DR. ANDREA RAVIDÀ (Orcid ID : 0000-0002-3029-8130)
DR. GIUSEPPE TROIANO (Orcid ID : 0000-0001-5647-4414)
PROF. LUCIO LO RUSSO (Orcid ID : 0000-0001-6127-6105)
PROF. WILLIAM V. GIANNOBILE (Orcid ID : 0000-0002-7102-9746)
PROF. HOM-LAY WANG (Orcid ID : 0000-0003-4238-1799)

Article type : Original Article Pre-Clinical Sciences

Development of a Nomogram for the Prediction of Periodontal Tooth Loss Using the Staging and Grading System: A Long-Term Cohort Study

Andrea Ravidà[1*], Giuseppe Troiano[2 *], Musa Qazi[1], Muhammad H. A. Saleh[1,3], Lucio Lo Russo[2], Henry Greenwell[3], William V. Giannobile[1], Hom-Lay Wang[1]

[1]Department of Periodontics and Oral Medicine, University of Michigan School of Dentistry, Ann Arbor, MI, USA

[2]Department of Clinical and Experimental Medicine, University of Foggia, Foggia, Italy.

[3]Department of Periodontics, University of Louisville School of Dentistry, Louisville, KY, USA.

* *Authors contributing equally*

Corresponding authors:

Hom-Lay Wang, DDS, MSD, PhD
Department of Periodontics and Oral Medicine,
University of Michigan School of Dentistry

1011 North University Avenue
Ann Arbor, Michigan 48109-1078, USA.

TEL: (734) 763-3325; FAX: (734) 936-0374
E-mail address: homlay@umich.edu

Abstract word count: 263

Total word count: 2945

Total number of tables and figures: 3 Figures, 2 Tables

Supplemental Material: 4 supplementary tables and 2 supplementary figures

Number of references: 31

Key words (MeSH): periodontitis; supportive periodontal therapy; tooth loss, disease risk, patient stratification

Running title: Nomogram for periodontal tooth loss prediction

ORCIDs & E-mail addresses:
Andrea Ravidà: 0000-0002-3029-8130; ravida@umich.edu
Giuseppe Troiano: 0000-0001-5647-4414; giuseppe.troiano@unifg.it
Musa Qazi: n/a; musaqazi@umich.edu
Muhammad H. A. Saleh: 0000-0001-5067-7317; muhsaleh@umich.edu
Lucio Lorusso: 0000-0001-6127-6105; lucio.lorusso@unifg.it
Henry Greenwell: N/a: henry.greenwell@louisville.edu.
William Giannobile: 0000-0002-7102-9746; wgiannob@umich.edu
Hom-Lay Wang: 0000-0003-4238-1799; homlay@umich.edu

Abstract

Aim: To develop and internally validate a nomogram built on a multivariate prediction model including parameters from the new classification of periodontal diseases, able to predict, at baseline, the occurrence of tooth loss due to periodontal reason (TLP).

Materials and methods: A total of 315 individuals diagnosed with periodontal disease and receiving a minimum of one annual supportive periodontal therapy visit were included in the study. Patients were staged and graded based upon baseline data. The population was divided into a development (254 patients) and a validation (61 patients) cohort to allow subsequent temporal validation of the model. According to the TLP at the 10-year follow-up, patients were categorized as "low tooth loss" (≤ 1 TLP) or "high tooth loss" (≥ 2 TLP). Bootstrap internal validation was performed on the whole dataset to calculate an optimism-corrected estimate of performance.

Results: The generated nomogram showed a strong predictive capability (AUC = 0.81) and good calibration with an intercept = 0 and slope =1. These findings were confirmed by internal validation using bootstrapping (average bootstrap AUC = 0.83).

Conclusions: The clinical implementation of the present nomogram guides the prediction of patients with high risk of disease progression and subsequent tooth loss for personalized care.

Clinical Relevance:

Scientific rationale for study:

The 2018 classification of periodontal disease has a potential applicability as a risk assessment tool through a nomogram built on a multivariate prediction model.

Principal findings: The present prediction model is generalizable to similar study populations. The model showed good discrimination and calibration capabilities in detecting patients with high risk of tooth loss.

Practical implications: A nomogram is a user-friendly graphical interface of a multivariate prediction model that can be readily used by clinicians to accomplish individualized predictions for periodontal tooth loss.

Introduction

Periodontitis is a chronic, polygenic infectious disease with various factors contributing to its development and progression (Kinane, Stathopoulou, & Papapanou, 2017). Analyses of the true prevalence of the disease are elusive due to sampling heterogeneity and inconsistencies in its definition across different clinical investigations. Nevertheless, periodontitis is considered one of the most common causative factors for tooth loss in adults (Kassebaum et al., 2014). The standard protocol for the treatment of periodontal disease is presented as the combination of non-surgical with/without surgical active periodontal therapy (APT) followed by routine supportive periodontal therapy (SPT). Although these treatment protocols can eliminate/reduce the bacterial burden and slow disease progression, a diverse set of local (tooth- and site- level) and systemic (patient-level) factors can still potentially contribute to tooth loss due to periodontal reason (TLP) in the long-term. Patient-related factors include non-modifiable traits (e.g., age, sex and genetics), systemic conditions (e.g., diabetes and metabolic syndrome) or habits (e.g., smoking and substandard compliance) can also alter response to treatment; in some cases leading to disease progression and speed-up the rate of TLP (Helal et al., 2019).

Several periodontal risk assessment methods have built upon the principles of personalized medicine with the goal of tailoring prevention and/or treatment strategies (Beck, 1994). Most prominently, a Periodontal Risk Assessment tool (PRA) was introduced in 2003 by Lang and Tonetti to evaluate the risk for recurrence of periodontitis at patient-level using six clinical parameters (Lang & Tonetti, 2003). Multiple score-based prediction models were introduced later with the same aim to determine patients' individual risk (Chandra, 2007; Lindskog et al., 2010; Page, Krall, Martin, Mancl, & Garcia, 2002). The general aim of such predictive modeling is to stratify patients into risk groups that take into account distinct differences in comprehensive patient profile; enabling individually tailored treatment options based on the risk of disease progression (Giannobile et al., 2013). This was shown to be even be more beneficial in periodontally compromised patients (Persson et al., 2003).

Among the important advances in patient risk identification is the application of the new classification of periodontal disease. The introduction of new components of patient assessment (stage, grade and extent) is similar to that already used in other medical fields such as oncology or endocrinology (Caton et al., 2018; Papapanou et al., 2018b). As a result, a nomogram was developed and has been used to detect an individual who may have a higher risk of developing cancer and design a program to monitor these patients closely. This has been one of the first drives towards personalized medicine. In periodontology, these dimensions (stage, grade and extent) could aid in evaluating disease complexity coupled with its risk of progression (Tonetti, Greenwell, & Kornman, 2018). The

pragmatic application of TLP prediction due to periodontal disease has been recently confirmed in a long-term retrospective investigation (Ravida et al., 2019). An additional hypothesis is that by combining parameters from the new classification in a prediction model, it can be possible to build a risk assessment tool for clinical use. Thus, the aim of this study was to develop and test the predictive performance of a nomogram built on a multivariate model that incorporates staging and grading of periodontitis to guide clinicians in the early detection patients with a high-risk of periodontal-related tooth loss.

Materials and Methods

This investigation was designed according to the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (Moons et al., 2015). In addition, this study was conducted in agreement with the 2008 revised Declaration of Helsinki and approved by the University of Michigan, School of Dentistry, Institutional Review Board for Human Studies (HUM00157260).

- *Study population*

This study was conducted on a cohort of patients who received treatment for periodontitis during the period between January 1966 and January 2007 at the University of Michigan School of Dentistry. The total patient population included individuals divided into a development cohort and a validation cohort based on diagnosis year and initiation of active therapy, in particular the temporal validation cohort consisted of patients who started treatment after the 1997. This approach allowed for subsequent temporal validation of the development model. Data from the development cohort as well as inclusion and exclusion criteria previously described (Ravida et al., 2019) were utilized. In brief, clinical data (periodontal chart, full mouth radiographs, and medical history) were collected during the first patient appointment at our center and patients were then subsequently staged and graded according to the 2018 classification scheme (Papapanou et al., 2018a). Follow-up calculation initiated from the baseline clinical session of scaling and root planing. Tooth loss during the hygienic phase (deemed as hopeless at the patient screening) was not considered in the calculation of teeth lost for periodontal disease, but was utilized to assess the baseline stage of the patients as recently suggested by Sanz and coworkers (Sanz, Papapanou, Tonetti, Greenwell, & Kornman, 2020). All the included individuals received a minimum of one annual SPT visit/year during the follow-up period.

- *Predictors and outcome definition*

TLP at the 10-year follow-up period from initial therapy was determined as the outcome of this study and was not influenced by altered clinical conditions to avoid changes in classification after the fixed categorization. TLP for other reasons (such as: caries, endodontic problems, fractures, etc.) was not considered in this model. Other than "new classification" - based predictors (Stage, Grade and Extent), a patient-based variable was implemented (patient age at the baseline) to build the multivariate prediction model. The age of the patient was included as a continuous variable, the extent of periodontitis as the dichotomous variable, while stage and grade as ordinal numerical categorical variables. Only patient-level predictors were included in the present prediction model since the tooth related variables are encompassed in the stage and grade. Because of the retrospective design, and the need for a sufficient amount of information in order to estimate the classification-based predictors at the time of initial treatment, a complete-case analysis was performed.

- *Statistical Analysis*

After assessing the number of TLP at the 10-year follow-up, patients were categorized as “low tooth loss” if they experienced ≤ 1 TLP at the end of the follow-up and for the “high tooth loss”, an individual had to lose ≥ 2 TLP over the observation. To note, all the analysis below reported were performed with different cut-off values (0, 1, 2 and 3); the final categorized classes (“low tooth loss” ≤1, “high tooth loss” ≥ 2) were chosen since the developed models displayed the best values of predictive performance (discrimination and calibration). A binary period (High/Low tooth loss patient) multivariate logistic prediction model was developed including the following predictors: Stage (1-2-3-4), Grade (A-B-C), Extent (Localized-Generalized) and Age. In addition, aiming to control for the confounding effect of the frequency of SPT sessions on the patients' risk class, the average number of sessions per year during the 10-years' follow-up was calculated and added as a new variable into the model, then testing its statistical significance.

Bootstrap internal validation (2000 iterations) was performed, on the development dataset, in order to calculate an optimism-corrected estimate of performance. The logistic regression formula, using coefficients calculated during model development, was applied on the temporal validation dataset in order to obtain the predicted probability for any patient in the temporal validation dataset and then compared with the actual outcome registered after follow-up. Discrimination was evaluated by calculating the mean area under the receiver operating characteristic (ROC) curve. Furthermore, from ROC analysis, sensitivity and specificity for each value of predicted probability was analyzed to

select the optimal cut-off point. In addition, aiming to evaluate the potential clinical utility of nomogram application we perform Decision Curve Analysis (DCA) calculating the net benefits for different threshold probabilities. Calibration for agreement between predicted- and observed-probabilities was analyzed using the Hosmer-Lemeshow goodness-of-fit test and calculating intercept and slope in the calibration plots.

Following the above procedures, to build the clinical nomogram, the two datasets were combined, and a new model was developed using the same methodologies, Stages 1 and 2 were clustered together as a single class given the strong similarity of this periodontal stage for application purposes. The statistical analyses were performed using Stata 16.0 system software (StataCorp, 4905 Lakeway Drive College Station, Texas 77845 US). The approach for nomogram construction was built using the Stata package *nomolog*. In addition, on the whole cohort a Bayesian regression with flat priors for coefficient on x1 and the intercept and with a Jeffreys prior on the variance parameter was performed in order to test if a new model built with a Bayesian approach would improve the classic logistic regression model.

Results

- *Population cohort*

A total of 315 patients treated and followed at the University of Michigan, School of Dentistry fulfilled the inclusion criteria were included in this investigation. The population was assembled into a “development cohort” (254 patients) and a "temporal validation" cohort (61 patients) on the basis of the year of initial treatment (after 1997 in the temporal validation cohort). Patients in the development cohort demonstrated a mean age of 46.7 ± 11.8 years, while the mean age in the validation cohort was 51.4 ± 13.0, with a significant higher age in the validation cohort ($p = 0.007$). At the end of the 10-year follow-up, patients in the development cohort lost 0.38 ± 0.96 teeth due to periodontal reasons and 0.64 ±1.78 in the temporal validation cohort, respectively. No differences were found in the frequencies of patients by membership in specific classes of: Stage, Grade, Extent and Gender. In particular, most of the patients had been classified in either Stage 2 and 3 at baseline in both cohorts (202/254, 79.53% in the development cohort and 48/61, 78.69% in the temporal validation cohort), similarly the major number of patients were Grade B (167/254, 65.75% in the development cohort and 41/61, 67.21% in the temporal validation cohort). More details about patients included in the two cohorts according to variables characteristics are reported on Table 1. At the 10-

year follow-up, the mean number of teeth lost was 0.17 ± 0.37 for patients in the "low-risk" group and 3.44 ± 2.31 for the "high-risk" group.

- *Model development, specification and performance*

As above, the first model was developed using binary logistic regression including 254 patients and temporally validated on 61 individuals from the same institution. After categorization, 21 patients (8.27%) were considered at high risk of TLP in the development cohort and 6 patients (9.63%) in the temporal validation cohort for a combined 27 (8.03%) in the entire population. The predictors correlated with being a “high risk” patient were primarily Grade categorization and Age. It's worth noting, the addition of the average number of maintenance session per year during the follow-up did not reach statistical significance when included as a variable in the multivariate model.

The model developed using coefficients extracted from the multivariate binary logistic regression (results of the univariate analysis are shown in sTable 1) are reported in sTable 2 demonstrated both good discrimination and calibration capabilities, as shown by an AUC of 0.8 (Figure 1A) and calibration plot (Figure1D). The logistic regression formula (using coefficients of the abovementioned model) was then implemented to calculate the predictive probabilities on the temporal validation cohort. This step was performed in order to verify the applicability of the previously developed model on the cohort of patients treated after the 1997. Based on TRIPOD, both discrimination and calibration were assessed to confirm the proper functioning of the model. Hence, on the smaller temporal validation cohort (61 patients), the model confirmed strong and consistent discrimination (AUC = 0.9) (Figure 1B) and calibration (Hosmer-Lemeshow p-value = 0.47; intercept = 0.025; slope =1.025 in the calibration plot reported on Figure 1E) demonstrating that this model can be generalized successfully to other similar study populations. At this point, in order to increase the power of the statistical findings, the entire dataset was unified and used to generate a new model aiming to build a patient-level nomogram for clinical application (Figure 2). As shown in Table 2, in this multivariate prediction model, categories correlated primarily with showing a high risk for TLP were Stage 4 severity and Grade C (univariate analysis shown in sTable 3). The generated nomogram showed a good predictive capability (AUC = 0.81, Fig. 1C) and good calibration (Hosmer-Lemeshow p-value = 0.59; and with intercept = 0 and slope =1 in the calibration plot reported on Fig. 1F). These findings were confirmed by internal validation using bootstrapping (average bootstrap AUC = 0.83). Furthermore, the Bayesian regression showed a predictive capability similar to the classical logistic regression model. In the final logistic model, the optimal cut-off point was chosen on the basis of ROC curve analysis. Specifically, we selected a value for predicted probabilities of 0.06 since this

value showed the best balance of sensitivity (80.8%) and specificity (71.2%), corresponding to a "Total Score" of 12.5 into the nomogram. The possible clinical utility of nomogram application was confirmed by means of DCA, in which the model showed improved net benefits compared to both the "Intervention for all" and the "Intervention for none" reference groups (Fig. 3).

Nomogram interpretation and application

The nomogram is a predictive tool to evaluate risk of TLP that is based on additive evaluation of individual risk factors of TLP (sFig. 1). The nomogram is composed of upper and lower components. In the upper portion contains the scale ("Score") that is used to compute the weight of each variable (Stage, Grade, Extent and Age). As an example, the nomogram can be applied to a hypothetical patient in order to predict whether a 53 years-old individual with Stage 3, Grade A generalized periodontitis is at high or low risk to subsequently experience TLP at 10 years following diagnosis (Fig. 2). The first step is to calculate the score for each variable: Age of 53 years-old corresponds to a score of 5 in the Age category, Stage 3 gives 3.9 points, Grade A gives 0 points, and Extent gives 1.5 points since the periodontitis is generalized. At this point, each single score for every variable is added to obtain the final "Total Score" of the patient. In this example, the total Score is calculated as sum of the scores obtained for each variable is 10.4 (3.9 (Stage 3) + 0 (Grade A) + 1.5 (Generalized) + 5 (53 years of age) = 10.4). The calculated "Total Score" is then applied to the lower component of the nomogram in order to predict if the patient is at high or low risk of losing >1 TLP during a 10-year follow-up period. If the final score falls below the cut-off point (12.5), the patient will be classified as "low-risk", while if it lands above the cut-off, the patient will be classified as having high risk. This model only applies to those individuals who present for maintenance visits at least 1 or more times annually since it has been built on a cohort of compliant patients.

Discussion

Predicting whether teeth can or cannot be retained in a periodontitis patient leads to less invasive treatment, better treatment outcomes, and more reduced overall treatment expenses (Schwendicke, Stolpe, & Graetz, 2017; Tan, Peres, & Peres, 2016), especially if this can be accomplished during the treatment planning stage. Multiple studies have attempted introducing/validating prediction models, some to calculate the probability of periodontitis development, rate of progression, tooth loss while others are more suited predicting periodontal health (Du, Bo, Kapellas, & Peres, 2018; Lang, Suvan, & Tonetti, 2015). Most prediction models do this retrospectively, by learning patterns from available

data and then applying the obtained model to other data, either on the same cohort (Internal validation), or a different cohort (external validation) (Schwendicke et al., 2018).

In 1996, Lang & Tonetti suggested the need for a multi-level risk assessment at the patient, tooth and tooth site level to improve predictive values (Lang and Tonetti, 1996). Few years later they introduced a risk assessment tool (Periodontal Risk Assessment or PRA) in a pivotal move towards personalized periodontal care (Lang & Tonetti 2003). In a position paper, the American Academy of Periodontology acknowledged that assessment of periodontal risk is a crucial element of comprehensive periodontal evaluations that may help dental professionals in predicting the potential disease progression and provide targeted treatment for patients at risk of progressive diseases (AAP 2008). Likewise, the World Workshop for the Classification of Periodontal Diseases and Conditions introduced a periodontal classification system with a built-in prognostic determination system, reinforcing the significance of risk assessment in comprehensive patient evaluation (Tonetti, Greenwell & Kornman 2018). Measurements of disease severity are used to evaluate periodontal patient status and classify them according to stage, grade and extent. Such a classification system resembles an established process of evaluation concerning neoplastic diseases, where stage is a measure of ensued damage prior to diagnosis; representing the primary aim of identifying disease phase at diagnosis. In contrast, the grade measures disease rate of progression based on the manifested histological and/or molecular features (Amin et al., 2017; Orucevic et al., 2015). These predictors are combined with the aim of tailoring the best treatment approach after stratifying patients into classes pertaining to risk of disease severity and progression (Giannobile et al., 2013; Tham et al., 2019). A nomogram is a user-friendly graphical interface of a multivariate prediction model that can be easily used by clinicians to perform individualized predictions (Iasonos, Schrag, Raj, & Panageas, 2008). We believe that the new classification can be utilized for the prediction and stratification of patients at high risk of disease progression and subsequent tooth loss due to periodontal disease. Adoption of the nomogram in clinical care will allow clinicians to quickly identify those at low- and high-risk of losing teeth right after diagnosis. It should be considered that only the teeth which were deemed as hopeless were extracted during the active therapy and were no considered as lost in the nomogram (only utilized to assess the baseline stage). This is an important methodological detail that should be considered at the time of employing this tool. Clinical judgment is the key at the time of extracting teeth during active therapy and providers should treat and assess response to treatment for those teeth that could be maintained after therapy. This tool has the potential be an important tool for dental students as well due to their limited experience. Furthermore, the nomogram can be utilized by general dentist to manage referrals to periodontists since high-risk patients will require more advanced procedures. Stage III or stage IV patients are more likely to have one or several local factors

(furcation involvement, tooth mobility, secondary trauma for occlusion) that should be identified and managed during the treatment.

The same should be attempted with grade. Behavioral modification (i.e., smoking cessation programs) should be conducted to help patient either quit smoking, or at least decrease the number of cigarettes smoked per day. This will help control the disease, and the long term might decrease the case grade. Ascertaining that the patient regularly monitors HbA1C levels and collaboration with a patient's endocrinologist to better achieve metabolic control of diabetes is strongly encouraged to provide a more predictable treatment outcome. We believe the user-friendly graphical interface of the nomogram can be utilized to explain to patients how risk factors such as smoking and diabetes (especially if uncontrolled) increases their probability of TLP. This is a valuable step toward personalized medicine, that guides patients on the impact of different risk factors on the overall disease and encourages them to more actively participate in their oral health care.

Our prediction model was built upon a cohort of patients who attended at least one annual SPT session throughout a 10-year follow-up period. This well-maintained population, treated in an academic setting, was comprised of a low percentage (8.03%) of high-risk patients. It is also sensible to compare the amount of teeth lost/year due to periodontal disease in the current retrospective study to other studies with long-term prospective designs (Costa et al., 2014; Isidor & Karring, 1986; Lindhe & Nyman, 1984). In the present study, a higher rate of TLP was found as compared to all the aforementioned studies. Interestingly, the rate of TLP over time did not differ throughout the timeframe (temporal validation) of the follow-up period despite significant clinical advancements in periodontal therapy during this period. Perhaps the increasing popularity and demand for implant therapy plays a role in influencing this trend since dental implants have emerged as a predictable treatment option for replacing missing teeth. Furthermore, it has to be considered that our study was conducted using data from patients treated at a dental school by a broad range of operators including undergraduate and graduate dental students, students of dental hygiene, general dentists, and periodontists. This may have led to different subjective criteria for the need for and the timing of extractions. The decision to extract a tooth hinges on multiple factors such as economic considerations and restorative planning that are not solely related to the periodontal status of individual teeth,

We had initially stratified the cohort by time, building a preliminary model constituting patients who had a start date for treatment before 1997 and analyzing the number of teeth lost for periodontal disease after that date. This process was exhibited in order to verify the absence of heterogeneity within the cohort concerning the variable of time. Based on the above, we used the cohort and built

a new predictive model including the entire group of 315 patients, subsequently building and internally validating a nomogram for clinical use (Moons et al., 2015).

In nomograms, the predicted probabilities for variables included into the model are mapped on a scale ("Score" on Figure 2), then the singular scores for each prediction are summed together ("Total Score" on Figure 2) to obtain the predicted probability ("Prob" on Figure 2) for a given patient. Different thresholds can be used on a nomogram to classify patients as high/low risk constituting at each threshold different values of sensitivity and specificity. For the reported nomogram, the threshold was selected by means of ROC analysis. In particular, setting the threshold of predicted probabilities at 0.06, the model shows 80.8% sensitivity and 71.2% specificity. Hence, if the total predicted probability ("Prob" on Figure 2) is higher than the threshold, hence the total score ("Total Score" on Figure 3) for the evaluated patient is higher than 12.5, clinicians should classify the patient as high risk – and correspondingly, if the value is lower, as a low risk – of exhibiting TL for periodontal disease. However, if clinicians prefer to have a more optimized model, they can use a more specific threshold (predictive results for all the threshold levels are reported in sTable 4).

Even if the nomogram showed a good predictive performance, its clinical utility may still be considered questionable. For such reasons, Decision Curve Analysis was applied. In this method, the clinical net benefits of the model are calculated in order to determine whether the nomogram could be warranted in selecting patients that will benefit from an additional intervention after non-surgical therapy. Depending on specific patient needs and suspected causes of deterioration further sessions of scaling and root planing, more frequent SPT, use of antibiotics for anti-infective treatment, medical consultation for diabetes control or smoking counseling, and even surgical therapy may be suggested to the patient. As reported in the graph (Figure 3), utilizing the nomogram will help to develop a personalized treatment plan aimed at reducing the rate of tooth loss due to periodontitis. Even if a single threshold of predicted probabilities has been chosen to simplify clinical use of the nomogram, in DCA, a wide range of threshold probabilities resulted in improved net benefits compared to the reference groups, supporting the clinical utility of the tool. A careful explanation for a deeper understanding of the clinical denotation of DCA for the application of models in the clinical practice has been recently published by (Vickers, van Calster, & Steyerberg, 2019).

Although bootstrap validation was implemented, the authors recognize that such developed models tend to perform better on the cohort they are built on. Indeed, the validity of a prediction model may be contingent on specific population or socioeconomic variables. It is worth mentioning that other

risk assessment models were successfully validated across different patient populations (Eickholz, Kaltschmitt, Berbig, Reitmeir, & Pretzl, 2008; Matuliene et al., 2010). In order to be able of assessing history of periodontal disease accurately in general populations, predictions models should be able to hold true in different settings or through different time frames. Hence, an external validation model is recommended over an internal; to determine whether a particular model can be generalized or not (Steyerberg & Harrell, 2016). Furthermore our predictive performance might have improved by taking into account more novel predictors, such as disease biomarkers markers (Giannobile et al., 2009).

In medicine, prediction usually includes both, the presence of disease (diagnosis) and an event in the future course of disease (prognosis) (Steyerberg & Vergouwe, 2014). This critical component is the foundation for this new classification system (Du et al., 2018). Also, a significant robustness of the present model is that only TLP were included, whereas most other prediction models relied on the overall tooth loss. This is because TLP is an indispensable criterion for a prediction model that has to be taken into consideration clinically (Krois et al., 2019; Martinez-Canut et al., 2018).

Conclusion:

The nomogram based on parameters contained in the 2018 classification of periodontal diseases demonstrated a strong predictive capability for identifying patients at risk to lose ≥ 2 tooth throughout 10-years of follow-up. Future studies should consider external validation of this study to explore the generalization of this model.

ACKNOWLEDGMENTS

We would like to thank Dr. Kenneth Kornman for his valuable contributions to this paper. This manuscript was partially supported by the University of Michigan Periodontal Graduate Student Research Fund. The authors do not have any financial interests, either directly or indirectly, in the products or information listed in the paper.

Author Contributions:

Andrea Ravidà: Contributed to the conception and design of the study, acquisition of the data and drafting of the article

Giuseppe Troiano: Contributed to the conception and design of the study. Data analysis and interpretation

Musa Qazi and Muhammad H. A. Saleh: Contributed to the acquisition of data

Lucio Lo Russo: Contributed do the drafting of the article

Henry Greenwell: Provided critical revision of the article

William V. Giannobile: Contributed to the conception, manuscript preparation and critical revision of the article

Hom-Lay Wang: Contributed to the conception, critical revision of the article, final approval of the version to be published

REFERENCES

Amin, M. B., Greene, F. L., Edge, S. B., Compton, C. C., Gershenwald, J. E., Brookland, R. K., . . . Winchester, D. P. (2017). The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin, 67*(2), 93-99. doi:10.3322/caac.21388

Beck, J. D. (1994). Methods of assessing risk for periodontitis and developing multifactorial models. *J Periodontol, 65*(5 Suppl), 468-478. doi:10.1902/jop.1994.65.5s.468

Caton, J. G., Armitage, G., Berglundh, T., Chapple, I. L. C., Jepsen, S., Kornman, K. S., . . . Tonetti, M. S. (2018). A new classification scheme for periodontal and peri-implant diseases and conditions - Introduction and key changes from the 1999 classification. *J Periodontol, 89 Suppl 1*, S1-S8. doi:10.1002/JPER.18-0157

Chandra, R. V. (2007). Evaluation of a novel periodontal risk assessment model in patients presenting for dental care. *Oral Health Prev Dent, 5*(1), 39-48.

Costa, F. O., Lages, E. J., Cota, L. O., Lorentz, T. C., Soares, R. V., & Cortelli, J. R. (2014). Tooth loss in individuals under periodontal maintenance therapy: 5-year prospective study. *J Periodontal Res, 49*(1), 121-128. doi:10.1111/jre.12087

Du, M., Bo, T., Kapellas, K., & Peres, M. A. (2018). Prediction models for the incidence and progression of periodontitis: A systematic review. *J Clin Periodontol, 45*(12), 1408-1420. doi:10.1111/jcpe.13037

Eickholz, P., Kaltschmitt, J., Berbig, J., Reitmeir, P., & Pretzl, B. (2008). Tooth loss after active periodontal therapy. 1: patient-related factors for risk, prognosis, and quality of outcome. *J Clin Periodontol, 35*(2), 165-174. doi:10.1111/j.1600-051X.2007.01184.x

Giannobile, W. V., Beikler, T., Kinney, J. S., Ramseier, C. A., Morelli, T., & Wong, D. T. (2009). Saliva as a diagnostic tool for periodontal disease: current state and future directions. *Periodontol 2000, 50*, 52-64. doi:10.1111/j.1600-0757.2008.00288.x

Giannobile, W. V., Braun, T. M., Caplis, A. K., Doucette-Stamm, L., Duff, G. W., & Kornman, K. S. (2013). Patient stratification for preventive care in dentistry. *J Dent Res, 92*(8), 694-701. doi:10.1177/0022034513492336

Helal, O., Gostemeyer, G., Krois, J., Fawzy El Sayed, K., Graetz, C., & Schwendicke, F. (2019). Predictors for tooth loss in periodontitis patients: Systematic review and meta-analysis. *J Clin Periodontol, 46*(7), 699-712. doi:10.1111/jcpe.13118

Iasonos, A., Schrag, D., Raj, G. V., & Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol, 26*(8), 1364-1370. doi:10.1200/JCO.2007.12.9791

Isidor, F., & Karring, T. (1986). Long-term effect of surgical and non-surgical periodontal treatment. A 5-year clinical study. *J Periodontal Res, 21*(5), 462-472. doi:10.1111/j.1600-0765.1986.tb01482.x

Kassebaum, N. J., Bernabe, E., Dahiya, M., Bhandari, B., Murray, C. J., & Marcenes, W. (2014). Global burden of severe periodontitis in 1990-2010: a systematic review and meta-regression. *J Dent Res, 93*(11), 1045-1053. doi:10.1177/0022034514552491

Kinane, D. F., Stathopoulou, P. G., & Papapanou, P. N. (2017). Periodontal diseases. *Nat Rev Dis Primers, 3*, 17038. doi:10.1038/nrdp.2017.38

Krois, J., Graetz, C., Holtfreter, B., Brinkmann, P., Kocher, T., & Schwendicke, F. (2019). Evaluating Modeling and Validation Strategies for Tooth Loss. *J Dent Res, 98*(10), 1088-1095. doi:10.1177/0022034519864889

Lang, N. P., Suvan, J. E., & Tonetti, M. S. (2015). Risk factor assessment tools for the prevention of periodontitis progression a systematic review. *J Clin Periodontol, 42 Suppl 16*, S59-70. doi:10.1111/jcpe.12350

Lang, N. P., & Tonetti, M. S. (2003). Periodontal risk assessment (PRA) for patients in supportive periodontal therapy (SPT). *Oral Health Prev Dent, 1*(1), 7-16.

Lindhe, J., & Nyman, S. (1984). Long-term maintenance of patients treated for advanced periodontal disease. *J Clin Periodontol, 11*(8), 504-514. doi:10.1111/j.1600-051x.1984.tb00902.x

Lindskog, S., Blomlof, J., Persson, I., Niklason, A., Hedin, A., Ericsson, L., . . . Blomlof, L. (2010). Validation of an algorithm for chronic periodontitis risk assessment and prognostication: analysis of an inflammatory reactivity test and selected risk predictors. *J Periodontol, 81*(6), 837-847. doi:10.1902/jop.2010.090483

Martinez-Canut, P., Alcaraz, J., Alcaraz, J., Jr., Alvarez-Novoa, P., Alvarez-Novoa, C., Marcos, A., . . . Zabalegui, I. (2018). Introduction of a prediction model to assigning periodontal prognosis based on survival time. *J Clin Periodontol, 45*(1), 46-55. doi:10.1111/jcpe.12810

Matuliene, G., Studer, R., Lang, N. P., Schmidlin, K., Pjetursson, B. E., Salvi, G. E., . . . Zwahlen, M. (2010). Significance of Periodontal Risk Assessment in the recurrence of periodontitis and tooth loss. *J Clin Periodontol, 37*(2), 191-199. doi:10.1111/j.1600-051X.2009.01508.x

Moons, K. G., Altman, D. G., Reitsma, J. B., Ioannidis, J. P., Macaskill, P., Steyerberg, E. W., . . . Collins, G. S. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med, 162*(1), W1-73. doi:10.7326/M14-0698

Orucevic, A., Chen, J., McLoughlin, J. M., Heidel, R. E., Panella, T., & Bell, J. (2015). Is the TNM staging system for breast cancer still relevant in the era of biomarkers and emerging personalized medicine for breast cancer - an institution's 10-year experience. *Breast J, 21*(2), 147-154. doi:10.1111/tbj.12367

Page, R. C., Krall, E. A., Martin, J., Mancl, L., & Garcia, R. I. (2002). Validity and accuracy of a risk calculator in predicting periodontal disease. *J Am Dent Assoc, 133*(5), 569-576. doi:10.14219/jada.archive.2002.0232

Papapanou, P. N., Sanz, M., Buduneli, N., Dietrich, T., Feres, M., Fine, D. H., . . . Tonetti, M. S. (2018a). Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Periodontol, 89 Suppl 1*, S173-S182. doi:10.1002/JPER.17-0721

Papapanou, P. N., Sanz, M., Buduneli, N., Dietrich, T., Feres, M., Fine, D. H., . . . Tonetti, M. S. (2018b). Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. *J Clin Periodontol, 45 Suppl 20*, S162-S170. doi:10.1111/jcpe.12946

Persson, G. R., Matuliene, G., Ramseier, C. A., Persson, R. E., Tonetti, M. S., & Lang, N. P. (2003). Influence of interleukin-1 gene polymorphism on the outcome of supportive periodontal therapy explored by a multi-factorial periodontal risk assessment model (PRA). *Oral Health Prev Dent, 1*(1), 17-27.

Ravida, A., Qazi, M., Troiano, G., Saleh, M. H. A., Greenwell, H., Kornman, K., & Wang, H. L. (2019). Using periodontal staging and grading system as a prognostic factor for future tooth loss: A long-term retrospective study. *J Periodontol*. doi:10.1002/JPER.19-0390

Sanz, M., Papapanou, P. N., Tonetti, M. S., Greenwell, H., & Kornman, K. (2020). Guest Editorial: Clarifications on the use of the new classification of periodontitis. *J Clin Periodontol*. doi:10.1111/jcpe.13286

Schwendicke, F., Schmietendorf, E., Plaumann, A., Salzer, S., Dorfer, C. E., & Graetz, C. (2018). Validation of multivariable models for predicting tooth loss in periodontitis patients. *J Clin Periodontol, 45*(6), 701-710. doi:10.1111/jcpe.12900

Schwendicke, F., Stolpe, M., & Graetz, C. (2017). Cost comparison of prediction-based decision-making for periodontally affected molars. *J Clin Periodontol, 44*(11), 1145-1152. doi:10.1111/jcpe.12796

Steyerberg, E. W., & Harrell, F. E., Jr. (2016). Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol, 69*, 245-247. doi:10.1016/j.jclinepi.2015.04.005

Steyerberg, E. W., & Vergouwe, Y. (2014). Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J, 35*(29), 1925-1931. doi:10.1093/eurheartj/ehu207

Tan, H., Peres, K. G., & Peres, M. A. (2016). Retention of Teeth and Oral Health-Related Quality of Life. *J Dent Res, 95*(12), 1350-1357. doi:10.1177/0022034516657992

Tham, T., Machado, R., Herman, S. W., Kraus, D., Costantino, P., & Roche, A. (2019). Personalized prognostication in head and neck cancer: A systematic review of nomograms according to the AJCC precision medicine core (PMC) criteria. *Head Neck, 41*(8), 2811-2822. doi:10.1002/hed.25778

Tonetti, M. S., Greenwell, H., & Kornman, K. S. (2018). Staging and grading of periodontitis: Framework and proposal of a new classification and case definition. *J Periodontol, 89 Suppl 1*, S159-S172. doi:10.1002/JPER.18-0006

Vickers, A. J., van Calster, B., & Steyerberg, E. W. (2019). A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res, 3*, 18. doi:10.1186/s41512-019-0064-7

Table and Figure Legends

Figure 1 (A) Discrimination and (D) calibration of the model built on the cohort of 254 patients starting treatment before the 1997, (B) receiver operating characteristic (ROC) curve and (E) calibration plot applying the first model on the temporal validation dataset. Results of discrimination (C) and (F) calibration for the final model developed on the whole dataset of 315 patients.

Figure 2: A 2018 classification-based nomogram able to classify patients at high or low risk tooth loss, was generated using Age, Stage, Grade and Extent.

Figure 3: Decision curve analysis was used to compare the clinical net benefit between the nomogram and the reference groups (Treat all and Treat None).

Supplementary Figure 1: Receiver Operating Characteristics (ROC) curve evaluating the performance of Bayesian regression analysis.

Supplementary Figure 2: Interpretation of the classification-based nomogram**.**

Table 1: Characteristics of distribution of variables in the development and validation cohorts

Table 2: Multivariate binary logistic regression analysis built on the whole cohort (315 patients) and used for the generation of the nomogram.

Supplementary Table 1: Univariate logistic regression analysis for variables included into the model on the whole development cohort of 254 patients starting treatment before the 1997.

Supplementary Table 2: Multivariate binary logistic regression analysis was performed on the development cohort of (254 patients) and used to build the first model.

Supplementary Table 3: Univariate logistic regression analysis for variables included into the model on the whole cohort of 315 patients.

Supplementary Table 4: Results of sensitivity and specificity for each threshold calculated from receiver operating characteristic (ROC) analysis.

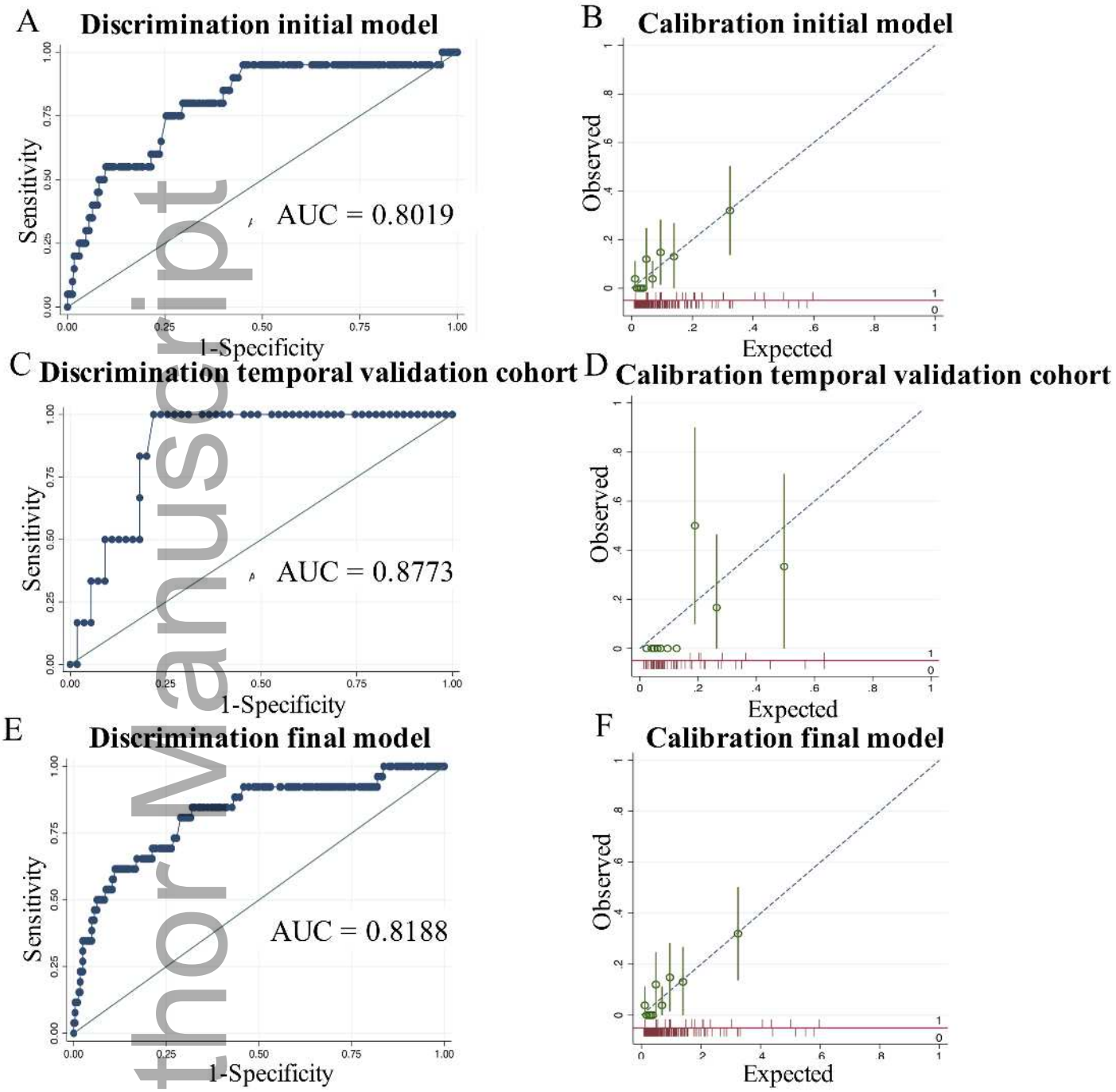
Table 1. Characteristics of distribution of variables in the development and validation cohorts

Variables	Development Cohort (n=254 Patients)	Validation Cohort (n=61 Patients)	p-value
Stage			
Stage 1	10.63% (27/254)	8.20% (5/61)	0.572
Stage 2	30.32% (77/254)	27.87% (17/61)	0.707
Stage 3	49.21% (125/254)	50.82% (31/61)	0.821
Stage 4	9.84% (25/254)	13.11% (8/61)	0.453
Grade			
Grade 1	11.02% (28/254)	11.48% (7/61)	0.919
Grade 2	65.75% (167/254)	67.21% (41/61)	0.828
Grade 3	23.23% (59/254)	21.31% (13/61)	0.749
Extent			
Localized	70.87% (180/254)	73.77% (45/61)	0.652
Generalized	29.13% (74/254)	26.23% (16/61)	
Sex			
Males	48.82% (124/254)	57.38% (35/61)	0.230
Females	51.18% (130/254)	42.62% (26/61)	
Age			
Mean Age	46.72 ± 11.84	51.41 ± 12.98	0.007*

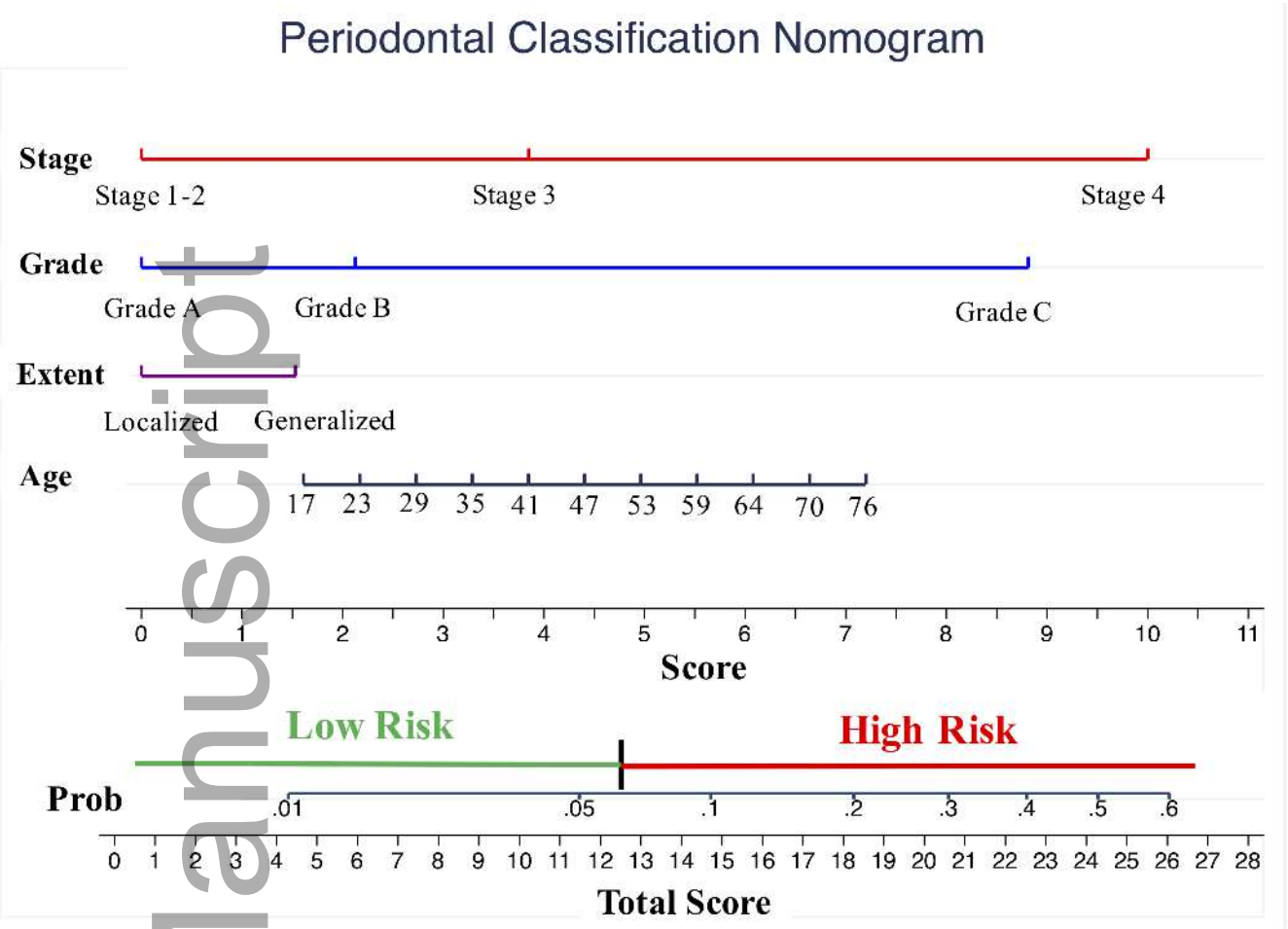
***: Statistically significant (p<0.05)**

Table 2: Multivariate Binary Logistic Regression Model for Nomogram Construction, based on both developmental and validation cohorts (n=315 patients)

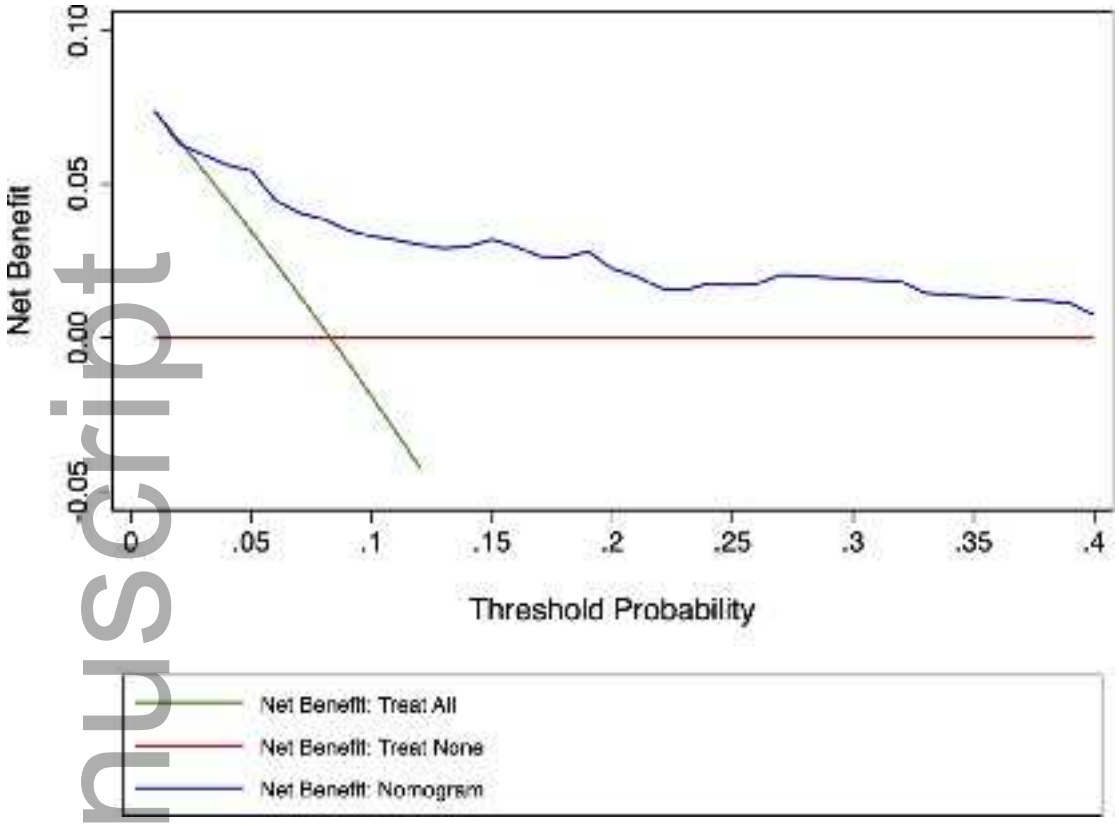
Variables		Multivariate		
		Coefficients	OR 95%(CI) -	p-value
Stage	1-2	-	1.00	-
	3	0.9255798	2.42 (0.63-9.68)	0.198
	4	2.269515	9.95 (2.36-41.89)	0.002*
Grade	A	-	1.00	-
	B	0.5013877	1.63 (0.14-14.58)	0.662
	C	2.138476	7.58 (0.80-71.8)	0.047*
Extent	Local	-	1.00	-
	Generalized	0.2907061	1.42 (0.54-3.71)	0.471
Age		0.0240872	1.02 (0.98-1.06)	0.274
Intercept		-5.711725		



jcpe_13362_f1.tiff



jcpe_13362_f2.tiff



jcpe_13362_f3.tiff