

# Enhancing an AI-Empowered Periodontal CDSS and Comparing with Traditional Perio-risk Assessment Tools

Jay S. Patel, BDS, MS, PhD<sup>1,3</sup>, Kajal Patel, BDS<sup>1</sup>, Hoa Vo, MPH, MS<sup>1</sup>, Liu Jiannan, MS<sup>2</sup>, Marisol M. Tellez, BDS, MS, PhD<sup>3</sup>, Jasim Albandar, DMD<sup>3</sup>, PhD, Huanmei Wu, PhD<sup>1</sup>

<sup>1</sup>College of Public Health, Temple University, Philadelphia, PA, USA; <sup>2</sup>School of Informatics and Computing, Indiana University Purdue University Indianapolis, Indianapolis, IN, USA;

<sup>3</sup>Temple University Kornberg School of Dentistry, Philadelphia PA, USA

## Abstract

*Periodontal disease (PD) is one of the most prevalent dental diseases. Fortunately, it can be prevented if identified early, especially for high-risk patients. Dental electronic health records (EHRs) could help develop a data-driven personalized prediction model using advanced machine learning development of clinical decision support system (CDSS) as in our Phase I, II AMIA-AI showcase. In phase II, we created a CDSS, the Perio-Risk Scoring system (PRSS), to help clinicians generate perio-scores and diagnoses and identify the influential factors. In Phase III (this study), we implemented and compared the patient's risk factors information in five periodontal risk assessment tools [periodontal risk assessment (PRA), PreViser, Sonicare, Cigna, and Periodontal Risk Scoring System (PRSS)]. We examined 1) agreement between the risk scores provided by each of the five risk assessment tools of 20 patients' information and 2) compare the risk scores provided by each tool to the original outcomes (five years outcomes). Fleiss Kappa, Cohen's Kappa, and percentage agreements were performed to determine the agreements between risk scores and original outcomes. We found a -1.24 Kappa value which indicates disagreement between the risk scores provided by five risk assessment tools. Compared to the original outcomes (five-year disease outcomes), PRSS provided the most accurate prediction (70%), followed by Previser (55%), PRA (35%), Phillips (35%), and Cigna (25%). We conclude that using advanced state-of-the-art informatics methods could help us utilize EHR data optimally to represent the current patient populations and their risk factors to provide the most accurate disease risk score. This may promote preventive strategies at the chairside, hoping to reduce PD prevalence, improve quality of life, and reduce healthcare costs.*

## Introduction

Advances in periodontal disease (PD) research and treatments showed that PD could be prevented if diagnosed and treated early. However, 80% of US adults still suffer from gingivitis (a mild form of PD with gum inflammation) and 42% from periodontitis (a more severe form of PD)<sup>1</sup>. PD leads to tooth loss, functional disabilities, and poor quality of life and is responsible for increased healthcare costs<sup>2,3</sup>. Unfortunately, the predominant dental care is still the reparative model, in which clinicians typically take care of the immediate pathology seen in the oral cavity and provide less focus on the preventive model. In the preventive model, clinicians take a step back and assess the risk of disease initiation and progression based on the etiological and risk factors<sup>4</sup>. The risk assessment approach can help dental clinicians identify high-risk periodontitis patients and take preventative measures to prevent or delay the disease progression, especially in its early stages<sup>5,6</sup>. As a result, a few PD risk assessment tools have been developed over the last two decades, including Periodontal Risk Assessment, Sonicare, Signa, PreViser, and PEMBRA. These tools utilize IF-ELSE conditional statements to generate PD risk scores into very low to high risks using the preassigned weigh of PD risk factors.

However, the effectiveness of these tools do not explain the risk scores with their underlying risk calculation algorithms<sup>7-13</sup>. Studies have demonstrated that the periodontal risk scores generated by two different risk assessment tools are significantly different when the same patient information is entered. As a result, it is difficult for clinicians to select the best tool that can provide patient-specific periodontal risk scores to take preventive measures<sup>6,12,14-16</sup>. Nevertheless, to our best knowledge, no studies have compared the risk scores provided by readily available tools and determined the similarities and differences between the provided risk scores. Moreover, limited research exists comparing the risk assessment tool-provided risk scores versus the original outcomes of patients' PD status.

It is critical to study the similarities and differences between the risk scores generated by different risk assessment tools for the following reasons. First, the risk scores highly depend on the inclusion of the risk factors involved in calculating risk scores. It is well understood that periodontitis is a multifactorial disease. Multiple risk factors are responsible for the disease initiation and progression<sup>3</sup>. Therefore, including comprehensive risk factors may help predict the risk of periodontitis with high accuracy. Second, the method used to create the tool that generates risk scores is also essential. For example, risk assessment tools developed using data-driven models have higher sensitivity but lower generalizability<sup>17</sup>. This is because the data is collected from the particular regional institute/clinics that could

best represent their sample population, and the prevalence of risk factors varies from region to region. On the other hand, the expert-driven risk assessment tools developed have less specificity but better generalizability. This is because these tools mainly include risk factors that are highly studied in the literature. For example, regardless of region, smoking, diabetes, and poor oral hygiene are strong predictors of periodontitis which have been very well known in the literature<sup>3,18</sup>. Both of these methods have advantages and limitations. The expert-driven model would help identify the risk factors that are highly prevalent in all populations but not patient-specific risk factors (community-level risk factors). The data-driven model utilizes machine learning and advanced statistical models to predict the risk scores that are specific to the patients and more personalized.

The expert-driven models are developed using well-known risk factors by utilizing simple rule-based algorithms such as IF and ELSE statements. There are five periodontal risk assessment tools (excluding our tool) available, which are either commercially or freely available to use to assess periodontal disease risks. These tools include 1) periodontal risk assessment (PRA): from hereon referred to as *tool A*, 2) Sonicare (Phillips): *Tool B*, 3) Cigna: *Tool C*, and 4) PreViser: *Tool D*. Only tool D is commercially available, and the remaining tools (tools A-C) are open-source tools<sup>6,12,14-16</sup>.

In the American Medical Informatics Association (AMIA) AI Stage I showcase, we developed a data-driven PD predictive model based on 74 candidate predictors from the dental EHR to provide patient-specific PD disease risk. We used dental records for 274,892 observations of 27,138 unique patients from the Temple University Kornberg School of Dentistry (TUKSoD) clinics. This model was developed using a robust tree-based machine learning model, XGBoost, to predict the risk of PD. Our model achieved good performance in predicting PDs (including severe PDs and mild PDs) from the healthy control patients.

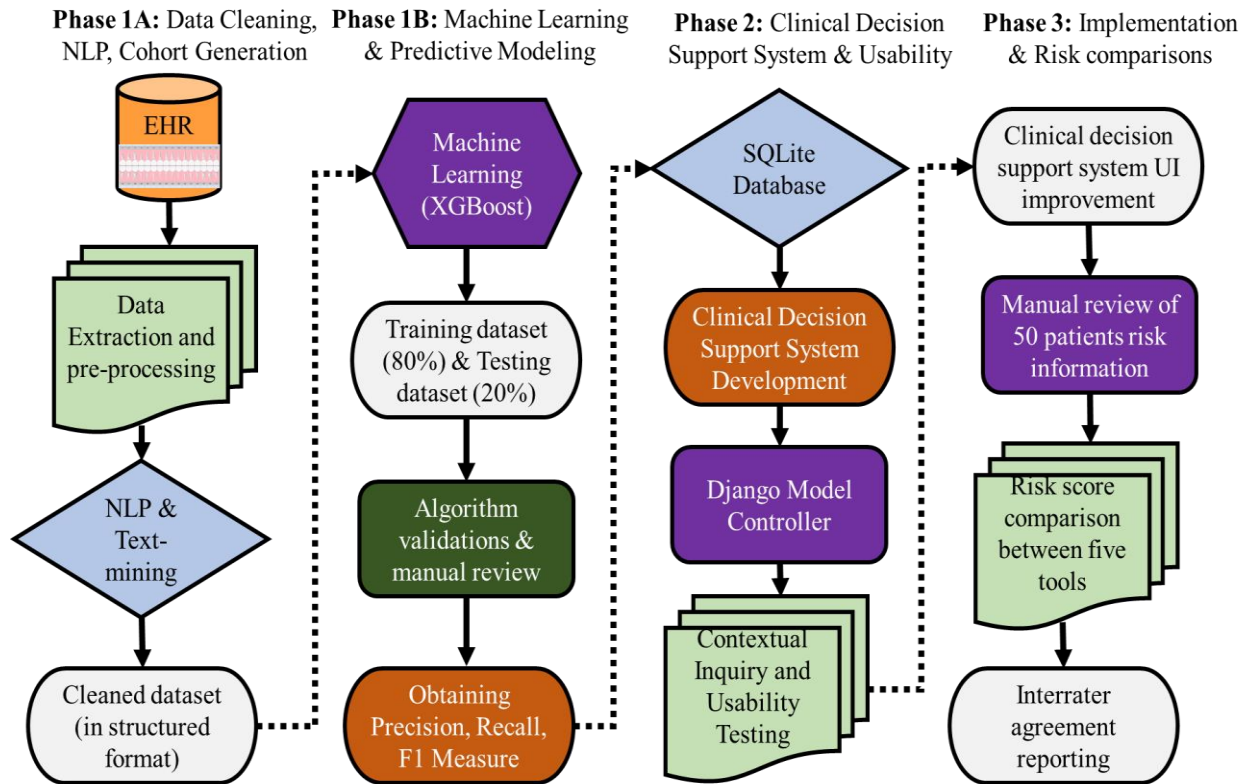
For the AMIA AI showcase phase II, we developed a user-friendly clinical decision support system (CDSS), *i.e.*, the Perio-Risk Scoring System (PRSS), based on our predictive model. We also conducted the CDSS usability study by interviewing periodontal residents and dentists. In addition, we performed a contextual inquiry to determine the clinical workflow for the deployment of the CDSS in postdoctoral periodontal clinics.

In this study (AMIA AI showcase Phase III), we compared the risk scores utilized by each of the five periodontal risk assessment tools (including PRA, Sonicare, Cigna, PreViser, PRSS). We compare and contrast the similarities and differences in risk assessment variables and outcomes. We also compared the risk scores provided by five PD risk assessment tools versus the five years disease progression outcomes. As per our best knowledge, this was the first study with the following three features: (i) utilizing a comprehensive list of variables (currently 74 features but can be extended to other factors) to develop the prediction model; (ii) applying advanced machine learning methods to determine the risk factors instead of the domain knowledge of experts as in previous studies; and (iii) developing a clinical decision support system with automated data entry from the structured and unstructured dental electronic health record (EHR) data using natural language processing algorithms.

## Methods

### Overall System Architecture

This study was reviewed and approved by our institutional review board (Temple University *IRB# 28321*). The overall workflow from the development and testing of the PRSS to the comparison of the five risk assessment tools is illustrated in *Figure 1*. It includes the information flow from the EHR to end-users, technical details from machine learning-based predictive modeling to web server development, user feedback for system improvement, contextual inquiry, and risk score comparisons. The overall project comprises 3 phases, (Phase 1 comprised of development, testing, and validation of the PD prediction model, Phase 2 consisted of the development of the CDSS, usability testing, and contextual inquiry, and Phase 3 consisted of implementation and risk scoring comparisons). The following describes the major steps in the workflow.



**Figure 1:** Overall workflow of dental EHR data extraction, processing, cohort generation, prediction model, clinical decision support systems, usability, and implantation

## 1) Phase 1A, B: Data Preprocessing, Cohort Generation, and Machine Learning Modeling

### Data Source, Data Retrieval, and Natural Language Processing (NLP)

Dental EHR (axiUm®, Exan software, Las Vegas, Nevada, USA) data of patients who received at least one comprehensive oral evaluation (COE) at TUKSoD between January 1, 2017, and August 31, 2021, was used ( $n = 27,138$  patients). More than half of risk factor information was recorded in free-text format; therefore, we created multiple NLP pipelines to convert unstructured data into a structured format. After iterations, our final NLP pipelines provided an average of 94.5% F-1 score, which is considered excellent. Detailed information on the steps involved in the development and testing of NLP pipelines (manual annotations, training & testing NLP applications, and validation) is described elsewhere<sup>19,20</sup>. We also used imputation methods to handle missingness in the dataset. First, we dropped the features that had a very high missing rate. For example, ethnicity was missing for more than 70% of patients, so we dropped the ethnicity feature. Features with limited missing values were imputed according to value types. Categorical features, such as medical histories, were imputed using the most frequent values. We also developed phenotyping algorithms to automatically generate patients' PD diagnoses from their periodontal charting information. These algorithms are heavily validated through a manual review process, and we received an average accuracy of 97%. We ran a machine learning model XGBoost on this clean dataset to develop a prediction model for PD. We grouped patients' outcomes into three classes, i.e., healthy control (HC), mild PD, and severe PD. To address the multi-label prediction task, we built the model using a "one-vs-rest" strategy. After identifying optimal hyperparameters based on which the model can achieve the best prediction performance, we retrained the predictive model using the whole training set. Model evaluation and statistical analysis were performed on the testing set, measured by the area under the receiver operating characteristics curve (AUC), and the confusion matrix to compute prediction accuracy, precision (PPV), recall (sensitivity), and F-1 measure (harmonic mean of precision and recall). We achieved the performance of 72% area under the curve that utilizes 74 unique features towards the prediction.

## 2) Phase 2: Build a Clinical Decision Support System

## **Developing a User Interface by Django Model Controller**

In the AMIA AI showcase stage II submission, PD-related features were transformed into the Python-Django deployment framework. First, the Django model controller will use the Form and View to have the record updated automatically. The Django Model queried patient data from the Perio-risk database and formatted it into Django classes. The Controller will apply the predictive model to patient data, generate statistics of model output, and calculate the patient risk score of PD. It also generates visualizations for the web server's graphical user interface. For example, the PRSS user interface will show the predictor importance, which illustrates the model evidence for predictions. In this work, we engaged the Shapley Additive Explanation (SHAP) values to assess the contributions of the predictors in distinguishing each class from the others. A SHAP values plot is generated to explain the predicted perio risk probabilities of each individual patient. Users (dentists) were recruited to test the CDSS and provide feedback for system enhancement. The feedback on data and features will be sent back to different modules of the overall workflow. For example, the user interface has been updated based on user input on the visualization types, text explanations, and additional features. In Stage II, we collected feedback from the users' experience in using the PRSS tool. We improved the user interface based on the feedback received. We also then tested the new interface on a small subset of the sample (two dentists) and collected their feedback.

### **3) Phase III: Risk score Comparisons Between Five Periodontal Risk Assessments tools**

First, we manually reviewed each PD risk assessment tool (Tools A-E) and documented mandatory variables to calculate risk scores in each tool (*See Table 1 & Table 2*). Two dentists and informaticists manually reviewed and documented 50 dental patients' COE that include 1) patient demographics, 2) medical histories, 3) dental histories, 4) periodontal charting, 5) clinical notes, 6) social determinates of health, 7) periapical radiographs, and 8) bone loss information. This information was then entered into each PD risk assessment tool. The output regarding PD risk of 20 patients was documented for each tool for comparison. Each tool provided different risk score outputs. For example, PRA provides risk outputs as either "low", "medium", or "high risk", while PreViser provides risk scores as "very low", "low", "moderate", "high", or "very high" risk. Therefore, we normalized risk scores for comparison into "low", "moderate," or "high" by merging these categories. The very low and low risk were categorized into low risk, and high and very high risk categories were merged into high risk. Our tool PRSS doesn't provide any risk categories; however, it provides risk scores between 0 and 100. Therefore, to be consistent with other risk assessment tools' risk scores, we categorized 0 to 30 scores risk into "low risk," 31 to 60 scores into "moderate risk," and more than 60 scores into "high risk."

### **4) Risk Score Comparisons between the Predicted Outcomes and Original Outcomes**

We used the longitudinal dental EHR data to compare the predicted risk scores by the five risk assessment tools versus the original outcomes. We obtained a dataset of those patients who had at least five years of consecutive visits to the dental school, and their complete COE data was used for the comparison. We entered the patient information obtained during their first visit in 2017 in all five risk assessment tools. We then compared the patient's actual diagnosis in 2021 to examine how accurately the risk assessment tools were able to predict the outcomes

### **5) Data Analysis**

Descriptive statistics with a 95% confidential interval were performed on patient demographics, medical history, and procedures for the patients who received at least one COE (Phase I). The prediction model's performance was determined by measuring sensitivity, specificity, and area under the curve (Phase II). In Phase III (this study), Fleiss' Kappa statistical test was performed to determine the interrater agreement between the risk scores provided by each risk assessment tool. Cohen's Kappa and percent agreement were performed to determine the agreement between the risk assessment tool's outcomes versus the original outcomes <sup>21</sup>.

## **Results**

### **1) Patient Demographics**

Our sample consisted of 20 unique dental patients who received at least one COE between January 1, 2017, and August 31, 2021. Our patients' most common age group was 58-67 years (n=12). African American was the most frequent race (n=13), followed by Whites. The majority of our patients were females (n=11).

**Table 1:** Included variables in two or more periodontal disease risk assessment tools

<b>Risk Factors</b>	<b>PRA (A)</b>	<b>Phillips (B)</b>	<b>Cigna (C)</b>	<b>PreViser (D)</b>	<b>PRSS (E)</b>
Age	√	√	√	√	√
% of bone loss	√	√	√	√	√
Systemic condition	√	√	√	√	√
Smoking status	√	√	√	√	√
# of BOP sites	√	√	×	√	√
# PPD>5 mm	√	√	×	√	√
# of dental visits	×	√	√	√	√
Past perio treatment	×	√	√	√	√
# of teeth present	√	×	×	√	√
# of missing teeth	√	×	×	√	√
Family history	×	√	√	×	√
Receding gums	×	√	√	×	√
Mobile teeth	×	√	√	×	√
Teeth extracted (perio)	×	√	√	×	√
# of brushing/day	×	√	√	×	√
# of flossing	×	√	√	×	√
Pregnancy	×	×	√	×	√
Bleeding gums	×	×	√	×	√
Time elapsed since extraction	×	×	√	×	√
Bad breath	×	×	√	×	√
# of times antibacterial mouthwash	×	√	√	×	×
Furcation involvement	×	×	×	√	√
Periopathogenic bacteria	×	×	√	×	×
Test (susceptibility to PD)	×	×	√	×	×
Oral-strain probiotics	×	√	×	×	×

**Table 2:** Variables used in our perio-risk scoring system only

Anxiety	Race/ethnicity	ASA classification	Bruxism/parafunctional
Alcohol	Tooth crowding	Self-image, uneasy	Injury/trauma to teeth
Chewing	Root proximity	Recreational drugs	Subgingival restoration
Calculus	Tooth mobility	Pathologic migration	Dental health condition
Bone loss	Floss frequency	Local drug delivery	Jaw muscle or joint pain
Insurance	Clench or grind	Vertical bone lesion	Orthodontic tooth movement
Gingivitis	Partial denture	Subgingival calculus	Usage of recreational drugs
Pain level	Speaking trouble	High caries activity	Decayed missing filled teeth
Radiographs	Oral surgery	High level of stress	Professionally dental hygiene
Gum trouble	Gingival pain	Scaling & root planing	Inadequate patient compliance
Medications	Healthy diet	Caries risk assessment	History of periodontal abscess
Oral hygiene	Open contacts	Defective restorations	Decayed missing filled surface
Plaque index	Diabetes mellitus	Abnormal tooth anatomy	

## 2) Prediction Model Performance

**Table 3:** Percent and Inter-rater agreement between five periodontal risk assessment tools versus five years disease progression outcome

Tool	Percent Agreement	Cohen's Kappa Agreement
PRSS	70	0.6
PreViser	55	0.4
PRA	35	0.3
Phillips	35	0.3
Cigna	25	0.2

We achieved an AUC of 0.72 (weighted average of three base models) in distinguishing HC, mild PD, and severe PD from each other (see Figure 2 and supplemental Table 3). When looking into the “one-vs-rest” base models, the models work well in distinguishing HCs from PDs (AUC = 0.69, F1-score = 0.66) as well as in distinguishing the severe PDs from HCs and mild PDs (AUC = 0.71, F1-score = 0.30).

## 3) CDSS User Interface Improvement

As demonstrated in the AMIA Stage II submission, the majority of feedback from the dentists includes 1) providing full names of the risk factors information, 2) Changing the color of the bar graphs (red color for risk factors and green color for the protective factors), 3) Adding username and passwords for data protections, 3) showing SHAP values in percentages, and 4) adding genomics and proteomics data towards the prediction model. We addressed these suggestions except the #4, as we do not collect patients' genomics and proteomics data in the EHR. We presented the updated user interface (<https://perio-risk-scoring.herokuapp.com/>) to two dentists, and they found the updates helpful and easier to navigate the risk factor information compared to the previous user interface (strongly agree = 2). One dentist suggested further improvement in the risk assessment model, including 1) a breakdown of each medical condition affecting the PD risk and 2) a list of medications that might be responsible for increased periodontal risks. We already have obtained this information while running the machine learning algorithms. We will update the user interface based on these further suggestions.

## 4) Comparison of Variables Necessary for PD Risk Calculation in Five Tools

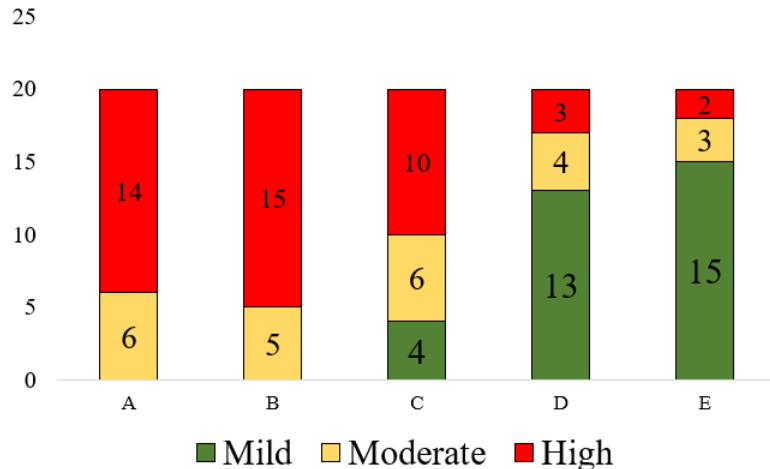
Table 1 demonstrates the variables included in each PD risk assessment tool included in this study. We found 74 unique variables when compiled from all five PD risk assessment tools. Out of 74 variables, only four variables (age, medical conditions, smoking habits, and periodontal bone loss) are common in all Tools (A-E). Next, variables such as the number of bleeding on probing sites (BOP), the number of sites with more than 5 mm of periodontal pocket depths, and past periodontal treatments were the second most common variables found in at least four out of five tools. PRSS had the maximum number of variables (N=74) that were included in the study. PRSS included 50 new variables that were highly associated with PD risk, which are missing from other tools (see Table 1). PRSS did not include four risk factors such as "use of mouth wash per day," "Periopathogenic bacteria," "susceptibility to PD," and "Oral-strain probiotics" because we do not collect this information in the EHR. Moreover, due to the retrospective nature of this study, it is not possible to contact these patients and obtain information about these variables. Our tool identified several medical conditions that have not obtained attention in dentistry. For example, we found that renal conditions, mental illnesses, neurodegenerative diseases, and hematological cancers were highly associated with the PD risk. More research studies are warranted to test the associations between these conditions and PD.

## 5) Comparison of Risk Scores Generated by Five PD Risk Assessment Tools

We obtained each of the variables described in Table 1 from a total of 20 patients' longitudinal dental EHR data. These twenty patients' information was entered into each risk assessment tool, and the risk scores were generated. Upon running the Kappa agreement, we found a Kappa value of -1.24 that indicates disagreement; this means that the risk scores provided by these five tools are significantly different. As shown in Figure 2, Tool A categorized 14 patients into high and six into moderate risk. Tool B categorized 15 patients into high and five patients into moderate risks. Tools D and E classified 13 and 15 patients into low, 4 and 3 patients into moderate, and 3 and 2 patients into high-risk groups, respectively.

## 6) Comparison of Risk Scores with Five Year Disease Progress Outcomes

As shown in Table 3, we found 70% agreement between PRSS and the original outcome, followed by Previser (55%), PRA (35%), Phillips (35%), and Cigna (25%). The Cohen's Kappa value between PRSS and the original outcome was [0.6 (moderate to high agreement)], followed by Previser [0.4 (low to moderate agreement)], PRA [0.3 (low agreement)], Phillips [0.3 (low agreement)], and Cigna [0.2 (no to low agreement)]. We also found the following risk and protective factors that are responsible for driving the risk of PD in these 20 patients (based on their influence on disease progression).



**Figure 2:** Risk assessment scores provided by five tools

The risk factors include high values of ASA classification (American Society of Anesthesiologists), higher caries risk assessment, number of teeth, smoking, bone loss, alcohol consumption, bleeding gums, multiple medical conditions, distance to the dental clinics, and such as. The protective factors include less DMFT/DMFS index, no to minimal attachment loss, higher brushing and flossing frequency, and presence of higher number of teeth (see Figure 3).

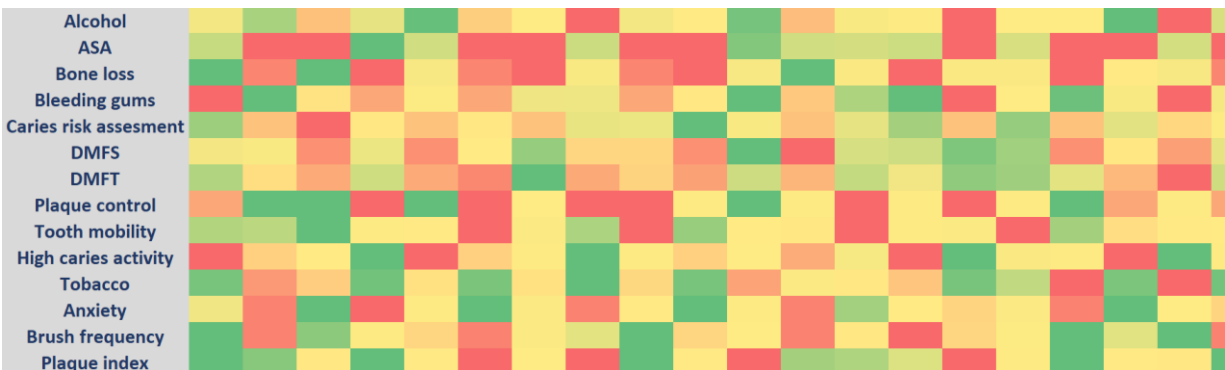
### Discussions

In the AMIA AI showcase series, we generated a cohort of patients with and without PD, cleaned and processed big dental EHR data, developed and tested various NLP pipelines, created a clean dataset, developed a machine learning-based prediction model, created CDSS, tested its usability, and implemented and compared the risk scores generated by different PD risk assessment tools. This was the first study that utilized big dental EHR data to develop a user-friendly CDSS to predict the risk of PD for clinical practice. In this study, we improved the user interface of our CDSS and implemented it to assess the risk of fifty selected patients. We also compared the risk scores generated by five PD risk assessment tools. The most significant findings of this study are 1) existing PD risk assessment tools lack involvement of comprehensive risk factors and variables to predict the risk of PD that have a positive association with PD, 2) disagreement between the risk scores provided by five tools, and 3) machine learning based prediction model (PRSS) provided the closest prediction compared to the original outcomes.

**Importance of adding multiple variables/risk factors in the prediction model for accurate risk scoring:** Variables and risk factors involved in the risk assessment tools are critical components affecting the accuracy of the model's performance. Assessing disease risk can be compared with the black box because it is challenging to study the number of all possible risk factors and their causations. For example, chemicals in the water we drink and the environment in which we live in could possibly affect periodontal health. However, it is not feasible to study all these variables. Hence, the inclusion of as many variables as possible may get us closer to the disease risk but may not provide 100% accuracy of the predictions due to many unknown factors/variables. The existing tools (tools A-D) are developed based on the risk factors studied more than two decades ago. However, the prevalence of risk factors has changed significantly; for example, smoking has been reduced, while obesity has increased significantly. Therefore, information from the current evidence needs to be utilized in developing and testing the prediction models. In our tool (PRSS), we utilized 74 unique variables, out of which 54 new variables were never utilized to predict PD risk (see Table 1). The associated risk factors can be modifiable or non-modifiable. The non-modifiable risk factors such as age, sex at birth, and genetics may play a role in the disease initiation and progression; however, we cannot control them.

On the other hand, modifiable risk factors are the risk factors that can be changed by the patient or by the clinicians. For instance, all new 54 associations and risk factors we determined as a result of this study are modifiable risk factors. So that dental clinicians can take a step back and provide counseling, recommendations, and procedures to address these risk factors. E.g., High decay missing filled teeth (DMFT) and decayed missing filled surface [DMFS (high caries/cavities)] should be treated, and patient education on diet, sealants, and brushing methods should be delivered to the patient to prevent the risk of PD. We also added many modifiable behavioral and social determinants of health factors, and these findings should be further confirmed by a larger sample, and policymakers should be informed to address some of the social barriers that patients face to receiving PD care.

**Moving from traditional regression and rule-based approaches to machine learning approaches:** The existing tools (tools A-D) are developed using rule-based algorithms. Based on the available evidence in the literature, each of the risk factors possesses a weight towards providing a risk score, and this information is obtained using simple statistical methods such as logistic regression models. However, there has been a significant advancement in statistical methods



**Figure 3:** Heatmap of risk factors and protective factors SHAP values in 20 patients included in this study

such as machine learning and deep learning that can provide better prediction and generalizability <sup>22</sup>. Moreover, the machine learning method handles confounding effects more accurately than the traditional statistical method. Unlike other tools, we utilized the XGBoost machine learning model to predict the risk of PD, which has been widely utilized and provided promising results. This method is highly flexible regarding big data analysis, power of parallel processing, gradient boosting, regulations, handling missing data, and cross-validation after each iteration <sup>23</sup>. This method also takes care about what risk factors are most relevant at the time and provides SHAP values based on their impact on the disease that other simple statistical methods do not provide.

**Advantages of our PRSS CDSS over other risk assessment tools:** As described earlier, existing literature demonstrates the development of expert-driven periodontal risk assessment tools such as periodontal risk assessment, Previser, and Miller-McEntire Periodontal Prognosis Scoring system. The PRA tool has a polygon surface that provides patients with periodontal risks into three categories that include low, moderate, or high-risk categories. The users are asked to enter patient information and the rule-based algorithm determines the risk of periodontitis using these manually entered risk factor information. Similarly, the PreViser risk assessment tool collects patients' similar demographics, periodontal findings, smoking, bone loss, and medical history information and provides risk into very low risk, low risk, moderate risk, high risk, or very high risk. PreViser is also a stand-alone tool as the periodontal risk assessment tool. Next, the Miller-McEntire Periodontal Prognosis Scoring system also provides periodontal risk using patients' objective and subjective criteria. This system is a paper-based system that provides the percentage of success of keeping the tooth. Our tool is significantly different than the existing systems/tools in the following ways. First, we developed automated processes to retrieve patient clinical variables from the dental EHR. Users do not need to manually input the parameters, thus greatly reducing the clinical labor in data manipulations. It also avoids any potential mistakes during the data entry process. Second, PRSS provides patients' risk scores between 0 and 100 which is more precise and personalized than categorizing patients into low, moderate, or high-risk categories. Next, PRSS also provides driving risk factors based on the weighting information, which is not offered by other tools. Lastly, PRSS provides both the risk and protective factors, while existing tools do not provide the protective features.

**Comparison of risk assessment tools studies:** Some studies have compared the performances of different tools by examining the same patients' PD risk by using different tools. These studies found that these tools had significantly different risk scores for the same patient. Hence, it was difficult for them to select one tool over the other to assess the patient's risk for PD<sup>6,12,14-16</sup>. They concluded that both (periodontal risk calculator versus periodontal risk assessment) tools help predict PD; however, it was necessary for them to use different tools to predict PD because of a lack of integration of all risk factors at a single place and time. They concluded that there is a need for a tool that has all possible risk factors in a single assessment system. Saleh et al. found that both periodontal risk assessment tools (PerioRisk and PRA) have the capability to predict the risk of PD; however, PerioRisk demonstrated the best discrimination and model fit, followed by PRA <sup>15</sup>. Moreover, we found that only our study compared the predictions of five risk assessment tools, while other studies compared two to three risk assessment tools' performances in predicting PD risk.

**Involvement of risk factors that are easy to collect in real life:** It is important to include those variables and risk factors information that is easy to collect during routine dental care. For example, medical history, dental history, dental anomalies, demographics, periodontal charting, social determinants of health, social vulnerable index, and oral hygiene are factors that are easily collected during routine patient care. However, some of the factors included in the



existing risk assessment tools like genetic test results, oral microbiome tests, periopathogenic bacteria, and oral-strain probiotics can be expensive for dental practitioners to perform. As a result, collecting this information to assess the risk of PD during routine dental care may not be feasible. PRSS only utilizes risk factors/variables collected in routine dental care, and the automated information extraction approach from dental EHR provides the feasibility of utilization of 74 risk factors in the risk assessment.

**Need for AI initiatives like AMIA AI showcase to develop, test, and implement AI tools:** We appreciate AMIA AI initiatives for organizing and managing AMIA AI showcase. This initiative calls for submissions for three AMIA conferences with different submission themes, including 1) Informatics Summit (AI technology development), 2) Clinical Informatics (system usability), and 3) Annual Symposium (Implementation). These themes provide a unique opportunity for researchers to develop AI systems, test their usability, and then implement the developed system in real-world clinical settings to determine its effectiveness. More opportunities like this should be provided to the informatics researchers and clinicians that allow them to complete and pilot test their systems. The critical feedback received from the committee, and peer reviewers have been extremely helpful in improving the system performance.

**Limitations:** Our study has several limitations. First, the sample used for comparison is small, and drawing solid conclusions from this small sample may not be appropriate. However, due to the time constraints (between each AMIA AI showcase submissions), we were only able to obtain 20 patients' information, including their radiographic findings and clinical charting (from all their visits), because the manual review was very time-consuming and required experts to review hundreds of pages of records. Second, the results of this study are not generalizable because we only have used one institute's dataset. Moreover, risk factors that are found to be significantly associated with PD are obtained from our patients. The weight assigned to each risk factor may vary depending on the geographical location. Despite this limitation, we were able to provide a complete framework on how to develop, test, and implement a CDSS system. By updating the dataset, we may be able to generate more generalizable results and identify additional risk factors based on different geographical locations. Next, there could be potential bias in the predicted outcome because we used the TUKSoD dataset to train and test our model. Then we used the same institute's data to validate the performance and compare risk scores provided by other tools. However, to reduce this bias, we compared only those patients' risk scores whose information was not included in the training or testing dataset.

**Future work:** We will use datasets from other institutions to train and test our prediction model to improve the generalizability of our model. Next, we will add more patient information to determine the differences in the prediction outcomes by five risk assessment tools. Due to the permission issues, we could not add Periodontal Management By Risk Assessment (PEMBRA) tool to this study and compare the risk score provided by PEMBRA. Therefore, we will obtain special permission and add risk assessment scores provided by this tool. Finally, we will implement PRSS at the chairside in the TUKSoD to determine the usefulness of our tool in real-world clinical settings.

## **Conclusion**

Five risk assessment tools (Tools A-E) provided significantly different risk scores for 20 unique patients. By examining the longitudinal outcome of the disease, PRSS provided the closest and more accurate risk assessment compared to the actual disease outcomes. We conclude that using advanced state-of-the-art informatics methods such as NLP and machine learning could help us utilize big EHR data optimally to represent the current patient populations and their risk factors, providing the most accurate disease risk score. This may lead to promoting preventive strategies at the chairside, hoping to reduce PD prevalence, improve quality of life, and reduce healthcare costs.

## **Acknowledgment**

We appreciate AMIA AI initiatives, especially the scientific program committee members, Dr. Gretchen Purcell Jackson, Dr. Laura Wiley, Dr. Paul Fu Jr., Dr. Rosemary Kennedy, and Ms. Dasha Cohen, for organizing and managing the AMIA AI showcase.

## **References**

1. Eke, P. I., Thornton-Evans, G. O., Wei, L., Borgnakke, W. S., Dye, B. A., & Genco, R. J. (2018). Periodontitis in US Adults: National Health and Nutrition Examination Survey 2009-2014. *Journal of the American Dental Association, 149*(7), 576-588.e6. <https://doi.org/10.1016/j.adaj.2018.04.023>
2. Tonetti, M. S., Greenwell, H., & Kornman, K. S. (2018). Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *Journal of Periodontology, 89*, S159–S172. <https://doi.org/10.1002/JPER.18-0006>
3. Bouchard, P., Carra, M. C., Boillot, A., Mora, F., & Rangé, H. (2017). Risk factors in periodontology: a conceptual framework. *Journal of Clinical Periodontology, 44*(2), 125–131.
4. Koshi, E., Rajesh, S., Koshi, P., & Arunima, P. R. (2012). Risk assessment for periodontal disease. *Journal of Indian Society of Periodontology, 16*(3), 324–328. <https://doi.org/10.4103/0972-124X.100905>

5. Eke, P. I., Zhang, X., Lu, H., Wei, L., Thornton-Evans, G., Greenlund, K. J., Holt, J. B., & Croft, J. B. (2016). Predicting periodontitis at state and local levels in the United States. *Journal of Dental Research*, *95*(5), 515–522.
6. Petersson, G. H., Åkerman, S., Isberg, P.-E., & Ericson, D. (2016). Comparison of risk assessment based on clinical judgement and Cariogram in addition to patient perceived treatment need. *BMC Oral Health*, *17*(1), 13.
7. Lang, N. P., Suvan, J. E., & Tonetti, M. S. (2015). Risk factor assessment tools for the prevention of periodontitis progression a systematic review. *Journal of Clinical Periodontology*, *42*(S16).
8. Meyer-Bäumer, A., Pritsch, M., Cosgarea, R., El Sayed, N., Kim, T.-S., Eickholz, P., & Pretzl, B. (2012). Prognostic value of the periodontal risk assessment in patients with aggressive periodontitis. *Journal of Clinical Periodontology*, *39*(7), 651–658. <https://doi.org/10.1111/j.1600-051X.2012.01895.x>
9. Page, R. C., & Eke, P. I. (2007). Case Definitions for Use in Population-Based Surveillance of Periodontitis - J Periodontal - July 2007 (SUPPL). *J Periodontol*. <https://doi.org/10.1902/jop.2007.060264>
10. Garcia, R., Compton, R., 2000, T. D.-P., & 2016, undefined. (n.d.). Risk assessment and periodontal prevention in primary care. *Wiley Online Library*. Retrieved November 12, 2018, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/prd.12124>
11. Shimpi, N., McRoy, S., Zhao, H., Wu, M., & Acharya, A. (2019). Development of a periodontitis risk assessment model for primary care providers in an interdisciplinary setting. *Technology and Health Care : Official Journal of the European Society for Engineering and Medicine*. <https://doi.org/10.3233/THC-191642>
12. Trombelli, L., Farina, R., Ferrari, S., Pasetti, P., & Calura, G. (2009). Comparison between two methods for periodontal risk assessment. *Minerva Stomatol*, *58*(6), 277–287.
13. Chandra, R. V. (2007). Evaluation of a novel periodontal risk assessment model in patients presenting for dental care. *Oral Health & Preventive Dentistry*, *5*(1), 39–48. <http://www.ncbi.nlm.nih.gov/pubmed/17366760>
14. Sai Sujai, G. V. N., Triveni, V. S. S., Barath, S., & Harikishan, G. (2015). Periodontal risk calculator versus periodontal risk assessment. *Journal of Pharmacy & Bioallied Sciences*, *7*(Suppl 2), S656-9. <https://doi.org/10.4103/0975-7406.163593>
15. Saleh, M. H. A., Dukka, H., Troiano, G., Ravidà, A., Qazi, M., Wang, H. L., & Greenwell, H. (2022). Long term comparison of the prognostic performance of PerioRisk, periodontal risk assessment, periodontal risk calculator, and staging and grading systems. *Journal of Periodontology*, *93*(1), 57–68. <https://doi.org/10.1002/JPER.20-0662>
16. Petsos, H., Arendt, S., Eickholz, P., Nickles, K., & Dannewitz, B. (2020). Comparison of two different periodontal risk assessment methods with regard to their agreement: Periodontal risk assessment versus periodontal risk calculator. *Journal of Clinical Periodontology*, *47*(8), 921–932. <https://doi.org/10.1111/JCPE.13327>
17. Ripan, R. C., Sarker, I. H., Hossain, S. M. M., Anwar, M. M., Nowrozy, R., Hoque, M. M., & Furhad, M. H. (2021). A Data-Driven Heart Disease Prediction Model Through K-Means Clustering-Based Anomaly Detection. *SN Computer Science 2021 2:2*, *2*(2), 1–12. <https://doi.org/10.1007/S42979-021-00518-7>
18. Douglass, C. W. (2006). Risk assessment and management of periodontal disease. *J Am Dent Assoc*, *137* Suppl, 27s-32s.
19. Patel J., Brandon R., Tellez M. Albandar M., Rao R., Joachim K., Su C., Wu, H. (2022). Developing Automated Computer Algorithms to Phenotype Periodontal Disease Diagnoses in Electronic Dental Records. *Journal of Public Health Dentistry*.
20. Patel J., Rao R., Tellez M. Albandar M., Joachim K., Wu, H. (2022). Natural Language Processing Applications to Extract Periodontal Disease and Medical History Information from Electronic Dental Records. *International Conference of Medical and Health Informatics*.
21. McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. <https://doi.org/10.11613/bm.2012.031>
22. Bzdok, D., Altman, N., & Krzywinski, M. (2018). Points of Significance: Statistics versus machine learning. *Nature Methods*, *15*(4), 233–234. <https://doi.org/10.1038/NMETH.4642>
23. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>