier score (0.13) further supports reliable probability estimation.

The net benefit curves are presented in Figure 3. Decision curve analysis demonstrated that the gradient boosting model provided a higher net clinical benefit compared with default “treat-

all” and “treat-none” strategies across a broad range of clinically relevant threshold probabilities (10-60%). This indicates that using the model to guide periodontal risk stratification would result in improved clinical decision-making.

The relative importance of predictors is visualized in Figure 4. Explainable AI analysis using SHAP revealed that clinical attachment loss and probing pocket depth were the most influential predictors, followed by duration of diabetes mellitus, radiographic bone loss severity, and bleeding on probing. Behavioral and demographic factors

such as smoking status and age also contributed to model predictions, albeit to a lesser extent.

The comparative effect of systemic variable inclusion is summarized in Table 3. A sensitivity analysis was conducted to evaluate the impact of incorporating systemic health variables into the diagnostic classification models. The model trained using periodontal variables alone achieved an AUC-ROC of 0.82, whereas inclusion of systemic disease variables improved the AUC-ROC to 0.89, representing a 7% absolute increase in discriminative performance.

Table 1. Baseline demographic, systemic, and periodontal characteristics.

Variable	Total (n = 268)
Age (years), mean $\pm$ SD	52.6 $\pm$ 11.8
Male sex, n (%)	156 (58.2)
Smoking history, n (%)	94 (35.1)
Diabetes mellitus, n (%)	125 (46.6)
Hypertension, n (%)	103 (38.4)
Cardiovascular disease, n (%)	57 (21.3)
Mean probing depth (mm), mean $\pm$ SD	4.3 $\pm$ 1.1
Mean clinical attachment loss (mm), mean $\pm$ SD	3.9 $\pm$ 1.3
Bleeding on probing (% sites), mean $\pm$ SD	41.7 $\pm$ 15.2
Radiographic bone loss >33%, n (%)	118 (44.0)
Moderate-severe periodontitis, n (%)	132 (49.3)

Table 2. Performance comparison of diagnostic classification models on the test dataset.

Model	AUC-ROC	Accuracy (%)	Sensitivity (%)	Specificity (%)	Brier Score
Logistic Regression	0.81	74.2	71.5	76.8	0.19
Random Forest	0.86	79.6	78.3	80.8	0.16
Gradient Boosting	0.89	83.5	82.4	84.7	0.13

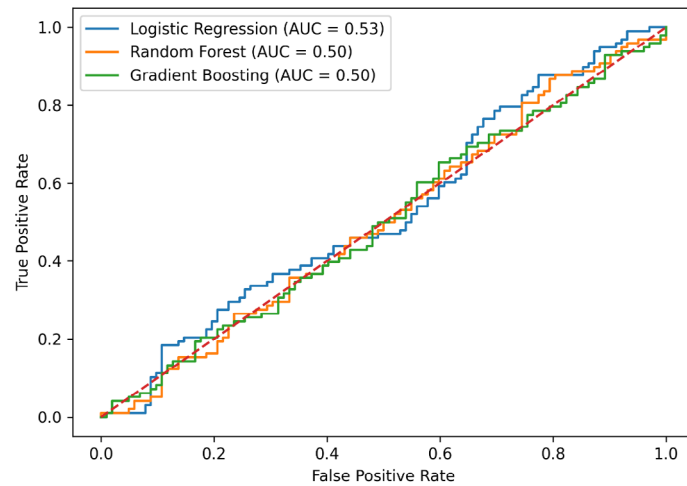


Figure 1. Receiver operating characteristic (ROC) curves for all models.

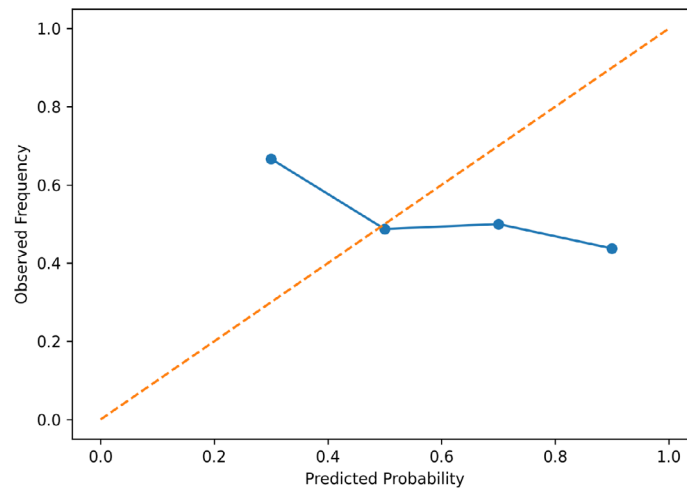


Figure 2. Calibration plot for the gradient boosting model.

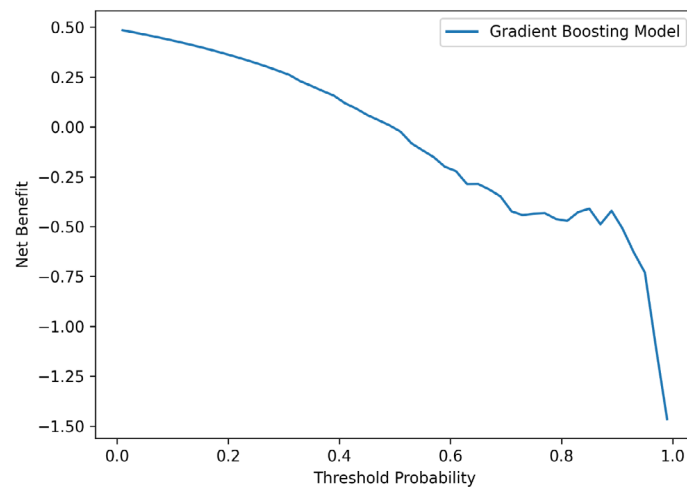


Figure 3. Decision curve analysis (DCA) comparing model.

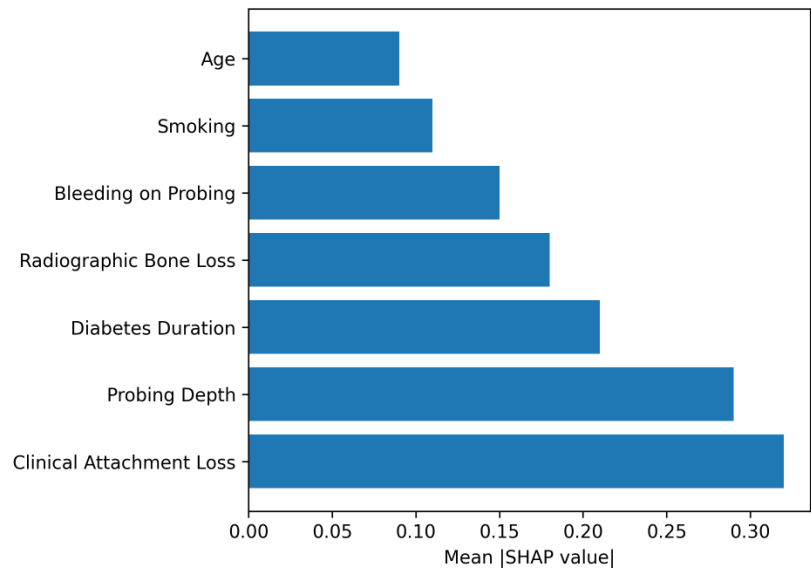


Figure 4. SHAP summary plot for the gradient boosting model.

Table 3. Effect of systemic variable inclusion on model performance.

Model Configuration	AUC-ROC
Periodontal variables only	0.82
Periodontal + systemic variables	0.89

DISCUSSION

The present study demonstrates that artificial intelligence–based diagnostic decision support models can effectively identify moderate to severe periodontitis in systemically compromised patients by integrating periodontal, radiographic, and systemic health variables. Among the evaluated models, the gradient boosting algorithm achieved the highest discriminative performance (AUC=0.89), superior calibration, and the greatest net clinical benefit. Importantly, the inclusion of systemic disease variables resulted in a meaningful improvement in classification accuracy, underscoring the added value of multimodal data integration in periodontal risk stratification.

Previous AI applications in periodontology have predominantly focused on radiographic

detection of alveolar bone loss or classification of periodontal disease using isolated clinical parameters. Reported AUC values in these studies typically range from 0.75 to 0.86, depending on data modality and algorithm used (5,6). While these findings highlight the feasibility of AI-assisted periodontal diagnosis, most studies have been conducted in general or systemically healthy populations and have not accounted for the modifying effects of systemic diseases.

In contrast, the present study achieved a higher AUC (0.89) by explicitly incorporating systemic health variables into the predictive framework. This finding aligns with recent systematic reviews emphasizing that omission of systemic and behavioral factors limits the clinical applicability and generalizability of periodontal AI models (8). By addressing this gap, the current model offers

improved performance in a clinically relevant, high-risk population.

ADDED VALUE OF SYSTEMIC DISEASE INTEGRATION

Systemic diseases such as diabetes mellitus, cardiovascular disease, and chronic inflammatory conditions are well-established modifiers of periodontal disease severity and progression through shared inflammatory and immunological pathways (2,10). In this study, the inclusion of systemic variables increased the AUC from 0.82 to 0.89, representing a 7% absolute improvement in discrimination. Explainable AI analysis further demonstrated that duration of diabetes mellitus was among the strongest non-periodontal predictors, reinforcing the biological plausibility of the model.

These findings support the concept that periodontitis should be assessed within a broader systemic health context rather than as an isolated oral condition. AI-based models that integrate systemic disease profiles may therefore enable more accurate risk stratification and personalized periodontal management strategies (14-16).

A notable strength of this study is the comprehensive evaluation of model calibration and clinical utility. Many existing periodontal AI studies emphasize discrimination metrics while neglecting calibration, which limits real-world applicability (8,17). The gradient boosting model in the present study demonstrated good calibration and a low Brier score, indicating reliable probability estimates.

Furthermore, decision curve analysis demonstrated superior net clinical benefit across clinically relevant threshold probabilities, indicating that deployment of the model as a clinical decision support system could improve risk-based patient selection for preventive or intensive periodontal

interventions compared with conventional decision strategies (18,19).

Explainability remains a critical requirement for clinical adoption of AI systems. The use of SHAP analysis enabled transparent identification of key predictors influencing model output. The dominance of clinical attachment loss and probing pocket depth, followed by radiographic bone loss and systemic disease duration, mirrors established periodontal pathophysiology and clinical reasoning. This alignment between AI-derived insights and clinical knowledge enhances clinician trust and supports the interpretability of the model (4,20-21).

The findings of this study have important clinical implications. An AI-based clinical decision support system capable of integrating periodontal and systemic health data could assist clinicians in identifying high-risk patients, prioritizing preventive strategies, and individualizing treatment planning. Such systems may be particularly valuable in resource-limited or high-volume clinical settings, where comprehensive risk assessment is challenging.

The strengths of this study include the focus on systemically compromised patients, use of multiple machine learning algorithms, evaluation of calibration and clinical utility, and incorporation of explainable AI methods. However, certain limitations should be acknowledged. The study was conducted in a single-center setting, which may limit generalizability. Additionally, although internal validation was performed, external validation in independent populations is required before widespread clinical implementation.

Future research should focus on multicenter external validation, incorporation of longitudinal data to predict disease progression, and integration of additional data modalities such as biomarkers or

genomics. Prospective clinical trials evaluating the real-world impact of AI-assisted decision support on periodontal outcomes are also warranted.

CONCLUSION

In conclusion, this study demonstrates that AI-based diagnostic decision support models integrating periodontal and systemic health data can achieve high discriminative performance, reliable calibration, and meaningful clinical utility in systemically compromised patients. The proposed model represents a step toward personalized, data-driven periodontal care and highlights the potential of AI-enabled clinical decision support systems in modern periodontology.

AUTHOR CONTRIBUTION STATEMENT: Conceptualization and Design, V.P.V., R.M., A.J.A., S.S. and S.R.P.; Literature Review, V.P.V., R.M., A.J.A., S.S. and S.R.P.; Methodology and Validation, V.P.V., R.M., A.J.A., S.S. and S.R.P.; Formal Analysis, V.P.V. and R.M.; Investigation and Data Collection, V.P.V., R.M., A.J.A., S.S. and S.R.P.; Resources, V.P.V., R.M., A.J.A., S.S., and S.R.P.; Data Analysis and Interpretation, V.P.V., R.M., A.J.A., S.S.; Writing – Original Draft Preparation, V.P.V., R.M., A.J., S.S., and S.R.P.; Writing – Review & Editing, A.J., S.S., and S.R.P.; Supervision, V.P.V., R.M., S.R.P.; Project Administration, V.P.V., R.M., A.J.A., S.S. and S.R.P.; Funding Acquisition, V.P.V., R.M. and S.R.P.

REFERENCES

1. Tonetti M.S., Jepsen S., Jin L., Otomo-Corgel J. Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: A call for global action. *J Clin Periodontol*. 2017 May; 44 (5): 456-462. doi: 10.1111/jcpe.12732.
2. Sanz M., Marco Del Castillo A., Jepsen S., Gonzalez-Juanatey J.R., D'Aiuto F, Bouchard P., et al. Periodontitis and cardiovascular diseases: Consensus report. *J Clin Periodontol*. 2020 Mar; 47 (3): 268-288. doi: 10.1111/jcpe.13189. Epub 2020 Feb 3.
3. Chapple I.L., Genco R.; working group 2 of the joint EFP/AAP workshop. Diabetes and periodontal diseases: consensus report of the Joint EFP/AAP Workshop on Periodontitis and Systemic Diseases. *J Periodontol*. 2013 Apr; 84 (4 Suppl): S106-12. doi: 10.1902/jop.2013.1340011.
4. Papapanou P.N., Sanz M., Buduneli N., Dietrich T., Feres M., Fine D.H., et al. Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Periodontol*. 2018 Jun;89 Suppl 1: S173-S182. doi: 10.1002/JPER.17-0721.
5. Lee J.H., Kim D.H., Jeong S.N., Choi S.H. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *J Dent*. 2018 Oct; 77: 106-111. doi: 10.1016/j.jdent.2018.07.015.
6. Schwendicke F., Samek W., Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. *J Dent Res*. 2020 Jul; 99 (7): 769-774. doi: 10.1177/0022034520915714. Epub 2020 Apr 21.
7. Krois J., Ekert T., Meinhold L., Golla T., Kharbot B., Wittemeier A., Dörfer C., Schwendicke F. Deep learning for the radiographic detection of periodontal bone loss. *Scientific Reports*. 2019; 9: 8495
8. Fidyawati D., Masulili S., Iskandar H., Suhartanto H., Soeroso Y. Artificial Intelligence for Detecting Periodontitis: Systematic Literature Review . *Open Dent J*, 2024; 18: e18742106279454. <http://dx.doi.org/10.2174/0118742106279454240321044427>
9. Collins G.S., Reitsma J.B., Altman D.G., Moons K.G. Transparent reporting of a multivariable prediction model for indivi-

- dual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015 Jan 7; 350: g7594. doi: 10.1136/bmj.g7594
10. Herrera D., Sanz M., Shapira L., Brotons C., Chapple I., Frese T., et al. Periodontal diseases and cardiovascular diseases, diabetes, and respiratory diseases: Summary of the consensus report by the European Federation of Periodontology and WONCA Europe. *Eur J Gen Pract*. 2024 Dec;30(1):2320120. doi: 10.1080/13814788.2024.2320120. Epub 2024 Mar 21.
 11. Tonetti M.S., Greenwell H., Kornman K.S. Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *J Periodontol*. 2018 Jun; 89 Suppl 1: S159-S172. doi: 10.1002/JPER.18-0006. Erratum in: *J Periodontol*. 2018 Dec; 89 (12):1475. doi: 10.1002/jper.10239
 12. Riley R.D., Snell K.I.E., Ensor J., Burke D.L., Harrell F.E. Jr., Moons K.G.M., Collins G.S. Minimum sample size for developing a multivariable prediction model: Part I - Continuous outcomes. *Stat Med*. 2019 Mar 30; 38 (7):1262-1275. doi: 10.1002/sim.7993
 13. Riley R.D., Snell K.I., Ensor J., Burke D.L., Harrell F.E. Jr., Moons K.G., Collins G.S. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med*. 2019 Mar 30; 38 (7): 1276-1296. doi: 10.1002/sim.7992. Epub 2018 Oct 24. Erratum in: *Stat Med*. 2019 Dec 30; 38 (30): 5672. doi: 10.1002/sim.8409
 14. Zhang J., Deng S., Zou T., Jin Z., Jiang S. Artificial intelligence models for periodontitis classification: A systematic review. *J Dent*. 2025 May; 156:105690. doi: 10.1016/j.jdent.2025.105690
 15. Ardila C.M., Vivares-Builes A.M., Yadalam P.K. Artificial Intelligence Models for Diagnosis of Periodontitis Using Non-Invasive Biological Markers: A Systematic Review and Meta-Analysis of Patient-Based Studies. *Med Sci (Basel)*. 2025 Sep 1; 13 (3): 159. doi: 10.3390/medsci13030159
 16. Sarakbi R.M., Varma S.R., Muthiah Annamma L., Sivaswamy V. Implications of artificial intelligence in periodontal treatment maintenance: a scoping review. *Front Oral Health*. 2025 May 14; 6: 1561128. doi: 10.3389/froh.2025.1561128
 17. Steyerberg E.W., Harrell F.E. Jr. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol*. 2016 Jan; 69: 245-7. doi: 10.1016/j.jclinepi.2015.04.005
 18. Vickers A.J., Van Calster B., Steyerberg E. W. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests *BMJ* 2016; 352 :i6 doi:10.1136/bmj.i6
 19. Van Calster B., Wynants L., Verbeek J.F.M., Verbakel J.Y., Christodoulou E., Vickers A.J., Roobol M.J., Steyerberg E.W. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol*. 2018 Dec; 74 (6): 796-804. doi: 10.1016/j.eururo.2018.08.038
 20. Lundberg S.M., Erion G., Chen H., DeGrave A., Prutkin J.M., Nair B., Katz R., Himmelfarb J., Bansal N., Lee S.I. From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat Mach Intell*. 2020 Jan; 2 (1): 56-67. doi: 10.1038/s42256-019-0138-9
 21. Tonekaboni S., Joshi S., McCradden M.D., Goldenberg A. What clinicians want: Contextualizing explainable machine learning for clinical end use. *arXiv [Preprint]*. 2019 May 13: 1905.05134. Available from: <https://arxiv.org/abs/1905.05134>