Proving the Associated Factor That Contributing the Caries Status among Pre-School Children in Bachok Kelantan

Wan Muhamad Amir W Ahmad, Ruhaya Hasan, Mohd Fadhli Khamis, Nor Azlida Aleng, Noraini Mohamad

Abstract. This paper focuses on the proving of the associated factor for the caries status among preschool children in Bachok, Kelantan. This research paper is mainly focused on the potential factor that most contributing to Early Childhood Caries (ECC). There are two methodologies approaches in this research paper which is Decision Tree Analysis (DTA) and Multi-Layer Perceptron (MLP). The results from both analyses can be used to assist the public and also the stakeholder to control the prevalence of ECC in the future. Results from both analyses are also very useful to redesign the health treatment among pre-school children, to educate the parents, teachers, and to improve the service which offered by the ministry of health from time to time by focusing the most influential factors which lead to ECC in the local community. According to the result of Decision Tree Analysis, the most factor that leads to caries status among preschool children are father's occupation, household income, children's weight and the type of water used in their house. While using Multi-Layer Perceptron (MLP) neural networks modeling, the factor can be summarized as household income factor, children's weight, father occupation and also the type of water used in their house. From the results of the Decision Tree and Multi-Layer Perceptron (MLP) reveals that the top three factors lead to ECC were household income factor, children's weight, father's occupation. This information will provide a very useful information to forecast ECC status among preschool children.

Keywords: Decision Tree Analysis, Multilayer Perceptron, Neural Network, Childhood Caries.

I. INTRODUCTION

ECC or Early childhood caries is one of the most prevalent chronic childhood diseases. It is triggered by a complex time-dependent interaction between fermentable carbohydrates and acid-producing bacteria. The latter cause's demineralization of inorganic component and dissolution of organic structure of the tooth, compromising the overall structure of tooth and often leading to cavitation [1]. The Australian national survey highlighted that low parental education, low family income and residence in rural area and / or remote areas were some crucial social contributing factor of higher prevalence of tooth decay in their children. Living in rural environments is an important risk factor for ECC due to the low socioeconomic status, poor nutritional habits, low level of parents' education and poor access to fluoridated water or dental services [2]. High ECC conditions in rural communities may have some consequences such as hospitalization and emergency care due to dental infection, affecting the quality of life of preschool children involved [3]. A South Australian study of children aged 4 to 9 years found positive associations between primary caries experience and the consumption of water from varied sources compared to the public water supply [4] In addition to water consumptions factor, nutritional status was also associated with the risk of ECC. There was a higher risk of developing dental caries in low-weight children, than overweight-obese children. While overweight children aged 6-7 years had a significantly lower dental caries severity than children of normal BMI for-age in [5] and consumption of sugar-sweetened beverages and feeding and eating patterns were more likely to develop ECC in [6], [7]. Dental preventive activity such as tooth brushing twice a day or more often reduced the odds of children having dental caries by half compared to those who brushed less often. On the other hand, children whose father's occupation was categorized as manual, the frequency of toothbrushes was not significantly associated with caries prevalence. Reference [8] proposed that tooth brushing perhaps more effective in the non-manual groups and supported this argument with proof that children in the manual occupation group were more likely to brush their teeth compared to the children in the non-manual occupation group. While [9] found that household income was inversely associated with the occurrence of dental caries which are regular with the previous findings that household income were significantly positively associated with dental caries in among children.

A recent study concluded that lower family income was associated with dental caries among children aged 0 to 6 years [9]. In addition, [10] supports the statement that household income were significantly related with caries in the primary dentition. From literatures, multiple factors have been proven associated with ECC status in different localities and statistical analyses methods. However, the factors in school children in Bachok, Kelantan have not been fully explored.

Revised Manuscript Received on November 15, 2019.

* Correspondence Author

Wan Muhamad Amir W. Ahmad*, School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM), 16150 Kubang Kerian, Kelantan, Malaysia.

Ruhaya Hasan, School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM), 16150 Kubang Kerian, Kelantan, Malaysia.

Mohd Fadhli Khamis, School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM), 16150 Kubang Kerian, Kelantan, Malaysia.

Nor Azlida Aleng, School of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030, Kuala Terengganu, Malaysia.

Noraini Mohamad, School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM), 16150 Kubang Kerian, Kelantan, Malaysia.

Thus this study aimed to determine the influential factors with two methods of analyses ie Decision Tree Analysis (DTA) dan Multi-Layer Perceptron (MLP).

II. MATERIAL AND METHODS

This study was conducted in Bachok, Kelantan which is held by 382 children which are 44.8% male and 55.2% female. The sample size was calculated by using G*power with effect size = 0.02, $\alpha = 0.05$, the power of the study = 0.68 and number of predictor were 2 (under multiple linear regression calculation). The minimum sample size requires is 372 respondents. Table 2.1 gives a description of the data which taken from preschool children.

Table 2.1 Description of data among preschool children

Variables	Code	Description
dft	Y	Carries Status 0 = No Caries, 1 = Low, 2 = Moderate, 3 = High
Water	x_1	Type of Water Used in House 1 = Using Well Water, 2 = Using Municipal Water
Weight	x_2	Weight of respondent
Father Occupational	x_3	Occupation of fathers 1 = Not working, 2 = Pension, 3 = Self employment, 4 = Hire worker, 5 = Employer with hire worker
Household Income	x_4	Household income
Sweet Tea	x_5	Having sweet tea every day 1 = Yes, 0 = No

FITTING DECISION TREE ANALYSIS

Decision tree is a graphical model techniques which describing decisions and their possible outcomes. A decision tree is a great and efficient method for classification, prediction and for facilitating decision making in sequential decision problems. Decision trees allow for an intuitive understanding of the problem and aid for the optimal decision making. A decision tree consists of three types of nodes (a) decision node (b) chance node (c) Endpoint node/Terminal node [11]. Figure 2.1 shows the concept of a decision tree.

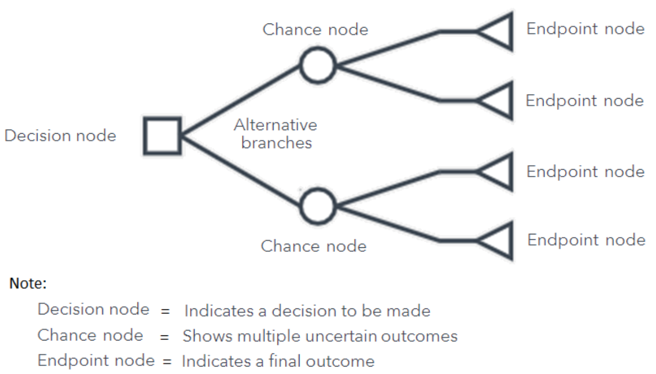

Figure 2.1. Decision trees are graphical models for describing sequential decision problems

MULTILAYER PERCEPTRON NEURAL NETWORK (MLP)

The architecture of neural network is consist of the number of input, hidden and output nodes. There are 6 selected variables, which were caries status, father's occupation, household income, children's weight, the type of water used in their house and having sweet tea every day. In this study, the dependent variable is referred to caries status (output node). The data was partitioned into two parts which are training (50%) and testing (50%).The Levenberg-Marquardt back-propagation is used as the training algorithm since it was claimed as the best training algorithm [12]. Below is the path which using SPSS Modeller to obtain the neural network analysis and decision tree analysis.

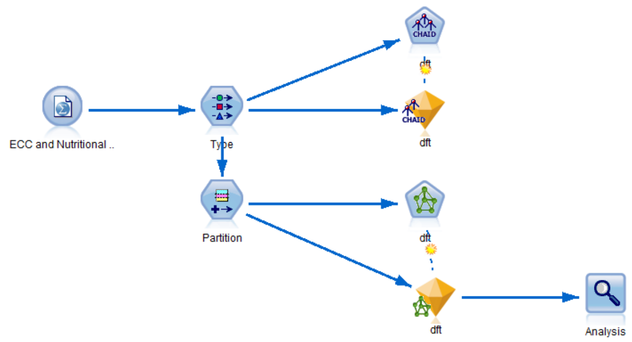

Figure 2.2. Path Analysis for Decision Tree Analysis and Multi-Layer Perceptron (MLP) Neural Networks

MLP contains of an input layer, hidden layers and an output layer. An essential factor of successes of the neural networks depends on the training network. Among the several learning algorithms available, back-propagation (BP) has been the most popular and most widely implemented [13]. The output node is fixed at one since there is only one independent variable which is referred to caries status. Thus, for the perceptron network with *N* input nodes, *H* hidden nodes, and one output node, the values $\hat{y}$ are given by:

$$\hat{y} = g_2\left(\sum_{j=1}^{H} w_j h_j + w_0\right) \qquad (1)$$

where w_j an output weight from hidden node *j* to the output node is w_0 the bias for the output node, *g* is an activation function. The values of the hidden node h_j, *j*=1…*H* are given by;

$$h_j = g_1\left(\sum_{i=1}^{N} v_{ji} x_i + v_{j0}\right),\ j = 1,\ldots,\mathrm{H} \qquad (2)$$

here, v_{ji} the output *weight* from input node I to hidden node j, v_{j0} is the bias for hidden node j, x_i is the independent variables where i=1…N and k is an activation function. The architecture of the multilayer perceptron neural network model is illustrated in Figure 2.2.

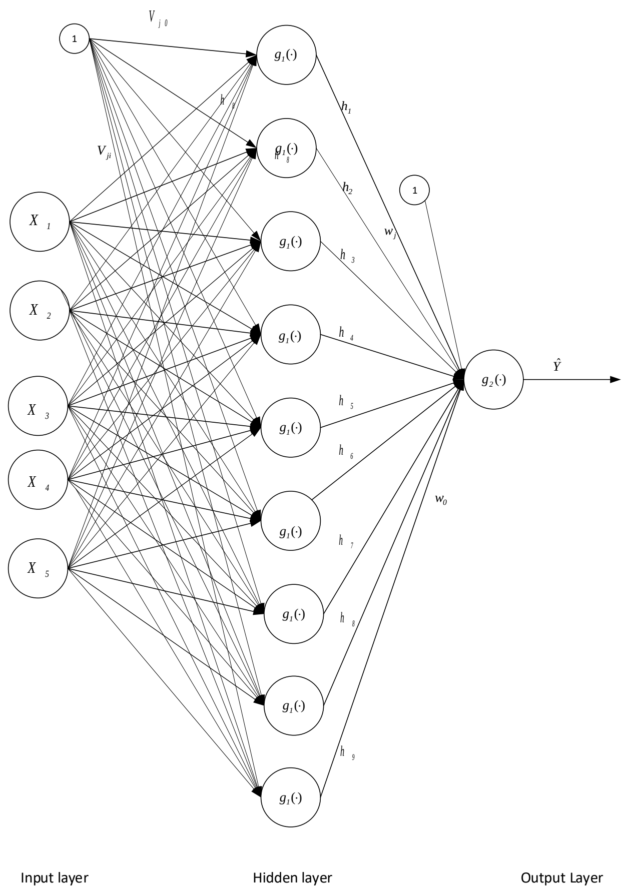

Figure 2.3. Multilayer perceptron neural network architecture model

III. RESULTS AND DISCUSSION

In this section, the output from the two methodologies will be discussed further based on their most important predictor and their contribution to the caries status. The first section, the decision tree analysis result will be discussed and the second section Multilayer Perceptron Neural Network (MLP). According to Figure 3.1, the top three predictors are ranking to their contribution is the occupation of father, household income, and children weight. Using the CHAID method, these three are the best predictors of caries status. In total there were four out of five predictors that the model deemed important.

DECISION TREE ANALYSIS

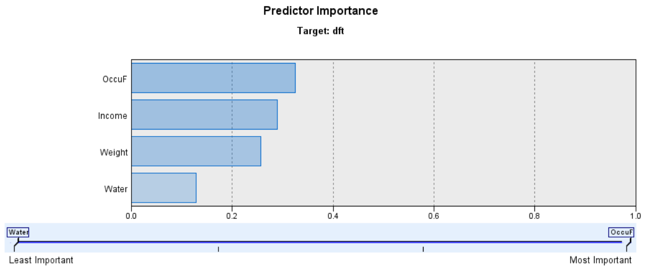

Figure 3.1. Predictor important (using decision tree analysis)

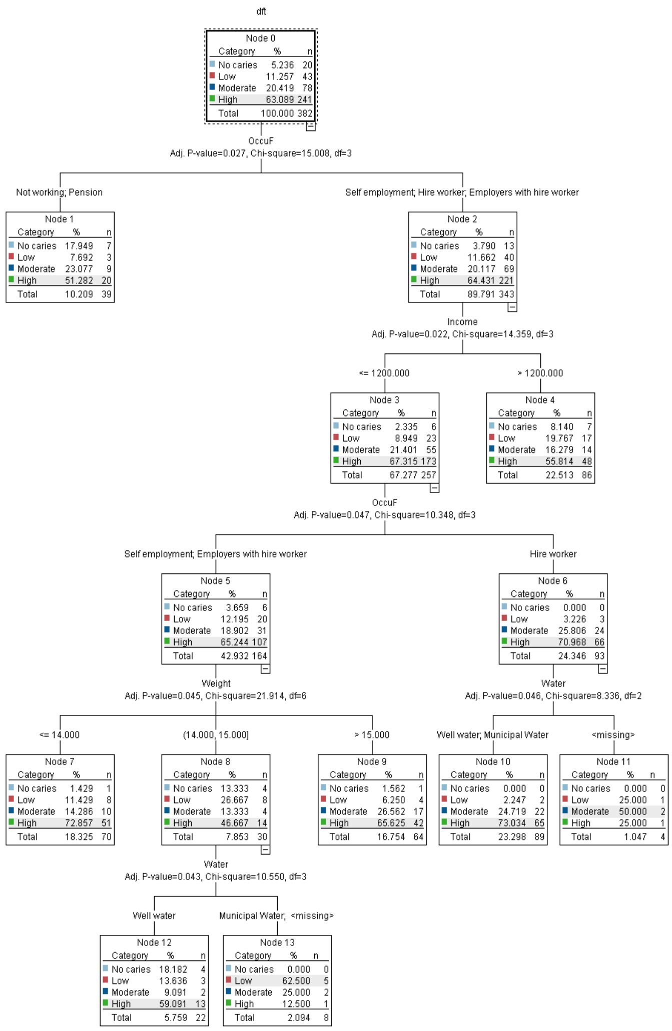

Figure 3.2. Decision tree for oral cancer classification

According to the predictor important of decision tree analysis in Figure. 4, the occurrence of high caries status is about 63.1%. SPSS modeler had suggested that first split on the decision tree analysis is determined by the father's occupation (which is the most important toward the stage of caries status). There are two groups had been suggesting by decision trees analysis. The first group is focusing on the father which are not working and pension. The first split had found that the case of caries status is lower compared to the second group. The second group is referring to that self-employment, hire worker and employers with hire worker.

Looking at the second group, the prevalence of high and moderate caries is mostly contributed by this group. It is reflection towards the father's education level and perhaps this lead to the time spent with their children. This contributes about 96.21% of the case from the second group of the first split. The second split is determined under the second group from the first split, which is referring to the household income. Most of the high and moderate caries status came from those who are having less than RM 1200 income. This focused group can be a target group for the stakeholder to educate about the oral health management among them. The third split shows that that self-employment and employers with hire worker having high and moderate caries compared to hire a worker. The fourth split is the split which is a branching under the third split. It shows that a pre-school which having the weight under 14kg is more mostly exposed to the caries status compared to pre-school between 14-15 kg and above 15 kg. The fifth split is the last split which emphasized on the type of water used. From the analysis, it is clearly stated that preschool children (having weight between 14-15kg) that used well water is much higher in caries status. It is about 59% from the fifth split.

MULTI-LAYER PERCEPTRON (MLP) NEURAL NETWORKS

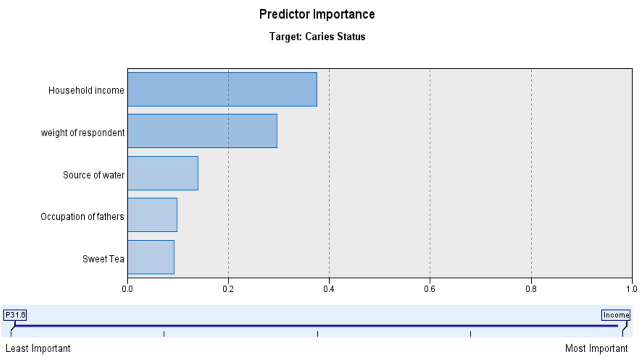

Figure 3.3. Predictor Important (Multi-Layer Perceptron (MLP) neural networks)

According to Figure 3.2. The top two predictors are ranking to their contribution are household income, and weight of preschool children (see the bold line produced in Figure 3.3). Using the CHAID method, these three are the best predictors of caries status. The MLP neural network architecture consist five input variable, one hidden node, and one output node. The MLP neural network architecture model is illustrated as follows.

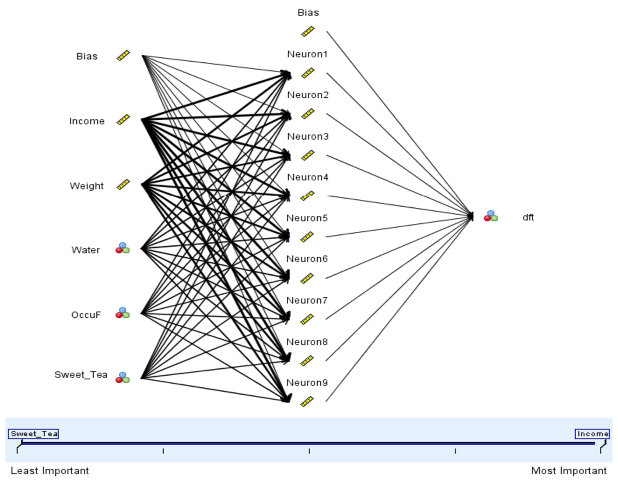

Figure 3.4. The Architecture of the best MLP model

The best number of hidden nodes for the MLP model is four nodes. Hence, the appropriate neural network architecture which results in the best multilayer perceptron ward neural network model for our case can be represented as follows:

$$\hat{y} = g_2\left(\sum_{j=1}^{9} w_j h_j + w_0\right) \tag{3}$$

where w_j is an output *weight* from hidden node *j* to the output node, w_0 is the bias for the output node and g_2 is the linear function. *h* is the values of the hidden layer nodes which can be represented as:

$$h_j = g_2\left(\sum_{j=1}^{9} v_{ji} x_i + v_{j0}\right) \qquad j=1,2,3,4,5,6,7,8,9 \tag{4}$$

where *v* is the input weight from input node *I* to hidden node *j*, *v* is the bias for hidden node *j* and *g* is an activation function. *X* is the independent variables where the independent variables are shown in Table 2.1. Equation (3) and (4) can also represent as follows:

$$\hat{Y} = w_0 + w_1 h_1 + w_2 h_2 \tag{5}$$

therefore;

$$h_i = \left[1 + exp\left[-\left(v_{j0} + v_{j1}x_1 + v_{j1}x_2 + v_{j1}x_3 + v_{j1}x_4 + v_{j1}x_5\right)\right]\right]^{-1} \quad = j=1,2 \tag{6}$$

Table 3.1. The result of Multilayer Perceptron Neural Network

Input Variables	Correct For Training *n* (%)	Correct For Testing *n* (%)
1. Water (x_1) 2. Weight (x_2) 3. Father Occupational (x_3) 4. Household Income (x_4) 5. Sweet Tea (x_5)	115(62.16%)	113(57.36%)

Overall the accuracy of this analysis is about 62.2%. The performance of MLP was evaluated using the partition of correct for testing/out-sample. During the training process, the output may fit the data very well, but it may produce poor results during the testing process. This situation implies that the neural network may not generalize well. Another way to see the neural network generalize well or not, it can be measured through the training and testing output. In this case, the performance of the MLP was evaluated as shown in Table 3.1. Training dataset (62.16%), the measures of accuracy are relatively similar to the testing dataset (57.36%).

IV. SUMMARY AND DISCUSSION

The main purpose of this paper is to prove such relationship and find the prognostics factor of caries status among the pre-school children. For this purpose, we analyzed the data through a decision tree analysis and MLP neural network models in order to gain the most potent factors that contributed to caries status among preschool children.

From the Decision Tree Analysis pint of view, the most factor that leads to caries status among preschool children are father's occupation, household income, children's weight and the type of water used in their house. Besides that, it shows that the most of the cases of high caries come from children that having less weight of 14 kg and they are coming from the family which self-employment or employers with hire worker. While using Multilayer Perceptron Neural Network, the most influential factor is household income, the weight of preschool children and sources of water. This information will provide a very useful information especially for the forecasting caries status among preschool children.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Universiti Sains Malaysia (USM) for providing the research funding RUI Grant no. 1001/PPSG/8012278, School of Dental Sciences, Kampus Kesihatan).

REFERENCES

1. N. J. Kassebaum, E. Bernabe, M. Dahiya, B. Bhandari, C. Murray, and W. Marcenes, "Global Burden of Untreated Caries: A systematic review and metaregression," *Journal Dent. Res*., vol. 94, 2015, pp. 650–658.
2. A. O. Arruda, K. R. Senthamarai, M. R. Inglehart, C. T. Rezende, and W. Sohn, "Effect Of 5% Fluoride Varnish Application on Caries among School Children in Rural Brazil: A Randomized Controlled Trial," *Community Dent Oral Epidemiol*, vol 40, 2012, pp. 267–276.
3. A. Blinkhorn, R. Byun, P. Mehta, and M. Kay, "A 4-year assessment of a new water fluoridation scheme in New South Wales, Australia," *International Dent Journal*, vol. 65(3), 2015, pp. 156-163.
4. J. M. Armfield, and A. Spencer, "Consumption of Nonpublic Water: Implications for Children's Caries Experience," *Community Dent Oral Epidemiol*, vol. 32, 2004, pp. 283-296.
5. F. Mojarad, and M. Maybodi, "Association Between Dental Caries and Body Mass Index Among Hamedan Elementary School Children," *Journal Dental*, vol. 5(2), 2011, pp. 170-107.
6. J. J. Warren, K. Weber-Gasparoni, T. A. Marshall, D. R. Drake, F. Dehkordi-Vakil, D. V. Dawson, and K. M. Tharp, "A Longitudinal Study of Dental Caries Risk among Very Young Low SES Children. Community *Dent. Oral Epidemiol*, vol. 37, 2009, pp. 116–122.
7. B. W. Chaffee, C. A. Feldens, P. H. Rodrigues, and M. R. Vítolo, "Feeding Practices in Infancy Associated with Caries Incidence in Early Childhood," *Community Dent. Oral Epidemiol*, vol. 43(4), 2015, pp. 338–48.
8. S. Gibson, and S. Williams, "Dental Caries in Pre-School Children: Associations with Social Class, Tooth Brushing Habit and Consumption of Sugars and Sugar-Containing Foods. Further analysis of data from the National Diet and Nutrition Survey of children aged 1.5-4.5 years," *Caries Res*, vol. 33, 1999, pp. 101-113.
9. H. Kato, K. Tanaka, K. Shimizu, C. Nagata, S. Furukawa, M. Arakawa, and Y. Miyake, "Parental Occupations, Educational Levels, and Income and Prevalence of Dental Caries in 3-Year-Old Japanese Children," *Environmental Health Preventive Med*., vol. 22(1), 2017, pp. 1-7.
10. D. H. Ha, K. F. Roberts-Thomson, P. Arrow, K. G. Peres, and L. G. Do, *Children's oral health status in Australia, 2012– 2014*, Adelaide University Press; Adelaide, Australia, 2016.
11. J. Mesaric, and D. Sebalj, "Decision Trees for Predicting The Academic Success of Students," Croatian *Operational Research Review*, vol. 7(2), 2016, pp. 367-388.
12. N. Mohamed, N. A. Aleng, W. M. A. Ahmad, and Maizah Hura, "Multilayer Feed-Forward Neural Network Approach to Lymphoma Cancer Data," *International Journal Contemporary Mathematical. Sciences*, vol. 7(35), 2012, pp. 1749 – 1756.
13. N. A. Aleng, N. Mohamed, N. and W. M. A. Ahmad, "Modelling of Tuberculosis Patients Using Multilayer Feed forward Neural Network," *Applied Mathematical Sciences*, vol. 6(50), 2012, pp. 2481 – 2488.

AUTHORS PROFILE

Wan Muhamad Amir W Ahmad is a lecturer at the school of Dental Sciences, Universiti Sains Malaysia (USM). He is a statistician by profession with BSc and MSc in Applied Statistics and PhD in Biostatistics from USM. His core academic teaching includes biostatistics, statistical software application, statistics computing, statistical consultant, epidemiology research designs in health sciences research, research methodology, elementary statistics, time series forecasting, operational research, advanced statistics and design of experiment.

Ruhaya Hasan is a lecturer at the school of Dental Sciences, Universiti Sains Malaysia (USM). She obtained her BSc (Nutrition and Community Health) (UPM), MSc (Nutrition) (USM), and PhD (Nutrition and Oral Health) (Health Promotion) (UM). Her research interest's include Nutrition and early childhood caries, Community nutrition and Health promotion

Mohd Fadhli Khamis has been appointed as the Deputy Dean (Academic, Student and Alumni), School of Dental Sciences, Universiti Sains Malaysia since January 2013. He graduated from the Universiti Malaya in 1994 with a BDS. After working with the Ministry of Health for six years, he pursued postgraduate studies and obtained PhD (Dentistry) from the University of Adelaide in 2006. Currently, Dr. Mohd Fadhli is also coordinating forensic dentistry teaching at the School of Dental Sciences and School of Health Sciences. His research interests inlude dental anotomy and anthropology, and forensic odontology. He also serves as the member of USM Animal Ethic Committee providing statistical advice.

Nor Azlida Aleng is currently pursuing her Ph.D. in Biostatistics. She is a lecturer at the School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu. She obtained her bachelor degree in Mathematics and a master degree in Management Mathematics. Her teaching experience for undergraduate students is in the field of biostatistics, neural network, mathematical modeling and Islamic financial in mathematics, and currently focusing on mathematical modeling in medical data using robust regression and multilayer feed forward neural network methods.

Noraini Mohamad is a medical lecturer at the School of Dental Sciences, Universiti Sains Malaysia. She obtained her bachelor's degree in medical Science from Universiti Putra Malaysia and master's Degree in family medicine from USM. Her research interest s include women's health and breastfeeding.