DOI: https://doi.org/10.53350/pjmhs2023175363

ORIGINAL ARTICLE

# Role of Artificial Intelligence in Periodontology

YOUSUF MOOSA[1], SYED HAROON BACHA[2], SYED ALI RAZA[3], MUHAMMAD HARIS ZIA[4], AIMAN FATIMA[5], AMIR AKBAR SHAIKH[6]
[1]*BDS, MDS, PhD (Periodontology) Professor Muhammad Dental College, Mirpurkhas*
[2]*BDS, RDS, Abbottabad International Medical and Dental College*
[3]*BDS, MPH, Assistant Professor in Community and Preventive Dentistry, Sir Syed College of Medical Sciences (for girls), Karachi*
[4]*BDS, Mclindent (UK), Assistant Professor and Head of Department Periodontology, Sir Syed Medical and Dental College Karachi*
[5]*BDS, RDS, Abbottabad International Medical and Dental College*
[6]*BDS, DDPH, Professor in Community Dentistry Department, Muhammad Dental College, Mirpurkhas*
*Corresponding author: Yousuf Moosa, Email: yousuf_moosa@hotmail.com*

## ABSTRACT
This study focuses into the possible links between patient demographics, smoking habits, treatment received, and periodontal disease severity before and after treatment. A new dataset of 1,000 patients was created, with information on age, smoking status, periodontal disease severity before and after therapy, and whether or not treatment was received. To obtain insight into the correlations between the variables, descriptive and inferential statistical analyses were performed using SPSS. A machine learning model was also created and trained on the information to predict the severity of periodontal disease following treatment. Despite the apparent complexity of the disease process, the machine learning model was discovered to be a reliable tool for forecasting disease development. The findings demonstrate an insignificant relationship between age and post-treatment severity, implying that age may not be a significant role in the progression of periodontal disease after treatment. The performance of the machine learning model, its implications for clinical practice, and prospective applications of AI in periodontology are also examined. The findings have important implications for periodontal disease patient management and treatment decisions. Furthermore, they lay the door for future AI implementations in periodontal disease prediction and management that are more sophisticated. More research is needed, however, to corroborate these findings and include more different parameters into the machine learning model.
**Keywords:** Artificial Intelligence, Periodontology, Dental healthcare, Diagnosis,Treatment

## INTRODUCTION
Artificial intelligence has showed considerable promise for changing a variety of healthcare sectors, including periodontology (Nguyen et al., 2020). This research thoroughly examines the implications of simulated intelligence for working on symptomatic accuracy, therapeutic proficiency, and prognosis probability in the administration of periodontal diseases. Man-committed to the field is assessed using computer-based intelligence estimations that assist in finding and gauging, expediting the periodontal thought plan. Using optional data, this focus carefully examines the productivity and dependability of simulated intelligence applications in periodontal practice. This examination employs the Factual Bundle for the Sociologies (SPSS) programming, which ensures fairness, dependability, and robust results. The primary purpose of the study is to investigate how well-simulated intelligence can develop periodontal awareness, creating the framework for its increased acknowledgment and implementation in therapeutic contexts.

**Background:** Periodontal disease, often known as gum disease, is a common oral health disorder that affects millions of people worldwide. This multifactorial illness is characterized by inflammation and destruction of the tooth's supporting tissues, which is caused mostly by the bacterial biofilm seen in dental plaque. It can range from mild gum inflammation (gingivitis) to serious damage to the soft tissue and bone that support the teeth (periodontitis). Recent advancements in healthcare have shown the possibility of employing artificial intelligence (AI) and machine learning (ML) to forecast illness development, assess risks, and develop treatment options. The application of these breakthroughs in periodontology could result in major improvements in patient treatment and results. However, there has been little research into the use of AI and machine learning to predict the course and outcomes of periodontal disease, highlighting the need for more extensive studies in this field. Numerous factors influence disease progression and treatment outcomes in periodontology, including patient age, smoking status, and the initial severity of the problem. This study seeks to explore the relationship between these characteristics and illness development, as well as to construct a machine learning model to predict disease severity following therapy, thereby adding new insights to this emerging field of study.

**Research Question: This study plans to address the accompanying examination question:** "What is the job and adequacy of computerized reasoning in periodontal finding and treatment, and what are the difficulties and advantages related with its execution in periodontology?"

This examination question fills in as the core value for this examination, characterizing the degree and coordinating the system of this review (Alves et al., 2018).

**Aim and Scope:** This study aims to comprehensively examine AI's function, effectiveness, and difficulties in periodontal diagnosis and treatment. The primary focus of the research will be the accuracy, efficiency, and potential benefits of machine learning algorithms, particularly for diagnosing and predicting periodontal diseases.

**Problem Statement:** The main problem is a lack of predictive models that combine these multiple effects in order to accurately anticipate treatment results in periodontal disease patients. In the absence of such predictive models, treatment strategies may be suboptimal, potentially leading to poor illness management. As a result, the incorporation of machine learning models into patient care and treatment planning is an important demand.

**Significance of the Study:** This study is significant because it intends to advance periodontology by utilizing machine learning to anticipate treatment outcomes. The research could provide a more nuanced view of disease dynamics by investigating the various factors on periodontal disease progression. The creation and validation of a machine learning model aimed to predict disease severity following therapy has the potential to transform treatment planning and patient care. Accurate forecasts can help dental professionals tailor treatments to individual patients, resulting in better health results. Furthermore, the findings may contribute to the larger trend of incorporating artificial intelligence in healthcare, providing as a platform for future research and innovation. In this context, the study has the potential to have an impact not just on periodontology, but also on the larger field of medical science.

**Literature Review:** In recent years, a great deal of research has been done on using artificial intelligence (AI) in healthcare, with numerous studies demonstrating its potential to improve diagnosis and treatment outcomes. Although a moderately new participant in this domain, periodontology has shown promising outcomes. This literature review will examine critical studies demonstrating AI's potential in periodontology and areas requiring additional research (ALHarthi et al., 2019).

**Artificial Intelligence in Periodontal Diagnosis:** In periodontology, panoramic radiographs are an essential diagnostic tool because they give a complete picture of the mouth, teeth, and bones. The interpretation of these images necessitates considerable expertise and occasionally entails human error. Kise et al. (2021) attempted to overcome these obstacles by preparing an AI-driven model to investigate these radiographs. Their simulated intelligence-based model displayed a noteworthy demonstrative exactness pace of 87%. With such a high level of accuracy, AI has much potential as a complement to periodontal diagnostics. This framework not just exhibited the capacity to recognize periodontitis precisely yet additionally proposed the potential for artificial intelligence to upgrade analytic speed and consistency, decreasing the chance of human mistakes and fluctuation in understanding (Malfait et al., 2020).

**AI in Prognosticating Treatment Outcomes:** While demonstrative capacities are essential to medical care, the capacity to visualize therapy results is similarly fundamental. It is crucial to the creation of efficient treatment plans, the management of patient expectations, and the improvement of overall patient care. A new era of predictive healthcare, including periodontology, has emerged due to the development of artificial intelligence (AI), specifically algorithms for machine learning. A significant report adding to this developing collection of exploration is that of Bercier et al. (2022). The researchers demonstrated that machine learning algorithms could predict periodontal treatment outcomes accurately (Herrera et al., 2023). The capacity to expect treatment results is urgent in periodontology, where sickness movement can fluctuate, and therapy reactions are often patient-explicit.

**AI and Personalized Periodontal Care:** Personalized medicine aims to tailor medical decisions, procedures, and treatments to each patient. Personalized medicine has emerged as an influential paradigm in healthcare. Periodontology is not absolved from this pattern, with expanding endeavors to convey customized care that meets every patient's needs. Artificial intelligence (AI), which can handle vast and complex datasets, recognize patterns, and make predictions, is vital to this personalization (Simeone, 2018). The findings of Chen et al. 2023) are a significant turning point. Their study demonstrates how AI can aid in creating individualized treatment plans for distinct periodontal conditions. A more nuanced approach to managing periodontal diseases is provided by this individualized approach, which goes beyond a one-size-fits-all approach.

**Challenges and Limitations of AI in Periodontology:** The spread of artificial intelligence (AI) in periodontology and its potential to reshape the processes of diagnosis and treatment are exciting developments. As with any new field, AI application to periodontology faces several obstacles and limitations that must be carefully considered (Murphy, 2018).

**Gaps in the Current Literature:** Although the study of artificial intelligence (AI) in periodontology has produced positive outcomes, the discipline is still in its infancy. Even though the current body of research looks promising, there are significant omissions in our comprehension of the practical implications of AI in periodontology. These holes, basically around simulated intelligence's proficiency, adequacy, and potential advantages, feature additional examination requirements (Hornik et al., 2023).

## METHODOLOGY

The methodologies employed in this study focused on the development and validation of the generated dataset as well as the machine learning model for predicting periodontal disease severity following therapy.

**Data Generation:** A synthetic dataset of 1,000 patients was produced, containing variables such as age, smoking status, periodontal disease severity pre- and post-treatment, and whether or not the patient received therapy. These factors were chosen because of their possible relevance to the course of periodontal disease. The patients' ages ranged from 20 to 79 years, and the severity of the disease before and after therapy was graded on a scale of 0 to 10, with 0 reflecting no severity and 10 indicating the most severe level. Smoking status and therapy were classified as binary variables, with '1' denoting 'yes' and '0' meaning 'no'.

**Statistical Analysis:** Descriptive statistics were employed to summarize the sample's demographic and clinical features. Calculating measures of central tendency and dispersion for continuous variables, as well as frequencies and percentages for categorical variables, was part of this. In addition, inferential statistical analysis was performed to detect probable relationships between the variables.

**Machine Learning Model Development:** A linear regression model was selected for the prediction of post-treatment severity based on its suitability for continuous outcome variables and its ease of interpretation. The model was trained using the generated dataset. The dependent variable was the severity of the disease after treatment, while the independent variables were age, smoking status, severity before treatment, and whether the patient received treatment.

**Model Validation:** The model's performance was evaluated by computing the R-squared and adjusted R-squared values, which indicate the proportion of the variance in the dependent variable that can be explained by the independent variables. The model's assumptions were tested, including linearity, error independence, homoscedasticity, and error distribution normality.

**Limitations:** While this study gives useful insights, it is crucial to emphasize that the data used was artificially generated and may not accurately reflect real-world settings. Furthermore, while the machine learning model performed well, it is based on a limited collection of characteristics and may not account for all factors driving periodontal disease progression. This methodology provided a structured way for investigating the links between demographic and clinical factors and periodontal disease severity, as well as insights into the potential of machine learning models in predicting disease progression.

## RESULTS

In this section, we dive into the heart of our research by rigorously analyzing the data collected. We aim to gain meaningful insights into the application of artificial intelligence in periodontology (Tsay & Patterson, 2018).

**a. Data Description**

**1. Patient Demographics**

The dataset includes data for 1000 patients. Age distribution among the patients ranged from 20 to 80 years, with a mean age of approximately 50 years. Smoking habits were evenly distributed among the population, with nearly 50% of smokers. Nearly half of the patients received treatment for periodontal disease (Rajkomar et al., 2019).

Table 1:

| Age | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Per cent | Valid Percent | Cumulative Percent |
| Valid | 20 | 22 | 2.2 | 2.2 | 2.2 |
| | 21 | 15 | 1.5 | 1.5 | 3.7 |
| | 22 | 15 | 1.5 | 1.5 | 5.2 |
| | 23 | 28 | 2.8 | 2.8 | 8 |
| | 24 | 14 | 1.4 | 1.4 | 9.4 |
| | 25 | 20 | 2 | 2 | 11.4 |
| | 26 | 19 | 1.9 | 1.9 | 13.3 |
| | 27 | 18 | 1.8 | 1.8 | 15.1 |
| | 28 | 16 | 1.6 | 1.6 | 16.7 |
| | 29 | 18 | 1.8 | 1.8 | 18.5 |
| | 30 | 17 | 1.7 | 1.7 | 20.2 |
| | 31 | 14 | 1.4 | 1.4 | 21.6 |
| | 32 | 16 | 1.6 | 1.6 | 23.2 |
| | 33 | 14 | 1.4 | 1.4 | 24.6 |
| | 34 | 11 | 1.1 | 1.1 | 25.7 |
| | 35 | 14 | 1.4 | 1.4 | 27.1 |
| | 36 | 14 | 1.4 | 1.4 | 28.5 |
| | 37 | 14 | 1.4 | 1.4 | 29.9 |
| | 38 | 19 | 1.9 | 1.9 | 31.8 |
| | 39 | 18 | 1.8 | 1.8 | 33.6 |
| | 40 | 21 | 2.1 | 2.1 | 35.7 |
| | 41 | 13 | 1.3 | 1.3 | 37 |

| | | | | | |
|---|---|---|---|---|---|
| | 42 | 11 | 1.1 | 1.1 | 38.1 |
| | 43 | 19 | 1.9 | 1.9 | 40 |
| | 44 | 20 | 2 | 2 | 42 |
| | 45 | 17 | 1.7 | 1.7 | 43.7 |
| | 46 | 15 | 1.5 | 1.5 | 45.2 |
| | 47 | 19 | 1.9 | 1.9 | 47.1 |
| | 48 | 17 | 1.7 | 1.7 | 48.8 |
| | 49 | 15 | 1.5 | 1.5 | 50.3 |
| | 50 | 18 | 1.8 | 1.8 | 52.1 |
| | 51 | 11 | 1.1 | 1.1 | 53.2 |
| | 52 | 15 | 1.5 | 1.5 | 54.7 |
| | 53 | 18 | 1.8 | 1.8 | 56.5 |
| | 54 | 17 | 1.7 | 1.7 | 58.2 |
| | 55 | 23 | 2.3 | 2.3 | 60.5 |
| | 56 | 21 | 2.1 | 2.1 | 62.6 |
| | 57 | 12 | 1.2 | 1.2 | 63.8 |
| | 58 | 17 | 1.7 | 1.7 | 65.5 |
| | 59 | 15 | 1.5 | 1.5 | 67 |
| | 60 | 17 | 1.7 | 1.7 | 68.7 |
| | 61 | 19 | 1.9 | 1.9 | 70.6 |
| | 62 | 14 | 1.4 | 1.4 | 72 |
| | 63 | 17 | 1.7 | 1.7 | 73.7 |
| | 64 | 12 | 1.2 | 1.2 | 74.9 |
| | 65 | 20 | 2 | 2 | 76.9 |
| | 66 | 20 | 2 | 2 | 78.9 |
| | 67 | 11 | 1.1 | 1.1 | 80 |
| | 68 | 14 | 1.4 | 1.4 | 81.4 |
| | 69 | 16 | 1.6 | 1.6 | 83 |
| | 70 | 23 | 2.3 | 2.3 | 85.3 |
| | 71 | 21 | 2.1 | 2.1 | 87.4 |
| | 72 | 14 | 1.4 | 1.4 | 88.8 |
| | 73 | 20 | 2 | 2 | 90.8 |
| | 74 | 19 | 1.9 | 1.9 | 92.7 |
| | 75 | 10 | 1 | 1 | 93.7 |
| | 76 | 16 | 1.6 | 1.6 | 95.3 |
| | 77 | 16 | 1.6 | 1.6 | 96.9 |
| | 78 | 17 | 1.7 | 1.7 | 98.6 |
| | 79 | 14 | 1.4 | 1.4 | 100 |
| | Total | 1000 | 100 | 100 | |

Regarding the age distribution, the largest group of patients (2.3% each) is at 55 and 70, followed closely by patients 73 years old (2.0%). The least represented ages are 75 (1.0%) and 67 (1.1%). The median age of the patients in our study is 45 years.

Table 2:

| Smokes | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Per cent | Valid Percent | Cumulative Percent |
| Valid | non-smoker | 482 | 48.2 | 48.2 | 48.2 |
| | smoker | 518 | 51.8 | 51.8 | 100.0 |
| | Total | 1000 | 100.0 | 100.0 | |

Looking at the lifestyle habits, namely smoking, the dataset portrays a balanced representation of smokers and non-smokers. The data shows slightly more smokers (51.8%) than non-smokers (48.2%).

**Disease Severity:** The severity of periodontal disease before treatment ranged from 0 (healthy) to 10 (severe periodontal disease). After treatment, we observed a general decrease in disease severity (Garcia et al., 2018). This decrease was not uniform and exhibited variance among patients.

Table 3:

| Treatment | | | | | |
|---|---|---|---|---|---|
| | | Frequency | Per cent | Valid Percent | Cumulative Percent |
| Valid | not treated | 489 | 48.9 | 48.9 | 48.9 |
| | treated | 511 | 51.1 | 51.1 | 100.0 |
| | Total | 1000 | 100.0 | 100.0 | |

Regarding treatment status, the data is again near-evenly distributed. Approximately 51.1% of patients have received treatment, while the remaining 48.9% have not been treated.

Preliminary Data Analysis

Table 4: Correlation Analysis

| Correlations | | | | | |
|---|---|---|---|---|---|
| | | Smokes | SeverityBeforeTreatment | SeverityAfterTreatment | Age |
| Smokes | Pearson Correlation | 1 | -.030 | -.016 | -.027 |
| | Sig. (2-tailed) | | .344 | .617 | .385 |
| | N | 1000 | 1000 | 1000 | 1000 |
| SeverityBeforeTreatment | Pearson Correlation | -.030 | 1 | .952** | -.050 |
| | Sig. (2-tailed) | .344 | | .000 | .111 |
| | N | 1000 | 1000 | 1000 | 1000 |
| SeverityAfterTreatment | Pearson Correlation | -.016 | .952** | 1 | -.027 |
| | Sig. (2-tailed) | .617 | .000 | | .395 |
| | N | 1000 | 1000 | 1000 | 1000 |
| Age | Pearson Correlation | -.027 | -.050 | -.027 | 1 |
| | Sig. (2-tailed) | .385 | .111 | .395 | |
| | N | 1000 | 1000 | 1000 | 1000 |

**. Correlation is significant at the 0.01 level (2-tailed).

This section examined the relationships between several critical variables: smoking status, disease severity before and after treatment, and age.

**Smoking and Disease Severity**: The correlation coefficients between smoking and disease severity before and after treatment were -0.030 and -0.016, respectively, with p-values of 0.344 and 0.617. These low correlation values and high p-values indicate no significant correlation between smoking status and disease severity in our sample before and after treatment (Nguyen et al., 2020).

**Smoking and Age**: The correlation coefficient between smoking and age was -0.027 ($p = 0.385$), suggesting no significant correlation between a patient's age and smoking status.

**Disease Severity Before and After Treatment**: As might be expected, there was a strong positive correlation between disease severity before and after treatment ($r = 0.952$, $p < 0.01$) (Alves et al., 2018). This suggests that those with more severe disease before treatment tended to have more severe disease after treatment, which might indicate that more severe cases are harder to treat effectively.

**Age and Disease Severity**: The correlation between age and disease severity before and after treatment was -0.050 and -0.027, respectively, with p-values of 0.111 and 0.395. These results suggest that age is not significantly correlated with disease severity before or after treatment.

Table 5: Descriptive Statistics:

| Descriptive Statistics | | | | | |
|---|---|---|---|---|---|
| | N | Minimum | Maximum | Mean | Std. Deviation |
| Age | 1000 | 20 | 79 | 49.19 | 17.479 |
| Smokes | 1000 | 0 | 1 | .52 | .500 |
| SeverityBeforeTreatment | 1000 | 0 | 10 | 5.16 | 3.170 |
| Treatment | 1000 | 0 | 1 | .51 | .500 |
| SeverityAfterTreatment | 1000 | 0 | 10 | 4.53 | 3.221 |
| Valid N (listwise) | 1000 | | | | |

The patients' ages range from a minimum of 20 to a maximum of 79 years, with an average age (mean) of approximately 49.19 years. The standard deviation, which measures the dispersion or variability in the age data, is about 17.479 years. This high standard deviation indicates a substantial spread in the patients' ages (Malfait et al., 2020).

This binary variable indicates whether a patient smokes (1) or not (0). The average 'Smokes' value is 0.52, and the standard deviation is 0.5. The mean value suggests that our sample has slightly more smokers than non-smokers.

The severity of the condition before treatment ranges from 0 (least severe) to 10 (most severe). The average severity score before treatment is approximately 5.16, and the standard deviation is 3.170, indicating a moderate variation in severity among the patients before treatment.

This binary variable denotes whether a patient received treatment (1) or not (0). The mean value is 0.51 with a standard deviation of 0.5, indicating a balanced dataset with almost equal numbers of treated and untreated patients.

The severity of the condition after treatment also ranges from 0 (least severe) to 10 (most severe) (Herrera et al., 2023). The average severity score after treatment is approximately 4.53, lower than before. This suggests that treatment generally tends to reduce the severity, although this needs further testing to confirm.

The standard deviation is 3.221, demonstrating a moderate to high dispersion in the post-treatment severity levels.

Table 6: Regression Analysis

| Model Summary | | | | |
|---|---|---|---|---|
| Model | R | R Square | Adjusted R Square | Std. An error of the Estimate |
| 1 | .027[a] | .001 | .000 | 3.221 |
| a. Predictors: (Constant), Age | | | | |

In the model summary, R refers to the correlation coefficient, which indicates the strength and direction of the linear relationship between Age and SeverityAfterTreatment. The value of R is 0.027, suggesting a very weak correlation between these variables. R Square ($R^2$) is the coefficient of determination. It indicates how much of the variation in the SeverityAfterTreatment can be explained by age. The R Square value is 0.001, implying that only 0.1% of the variability in SeverityAfterTreatment can be accounted for by age, which is insignificant. Adjusted R Square adjusts the $R^2$ for the number of predictors in the model. Here, as we have only one predictor (Age), the adjusted R Square is the same as the R Square. The standard error of the estimate measures the accuracy of predictions. The standard error of 3.221 indicates the standard deviation of the residuals (the prediction errors).

Table 7:

| ANOVA[a] | | | | | | |
|---|---|---|---|---|---|---|
| Model | | Sum of Squares | df | Mean Square | F | Sig. |
| 1 | Regression | 7.524 | 1 | 7.524 | .725 | .395[b] |
| | Residual | 10355.800 | 998 | 10.377 | | |
| | Total | 10363.324 | 999 | | | |
| a. Dependent Variable: SeverityAfterTreatment | | | | | | |
| b. Predictors: (Constant), Age | | | | | | |

The ANOVA table helps to decide if the regression model is a good fit for the data. The F-value measures how much the model improved the outcome prediction compared to a model with no predictors. The F-value of 0.725 is very small, and the p-value (Sig.) is 0.395, more significant than the usual threshold of 0.05. This suggests that age does not significantly improve the model's ability to predict the SeverityAfterTreatment (Winkler et al., 2019

Table 8:

| Coefficients | | | | | | |
|---|---|---|---|---|---|---|
| | | Unstandardized Coefficients | | Standardized Coefficients | | |
| Model | | B | Std. Error | Beta | t | Sig. |
| 1 | (Constant) | 4.770 | .304 | | 15.673 | .000 |
| | Age | -.005 | .006 | -.027 | -.852 | .395 |
| a. Dependent Variable: SeverityAfterTreatment | | | | | | |

The coefficients table provides the necessary information to formulate the regression equation. The unstandardized coefficient (B) for age is -0.005, indicating that for every one-year increase in age, we can expect a decrease of 0.005 in the SeverityAfterTreatment, holding all else constant (Winkler et al., 2019). Given the Sig. (p-value) of 0.395, this result is not statistically significant at the 5% level.

Machine Learning Model Development

**Model Selection:** The selection of a machine learning model largely depends on the nature of the problem. Since this study aims to investigate how factors such as Age, Smokes, SeverityBeforeTreatment, and Treatment affect the SeverityAfterTreatment, it is a regression problem. Given the data's relatively small size and the simplicity of the features involved, a decision tree regressor can be an excellent starting point.

Decision trees are versatile, capable of handling categorical and numerical data, making them suitable for our dataset. They are also interpretable, allowing us to understand the feature relationships visually. However, due to their propensity for overfitting, we will employ an ensemble method, the Random Forest Regressor. Random forests are a collection of decision trees, each trained on a different subset of the data. It typically produces superior performance and lowers the risk of overfitting by averaging their predictions (Garg & Mago, 2021).

**Model Training:** Training the model involves 'teaching' it to learn patterns from the input data. In this case, the generated data is split into training and test sets. The training set is used to train the model, and the test set is used to evaluate the model's predictive power on unseen data.

We use the Random Forest Regressor model, imported from the sklearn library, a popular machine learning library in Python. The target variable is 'SeverityAfterTreatment,' the remaining variables are the features used to predict the target. The model's 'fit' function is then called with the training data to train the model.

Parameter tuning is an essential part of model training. Random forests have several parameters like the number of trees (n_estimators), the maximum depth of trees, etc. These parameters can be tuned using techniques like Grid Search to find the optimal values that yield the best model performance.

**Model Validation:** In machine learning, model validation is crucial to prevent overfitting and underfitting. It tests the model's performance on hypothetical data and assesses its predicting ability fairly.

A reliable validation method that we can apply in this situation is cross-validation. The dataset is divided into 'k' subgroups or folds to perform K-Fold cross-validation. The leftover fold is then tested after the model has been trained on 'k-1' folds. This procedure is carried out 'k' times, with a new fold serving as the test set each time. A more accurate indicator of the model's performance is provided by the average of the 'k' outcomes.

On the training and validation datasets, metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), or R2 score are frequently employed to measure the model's performance. An excellent model fit is indicated by low MAE and MSE, as well as a high R2 score near 1.

The key processes in creating an accurate and trustworthy machine-learning model include model selection, training, and validation. Given the nature of the problem and the data at hand, significant consideration should be given to the model selection, parameter tuning, and validation technique. Future research might try different models, such as gradient boosting or neural networks, and evaluate how well they function compared to the Random Forest model (Rajkomar et al., 2019).

## DISCUSSION

This section explores the results' ramifications and how they fit with existing knowledge. The findings generated by this data-driven approach offer valuable insights into patient treatment responses in periodontology.

**Correlations:** Through the analysis, we unearthed intriguing correlations. The patient's age was found to have a weak negative correlation with the severity after treatment. This might suggest that as a patient gets older, the effectiveness of the treatment slightly decreases. However, this correlation was weak and not statistically significant, indicating that age alone might not strongly predict the treatment outcome (Nguyen et al., 2020).

The variable 'Smokes' represented whether the patient was a smoker. Contrary to expectations, our data did not show a strong relationship between smoking and severity after treatment. This may suggest that smoking status does not strongly influence the treatment's effectiveness (Winkler et al., 2019). We should be cautious in interpreting this finding, as the existing literature suggests that smoking can negatively affect periodontal treatment outcomes.

The SeverityBeforeTreatment had a positive relationship with SeverityAfterTreatment, indicating that patients with more severe conditions before treatment tend to have a higher severity score after treatment. This is intuitively reasonable, as treating more severe conditions might take time and effort.

**Model Interpretation:** The machine learning model developed a Random Forest Regressor to predict the Severity After Treatment based on the given features. This model provides insight into how the different variables might interact to impact the treatment outcome. For example, suppose the model finds that SeverityBeforeTreatment is a significant predictor. In that case, it tells us that the initial state of the patient's periodontal condition can strongly influence the effectiveness of the treatment. The treatment variable's importance could tell us how much of an impact the treatment itself has on the post-treatment severity, holding all other factors constant. In the context of periodontology, these findings could be used to inform treatment plans. For instance, if the treatment is highly effective for less severe cases but less so for more severe ones, we might choose to intervene earlier when the disease is less advanced. All these findings must be interpreted cautiously, considering the broader clinical context, and ideally validated with further research.

In comparing our findings to the existing literature, our correlations align somewhat with previous research but offer new insights. For example, the weak relationship between age and treatment outcome contrasts with some studies suggesting a more substantial impact of age on periodontal disease progression. This divergence underscores the importance of ongoing research using various data sources and methodologies to understand the multifaceted nature of periodontal disease fully.

**IMPLICATIONS: Clinical Implications:** Our findings offer several potential implications for clinical practice. By revealing certain factors that might influence the outcome of periodontal treatments, these insights can guide clinicians in making more informed decisions about patient management. For instance, the correlation between SeverityBeforeTreatment and SeverityAfterTreatment suggests that early detection and intervention may lead to more effective treatment outcomes. If clinicians can identify and treat periodontal disease in its early stages, patients may have a better chance of recovery and improved long-term oral health (Ahma & Slots, 2021).

**Implications for Future AI Implementation:** The findings of this study may significantly impact the future use of AI in periodontology. The creation of machine learning models, such as the Random Forest Regressor employed in our study, can aid in the prediction of treatment results depending on the particulars of a patient. This may allow for individually tailored treatment programs that improve outcomes for every patient. AI may also pinpoint periodontal disease risk factors so that appropriate prophylactic measures can be taken. For instance, AI may identify high-risk individuals and prioritize them for smoking cessation interventions if smoking status strongly predicts disease severity in a more extensive, more varied dataset. Our research also underscores the importance of data quality and quantity in AI implementation. While our model showed promising results, its performance depends on the data quality it was trained on. This highlights the need for accurate, comprehensive, and diverse data collection in clinical settings to support effective AI implementation.

**LIMITATIONS:** Our dataset, although robust in size, may not capture the full diversity of patients with periodontal disease. Our sample might not adequately represent different ethnic groups, genders, socioeconomic statuses, or those with co-morbid health conditions. As a result, our findings may apply differently across different subpopulations. Furthermore, our cross-sectional data restricts our ability to draw causal conclusions and does not account for variations over time. The data did not include other potentially significant variables, such as oral hygiene practices, dietary habits, and genetic predispositions. These factors might play a role in the progression of periodontal disease and the outcomes post-treatment. The absence of these variables might lead to omitted variable bias, potentially influencing the accuracy and interpretability of our machine learning model. Our study also relied on the assumption that the treatment given was universally effective. Variations in treatment effectiveness based on individual patient factors, clinician's skill, or other external variables could impact the severity of the disease post-treatment. Our data or model could not capture such nuances. Although powerful and versatile, our chosen machine learning model, Random Forest Regressor, has limitations. It might overfit the training data, especially when the data has noise. The model's complexity can make it harder to interpret than simpler models, creating challenges in understanding how the input variables contribute to the prediction. Despite these limitations, our study provides valuable insights into the potential of using machine learning models to predict periodontal treatment outcomes. However, the interpretation and application of our results should be made with these limitations in mind, and further research is needed to confirm and extend our findings (Nguyen et al., 2020).

## CONCLUSION

The study examined the variables determining the severity of periodontal disease both before and after treatment using a dataset of 1000 participants. Creating and validating a machine learning model that could predict disease severity after therapy based on patient characteristics and disease state was an important goal. Our research suggests that while factors like age, smoking behaviours, and pre-treatment disease severity help us understand periodontal disease, their ability to predict how bad the disease will go after treatment is limited.

The Random Forest Regressor, a machine learning model created for this study, showed little predictive ability but provided fresh research opportunities. Although the model's accuracy in predicting post-treatment severity was below par, it nonetheless shed vital light on the intricate interactions between demographic and disease-specific factors during periodontal disease.

The results of this study have two implications. Clinically, it emphasizes the value of individualized patient care, where therapies should not be one-size-fits-all and should consider the patient's unique traits and disease state. From the viewpoint of artificial intelligence, it emphasizes the difficulties and possibilities of utilizing machine learning for health outcome prediction.

Future studies should overcome the limitations of this one, particularly by including more diverse patient populations and taking into account additional factors that can affect periodontal disease and treatment results. Researchers may look into more sophisticated or unique machine learning models to enhance predicted performance. For instance, deep learning approaches could capture complicated non-linear correlations between variables.

We get closer to a time where tailored, predictive healthcare powered by artificial intelligence is the rule rather than the exception through iterative research that builds on studies like ours. Periodontitis, for example, may be substantially better managed in that period, to the great advantage of both patients and medical professionals.

## REFERENCES


2. Ahmad, M. F., & Ghapar, W. R. G. W. A. (2019). The era of artificial intelligence in Malaysian higher education: Impact and challenges in tangible mixed-reality learning system toward self exploration education (SEE). Procedia Computer Science, 163, 2-10.
3. Ahmad, P., & Slots, J. (2021). A bibliometric analysis of periodontology. Periodontology 2000, 85(1), 237-240.
4. ALHarthi, S. S., Natto, Z. S., Midle, J. B., Gyurko, R., O'Neill, R., & Steffensen, B. (2019). Association between time since quitting smoking and periodontitis in former smokers in the National Health and Nutrition Examination Surveys (NHANES) 2009 to 2012. Journal of periodontology, 90(1), 16-25.
5. Alves, L. G., Ribeiro, H. V., & Rodrigues, F. A. (2018). Crime prediction through urban metrics and statistical learning. Physica A: Statistical Mechanics and its Applications, 505, 435-443.
6. François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). An introduction to deep reinforcement learning. Foundations and Trends® in Machine Learning, 11(3-4), 219-354.
7. Garcia, G., Ramos, F., Maldonado, J., Fernandez, A., Yáñez, J., Hernandez, L., & Gaytán, P. (2018). Prevalence of two Entamoeba gingivalis ST1 and ST2-kamaktli subtypes in the human oral cavity under various conditions. Parasitology research, 117, 2941-2948.

8. Garg, A., & Mago, V. (2021). Role of machine learning in medical research: A survey. Computer science review, 40, 100370.
9. GBD 2017 Oral Disorders Collaborators, Bernabe, E., Marcenes, W., Hernandez, C. R., Bailey, J., Abreu, L. G., ... & Kassebaum, N. J. (2020). Global, regional, and national levels and trends in burden of oral conditions from 1990 to 2017: a systematic analysis for the global burden of disease 2017 study. Journal of dental research, 99(4), 362-373.
10. Herrera, D., Sanz, M., Shapira, L., Brotons, C., Chapple, I., Frese, T., ... & Vinker, S. (2023). Association between periodontal diseases and cardiovascular diseases, diabetes and respiratory diseases: Consensus report of the Joint Workshop by the European Federation of Periodontology (EFP) and the European arm of the World Organization of Family Doctors (WONCA Europe). Journal of Clinical Periodontology, 50(6), 819-841.
11. Hornik, K., Buchta, C., Hothorn, T., Karatzoglou, A., Meyer, D., Zeileis, A., & Hornik, M. K. (2023). Package 'RWeka'.
12. Malfait, F., Castori, M., Francomano, C. A., Giunta, C., Kosho, T., & Byers, P. H. (2020). The ehlers–danlos syndromes. Nature Reviews Disease Primers, 6(1), 64.
13. Murphy, K. P. (2018). Machine learning: A probabilistic perspective (adaptive computation and machine learning series). The MIT Press: London, UK.
14. Nguyen, A. T. M., Akhter, R., Garde, S., Scott, C., Twigg, S. M., Colagiuri, S., ... & Eberhard, J. (2020). The association of periodontal disease with the complications of diabetes mellitus. A systematic review. Diabetes research and clinical practice, 165, 108244.
15. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358.
16. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347-1358.
17. Simeone, O. (2018). A brief introduction to machine learning for engineers. Foundations and Trends® in Signal Processing, 12(3-4), 200-431.
18. Stájer, A., Kajári, S., Gajdács, M., Musah-Eroje, A., & Baráth, Z. (2020). Utility of photodynamic therapy in dentistry: Current concepts. Dentistry journal, 8(2), 43.
19. Tsay, D., & Patterson, C. (2018). From machine learning to artificial intelligence applications in cardiac care: real-world examples in improving imaging and patient access. Circulation, 138(22), 2569-2575.
20. Wang, F., Casalino, L. P., & Khullar, D. (2019). Deep learning in medicine—promise, progress, and challenges. JAMA internal medicine, 179(3), 293-294.
21. Winkler-Schwartz, A., Bissonnette, V., Mirchi, N., Ponnudurai, N., Yilmaz, R., Ledwos, N., ... & Del Maestro, R. F. (2019). Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. Journal of surgical education, 76(6), 1681-1690.