International Journal of Community Medicine and Public Health
Javali SB et al. Int J Community Med Public Health. 2018 Dec;5(12):5301-5309
http://www.ijcmph.com
pISSN 2394-6032 | eISSN 2394-6040

DOI: http://dx.doi.org/10.18203/2394-6040.ijcmph20184807

Original Research Article

Prediction of risk factors of periodontal disease by logistic regression: a study done in Karnataka, India

Shivalingappa Basavantappa Javali*, Mohan Anantarao Sunkad, Appa Saheb Wantamutte

Department of Community Medicine, USM-KLE-International Medical Program, Nehrunagar, Belgaum, Karnataka, India

Received: 11 July 2018
Revised: 11 November 2018
Accepted: 13 November 2018

***Correspondence:**
Dr. Shivalingappa Basavantappa Javali,
E-mail: javalimanju@rediffmail.com

ABSTRACT

Background: The purpose of the study was to analyze the dependence of oral health diseases i.e. periodontal disease by Community Periodontal Index of Treatment Needs (CPITN) by considering the number of risk factors through the applications of logistic regression model.
Methods: This cross sectional study involves a systematic random sample of 600 permanent dentition aged between 18-40 years in Karnataka, India. The mean age was 34.26±7.28. The risk factors of periodontal disease were established by multiple logistic regression models using SPSS 21.0 statistical software.
Results: The factors like frequency of brushing, timings of cleaning teeth and type of toothpastes are significant persistent predictors of periodontal disease. The log likelihood value of full model is –1085.7876 and AIC is 1.2577 followed by reduced regression model are -1019.8106 and 1.1748 respectively for periodontal disease. The area under receiver operating characteristic (ROC) curve for the periodontal disease is 0.6128 (full model) and 0.5821 (reduced model).
Conclusions: The logistic regression model is useful in predicting risk factors like-frequency of brushing, timings of cleaning teeth and type of toothpastes for periodontal disease. The fitting performance of reduced logistic regression model is slightly a better fit as compared to full logistic regression model in identifying the these risk factors for both dichotomous periodontal disease.

Keywords: Akaike information criterion, Full model, Periodontal disease

INTRODUCTION

Periodontal disease remains a significant oral health problem and is a major cause of tooth loss.[1-3] It is seen in all the populations across the globe. The prevalence of severe periodontal disease may be low with its destructive element may be slow in pace but it offers unique opportunity to understand the causative factors and design preventive interventions.[4,5] Though a convincing natural history of this disease is still elusive, several factors are attributed as causatives. However those who have focused on the disease, brought in continuing scientific advances, geared towards the treatment of this disease, foresee early diagnosis is essential to limit the extent of loss and increase the potential for success of any definitive therapy provided.

In developing countries like India, the present trend indicates that there is an increase in prevalence of this disease and it plays an important role in deciding the oral health status. There have been quite good number of studies in the literature review however the issue needs some more insights from different angles. There are challenges in understanding the risk factors leading to the

disease, natural history of the disease, individual exposure cum resistance and the role of social determinants. Attempts are also made to find out the significant risk factors, which are influencing periodontal disease.[6,7] Thus, the risk assessment has become increasingly important in the prevention of caries and periodontal disease. However, the changes in our knowledge of the etiology and the reorganization of the probable significance of susceptibility factors as they affect initiation and progression, have led to intense study of specific risk factors. The clinical disease with its implicit care model may not offer much, while efforts towards functional periodontal tissue and its associative factors have abundant dividends.[8]

To investigate the risk factors associated with a periodontal disease, the regression methods have become an integral component of any data analysis concerned with the explanation of relationship between a response variable and one or more explanatory variables called factors.[9,10] Many different types of linear models have been seen in the literature and its use is discussed in many areas including dental epidemiology. The use of logistic regression modeling has been explored during the past few decades.[11,12] This method is now commonly applied in many fields of research including biomedical research, business and finance, criminology, economics, ecology, engineering, health policy, medicine, agriculture and dental epidemiology, etc. The logistic regression model is an important method to understand the principle that the goal of an analysis is the same as that of traditional model building technique used in statistical theory to find suitable description of relationship between response variable and a set of factors. In traditional linear regression techniques we assumed that dependent variable must be continuous or quantitative.

In logistic model, we consider the response variable is a categorical random variable, attaining only two possible outcomes called binary or dichotomous. This difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. In this article, the periodontal disease (CPITN index) is considered as dichotomized response variable. Since, the response variable is dichotomous, it is inappropriate to assume that it is normally distributed. Thus, the data cannot be analyzed using the traditional linear regression methods. It is convenient to denote one of the outcomes of response as without and with periodontal disease.

It is a standard practice to let the Y (Periodontal disease) to be two binary response variables, which are defined as

$$Y = \begin{cases} 1, & if\ Y\ (CPITN) > 0 \\ 0, & if\ Y\ (CPITN) = 0 \end{cases}$$

To explore the relationships between a set of factors i.e.

$$X' = (x_1, x_2, ..., x_p)$$

with response variable and the proportion of subjects with periodontal disease (CPITN>0) versus midpoint of each age intervals are presented in Figure 1.

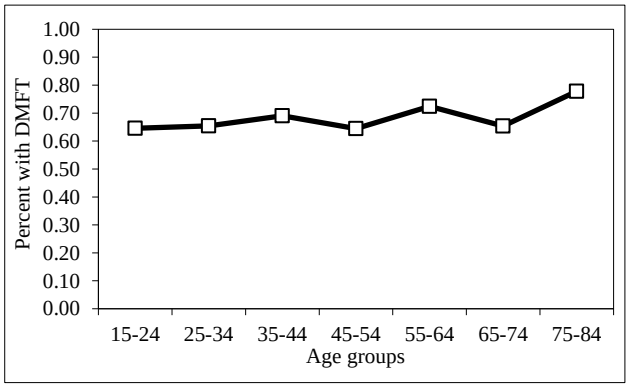

Figure 1: Plot of the percentage of subjects with periodontal disease (CPITN>0) in each age group.

Table 1: Code sheet for the independent factors.

No	Description	Code/values
1	Gender	Male=0, Female=1
2	Age (in years)	As a continuous
3	Religion	Hindu=1, Non-Hindu =2
4	Caste	SC/ST/OBC=1, GM=2
5	Socio-economic status	As a continuous
6	Family size	≤5 =0, >5=1
7	Staple food	Wheat/Rice/Jower=1, Others=2
8	Sources of drinking water	Pipeline/River/Pound=1, Tube well/Hand pump=2
9	Types of diet	Vegetarian=0, Non-vegetarian=1
10	Time for sweet consumption	During/Between meals=0, During and between meals=1
11	Frequency of sweet consumption	≤2 times =0, >2 times =1
12	Oral hygiene habits	Finger/datun/others =0, Tooth Brush =1
13	Frequency of brushing	Once=1, Twice or more=2
14	Timings of cleaning the teeth	Morning or Night=1, Morning and Night=0
15	Methods of brushing	Circular/ Vertical=1, Horizontal=2
16	Materials used for brushing their teeth	Paste/powder=1, Others=2
17	Types of toothpaste	Non-fluoridated=0, Fluoridated=1
18	Mouth rinsing habit	No=0, Yes=1
19	Smoking habit	No=0, Yes=1
20	Chewing habit	No=0, Yes=1
21	Alcohol habit	No=0, Yes=1

The main aim of this study is to find the significant risk factors of periodontal disease dichotomized data by utilizing the applications of multiple logistic regression models.

Response variable and independent factors

The periodontal disease by CPITN Index is the response variable. For analysis purpose, the CPITN index is based on ordered scoring criteria ranging from 0 to 4 that are described as below:

1. 0 if no periodontal disease (healthy gums),
2. 1 if bleeding on gentle probing (gums)
3. 2 if calculus felt during the probing
4. 3 if periodontal pocket depth between 3.5 to 5.5 mm
5. 4 if periodontal pocket depth of 6 mm or more.

For convenience of statistical modeling, the CPITN scores are dichotomized as 0 if CPITN=0 and 1 if CPITN>0. Apart from response variables, the data set on independent factors like socio-economic-demographic, food habits, eating habits, oral hygiene practices and deleterious habits are obtained by structured questionnaire with personal interview method. The detailed code sheet for the independent factors which are binary in nature except age and socio-economic status is presented in Table 1.

METHODS

Study area, study population and sampling procedure

This is a cross sectional study, the study involves a systematic random sample of 600 permanent dentition aged between 18-40 years. Sample size was determined based on results of pilot study, with precision of 5% and 99% of confidence level, the sample size was estimated to be 556 ≅600. The mean age of the study subject was 34.26±7.28.

Clinical examination

The selected subjects were called for free dental examination. The dental examination of periodontal disease was carried out by two well qualified community dentists with plane mouth mirror, CPITN Probe, dental explorer, disposable gloves and sterilized instruments under artificial light for each subject. The findings of the periodontal disease examination by CPITN index were recorded according to diagnostic criteria recommended by World Health Organization (3, 31). A pilot study was conducted to assess the intra and inter examiner agreement for recording CPITN index scores on convenient samples of 50 study subjects. The intra-examiner agreement was 0.8984 (first examiner) and 0.8810 (second examiner) for CPITN index. The inter examiner (between the two examiners) agreement in assessment of periodontal disease by CPITN index was found to be 0.8969. The data on selected explanatory variables were collected and recorded by structured questionnaire with interview method.

Formulation of multiple logistic regression models

Nelder and Weldderburn introduced generalized linear models which are a class of statistical models.[13] It includes response variables that follow any probability distribution in the exponential family of distributions. An excellent treatment of generalized linear models is presented in Agresti.[14] In this study the periodontal disease (CPITN Index) is binary; therefore, the logistic regression model is an appropriate model, which is a part of generalized linear models.

The response variable in logistic regression is usually dichotomous, that is, the response variable can take the value 1 with a probability of success P, or the value 0 with probability of failure 1-P. This type of variable is called a Bernoulli (or binary) variable. The relationship between the predictor and response variables is not a linear function in logistic regression; instead, the logistic regression function is used, which is given as

$$P(x)=\frac{e^{(\alpha+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k)}}{1+e^{(\alpha+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k)}}$$

We now find the link function for which the logistic regression model is a generalized linear model (GLM). For this model the odds of making response 1 are

$$\frac{P(x)}{1-P(x)}=e^{(\alpha+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k)}$$

$$\log\left[\frac{P(x)}{1-P(x)}\right]=\alpha+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k$$

Thus the appropriate link is the log odds transformation, the log it. The logistic regression model is given by

$$\log it\ [P(x)]=\frac{P(x)}{1-P(x)}=e^{(\alpha+\beta_1x_1+\beta_2x_2+\ldots+\beta_kx_k)}$$

The parameters in this model, α, β_1, $\beta_2\ldots$ β_k can no longer be estimated by least squares, but are found using the maximum likelihood method.[15] Logistic regression calculates the probability of success over the probability of failure; therefore, the results of the analysis are in the form of an odds ratio. Logistic regression also provides knowledge of the relationships and strengths among the variables.

Fitting the multiple logistic regression model

Suppose assuming a sample of n independent observations (x_i, y_i) i=1, 2, 3… n., fitting the model

requires estimates of the vector $\beta' = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)$. The likelihood of β is given by

$$l(\beta) = \prod_{i=1}^{n} P(x_i)^{y_i}[1 - P(x_i)]^{1-y_i}$$

where $P(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$

The likelihood function is given by

$$\log l(\beta) = \sum_{i=1}^{n} \{y_i \log[P(x_i)] + (1 - y_i)\log[1 - P(x_i)]\}$$

Here we get (k+1) likelihood equations that are obtained by differentiating the log likelihood function with respect to the k+1 coefficient. The likelihood equations obtained may be expressed as follows:

$$\sum_{i=1}^{n}[y_i - \pi(x_i)] = 0 \quad \text{and} \quad \sum_{i=1}^{n} x_{ij}[y_i - \pi(x_i)] = 0$$

for j=1, 2, …, k.

Let $\hat{\beta}$ denote the solution to these equations. Thus, the fitted values of multiple logistic regression model are $\hat{P}(x_i)$. Then, the method of estimating the variance and co-variances of these estimated coefficients follows from the well-developed theory of maximum likelihood estimation.[16] This theory states that estimators are obtained from the matrix of partial derivatives of the log likelihood function. Further, the estimated standard errors of the estimated coefficients of logistic regression model is given by $S\hat{E}(\hat{\beta}_j) = \sqrt{\hat{Var}(\hat{\beta}_j)}$ for j = 0, 1, 2, …,p. Alternatively, the Wald or Z statistic is commonly used to test the significance of individual logistic regression coefficients for each independent variable. The test statistic is given by $W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}, j = 0,1,2,..., p$

The multivariable analog of the Wald test is given by

$$N = \hat{\beta}_j\left[\hat{Var}(\hat{\beta})\right]^{-1}\hat{\beta} = \hat{\beta}^I(XVX)\hat{\beta}$$

The statistic N is distributed as chi-square with (k+1) degrees of freedom under the hypothesis that each of the k+1 coefficient is equal to zero. The multivariable analog of the score for the significance of the model is based on the distribution of p derivatives of L (β) with respect to β.

It is distributed as chi-square with 1 degree of freedom. The Wald statistic is simply the square of the (asymptotic) *t*-statistic. The Wald statistic can be used to calculate a confidence interval for β_j. We can assert with 100(1−α)% confidence interval that the true parameter lies in the interval with boundaries $\hat{\beta} \pm Z_{\alpha/2}(ASE)$, where ASE is the asymptotic standard error of logistic $\hat{\beta}$. Parameter estimates are obtained using the principle of maximum likelihood; therefore hypothesis tests are based on comparisons of likelihoods or the deviances of nested models. The likelihood ratio test uses the ratio of the maximized value of the likelihood function for the full model (L1) over the maximized value of the likelihood function for the simpler model (L_0). The likelihood-ratio test statistic equals:

$$-2\log\left(\frac{L_0}{L_1}\right) = -2[\log(L_0 - L_1)] = -2(L_0 - L_1)$$

This log transformation of the likelihood functions yields a chi-squared statistic. This is the recommended test statistic to use when building a model through backward elimination procedure. Once $\hat{\beta}$ has been obtained, the estimated value of the linear systematic component (also known as linear predictor) of the model is $\hat{\eta}_i = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_{ki}$ From this, the fitted probabilities $\hat{P}_i$ can be found using $\hat{P}_i = \frac{e^{\hat{\eta}_i}}{1 + e^{\hat{\eta}_i}}$.

The purpose of logistic regression is to correctly predict the category of outcome for individual cases using the most parsimonious model. To accomplish this purpose, a model is created that includes all predictor factors that are useful in predicting the response variable. Several different options are available during model creation. Factors can be entered into the model in the order specified by the researcher or logistic regression can test the fit of the model after each coefficient is added or deleted (backward elimination procedure), called stepwise procedure. Where the analysis begins with a full or reduced model and factors are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each factor to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed

Sensitivity, specificity and receiver operating characteristic curve

The sensitivity and specificity of the CPITN index was calculated. Further, the diagnostic performance of a test or the accuracy of a test to discriminate diseased cases from normal cases is evaluated using Receiver Operating Characteristic (ROC) curve analysis.[17-21] ROC curve also be used to compare the diagnostic performance of two or more laboratory/diagnostic tests. A test with perfect discrimination (no overlap in the two distributions) has a ROC plot that passes through the upper left corner (100% sensitivity, 100% specificity). Thus, closer the ROC plot to the upper left corner, the higher is the overall accuracy of the test.

Analysis and interpretation

The dichotomized periodontal disease data is analyzed. We have computed measures of central tendency and

dispersion for both continuous and nominal variables. The multiple logistic regression models are constructed between the binary response variables i.e. periodontal disease (CPITN) with factors independently.

The model estimation, in the first step, the multiple logistic full model is constructed for considering all factors and in the second step, the stepwise called multiple logistic reduced model is performed by considering only significant factors from the full model. In the selection procedure using the stepwise multiple logistic model analysis, we first select the variable having a greatest influence power. Then the effect of this variable is eliminated from the information content of all the other variables. The variable, which then has the greatest power of influence after the above elimination procedure, is ranked as the second etc. Thus, the variables are listed in decreasing order with respect to probability of additional information gained from including further variables was less than 0.05. Lastly, the variable having the weaker influence power may be dropped from the final analysis. In order to weigh the significance of each chosen variable with respect to their influence, its correlation with multiple logistic regression model and parameter estimates, standard error of estimates, odds ratios, 95% confidence intervals and p-values of each factor computed.

RESULTS

Comparisons in terms of parameter estimates, log likelihood and AIC values in particular are carried, discussed and presented. A total of 600 subjects are included in the study (50.00% are males and 50.00% are females) which has mean age as 34.26 and mean family size as 2.94. Similarly, 409 (68.18%) are Hindus, 575 (95.85%) are backward castes, 350 (58.24%) are with high socio-economic status, 100 (16.59%) are users of wheat or rice or Jower as a main staple food, 488 (81.36%) are drinking tube well / hand pump water, 342 (56.93%) are non-vegetarian, 560 (93.41%) are eating sweet during or between meals, 571 (95.11%) are taking sweet consumption at least twice in a day, 317 (52.78%) are brushing their teeth with tooth brushes as an oral hygiene habit, 500 (83.30%) are brushing their teeth only once in a day, 486 (80.97%) are brushing their teeth both in morning and night, 498 (83.01%) are brushing their teeth by horizontal method, 434 (72.33%) are brushing their teeth by paste/powder, 392 (65.34%) are users of non-fluoridated toothpastes, 328 (54.66%) are changing their toothbrush once in four months, 393 (65.57%) are not rinsing their mouth after every meal with water, 461 (76.82%) are smokers, 286 (47.73%) are chewers and 328 (54.66%) are alcohol drinkers as compared to their counterparts.

Analysis of periodontal disease (CPITN index)

Out of 21 factors, only five factors are found to be statistically significant factors of periodontal disease ($p<0.05$) such as gender, frequency of brushing, timings of cleaning the teeth and type of toothpastes and family size. These significant factors have different signs of influence (Table 2). Further, a log likelihood value of model is –1085.7876. The Akaike's information criterion (AIC) value is 1.2577. But, our goal is to estimate the best fitting model for periodontal disease while minimizing the number of factors. The reduced regression model is obtained by excluding the factors which are not significant in full model. But only 5 factors are found to be significant predictors of periodontal disease in reduced modal (Table 2). The log likelihood and AIC value of the reduced regression model are -1098.4320 and 1.2539 respectively. Based on log likelihood and AIC values, the full and reduced logistic regression models have approximately similar fit. Thus, there is no advantage in excluding some factors from the model for assessment of significant determinants of occurrence of periodontal disease.

Table 2: The estimated coefficients of factors from full and reduced logistic regression model to periodontal disease dichotomous data.

Factors	Full model		Reduced model	
	Estimate	Std. Err.	Estimate	Std. Err.
Constant	1.7142	0.8427	1.1382*	0.2185
Gender	0.4064*	0.1168	0.3980*	0.1035
Age (in years)	0.0531	0.0391		
Religion	-0.0246	0.1125		
Caste	-0.0619	0.0644		
Socio-economic status	-0.1165	0.0827		
Family size	0.2981*	0.1598	0.2881*	0.1579
Staple food	0.0481	0.1826		
Sources of drinking water	-0.3023	0.1990		
Dietary habits	-0.1461	0.1078		
Time for sweet consumption	0.0978	0.3221		
Frequency of sweet consumption	0.3592	0.3746		
Oral hygiene habits	-0.0722	0.1046		
Frequency of brushing	-0.3533*	0.1430	-0.2769*	0.1368

Continued.

Factors	Full model		Reduced model	
	Estimate	Std. Err.	Estimate	Std. Err.
Timings of cleaning the teeth	-0.3069*	0.1431	-0.1303	0.1326
Methods of brushing	-0.0174	0.1189		
Materials used for brushing their teeth	0.1733	0.1684		
Type of toothpastes	-0.4708*	0.1418	-0.3888*	0.1110
Mouth rinsing habit	0.0093	0.1173		
Smoking habit	0.0199	0.1325		
Chewing habit	0.2154	0.2095		
Alcohol habit	0.3237	0.2130		
Log likelihood	-1085.7876		-1098.4320	
AIC	1.2577		1.2539	

*Significant at 5%level of significance (p<0.05).

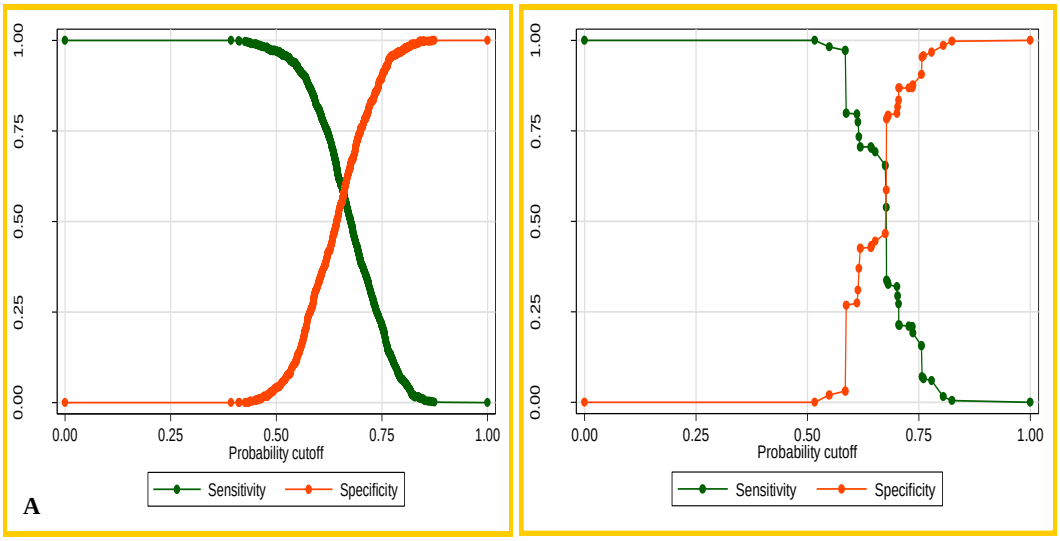

Figure 2: The Plot of sensitivity and specificity versus criterion value for the response variable (CPITN Index) in the full and reduced model: (A) Full model; (B) reduced model.

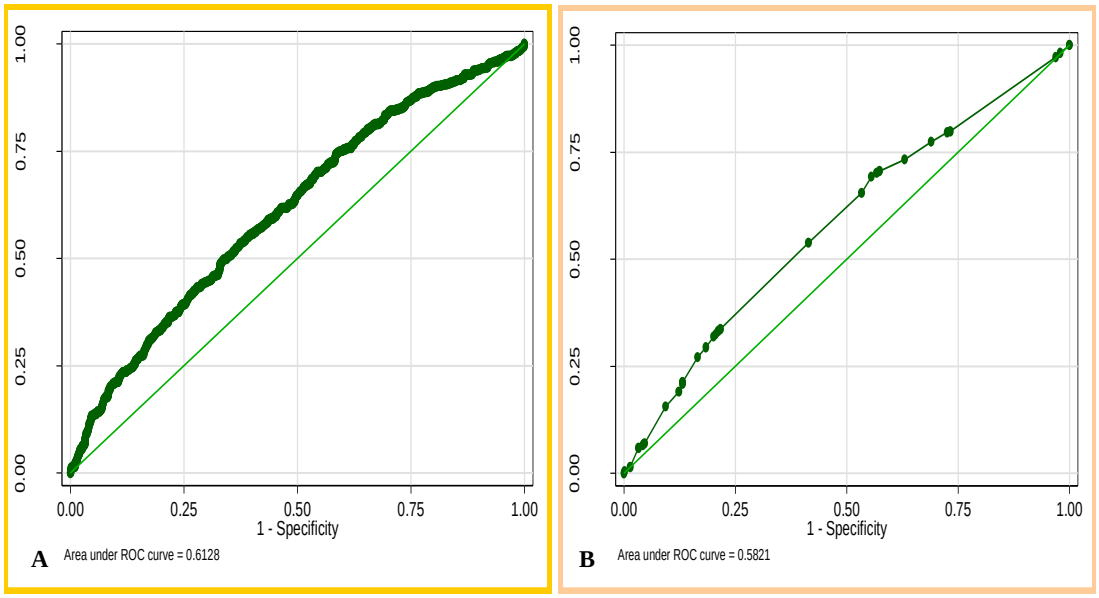

Figure 3: The accuracy of the test in the means of ROC (CPITN Index) for full and reduced model: (A) Full model; (B) reduced model.

The estimated odds ratio of gender (OR=1.5015, 95% CI: 1.1942-1.8878), family size (OR=1.3474, 95% CI: 0.9851-1.8428), frequency of brushing (OR=0.7024, 95% CI: 0.5307-0.9296), timings of cleaning the teeth (OR=0.7357, 95% CI: 0.5558-0.9740) and type of toothpastes (OR=0.6245, 95% CI: 0.4730-0.8246) have found to be significant ($p<0.05$). It means that, the gender, family size, frequency of brushing, timings of cleaning the teeth and type of toothpastes have a significant influence on periodontal disease. In other words, the women living in a larger family (>5 members in a family) brushing their teeth only once a day in a morning, without pastes/powder have higher prevalence of periodontal disease as compared to their counterparts. However, there is an improvement in the strength of association among some covariates seen in reduced model.

The plot of sensitivity and specificity versus criterion value for the response variable (CPITN Index) in the full and reduced model is presented in the following Figure 2.

The area under ROC curve of the response variable (CPITN Index) for the full model is 0.7509 and for the reduced model is 0.5821. It provides a summary of the accuracy of the diagnostic test which is approximately 61% in full model and 58% in reduced model. The accuracy of the test in terms of ROC (CPITN index) is presented in the following Figure 3.

DISCUSSION

Different statistical methods i.e. regression analysis, multilevel modeling, logistic regression and ordinal regression techniques are employed to analyze the periodontal disease (CPITN index), data yield results having different focuses. The regression methods allow researchers to identify factors related to social-economic-demographic and other factors related to oral health that contribute to overall status of response variables. These methods also permit researcher to estimate the magnitude of the effect of factors. Therefore, the periodontal disease data is converted into binary or dichotomous outcomes (with and without disease). Use of a dichotomous outcomes in traditional multiple linear regression model violates the assumption of normality and homoscedasticity. Hence, the main aim of the present study is to utilize the applications of logistic regression model to assess the effect and relationship between factors and binary response variable. Changes in our knowledge of the etiology of periodontal disease and the recognition of the potential importance of susceptibility factors as they affect initiation and progression of periodontal disease, have led to an intense study of specific risk factors for periodontal disease. Impact of different factors (determinants) on a periodontal disease has already been investigated in several studies. The gender factor is associated with periodontal disease. It means that, the periodontal disease is more prevalent in males than in females at any comparable ages. This result coincides with several studies done by Miller et al and Grossi et al.[22-26] Males usually exhibit proper oral hygiene than females. The reasons for these gender differences are not clear and their elucidation may reveal important destructive or protective mechanism.

The age is an insignificant factor having positive association with periodontal disease in the study. However, the studies on periodontal disease prevalence with extent and severity show that disease is more prevalent in older age groups as compared to younger groups.[27-29] Also it is found that the severity of the disease is more with respect to plaque development and gingivitis in elderly persons as compared to younger persons.[30-32]

The relationship of periodontal disease and socioeconomic status can be viewed globally, where wide variations in socio-economic status among different populations are compared. These studies compare populations from developing countries with those from industrialized countries which suggest that periodontal disease may be associated with nutritional deficiencies.[33-35] However, in this study an association is not found to be statistically significant. But, the Ramfjord et al found that the periodontal condition of young men in India who exhibited clinical symptoms of general malnutrition is not different from that of the periodontal condition of well-nourished individuals.[36] Non-Hindus showed that they have apparently more periodontal destruction compared to Hindus. No studies are found in relation to religion on Indian population with respect to periodontal disease.

There is a history of association between tobacco smoking and periodontal disease and prevalence of acute ulcerative gingivitis (ANUG) was demonstrated as early in 1946.[37-39] However, the perception that greater levels of plaque and calculus is more in smokers than that in non-smokers. In this study, it is shown that smoking tobacco is not significantly associated with periodontal disease. This result coincides with some of the earlier studies.[40-42] It is likely that smoking is a major factor for destructive periodontal disease in man. Hence the modification of this factor is important in the treatment and prevention of periodontal disease.

Further, in this study, we compared performance of full logistic model with that of reduced logistic model using log likelihood estimate for CPITN index data. The results show that, the fitting performance of full logistic regression model is slightly better as compared to reduced logistic regression model applied to dichotomized CPITN index data. Studies using logistic regression tool are few and scarce, perhaps dentists need encouragement.

CONCLUSION

The fitting performance of reduced logistic regression model is slightly a better fit as compared to full logistic

regression model in identifying the these risk factors for both dichotomous periodontal disease. Our observations point towards the frequency of brushing, timings of cleaning teeth and type of toothpastes as the significant risk factors of periodontal disease.

ACKNOWLEDGEMENTS

We acknowledge the support, encouragement and cooperation provided by our Institutional head to do this study. We are indebted to the entire faculty who whole heartedly strengthened our hands for clinical work in the community. We thank and cherish the memory of all the community members for their enthusiastic participation in this study. We are also thankful to many office staff for secretarial assistance in preparing the manuscript.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. LaMonte MJ, Genco RJ, Hovey KM, Wallace RB, Freudenheim JL, Michaud DS, et al. History of Periodontitis Diagnosis and Edentulism as Predictors of Cardiovascular Disease, Stroke, and Mortality in Postmenopausal Women. J Am Heart Assoc. 2017;6(4):e004518.
2. Zhang Q, Li Z, Wang C, Shen T, Yang Y, Chotivichien S, et al. Prevalence and predictors for periodontitis among adults in China, 2010. Glob Health Action. 2014;7:24503.
3. Hess RF, Gililland CS, Dembélé J. Prevalence and Predictors of Periodontal Disease among Pregnant Women in Mali, West Africa. Ann Med Health Sci Res. 2017;7:263-70.
4. Almerich-Silla J-M, Almiñana-Pastor PJ, Boronat-Catalá M, Bellot-Arcís C, Montiel JM. Socioeconomic factors and severity of periodontal disease in adults (35-44 years). Clin Exp Dent. 2017; 9(8):988–94.
5. Bokhari SA, Suhail AM, Malik AR, Imran MF. Periodontal disease status and associated risk factors in patients attending a Dental Teaching Hospital in Rawalpindi, Pakistan. J Indian Soc Periodontol 2015;19:678-82.
6. Najeeb S, Zafar MS, Khurshid Z, Zohaib S, Almas K. The Role of Nutrition in Periodontal Health. Nutrients. 2016;8(9):530.
7. Pandit PV, Javali SB. Multiple logistic regression model to predict risk factors of oral health diseases. Revista Română de Statistică. 2012;60(05):73-86.
8. AlJehani YA. Risk Factors of Periodontal Disease: Review of the Literature. Int J Dent. 2014;2014:182513.
9. Martinez-Herrera M, Silvestre-Rangil J, Silvestre F-J. Association between obesity and periodontal disease. A systematic review of epidemiological studies and controlled clinical trials. Med Oral Patol Oral Cir Bucal. 2017;22(6):708-15.
10. Thyvalikakath T, Song M, Schleyer T. Perceptions and attitudes toward performing risk assessment for periodontal disease: a focus group exploration. BMC Oral Health. 2018;18(1):90.
11. Stoykova M, Musurlieva N, Boyadzhiev D. Risk factors for development of chronic periodontitis in Bulgarian patients (pilot research). Biotechnol Biotechnol Equip. 2014;28(6):1150–4.
12. Pindborg JJ. Tobacco and gingivitis: statistical examination of the significance of tobacco in the development of ulcero membranous gingivitis and in the formation of calculus. J Dent Res. 1947;26:261-4.
13. Agresti A. Tutorial on modeling ordered categorical response data. Psychol Bull. 1989;105(2):290-301.
14. David WH, Stanley. Applied Logistic Regression, Second edition, A Wiley Inter science Publication, John Wiley and Sons, Inc. New York; 2000.
15. Metz CE. Basic principles of ROC analysis. Seminars in Nuclear Medicine. 1978;8:283-98.
16. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots, a fundamental evaluation tool in clinical medicine. Clinical Chemistry. 1993;39(4):561-77.
17. Grossi SG, Genco RJ, Machtei EE, Ho AW, Koch G, Dunford R, et al. Assessment of risk for periodontal disease. II. Risk indicators for alveolar bone loss. J Periodontol. 1995;66(1):23-9.
18. Herulf G. On the marginal alveolar ridge in adults. Sven Tandlak Tidsk. 1968;63:675-703.
19. Macgregor ID, Edgar WM, Greenwood AR. Effects of cigarette smoking on the rate of plaque formation. J Clin Periodontol. 1985;12(1):35-41.
20. Marshal-Day CD, Stephens RG, Quigley LF. Periodontal disease prevalence and incidence. J Periodontol. 1955;26:185-203.
21. Brunelle JA, Carlo JP, Brown LJ, Loe H. Oral Health of United States Adults. National findings. Bethesda MD. National Institute of Dental Research. NIH publication No. 87, 2868, 1987.
22. Nelder JA, Weldderburn RWM. Generalized linear models. Journal of the Royal Statistical Society A. local. 1972;135:370-84.
23. Grossi SG, Zambon JJ, Ho AW, Koch G, Dunford RG, Machtei EE, et al. Assessment of risk for periodontal disease-I, Risk indicators for attachment loss. J Periodontol. 1994;65(3):260-7.
24. Grossi SG, Genco RJ, Machtei EE, Ho AW, Koch G, Dunford R, et al. Assessment of risk for periodontal disease. II. Risk indicators for alveolar bone loss. J Periodontol. 1995;66:23-9.
25. Griner PF, Mayewski RJ, Mushlin AI, Greenland P.. Selection and interpretation of diagnostic tests and procedures. Ann Internal Med. 1981;94;557-92.
26. Abdellatif HM, Burt BA. An epidemiological investigation into the relative importance of age and oral hygiene status as determinants of periodontitis. J Dent Res. 1987;66;13-8.

27. Agresti A. Tutorial on modeling ordered categorical response data. Psychological Bulletin, 1989;105:290-301.
28. Ainamo J, Barmes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). Int Dent J. 1982;32(3);281-91.
29. Bergstrom J, Floderus-Myred B. Co-twin control study of the relationship between smoking and some periodontal disease factors. Community Dent. Oral Epidemiol. 1983;11:113-6.
30. Bergstrom J. Oral hygiene compliance and gingivitis expression in cigarette smokers. Scand J Dent Res. 1990;98:497-503.
31. Bergstrom J. Short term investigation on the influence of cigarette smoking upon plaque accumulation. Scand J Dent Res. 1981;89:235-8.
32. Cox DR, Snell EJ. Analysis of Binary Data. Second Edition. London: Chapman and Hall; 1970.
33. David WH, Stanley L. Applied Logistic Regression, Second edition, A Wiley Inter science Publication. New York: John Wiley and Sons Inc.; 2015.
34. Eija KK, Leena R. Relationship of oral hygiene and sugar consumption to risk of caries in children. Community Dent Oral Epidemiol. 1982;10:224-33.
35. Ramfjord SP, Emslie RD, Greene JC, Held AJ, Waerhaug J. Epidemiological studies of periodontal disease. Am J Public Health. 1968;28:1713-22.
36. Rao CR. Linear Statistical Inference and its application. Second edition. New York: Wiley Inc; 2001.
37. Russell AL. Geographical Distribution and Epidemiology of Periodontal Disease. Geneva. World Health Organization, (WHO/DH/33/34).
38. Schei O, Waerhaug J, Lovdal A, Arno A. Alveolar bone loss as related to oral hygiene and age. J Periodontol. 1959;30:7-16.
39. Sheiham A. Periodontal disease and oral cleanliness in tobacco smokers. J Periodontol. 1971;42:259-63.
40. Solomon HA. Cigarette smoking and periodontal disease. J Am Dent Assoc. 1968;77:1081-4.
41. U. S. Public Health Service. National center for Health Statistics. Basic data on Dental examination findings of persons 1-75 years, United States 1971-1974. DHEW publication No. (PHS) 79-1662, Series 11. No. 214. Washington DC. Government printing office, 1974.

Cite this article as: Javali SB, Sunkad MA, Wantamutte AS. Prediction of risk factors of periodontal disease by logistic regression: a study done in Karnataka, India. Int J Community Med Public Health 2018;5:5301-9.