Collaborative machine learning-guided overall survival prediction of oral squamous cell carcinoma

Author(s): Alabi, Rasheed Omobolaji; Elmusrati, Mohammed; Leivo, Ilmo; Almangush, Alhadi; Makitie, Antti A.

Title: Collaborative machine learning-guided overall survival prediction of oral squamous cell carcinoma

Year: 2024

Version: Accepted Manuscript

Please cite the original version:

Alabi, R.O., Elmusrati, M., Leivo, I., Almangush, A. & Makitie, A.A., (2024). Collaborative machine learning-guided overall survival prediction of oral squamous cell carcinoma. *Acta oto-laryngologica*. https://doi.org/10.1080/00016489.2024.2437012

Collaborative Machine Learning-Guided Overall Survival Prediction of Oral Squamous Cell Carcinoma

Rasheed Omobolaji Alabi PhD [a,b], Mohammed Elmusrati D.Sc [b], Ilmo Leivo MD, PhD[c], Alhadi Almangush DDS, PhD [a,c,d*], Antti A. Mäkitie MD, PhD [a,e,f*]

[a] Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland.

[b] Department of Industrial Digitalization, School of Technology and Innovations, University of Vaasa, Vaasa, Finland.

[c] University of Turku, Institute of Biomedicine, Pathology, Turku, Finland.

[d] Department of Pathology, University of Helsinki, Helsinki, Finland.

[e] Department of Otorhinolaryngology – Head and Neck Surgery, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.

[f] Division of Ear, Nose and Throat Diseases, Department of Clinical Sciences, Intervention and Technology, Karolinska Institute and Karolinska University Hospital, Stockholm, Sweden.

*The last two authors have equal contribution.

Word count: 3999

Corresponding Author: Rasheed Omobolaji Alabi
Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki, Helsinki, Finland. **E-mail address:** rasheed.alabi@helsinki.fi

Abstract

Background: there is a lack of prognosticators of overall survival (OS) for Oral squamous cell carcinoma (OSCC). **Objectives**: We examined collaborative machine learning (cML) in estimating the OS of OSCC patients. The prognostic significance of the clinicopathological parameters was examined. **Methodology**: Altogether, 9439 OSCC patients were extracted from Surveillance, Epidemiology, and End Results database (Us). Five ML models – voting ensemble, stacked ensemble, extreme gradient boosting, light boosting, and logistic regression were used to predict OS. Three of these ML algorithms were combined to form a cluster of cML models. The performance of the cML was compared with the best performing individual ML algorithm following model training. **Results**: the performance accuracy of the voting ensemble, stacked ensemble, extreme gradient boosting, light boosting, and logistic regression models were 70.2%, 69.9%, 69.1%, 69.4%, and 69.5%respectively, following model training. When the voting ensemble model was compared with cML using temporal validation, the cML showed a comparable performance accuracy. The most significant prognostic factors were age of the patient at diagnosis, t stage, tumor grade, marital status, gender, primary site, surgery, N stage, radiotherapy, ethnicity, chemotherapy, and M stage. **Conclusions:** cML appears to give reliability to the final prediction and thereby may mark a paradigm shift from model individualism to a more cooperative paradigm. This approach may aid in determining enhanced individualized treatment for OSCC patients.

KEYWORDS: Artificial Intelligence; Cooperative Machine Learning; Oral Cancer; Overall survival; Ensemble Algorithms; SEER

Introduction

Oral cancer has attracted global attention due to its high incidence, diagnosis at a late stage, significant morbidity, and low survival rate (below 50%) [1]. The incidence and prevalence of oral cancer vary significantly due to geographic variability in the associated risk factors globally. In 2020, the global annual incidence was about 380,000 new cases with total death ratio estimated to be around 178, 000 [2,3]. Thus, it represents one of the most common malignancies worldwide [1]. In general, the most common associated risk factors include excessive alcohol consumption, chewing of areca nut products, and the use of tobacco products.

Remarkably, most oral cancer patients are presented with locally advanced disease, requiring multimodality therapy [4]. While this approach has a curative intent, a significant subset of these patients may develop locoregional recurrences and/or distant metastases, which reduce their chance of overall survival. Therefore, a concise effort is needed to predict oral tumor behavior, but the lack of specific prognostic indicators still constitutes a major challenge [5]. In addition, the decision-making regarding the best treatment approach still is challenging for many cases of oral cancer [4]. As a result, the prediction of the chance of overall survival for oral cancer patients is of utmost interest to both clinicians and patients [6].

Several studies have emphasized the potential of machine learning in estimating important outcomes such as locoregional recurrence, lymph node metastasis, and the chance of overall survival in different subsites of oral cancer [4,7–9]. However, most of these studies have emphasized the significance of a single machine learning algorithm. Interestingly, machine learning algorithms generally possess certain features that differentiate them when compared [10]. For example, some of these algorithms perform well with small datasets, utilize little memory, and show excellent accuracy. Ensemble algorithms have been reported to have two main advantages – performance and robustness [4,9,10]. Thus, several variants of ensemble methods such as decision tree, decision forest, extreme random forest, and voting ensemble exist.

Interestingly, each of these variants has its unique features that can aid in outcome prognostication. Rather than considering ML models from the methodological individualism perspective, a common ground of cooperation among these predictive models should be encouraged to properly harness their unique features. The common ground is expected to be non-competitive with zero gains at the expense of the other participating algorithms. To this end, our study aims to explore the collaborative machine learning (cML) paradigm for the prediction of overall survival in oral cancer. The methodology of the expected output of the participating ensemble algorithms was determined through a democratic voting approach to avoid conflicts. Furthermore, we aim to combine the predictions of each of the participating algorithms to demonstrate their cooperativeness and collaborative nature in estimating the overall survival in oral cancer patients. To the best of our knowledge, this is the first study that leverages the collaborative machine learning models paradigm for the chance of overall survival prediction in oral cancer.

2. Material and Methods

2.1. Data source

The Surveillance, Epidemiology, and End Results (SEER) program database was used in this study. Access to this database was approved by identification number 17247-Nov2020 for research plus data for November 2020 Submission.

2.2. Data processing and selection of attributes

The SEER database was queried for (I) oral cancer (lip and oral cavity) from the following main subsites – C00.0-C00.9, C02.0-C02.3, C02.8- C02.9, C03.0-C03.9, C04.0-C04.9, C05.0, C06.0-C06.9), as shown in Table 1, and (II) Patients with histologically diagnosed SCC (Third Edition (ICD-O-3) site code, Hist/behave: 8050–8089). The database was queried for Research

Plus Dataset for 17 Registries, Nov 2023 submission from 2010 – 2021. The included clinicopathological parameters available were age at diagnosis, ethnicity, gender, marital status, tumor grade, and stage classification according to the American Joint Committee on Cancer (AJCC) tumor-nodal-metastasis (TNM) 7th edition (Table 1), and treatment parameters. Overall survival was the primary endpoint and target variable. This extracted clinicopathological parameters of oral cancer patients were further differentiated into categorical and quantitative variables (Table 1). The exclusion criterion includes all other tumor sites than oral cancer. All cases of unknown and incomplete data were excluded. The results of this query produced a total of 9439 pathologically confirmed oral cancer patients. The categorical parameters were normalized for the machine learning training phase (**sub-section 2.3**).

2.3. Collaborative machine learning model development

Ensemble and tree-based machine learning models have shown promising results in cancer prognostication tasks [11,12]. Therefore, we have used ensemble methods for the models' development. These models have been chosen since they have been reported to perform better with a relatively small number of datasets, thereby producing better and more accurate predictions. Voting ensemble, stacked ensemble, extreme gradient boosting (XGBoost), light gradient boosting (light GBM), and logistic regression were used as participating models for the chance of overall survival estimation of oral cancer patients. In forming the cML, we used voting ensemble, stacked, and Light GBM. These variants of the ensemble method have been chosen since they follow the three main paradigms – bagging, stacking, and boosting paradigms, respectively. Microsoft Azure automated learning studio was used for model development. We have used a traditional machine learning algorithm, such as logistic regression as a reference to evaluate the prediction of OS in OSCC.

The extracted data (**sub-section 2.2**) were imported into the Azure machine learning studio (Figure 1). Following the loading of data, the necessary training parameters such as the algorithms, learning rate, k-fold cross-validation, and performance metrics were defined. For the models' training, we used a 5-fold cross-validation. We adjusted various training hyperparameters tunning to maximize the performance of the model (Figure 1). The hyperparameters that give the best accuracy for each model were selected. The models were primarily evaluated based on accuracy. Other secondary metrics used were Sensitivity, Specificity, Precision, Negative Predictive Value, False Positive Rate, False Discovery Rate, False Negative Rate, F1 Score, and Mathew's Correlation Coefficient (MCC) (Table 2).

The developed models were integrated as a web-based tool to demonstrate the collaborative approach of the participating models through voting. These integrated models were further programmed to demonstrate cooperation among them through democratic voting of the most widely supported prediction among the participating models (Figure 2). For example, if at least *two (2) out of the three (3)* participating models predict the oral cancer patient as having overall survival, then the final prediction of these models would be "*overall survival"* through democratic voting. Similarly, if at least *two (2) out of the three (3)* participating models predict the oral cancer patient as having low overall survival, then the final prediction of all the participating models for the outcome would be "*dead*". We have deliberately selected an odd number of participating algorithms to avoid a tight (equilibrium) in terms of the output from participating algorithms.

2.4. Independent validation for cooperative machine learning models

We used a temporal validation (N = 1201) approach due to the unavailability of geographic independent validation. The result of the temporal validation is to determine the best performing individual ML model. To validate the performance of the best performing individual model with the cML, we randomly selected 20 cases from the temporal validation cohort. The result of this was considered for reporting the performance of the best individual model and cML for OS prediction.

Table 1. Baseline demographic and tumor characteristics of oral cancer patients in the SEER database.

Variables	Total ($N=9439$) (%)	Categorization for machine learning analysis	Data type after categorization
Age at incident			
<40 years old (Young)	326 (3.5%)	No categorization	Discrete
>=40 years old (Old)	9113 (96.5%)		
Gender			
Female	3718 (39.4%)	0 = Female	Numeric
Male	5721 (60.5%)	1 = Male	
Ethnicity			
White	8292 (87.8%)	0 = White	
Black	379 (4.0%)	1 = Black	
Others (American Indian/AK Native, Asian/Pacific Islander)	768 (8.1%)	2 = Other	
Marital status			
Married	5514 (58.4%)	1 = Married	Numeric
Unmarried	3925 (41.6%)	2 = Unmarried	
Primary site*			
Lip	1393 (14.8%)	0 = Lip	Numeric
Tongue	4764 (50.5%)	1 = Tongue	
Gum	857 (9.1%)	2 = Gum	
Mouth	1157 (12.3%)	3 = Mouth	
Hard palate	166 (1.8%)	4 = Hard palate	
Others	1102 (11.7%)	5 = Others	
Tumor grade			
Well differentiated, Grade I	3044 (32.3%)	1 = Grade I	Numeric
Moderately differentiated, Grade II	5082 (53.8%)	2 = Grade II	
Poorly differentiated, Grade III	1285 (13.6%)	3 = Grade III	
Undifferentiated (Anaplastic), Grade IV	28 (0.3%)	4 = Grade IV	
T stage			
T1	5816 (61.6%)	T1 = 1	
T2	2825 (29.9%)	T2 = 2	Numeric
T3	798 (8.5%)	T3 = 3	
T4	0 (0%)	T4 = 4	
N stage			
N0	8204 (86.9%)	N0 = 0	Numeric
N1	1179 (12.5%)	N1 = 1	
N2	0 (0%)	N2 = 2	
N3	56 (0.6%)	N3 = 3	
M stage			
M0	9385 (99.4%)	M0 = Absent	Numeric
M1	54 (0.6%)	M1 = Present	
Surgery			
No	698 (7.4%)	0 = No surgery	
Yes	8741 (92.6%)	1 = Surgery	Numeric
Radiotherapy			
No	6913 (73.2%)	0 = No radiotherapy	Numeric
Yes	2526 (26.8%)	1 = Radiotherapy	
Chemotherapy			
No	8548 (90.6%)	0 = No chemotherapy	Numeric
Yes	891 (9.4%)	1 = Chemotherapy	
Overall survival			
Alive	4619 (48.9%)	0 = Alive at the end of follow-up	Numeric
Dead	4820 (51.1%)	1 = Dead of this cancer or other causes	

*Others ($n=1102$) include cheek mucosa ($n=584$), vestibule of mouth ($n=34$), retromolar area ($n=347$), overlapping lesion ($n=41$), and mouth ($n=96$).

3. Results

3.1. Patient characteristics

The study cohort included 9439 patients with oral cancer; 5721 males and 3718 females in a male-to-female ratio of 1.5:1. The mean age at diagnosis was 65.3 years (SD ± 13.7: range 12 – 90) and the median age was 65.0 years. The oral cavity subsites in the series were lip (14.8%), tongue (50.5%), gum (9.1%), hard palate (1.8%), and others (11.7%) (Table 1). Considering the tumor, the AJCC 7th TNM staging scheme showed that 5816 (61.6%) patients had stage T1, 2825 (29.9%) stage T2, 798 (8.5%) stage T3, and no one stage T4. Most of the cases had N0 nodal classification. Specifically, in terms of the nodal parameter, 8204 (86.9%) were N0, 1179 (12.5%) were N1, none had N2, and 54 (0.6%) were N3; while 9385 (99.4%) had M0 and 54 (0.6%) M1. Regarding tumor grading, 3044 (32.3%) tumors were well-differentiated, 5082 (53.8%) moderately differentiated, 1285 (13.6%) poorly differentiated, and 28 (0.3%) were undifferentiated. The clinicopathologic characteristics are briefly summarized in Table 1.

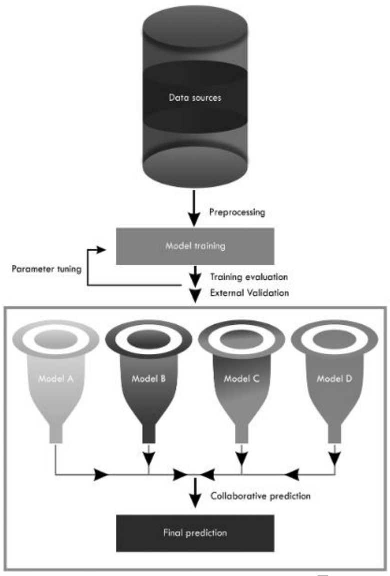

Figure 1. A framework for collaborative model training

In terms of oncological primary treatment, a total of 8741 (92.6%) oral cancer patients received surgery, and 2526 (26.8%) received radiotherapy and beam radiotherapy (n = 2441, 96.6%) being the most common type of radiation. A total of 891 (9.4%) received chemotherapy. Ethnicity formed another important parameter, 801 (48.8%) being of white origin, 189 (11.5%) were black, and 651 (39.7%) were from other origins including American Indian/AK Native, Asian/Pacific Islander. Considering marital status, 5514 (58.4%) were married (married and domestic partners) while 3925 (41.6%) were considered unmarried (single, divorced, widowed, and separated) at the time of diagnosis (Table 1). The follow-up time ranged from 0 to 143 months (Mean 69.6; Median 77; SD $\pm$ 41.5).

3.2. Performance metrics for the algorithms

The individual performance accuracy of the individual participating algorithms was 70.2%, 69.9%, 69.1%, 69.4%, and69.5% for voting ensemble, stacked ensemble, XGBoost, Light GBM, and logistic regression, respectively following model training. (Figure 2). The performance accuracy based on the temporal validation is shown in Table 2. When the predictive outcomes of three of these algorithms were combined for collaborative decision-making, the overall performance accuracy of the cML showed comparable performance with the voting ensemble. (Figure 2).

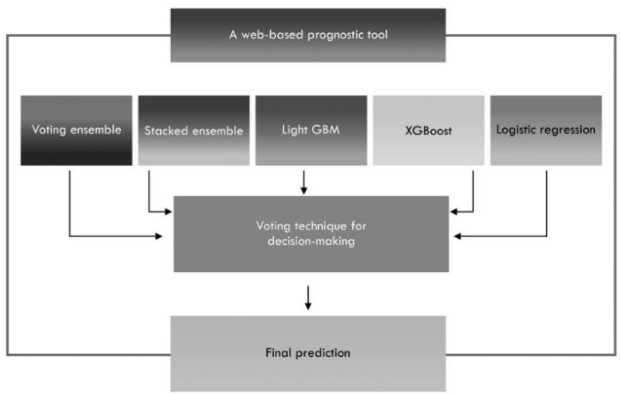

Figure 2. A typical scenario of participating models for collaborative training.

3.4. Evaluating the input variables for importance

The feature importance of the input variables showed that the age of the patient at diagnosis, T stage, tumor grade, marital status, gender, primary site, surgery, N stage, radiation treatment, ethnicity, chemotherapy, and M stage, in decreasing order of importance, were significant for the model's ability to predict the overall survival of oral cancer patients (Figure 3). Specifically, the individual examination of some of the input variables showed that the survival trends between males and females are comparable. Similarly, Caucasian oral cancer patients had better overall survival than oral cancer patients from Black origin or other (American Indian/AK Native, Asian/Pacific Islander) ethnicity. Married cancer patients were found to have better survival than unmarried oral cancer patients.

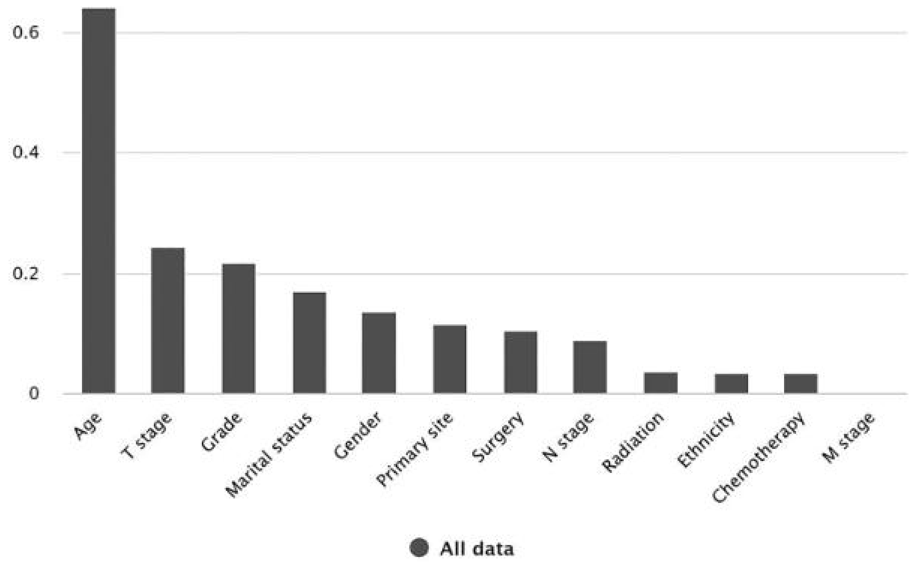

Figure 3. Permutation feature importance.

Table 2. Performance metrics from the five algorithms evaluated during temporal validations.

	Performance metrics	Voting Ensemble	Stacked Ensemble	Light GBM	XGBoost Classifier	Logistic regression
Confusion matrix parameters	True positive	430	432	424	434	426
	False positive	154	158	144	155	162
	False negative	202	200	208	198	206
	True negative	415	411	425	414	407
Other metrics	Sensitivity (recall)	0.68	0.68	0.67	0.69	0.67
	Specificity	0.73	0.72	0.75	0.73	0.72
	F1 score	0.71	0.70	0.71	0.71	0.70
Accuracy (%)	Accuracy	70.4	70.1	70.6	70.6	69.4
	Balanced accuracy	70.5	70.3	70.9	70.7	69.5
AUC	Weighted AUC	0.79	0.79	0.79	0.79	0.78

4. Discussion

In this cohort study of 9439 oral cancer patients, we evaluated the predictive performance of five different machine learning models such as voting ensemble, stacked ensemble, extreme gradient boosting, light boosting, and logistic regression for OS in OSCC patients. In addition, we thoroughly combined three (voting ensemble, stacked ensemble, light gradient boosting) out of these machine learning models to form a collaborative decision-making model. In addition, we performed feature importance to evaluate the order of significance of the input variables for overall survival prediction in oral cancer. Also, by extension through sub-parameter analysis using feature importance, we checked the parameters contained in each of these input variables to evaluate their overall survival trends for oral cancer. Meanwhile, to the best of our knowledge, this is the first study that leverages the paradigm of collaborative machine learning (cML) models for the chance of overall survival prediction in oral cancer.

We noticed that the performance accuracy of the cML for predicting overall survival showed comparable performance with the voting ensemble. The performance accuracy may be significant clinically, considering the sensitive nature of medical applications such as cancer management where reliability of the prediction made ML model is desirable. Therefore, the idea of collaborative models may not necessarily be hinged on performance enhancement but on providing reliability to the prediction made for a particular patient. Remarkably, the idea of a collaborative prediction paradigm was not intended to provide a platform for competition between collaborative models. Rather, each model gives a prediction, and the outcome of the prediction is determined through a collaborative voting method. Essentially, this approach is poised to enhance the reliability of the prediction made by these models. Of note, the performance accuracy of the voting ensemble was on par with the cML since the paradigm of voting ensemble involves the use of different machine learning models to train on the same dataset to make predictions. The final accuracy depends on the prediction probabilities of each model. In this way, the effect of poor performance by one model can be managed by the strong model to produce a reliable final prediction. This also is simi-lar to the cML paradigm; hence, they have shown similar performance metrics in this study

Traditionally, most machine learning models have been considered as a single entity for making predictions. This is more of an individualistic approach to making predictions. However, solo or individual machine learning models may experience certain limitations such as being built on biased datasets or having overfitted data. Therefore, the collaborative methodology gives learners the chance to work together for a final prediction. This approach may preclude encountering the inherent limitations of individual models for predictions.

The feature importance of the input variables showed that the age of the patient at diagnosis, T stage, tumor grade, marital status, gender, primary site, surgery, N stage, radiation treatment, ethnicity, chemotherapy, and M stage, in decreasing order of importance, were significant for model's ability to predict the overall survival of oral cancer patients. Furthermore, the feature importance analysis of some of the input parameters revealed the survival trends. For example, we found that there was no clear survival difference between the genders (Figure 4a). This result was supported by some of previous studies where it was found that there was no survival difference between males and females for oral cancer [13]. For instance, the studies by Lin et al. (2020) and Lee et al. (2021) found no difference in the Taiwanese population [14,15]. A slight difference may arguably be observed between the survival trends of the genders for this US dataset (Figure4(a)). It appeared that the males had slightly better survival than the females. This observation is in tandem with a study of Benchetrit et al. that was conducted using dataset from the same country (United States data series), where it was suggested that the reason for better male survival trend may be attributed to the fact, that women are less likely than men to receive definitive chemoradiotherapy as opposed to definitive radiotherapy [16].

Also, Caucasians and Blacks with oral cancer show better survival than other ethnic groups such as Asia (Figure 4(b)). This may be due to later-stage presentation, socioeconomic (health insurance, living condition), clinical, and treatment-related factors for Asian oral cancer patients [17, 18]. Married patients showed better survival than unmarried patients (Figure 4(c)). Marriage and living with a partner may offer social and spousal support mechanisms to oral cancer patients [19]. Of note, tumor grade classified as either moderately, poorly, or undifferentiated was shown to have higher survival than tumor grade I (Figure 4(d)). This may be due to misclassification or the possibility of offering more aggressive treatment alternatives for oral cancer patients at this stage. The current survival trends observed for tumor grade warrants further research.

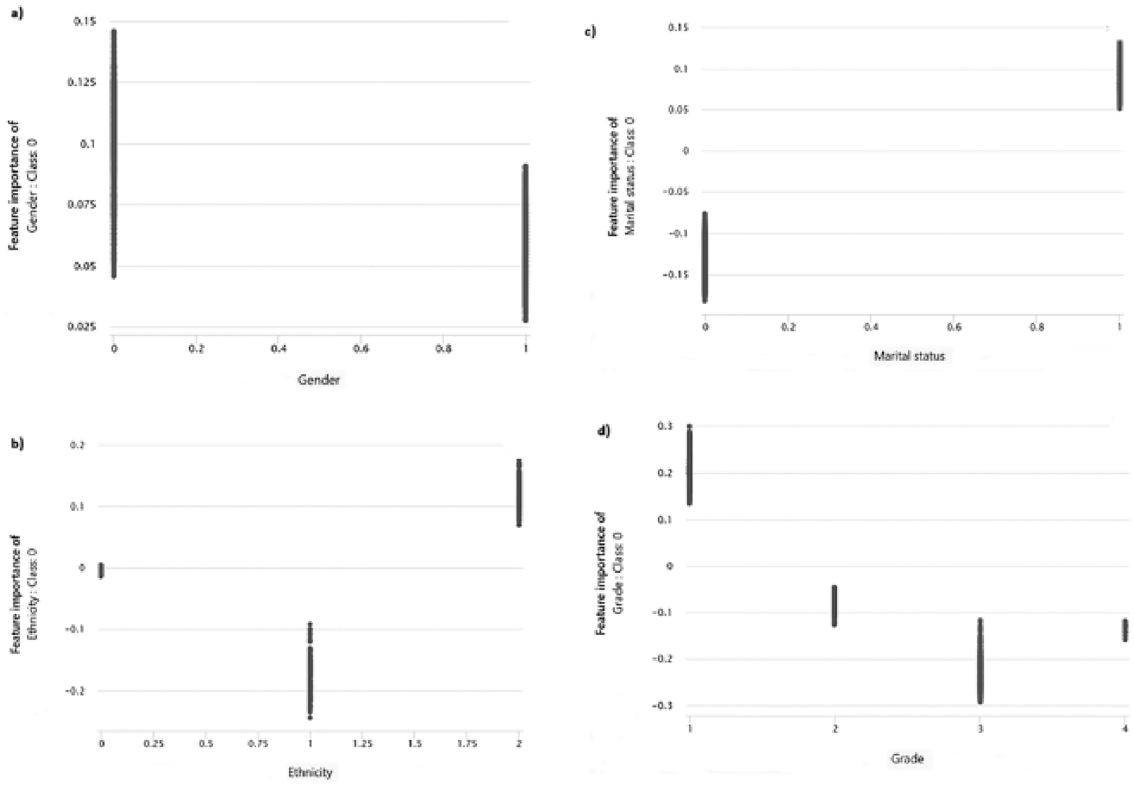

Figure 4. The survival trends for gender, ethnicity, marital status, tumor grade and tumor stage.

In terms of the survival trends for the treatment modalities, our findings further demonstrated surgery as an important treatment approach. Patients who received surgical treatment showed comparable overall survival. This finding is supported by the study of Spiotto et al. that supported the use of surgery as the primary treatment modality for operable oral cancer [20]. Similarly, radiotherapy treatment was associated with improved survival (Supplementary IIa). This is particularly useful in combined treatment approach with surgery for patients with locally advanced oral cancer [20]. Chemotherapy treatment improved survival trends and this remains obvious for the advanced cases. Patients with less advanced tumors did not receive chemotherapy and thus showed relatively comparable overall survival (Supplementary IIb).

Our study has some limitations. Firstly, the data used for model development was of retrospective nature. As with all retrospective studies, selection bias cannot be completely eliminated. Secondly, even though our model was developed using data obtained from the SEER database (one of the largest publicly available cancer databases for the United States population), it did not contain such parameters as smoking history, comorbidities, treatment intent, type of chemotherapy, among other important information. These informant parameters, which extend beyond mere cancer-related parameters, often play a significant role in developing a robust model clinical decision-decision. This unfortunately could introduce a major limitation in the interpretation of our findings. Thirdly, uncertainties in terms of the sequence of radiation treatment and completeness of chemotherapy and radiotherapy data in the SEER dataset form a limitation.

In conclusion, the use of cML may increase the accuracy of estimating the overall survival of oral cancer patients. While enhancing the accuracy may be desirable, improving the reliability of the prediction made by the model is equally important for sensitive applications such as in cancer management. Thus, insightful, targeted, and personalized treatment planning can be enhanced. The collaborative machine-learning methodology was enhanced through democratic voting among the participating machine learning models. In the future, an end-to-end iterative workflow allowing the participating algorithms to train data and collaborate in using other approaches than voting should be explored. Considering the various concerns such as privacy and security regarding the availability of data, a federated learning approach for model improvement should be explored in future studies. In the federated learning approach, model performance is jointly improved collaboratively without compromising data security and privacy. In the future, these methodologies will be extended to include other types of medical data formats such as imaging data. In addition, cML approach may be validated with relatively large amount of data to visualize the significance of cML.

Acknowledgment

The Sigrid Jusélius Foundation. Finska Läkaresällskapet.

Disclosure statement

The authors report that there are no competing interests to declare.

Figure Legend

Figure 1. A framework for collaborative model training.

Figure 2. Participating models for collaborative training.

Figure 3. Permutation feature importance.

Figure 4. The survival trends for gender, ethnicity, marital status, tumor grade and tumor stage.

Supplementary I. The survival trends for tumor stage.

Supplementary II. The survival trends for treatment modalities – a) radiotherapy and b) chemotherapy.

Tables

Table 1. Baseline demographic and tumor characteristics of oral cancer patients in the SEER database.

Table 2. Performance metrics from the five algorithms evaluated during temporal validations.

References

[1] Adorno-Farias D, Morales-Pisón S, Gischkow-Rucatti G, et al. Genetic and epigenetic landscape of early-onset oral squamous cell carcinoma: Insights of genomic underserved and underrepresented populations. Genet Mol Biol [Internet]. 2024 [cited 2024 Aug 9]; Available from: https://www.scielo.br/j/gmb/a/pPnHv5zwjWLn8qZm5mjqKMp/?lang=en.

[2] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;

[3] Barsouk A, Aluru JS, Rawla P, et al. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. Medical Sciences. 2023;11:42.

[4] Alabi RO, Elmusrati M, Leivo I, et al. Advanced-stage tongue squamous cell carcinoma: a machine learning model for risk stratification and treatment planning. Acta Oto-Laryngologica. 2023;1–9.

[5] Prognostic factors in tongue cancer – relative importance of demographic, clinical and histopathological factors. British Journal of Cancer. 2000;83:614–619.

[6] Alabi RO, Mäkitie AA, Pirinen M, et al. Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer. International Journal of Medical Informatics. 2021;145:104313.

[7] Alabi RO, Elmusrati M, Sawazaki-Calone I, et al. Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: a Web-based prognostic tool. Virchows Archiv. 2019;475:489–497.

[8] Alabi R, Almangush A, Elmusrati M, et al. An interpretable machine learning prognostic system for risk stratification in oropharyngeal cancer. International Journal of Medical Informatics. 2022;168:104896.

[9] Alabi RO, Elmusrati M, Leivo I, et al. Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP. Sci Rep. 2023;13:8984.

[10] Alabi RO, Elmusrati M, Sawazaki-Calone I, et al. Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer. International Journal of Medical Informatics. 2019;104068.

[11] Mäkitie AA, Alabi RO, Ng SP, et al. Artificial Intelligence in Head and Neck Cancer: A Systematic Review of Systematic Reviews. Adv Ther. 2023;40:3360–3380.

[12] Alabi RO, Youssef O, Pirinen M, et al. Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. Artificial Intelligence in Medicine. 2021;115:102060.

[13] Honorato J, Rebelo MS, Dias FL, et al. Gender differences in prognostic factors for oral cancer. International Journal of Oral and Maxillofacial Surgery. 2015;44:1205–1211.

[14] Lin N-C, Hsu J-T, Tsai K-Y. Difference between Female and Male Patients with Oral Squamous Cell Carcinoma: A Single-Center Retrospective Study in Taiwan. IJERPH. 2020;17:3978.

[15] Lee Y-C, Young C-K, Chien H-T, et al. Characteristics and outcome differences in male and female oral cavity cancer patients in Taiwan. Medicine. 2021;100:e27674.

[16] Benchetrit L, Torabi SJ, Tate JP, et al. Gender disparities in head and neck cancer chemotherapy clinical trials participation and treatment. Oral Oncology. 2019;94:32–40.

[17] Yu AJ, Choi JS, Swanson MS, et al. Association of Race/Ethnicity, Stage, and Survival in Oral Cavity Squamous Cell Carcinoma: A SEER Study. OTO Open. 2019;3:2473974X1989112.

[18] Baliga S, Yildiz VO, Bazan J, et al. Disparities in Survival Outcomes among Racial/Ethnic Minorities with Head and Neck Squamous Cell Cancer in the United States. Cancers. 2023;15:1781.

[19] Inverso G, Mahal BA, Aizer AA, et al. Marital status and head and neck cancer outcomes. Cancer. 2015;121:1273–1278.

[20] Spiotto MT, Jefferson G, Wenig B, et al. Differences in Survival With Surgery and Postoperative Radiotherapy Compared With Definitive Chemoradiotherapy for Oral Cavity Cancer: A National Cancer Database Analysis. JAMA Otolaryngol Head Neck Surg. 2017;143:691.