Clinical Prediction of Chronic Periodontitis

by

Htun Teza

M.Sc. Data Science for Healthcare

Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

Email : htun.tez@student.mahidol.edu

Anuchate Pattanateepapon

Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

Email : anuchate.gab@mahidol.ac.th

Attawood Lertpimonchai

Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand.

Email : Attawood.L@chula.ac.th

Prin Vathesatogkit

Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

Email : prin.vat@mahidol.ac.th

And

Ammarin Thakkinstian

Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.

Email : ammarin.tha@mahidol.ac.th

Abstract

Diagnosis of chronic periodontitis is a time and resource-consuming procedure that require exhaustive periodontal probing for each present tooth. Risk prediction models are developed to remove the need for such parameter. Those models are commonly cross-sectional logistic regression models, enhanced by applying demographic and risk factors data. With present study, we further augment the performance by applying longitudinal data to develop the model. Considering population average and subject specific effects from longitudinal data, mixed effects model result in higher performance. Sensitivity and specificity are 89.5% and 92.5% while performing upon validation data. The model is 91.5% accurate with 0.91 discriminative power. Positive predictive value and negative predictive value are 86.2% and 94.4%. The positive likelihood ratio of the model is 11.9.

Keywords: Chronic periodontitis, Predictive modeling, Mixed effects logistic regression.

1. Introduction

Periodontitis is one of the most common oral diseases and causes of tooth loss in adults. It is the world's 6th most prevalent oral disease, affected around 743 million people worldwide.[9] The prevalence was at 11.2% globally, and 15.0-20.0% of Asians.[1] According to 8th Thailand national oral health survey (2017), the prevalence of periodontitis in Thai adults is 26%, and for the elderly, it is 36%. Periodontitis is a complex inflammatory disease that leads to the destruction of the supporting structures around the tooth, resulting in the loosening of the teeth and eventual tooth loss. This leads to decreased occlusal ability, digestive ability and effectively the patient's quality of life. In addition to oral manifestations, previous studies have found association of chronic periodontitis with systemic diseases and conditions such as atherosclerotic vascular diseases (ASVD), Diabetes Mellitus, Chronic Kidney Disease (CKD). Other systemic diseases such as chronic obstructive pulmonary disease (COPD), rheumatoid arthritis (RA), Alzheimer's disease and erectile dysfunction, also have been reported to have relationships with chronic periodontitis. [6, 7, 8]

In chronic periodontitis, the loss of clinical attachment level is the major characteristic therefore the diagnosis requires full-mouth periodontal probing which is the manual measurement of the distance between the cementoenamel junction and the base of the periodontal pocket for all present teeth. Such measure is gold-standard, but it is time and resource-consuming since it requires exhaustive measurements by trained personnel such as a dentist or dental hygienist. Such a scenario can be more efficiently addressed by the presence of a risk prediction system.

Risk prediction systems are usually statistical models, most commonly logistic regression models, where the log odds of having interested event (dependent variable) are modeled as the linear combination of the predictor variables (independent variables). They are developed by applying cross-sectional data and the predictor variables are selected in terms of significant statistical relationships with the outcome. Periodontal parameters, especially clinical attachment level, are the golden standard of diagnosis therefore Cyrino et.al.[2] have observed that not including them within the prediction system reduces the performance of the model. However, to the remove the need for exhaustive periodontal probing is one of the major goals, the prediction systems are developed by including other parameters such as demographics and risk behaviors to improve their performance of risk assessment.[5, 10] With present study, we further augment the performance of the models by utilizing longitudinal data.

2. Materials and Methods

2.1. Data Description

This study was a sub-cohort of prospective cohort study, namely Electric Generating Authority of Thailand (EGAT) cohort, by retrieving 5-years follow up period. Details about EGAT cohort are referenced [11], but in short, EGAT project contains three parallel cohorts, also known as EGAT1, EGAT2 and EGAT3. This study was conducted applying EGAT2 cohort; both the 3rd survey (2/3: 2008) and the 4th survey (2/4: 2013) were used for training and validation of the models. All subjects were included unless they meet exclusion criteria. Some subjects were not present in ALL periodontal examinations due to (1) refusal to participate, (2) systemic conditions which required antibiotic prophylaxis before dental procedure including congenital heart disease or valvular heart disease, previous history of bacterial endocarditis or rheumatic fever, total joint replacement and end-stage renal disease, and (3) fully edentulous subjects. Such subjects were excluded for all models.

2.2. Data Collection

In each survey, general demographic data (gender, educational level), behavioral data (smoking status), underlying diseases (diabetes mellitus) were collected by self-administered questionnaires. Oral examinations included number of teeth, plaque score, periodontal pocket depth, and gingival recession which were carried out on all fully erupted teeth, except third molars and retained roots. Periodontal pocket depth is the measurement from coronal margin of gingival margin to the tip of a periodontal probe, and gingival recession is the measurement from coronal margin of gingival margin to the cementoenamel junction. The parameters were measured applying a periodontal probe - University of North Carolina 15 (PCP-UNC15) on six sites, i.e., mesial, mesio-buccal, mesio-lingual, disto-buccal, disto-lingual and lingual site of the gingival sulcus per tooth. These measurements were made in millimeters and were rounded to the nearest whole millimeter.

2.3. Periodontitis case classification

We categorized our samples into two, severe periodontitis and non-severe periodontitis (none, mild and moderate) according to criteria proposed by Centers for Disease Control and Prevention/American Academy of Periodontology (CDC/AAP) working group. Severe periodontitis was defined as subject with ≥2 interproximal sites with CAL ≥6 mm in different teeth and one site with PD ≥5 mm.

2.4. Statistical analyses

All analyses were done using STATA version 16.0. Using Pareto principle, the data was cautiously split into training (80%) and validation (20%) datasets, to avoid situations where the same individual appeared in both.

2.4.1. Mixed Models

Mixed effects models are statistical models which include both fixed effects (population average) and random effects (subject specific).[3] Mixed models are applied when multiple correlated measurements are made on each unit of interest. In matric notation, linear mixed effects models can be represented as –

$$y = X\beta + Zu + \epsilon$$

where –

y = known vector of observations (dependent variable)

X, Z = known design matrices relating the observations y to β and u, respectively (independent variables)

β = unknown vector of fixed effects

u = unknown vector of random effects.

Since our outcome of interest was dichotomous (severe and non-severe), we applied mixed effects logistic regression, where y was the log odds of interested outcome. [4]

For variable selection, stepwise method with forward selection was performed on the training dataset. The output of the model was dichotomized using the prevalence of severe periodontitis (35%). The final model was evaluated using the following metrics: sensitivity, specificity, accuracy, discrimination, positive likelihood ratio, positive and negative predictive value.

2.5. Results

In our data, 71% of the subjects were men. 8% of the subjects were educated higher than bachelor's degree and 38% hold bachelor's degree while the rest were educated less. 46% of the subjects were current or ex-smoker. 13% had underlying diabetes mellitus. The median number of teeth was 25 and the mean plaque score was 70. Fixed effects coefficients for significant variables retained in final multivariate models are reported in Table 1. To assess the risk score for developing severe periodontitis,

Risk score = -3.93 + (0.97 x male)
+ (2.04 x education < High school)
+ (1.35 x education Vocational School)
+ (0.29 x education Bachelor's degree)
+ (0.73 x Ex-smoker) + (1.68 x Current smoker)
+ (0.50 x diabetes mellitus)
+ (-0.06 x number of teeth) + (0.03 x plaque score)

– where the covariate should be replaced with 1 if applicable and 0 if else. From the risk score, the subject's risk of developing the condition can be calculated as $\frac{e^{Risk\ score}}{1+e^{Risk\ score}}$.

The odds of having severe periodontitis for male is 2.63 times that of female. It can be associated with the fact that 23.31% of male are current smokers and 40.12% used to be while only 5.85% of female subjects are current or ex-smokers in our survey. This is consistent that the odd ratio of ex-smokers to non-smokers is 2.09 and the odd ratio for current smokers is 5.28 times of non-smokers. Education levels also affects since the odds for non-high school graduates, vocational school graduates and bachelor's degree holders are 2.04, 1.35 and 1.34 respectively in comparison to people with post-graduate degrees. The effects of diabetes mellitus have been established before and it is observed that underlying diabetes mellitus results in 1.66 times higher odds. Dental plaque is the major predisposing factor of periodontitis and increasing the plaque score by one increase odd by 1.03. At the same time, it is interesting to observe that retaining more teeth in the oral cavity reduces the odds by 0.94. This effect can be explained by good occlusal force distribution since tooth loss result in unbalanced occlusal force and traumatic occlusion in alveolar regions with present teeth.

Application of the final model as the risk prediction model has great performances. Sensitivity and specificity of the model is 89.5% and 92.5%, with the overall accuracy of 91.5%. Higher specificity is preferred than higher sensitivity because false positive cases can be followed with further investigations. Discriminative power is evaluated using area under receiver operating characteristic curve. While values over 0.9 is considered outstanding, our model has the value of 0.91. The comprehensive performance metrics upon training and validation data are stated in Table 2.

While our model performs outstandingly well, it should be noted that our model was trained and validated by the same survey. External validation using other survey or population should be done to further evaluate. While our model removes the need for exhaustive periodontal probing, the plaque scores still require a trained person to examine therefore the model cannot be applied by the patients using self-reportable questionnaires. High performing algorithms such as machine learning and neural networks taking advantage of longitudinal data can be applied to further improve the performance.

2.6. Conclusion

While previous studies applied logistic regression, the use of mixed effects model accounts for subject specific effects of latent variables resulting in better estimation of population average effects and higher performance despite using only fixed effects for classification. By removing exhaustive periodontal probing, use of such model can act as a screening tool and reduce the workload of dentists and dental hygienists, which in turn would reduce time and resource requirements.

Table 1. Fixed Effects Coefficients and Odds Ratio Estimates for Significant Variables Retained in the Final Multivariate Mixed Effects Logistic Regression Model

Variables	Covariates	Coefficient (SE)	Odd ratios (95% CI)	P-value
Gender	Male	0.97 (0.23)	2.63 (1.68 to 4.10)	< 0.001
	Female	ref	ref	
Education	< High school	2.04 (0.38)	7.68 (3.62 to 16.30)	< 0.001
	Vocational School	1.35 (0.35)	3.86 (1.93 to 7.72)	< 0.001
	Bachelor's degree	0.29 (0.35)	1.34 (0.68 to2.64)	< 0.001
	> Bachelor's degree	ref	ref	0.393
Smoking	Non-smoker	ref	ref	
	Ex-smoker	0.73 (0.21)	2.09 (1.38 to 3.17)	0.001
	Current smoker	1.68 (0.25)	5.38 (3.28 to 8.83)	< 0.001
Diabetes Mellitus	Positive	0.50 (0.22)	1.66 (1.07 to 2.57)	0.024
	Negative	ref	ref	
Number of teeth	-	-0.06 (0.02)	0.94 (0.91 to 0.97)	< 0.001
Plaque score	-	0.03 (0.004)	1.03 (1.02 to 1.03)	< 0.001

Abbreviation: CI: Confidence Interval; SE: Standard Error; ref: Reference covariate group.

Table 2. Performance Metrics of Final Multivariate Mixed Effects Logistic Regression Model

Metrics	Training data	Validation data
%Sensitivity	91.4 (89.5 to 93.0)	89.5 (85.1 to 92.9)
%Specificity	90 (88.7 to 91.3)	92.5 (89.9 to 94.6)
%Accuracy	90.5 (89.4 to 91.5)	91.5 (89.3 to 93.3)
AUC	0.91 (0.90 to 0.92)	0.91 (0.89 to 0.93)
Positive likelihood ratio	9.18 (8.05 to 10.50)	11.9 (8.77 to 16.30)
%Positive predictive value	82.9 (80.7 to 85.0)	86.2 (81.6 to 90.1)
%Negative predictive value	95.2 (94.1 to 96.1)	94.4 (92.0 to 96.2)

References

1. Corbet EF, Leung WK. (2011), Epidemiology of periodontitis in the Asia and Oceania regions, Periodontology 2000, Vol.56, No.1, pp. 25-64.
2. Cyrino RM, Miranda Cota LO, Pereira Lages EJ, Bastos Lages EM, Costa FO. (2011), Evaluation of Self-Reported Measures for Prediction of Periodontitis in a Sample of Brazilians, Journal of Periodontology, Vol.81, No.12, pp. 1693-1704.
3. Demidenko E. (2004), Mixed Models: Theory and Applications, Hoboken, NJ: Wiley.
4. Guo, G, H. Zhao. (2000), Multilevel modeling of binary data, Annual Review of Sociology, Vol.26, pp. 441–462.
5. Leite FRM, Peres KG, Do LG, Demarco FF, Peres MAA. (2017), Prediction of Periodontitis Occurrence: Influence of Classification and Sociodemographic and General Health Information Journal of Periodontology, Vol.88, No.8, pp. 731-743.
6. Lertpimonchai A, Rattanasiri S, Tamsailom S, Champaiboon C, Ingsathit A, Kitiyakara C, et al. (2019), Periodontitis as the risk factor of chronic kidney disease: Mediation analysis, Journal of Clinical Periodontology, Vol.46, No.6, pp. 631-639.
7. Linden GJ, Lyons A, Scannapieco FA. (2013), Periodontal systemic associations: review of the evidence. Journal of Clinical Periodontology. 2013; Vol. 40, Supplementary 14, pp. S8-S19.
8. Monsarrat P, Blaizot A, Kémoun P, Ravaud P, Nabet C, Sixou M, et al. (2016), Clinical research activity in periodontal medicine: a systematic mapping of trial registers. Journal of Clinical Periodontology. Vol.43, No.5, pp. 390-400.
9. Phipps KR, Stevens VJ. (1995), Relative contribution of caries and periodontal disease in adult tooth loss for an HMO dental population, Journal of Public Health Dentistry, Vol.55, No.4, pp. 250- 252.
10. Torrungruang K, Tamsailom S, Rojanasomsith K, Sutdhibhisal S, Nisapakultorn K, Vanichjakvong O, et al. (2005), Risk Indicators of Periodontal Disease in Older Thai Adults. Journal of Periodontology, Vol.76, No.4, pp. 558-565.
11. Vathesatogkit P, Woodward M, Tanomsup S, Ratanachaiwong W, Vanavanan S, Yamwong S, et al. (2012), Cohort profile: the electricity generating authority of Thailand study, International Journal of Epidemiology, Vol.41, No.2, pp. 359-65.