

Predictive Analysis of Oral Squamous Cell Carcinoma Patient Status Using Machine Learning: A Focus on Lifestyle and Demographic Factors

Wan Muhamad Amir W Ahmad¹, Mohamad Nasarudin Adnan¹, Farah Muna Mohamad Ghazali², Nor Azlida Aleng³, Nurfadhline Abdul Halim⁴, Mohamad Shafiq Mohd Ibrahim⁵

¹School of Dental Sciences, Health Campus, Universiti Sains Malaysia, Kota Bharu, Kelantan, Malaysia.

²Faculty of Computer and Mathematics Sciences, Universiti Teknologi MARA, Terengganu, Terengganu, Malaysia.

³Faculty of Computer Sciences and Mathematics, Universiti Malaysia Terengganu, Terengganu, Malaysia.

⁴Risk and Analytics Research Group, Faculty of Science and Technology, University Sains Islam Malaysia, Negeri Sembilan, Malaysia.

⁵Kulliyah of Dentistry, International Islamic University Malaysia, Kuantan Campus, Pahang, Malaysia.

Corresponding author: Wan Muhamad Amir W Ahmad

E-mail: wmamir@usm.my

Academic Editor: Fátima Regina Nunes de Sousa

Received: January 14, 2025 / **Review:** March 18, 2025 / **Accepted:** April 11, 2025

How to cite: Ahmad WMAW, Adnan MN, Ghazali FMM, Aleng NA, Halim NA, Ibrahim MSM. Predictive analysis of oral squamous cell carcinoma patient status using machine learning: A focus on lifestyle and demographic factors. *Pesqui Bras Odontopediatria Clín Integr*. 2026; 26:e250008. <https://doi.org/10.1590/pboci.2026.045>

ABSTRACT

Objective: To develop and validate a predictive model for oral squamous cell carcinoma (OSCC) survival outcomes, assessing the impact of key risk factors on survival probability. **Material and Methods:** A retrospective analysis was performed using a Multilayer Feedforward Neural Network (MLFFNN) in R. The outcome variable was survival status (live or dead), while predictors included smoking status, age, betel quid use, alcohol consumption, and sex. Data were normalized and partitioned into training, testing, and validation sets using bootstrap resampling. The MLFFNN model was optimized with hidden layers and a logistic activation function to enhance predictive performance. **Results:** Smoking status was the most influential predictor of OSCC survival (24%), followed by age (16.91%), betel quid use (15.37%), sex (10.21%), and alcohol consumption (8.57%). Performance metrics included a Mean Absolute Error of 0.2628, Root Mean Squared Error of 0.3466, and validation accuracy of 73.715%. **Conclusion:** This study highlights the potential of neural network-based models in predicting OSCC survival outcomes. Smoking emerged as the dominant risk factor, providing critical insights for data-driven clinical management strategies.

Keywords: Carcinoma, Squamous Cell; Predictive Learning Models; Machine Learning; Life Style.

■ Introduction

Oral squamous cell carcinoma (OSCC) represents a significant global health concern, particularly in regions with prevalent risk factors such as tobacco use, alcohol consumption, and betel quid chewing [1,2]. These lifestyle choices, along with demographic variables like age and sex, have been extensively studied for their association with OSCC incidence and patient outcomes [3,4]. For instance, smoking and alcohol use are well-established risk factors for OSCC, with studies indicating that combined usage significantly elevates cancer risk [5,6]. Similarly, betel quid chewing, common in Southeast Asia, has been linked to increased OSCC incidence [7,8].

Despite advancements in treatment, the prognosis for OSCC patients remains variable, often influenced by the aforementioned factors [9]. Traditional statistical methods have provided insights into these associations; however, they may not fully capture the complex, non-linear interactions between multiple variables that contribute to patient outcomes [10]. In this context, machine learning (ML) approaches, particularly Multilayer Feedforward Neural Networks (MLFFNN), offer a promising avenue for predictive analysis. MLFFNNs are capable of modeling intricate patterns within data, making them suitable for predicting patient status based on a combination of lifestyle and demographic factors [11-15].

This study aims to leverage MLFFNNs, enhanced with bootstrap resampling techniques, to predict the survival status (alive or deceased) of OSCC patients. Bootstrap methods are employed to improve the robustness and generalizability of the predictive model by mitigating overfitting and providing more reliable estimates [12]. By integrating key independent variables-smoking, age, betel quid usage, alcohol consumption, and sex. This research seeks to develop a predictive framework that can assist clinicians in identifying high-risk patients and tailoring personalized treatment strategies.

The application of machine learning in oncological prognostics represents a significant step toward precision medicine [16,17]. By accurately predicting patient outcomes based on individual risk profiles, healthcare providers can enhance decision-making processes, optimize resource allocation, and ultimately improve patient survival rates. This study contributes to the growing body of literature advocating for the integration of advanced computational techniques in medical research, particularly in the realm of cancer prognosis [13,14].

■ Material and Methods

Study Design and Ethical Clearance

This research adopts a retrospective study design, utilizing data collected from the Hospital Pakar Universiti Sains Malaysia (HPUSM) dental clinic. Ethical approval was granted by the Universiti Sains Malaysia Research Ethics and Human Research Committee (USM/JEPeM/16050184), ensuring compliance with ethical standards, including patient privacy and confidentiality.

Data Collection

The dataset used includes 54 participants with associated lifestyle and demographic attributes. The dependent variable, "Status," represents the survival outcome of oral squamous cell carcinoma (OSCC) patients, coded as 1 for alive and 0 for dead. The independent variables include age, a continuous variable representing the patient's age in years, and four binary variables: smoking status (Yes, No), betel quid usage (Yes, No), alcohol consumption (Yes, No), and sex (Male, Female). These variables capture key lifestyle factors and demographic characteristics influencing OSCC outcomes, providing a comprehensive dataset for predictive modeling. The

integration of these variables allows for an in-depth analysis of their impact on the survival status of OSCC patients, forming the basis for the machine-learning approach employed in this study.

Computational Biometry Modeling

The computational modeling in this study utilized a Multilayer Feedforward Neural Network (MLFFNN) framework implemented using R Studio and its associated machine-learning libraries, such as *caret*, *nnet*, and *neuralnet*. This approach integrated bootstrap resampling techniques (e.g., via the *boot* or *rsample* packages) to enhance model stability and mitigate overfitting. The dataset was systematically divided into training (70%) and testing (30%) subsets using functions such as `createDataPartition` from the *caret* package to develop and validate the predictive model.

The MLFFNN architecture was constructed using R packages such as *nnet* or *neuralnet*, which allowed for defining an input layer corresponding to the five independent variables, one or more hidden layers optimized for non-linear interactions, and an output layer with a single neuron for binary classification. The output layer utilized a sigmoid activation function to produce probabilities for classification, which is standard in binary classification tasks. The model was trained using backpropagation algorithms with an adaptive learning rate, implemented through the *neuralnet* package, and minimized the binary cross-entropy loss function.

Multilayer Feed-Forward Neural Network (MLFNN) with Two Hidden Layers Approach

Artificial Neural Networks (ANNs) are computational systems modeled after the structure and function of biological neural networks, often referred to as neural networks (NNs). This study employs a Multilayer Feed-Forward (MLFF) architecture, a specific type of neural network that features one or more layers between the input, hidden, and output layers. As the focus of this study is on a single dependent variable, the MLFF model is designed with a single output node. Figure 1 demonstrates the MLFF model, comprising N input nodes, H hidden nodes, and a single output node. The values of the hidden node h_j , $j = 1 \dots 3$ are given by

$$h_j = g_1 \left(\sum_{i=1}^3 v_{ji} x_i + E_1 \right) \text{ where } v_{ji} \text{ the output weight, } E_1 \text{ is the bias. The values of the hidden node } n_j, j = 1 \dots 3$$

are given by $n_j = g_2 \left(\sum_{i=1}^3 v_{ji} h_i + E_2 \right)$ where v_{ji} the output weight, E_2 is the bias. The values of the hidden node

$$Y_j, j = 1, 2 \text{ are given by where } Y_i = g_3 \left(\sum_{j=1}^3 v_{ji} n_i + E_3 \right) \text{ } v_{ji} \text{ the output weight, } E_3 \text{ is the bias.}$$

Bootstrap

The methodology for developing the predictive model for patient status begins with data collection, variable selection, and screening to ensure data quality and relevance. A bootstrap sampling technique ($n = 1000$) is employed to enhance model accuracy and generalizability. The data is then split into 70% for training and 30% for testing to facilitate model development and validation. The modeling process utilizes a Multi-Layer Feedforward Neural Network (MLFFNN) with a logistic activation function, which effectively captures complex relationships among variables and handles binary outcomes to classify patient survival status accurately. The model's performance is evaluated using metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Median Absolute Error (MedAE), providing a comprehensive assessment of the model's predictive capability and robustness in determining patient survival outcomes (Figure 2).

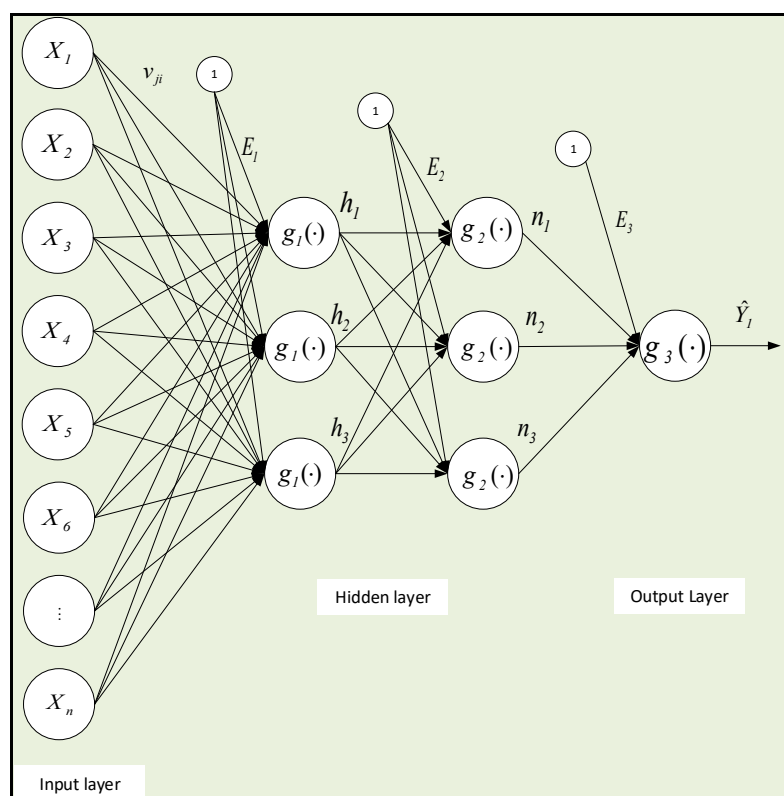


Figure 1. The general architecture of the MLFFNN with two hidden layers, N input nodes, and one output node.

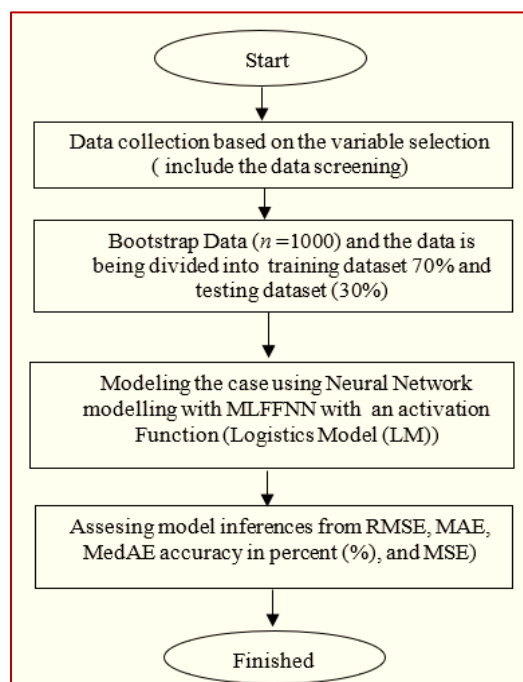


Figure 2. Flowchart of the proposed hybrid methodology for predictive modeling using MLFFNN with bootstrap resampling.

Results

The model demonstrated robust predictive capabilities, achieving a Mean Absolute Error (MAE) of 0.2628 and a Root Mean Squared Error (RMSE) of 0.35 on the validation dataset, underscoring its reliability.

Furthermore, the model attained an accuracy of 73.72% on the validation dataset. The Mean Squared Error (MSE) of 0.17 observed on the testing dataset further highlights the model's effectiveness and precision in predictive analysis. Figure 3 illustrates the model architecture visually.

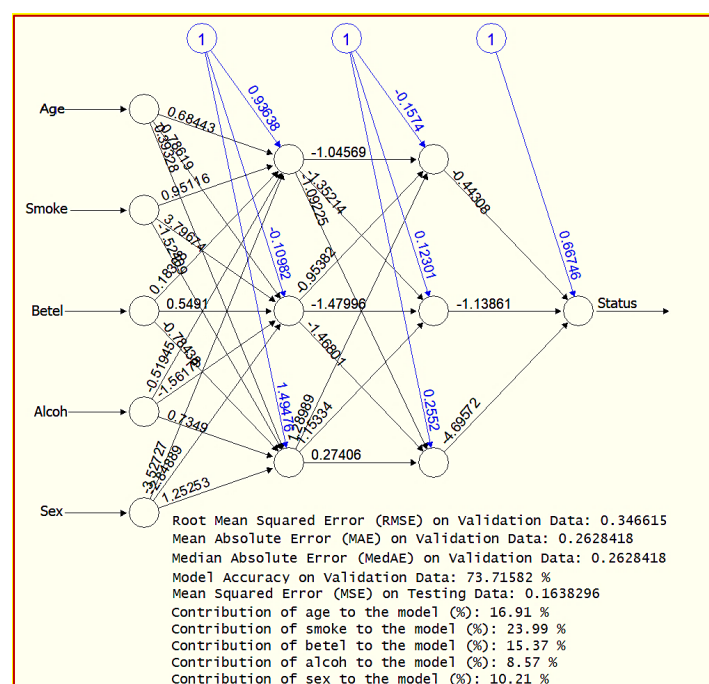


Figure 3. The architecture of the optimized MLFFNN model with five input variables, one hidden layer, and a single output node for predicting patient survival status.

Contribution of Age

The model reveals that age accounts for 16.91% of the predictive capability for patient survival status, emphasizing its significant role in the prognosis of oral squamous cell carcinoma (OSCC). Age is a well-established factor in cancer outcomes, as older patients often experience compromised immune function, a higher prevalence of comorbidities, and slower recovery rates. These age-associated physiological and health challenges contribute to poorer survival outcomes in elderly patients. The substantial contribution of age in the model underscores its predictive value, advocating for age-specific strategies, such as enhanced screening for older individuals and personalized treatment plans, to improve OSCC management. This observation aligns with the broader medical understanding that older patients necessitate more intensive monitoring and interventions to mitigate their distinct health risks.

Contribution of Smoking

Smoking emerged as the most significant factor influencing the model's predictions, accounting for 23.99% of its contribution. This finding aligns with extensive research over decades, which has consistently identified smoking as a primary risk factor for OSCC and a determinant of poorer survival outcomes. Tobacco is known to contain numerous carcinogens that promote cellular mutations, chronic inflammation, and impaired tissue repair, thereby accelerating disease progression. The substantial contribution of smoking highlights the urgent need to integrate smoking cessation initiatives into OSCC prevention and treatment strategies. Additionally, the model's ability to quantify the impact of smoking underscores its pivotal role in shaping

survival outcomes, reinforcing the imperative for public health interventions to reduce smoking prevalence among high-risk populations.

Contribution of Betel Quid Chewing

Betel quid chewing contributes 15.37% to the predictive model, reflecting its significant role in OSCC prognosis, particularly in regions where this practice is culturally ingrained. Betel quid contains areca nut, lime, and often tobacco, all of which have been shown to have carcinogenic effects on the oral mucosa. Prolonged exposure to these substances can lead to chronic irritation and precancerous conditions such as leukoplakia, ultimately increasing the risk of OSCC. The model's inclusion of this variable highlights the need for targeted interventions in regions where betel quid chewing is prevalent. Public health campaigns focusing on education and cessation programs can play a vital role in reducing OSCC incidence and improving survival rates. This result also underscores the importance of culturally sensitive approaches in addressing lifestyle-related cancer risks.

Contribution of Alcohol Consumption

Alcohol consumption accounts for 8.57% of the model's predictive capacity, underscoring its role as a moderate yet significant risk factor in OSCC prognosis. Its synergistic interaction with other risk factors, such as smoking and betel quid chewing, exacerbates cellular damage and carcinogenesis. Alcohol also compromises immune defenses and impairs tissue repair mechanisms, thereby increasing the likelihood of cancer progression. Although its contribution is less pronounced compared to smoking or betel quid chewing, alcohol remains a critical factor that warrants attention. Including alcohol consumption in the predictive model provides a comprehensive understanding of OSCC risk factors and emphasizes the necessity of addressing alcohol use alongside other behavioral interventions to enhance patient outcomes.

Contribution of Sex

The model identifies sex as contributing 10.21% to OSCC survival outcomes, underscoring its significant role. Males are generally at greater risk for OSCC, attributed to the higher prevalence of risk behaviors such as smoking, alcohol consumption, and betel quid chewing. Additionally, biological and hormonal differences may influence disease progression and treatment response between sexes. Incorporating sex as a variable enhances the model's ability to account for these complexities, enabling more precise predictions across demographic groups. These findings highlight the importance of personalized treatment strategies that consider sex-based differences, fostering equitable healthcare for both male and female patients.

■ Discussion

To further substantiate the impact of these factors on patient survival outcomes, it would be beneficial to incorporate specific clinical case examples that demonstrate how lifestyle factors such as smoking or alcohol consumption directly influence the progression of OSCC in real-world scenarios. Previous studies have consistently shown that tobacco use accelerates tumour progression and adversely affects survival rates, even in the presence of medical interventions [5,6]. Similarly, age has been identified as a significant prognostic factor, with longitudinal cohort studies reporting diminished survival outcomes in older OSCC patients compared to younger groups [3,4]. These findings highlight how behavioural and demographic risk factors contribute to survival disparities, supporting the relevance of the predictive model developed in this study.

In this context, the proposed hybrid methodology has demonstrated notable effectiveness in estimating event probabilities, particularly when predicting comorbidity risks such as hypertension in patients with dyslipidemia and type 2 diabetes mellitus [12]. By combining multiple logistic regression (MLR) with a Multi-Layer Feed-Forward Neural Network (MLFFNN), the approach has yielded a robust and precise predictive model. Traditional non-linear regression models often face significant hurdles, such as computational complexity and diminished accuracy in capturing predictor-outcome relationships [10,11]. In contrast, the hybrid methodology capitalizes on the complementary strengths of both bootstrap sampling and logistic regression, ensuring enhanced validation and reliability across both training and testing datasets [12]. Notably, the MLFFNN model exhibited impressive predictive power for OSCC survival outcomes, with smoking status emerging as the most influential factor, contributing 23.99% to the model's performance, followed by age (16.91%), betel quid use (15.37%), sex (10.21%), and alcohol consumption (8.57%). This hierarchy underscores the multifactorial nature of OSCC outcomes, highlighting the pivotal roles that behavioural and demographic factors play in determining survival [2,7,8]. The model demonstrated robust predictive accuracy, evidenced by a Mean Absolute Error (MAE) of 0.2628 and a Root Mean Squared Error (RMSE) of 0.3466 on the validation data. Its overall accuracy of 73.72%, coupled with a Mean Squared Error (MSE) of 0.1638 on the testing data, further affirms the reliability of the model. The architecture of the MLFFNN, as depicted in Figure 2, effectively captured complex non-linear relationships between the input variables, hidden layers, and the output node. This hybrid model serves as a valuable tool for healthcare providers, facilitating early detection and personalized intervention strategies for patients at high risk of OSCC-related mortality. By identifying critical predictors, the model not only informs targeted prevention efforts but also underscores the necessity of incorporating behavioural and demographic considerations into clinical decision-making [13,14].

From a methodological perspective, the integration of bootstrapping with logistic regression represents an advancement in medical statistical modelling. The process of variable selection, guided by clinical expertise, ensured that the predictors included were both relevant and interpretable within the clinical context. The application of bootstrapping played a crucial role in enhancing the robustness of the dataset, generating a comprehensive 'mega' file through repeated resampling. This technique mitigated variability and improved the generalizability of the model [12]. Subsequently, the dataset was partitioned into training (70%) and testing (30%) subsets to facilitate rigorous model development and evaluation. The statistical computations, performed using R programming, enabled the seamless integration of analytical techniques, yielding an effective logistic regression model. Despite its successes, the methodology faced challenges, including the appropriate selection of input variables, data preparation, and the standardization required for logistic modeling. Addressing these challenges was vital to ensuring reliable and clinically meaningful outcomes. The findings underscore the potential of hybrid predictive models to advance personalized medicine, foster equitable healthcare delivery, and improve clinical outcomes for OSCC patients [16,17].

The hybrid machine learning methodology demonstrates promising results, but several limitations merit further exploration. First, the relatively small sample size of 54 participants could limit the generalizability of the findings, as it may not fully reflect the broader OSCC patient population. Additionally, although the study incorporated key lifestyle and demographic factors, the exclusion of potentially significant variables, such as genetic predispositions, comorbid conditions, and treatment history, limits the model's ability to account for all relevant influences on survival outcomes [9]. Furthermore, despite the robust performance of the MLFFNN model, overfitting remains a concern, particularly due to the inherent complexities of neural network models [11]. Future research should focus on expanding the dataset to include a more diverse and larger sample, as well

as integrating additional variables, to further refine the model's predictive accuracy [15]. Lastly, while bootstrap resampling mitigates overfitting to some extent, the reliance on this technique underscores the need for further validation with independent external datasets to confirm the model's generalizability [13].

Despite its promising results, the study is not without limitations. The relatively small sample size limits the generalizability of the model, and the exclusion of potentially influential variables-such as genetic factors and comorbid conditions- suggests that future work should integrate these aspects for a more comprehensive prediction model [1]. Further validation with larger, more diverse datasets is necessary to confirm the model's applicability in broader clinical settings. Future studies should explore the integration of additional biomarkers and treatment modalities to refine the predictive framework, enhancing its clinical relevance [11,17].

Overall, this study highlights the potential of hybrid predictive modelling to advance personalized medicine, inform targeted prevention strategies, and enhance patient management in OSCC. By combining statistical and machine learning approaches, the methodology contributes to the growing body of evidence supporting AI-driven innovations in oral cancer research [15-17].

■ Conclusion

This study demonstrates the power of a hybrid machine-learning methodology, combining multilayer feedforward neural networks with bootstrap resampling, to predict survival outcomes in oral squamous cell carcinoma patients. The model effectively integrates key lifestyle and demographic factors such as smoking, age, betel quid usage, alcohol consumption, and highlights their multifactorial impact on oral squamous cell carcinoma prognosis. Smoking was identified as the most significant predictor, contributing 23.99% to the model's predictive power, followed by age, betel quid usage, sex, and alcohol consumption. These findings underscore the importance of considering both behavioural and demographic factors in clinical decision-making for oral squamous cell carcinoma management.

With an accuracy of 73.72% on the validation dataset and robust performance metrics, including a mean absolute error of 0.2628 and a root mean squared error of 0.3466, the multilayer feedforward neural networks framework proved adept at capturing complex, non-linear relationships in oral squamous cell carcinoma data. Bootstrap resampling further enhanced the model's generalizability and stability, mitigating overfitting concerns typical in machine-learning applications. Clinically, this model serves as a valuable tool for early detection and personalized treatment strategies, emphasizing the need for a data-driven approach to optimize patient care. Additionally, the findings support the development of public health initiatives targeting high-risk behaviours, such as smoking cessation programs and culturally sensitive interventions.

This research not only advances the field of oncological prognostics through the application of hybrid machine learning but also underscores the critical need for interdisciplinary collaboration in addressing the complexities of oral squamous cell carcinoma. By extending this model to encompass a wider range of predictive variables and larger patient populations, we can better equip clinicians with the tools needed to improve patient outcomes and personalize care in the fight against oral squamous cell carcinoma.

■ Authors' Contributions

WMAA  <https://orcid.org/0000-0003-2366-3918>

Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review and Editing, Visualization, Supervision, Project Administration and Funding Acquisition.

MNA		https://orcid.org/0009-0008-4759-4869	Conceptualization, Methodology, Software, Validation, Resources, Data Curation, Writing - Original Draft and Writing - Review and Editing.
FMMG		https://orcid.org/0000-0001-6772-6779	Conceptualization, Methodology, Software, Validation, Data Curation, Writing - Original Draft, Writing - Review and Editing and Visualization.
NAA		https://orcid.org/0000-0003-1111-3388	Conceptualization, Formal Analysis, Resources, Data Curation, Writing - Original Draft and Writing - Review and Editing.
NAH		https://orcid.org/0000-0003-4417-3096	Conceptualization, Methodology, Validation, Formal Analysis, Writing - Original Draft and Writing - Review and Editing.
MSMI		https://orcid.org/0000-0003-1356-2902	Conceptualization, Methodology, Data Curation, Writing - Original Draft, Writing - Review and Editing and Visualization.
All authors declare that they contributed to critical review of intellectual content and approval of the final version to be published.			

■ Financial Support

The authors express their gratitude to Universiti Sains Malaysia (USM) for their support in funding this study through the Ministry of Higher Education (MOHE) Fundamental Research Grant Scheme (FRGS/1/2022/STG06/USM/02/10).

■ Conflict of Interest

The authors declare no conflicts of interest.

■ Data Availability

The data used to support the findings of this study can be made available upon request to the corresponding author.

■ References

- [1] Borse V, Konwar AN, Buragohain P. Oral cancer diagnosis and perspectives in India. *Sens Int* 2020, 1:100046. <https://doi.org/10.1016/j.sintl.2020.100046>
- [2] Warnakulasuriya S, Chen THH. Areca nut and oral cancer: Evidence from studies conducted in humans. *J Dent Res* 2022; 101(10):1139-1146. <https://doi.org/10.1177/00220345221092751>
- [3] Shenoi R, Devrukhkar V, Chaudhuri, Sharma BK, Sapre SB, Chikhale A. Demographic and clinical profile of oral squamous cell carcinoma patients: A retrospective study. *Indian J Cancer* 2012; 49(1):21-26. <https://doi.org/10.4103/0019-509X.98910>
- [4] Xu Q, Wang C, Li B, Kim K, Li J, Mao M, et al. The impact of age on oral squamous cell carcinoma: A longitudinal cohort study of 2,782 patients. *Oral Dis* 2019; 25(3):730-741. <https://doi.org/10.1111/odi.13015>
- [5] Eloranta R, Vilén ST, Keinänen A, Salo T, Qannam A, Bello IO, et al. Oral squamous cell carcinoma: Effect of tobacco and alcohol on cancer location. *Tob Induc Dis* 2024; 22. <https://doi.org/10.18332/tid/189303>
- [6] Mello FW, Melo G, Pasetto JJ, Silva CAB, Warnakulasuriya S, Rivero ERC. The synergistic effect of tobacco and alcohol consumption on oral squamous cell carcinoma: A systematic review and meta-analysis. *Clin Oral Invest* 2019; 23(7):2849-2859. <https://doi.org/10.1007/s00784-019-02958-1>
- [7] Cirillo N, Duong PH, Er WT, Do CTN, De Silva MEH, Dong Y, et al. Are there betel quid mixtures less harmful than others? A scoping review of the association between different betel quid ingredients and the risk of oral submucous fibrosis. *Biomolecules* 2022; 12(5):664. <https://doi.org/10.3390/biom12050664>
- [8] Jasim A, Li X, Octavia A, Gunardi I, Crocombe L, Sari EF. The association between betel quid use and oral potentially malignant and malignant disorders in Southeast Asian and Pacific regions: A systematic review and meta-analysis with GRADE evidence profile. *Front Oral Health* 2024; 5:1397179. <https://doi.org/10.3389/froh.2024.1397179>
- [9] Kim MJ, Ahn KM. Prognostic factors of oral squamous cell carcinoma: The importance of recurrence and pTNM stage. *Maxillofac Plast Reconstr Surg* 2024; 46(1):8. <https://doi.org/10.1186/s40902-024-00410-3>
- [10] Cheung HMC, Rubin D. Challenges and opportunities for artificial intelligence in oncological imaging. *Clin Radiol* 2021; 76(10):728-736. <https://doi.org/10.1016/j.crad.2021.03.009>
- [11] Dixit S, Kumar A, Srinivasan K. A current review of machine learning and deep learning models in oral cancer diagnosis: recent technologies, open challenges, and future research directions. *Diagnostics* 2023; 13(7):1353. <https://doi.org/10.3390/diagnostics13071353>
- [12] Reeves M, Bhat HS, Goldman-Mellor S. Resampling to address inequities in predictive modeling of suicide deaths. *BMJ Health Care Inform* 2022; 29(1):e100456. <https://doi.org/10.1136/bmjhci-2021-100456>
- [13] Zhang B, Shi H, Wang H. Machine learning and AI in cancer prognosis, prediction, and treatment selection: A critical approach. *J Multidiscip Healthc* 2023; 16:1779-1791. <https://doi.org/10.2147/JMDH.S410301>
- [14] Alabi RO, Youssef O, Pirinen M, Elmusrati M, Mäkitie AA, Leivo I, et al. Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. *Artif Intell Med* 2021; 115:102060. <https://doi.org/10.1016/j.artmed.2021.102060>

- [15] de Chauveron J, Unger M, Lescaille G, Wendling L, Kurtz C, Rochefort J. Artificial intelligence for oral squamous cell carcinoma detection based on oral photographs: A comprehensive literature review. *Cancer Med* 2024; 13(1):e6822. <https://doi.org/10.1002/cam4.6822>
- [16] Beristain-Colorado MDP, Castro-Gutiérrez MEM, Torres-Rosas R, Vargas-Treviño M, Moreno-Rodríguez A, Fuentes-Mascorro G, et al. Application of neural networks for the detection of oral cancer: A systematic review. *Dent Med Probl* 2024; 61(1):121-128. <https://doi.org/10.17219/dmp/159871>
- [17] Viet CT, Zhang M, Dharmaraj N, Li GY, Pearson AT, Manon VA, et al. Artificial intelligence applications in oral cancer and oral dysplasia. *Tissue Eng Part A* 2024; 30(19-20):640-651. <https://doi.org/10.1089/ten.tea.2024.0096>