This is a self-archived – parallel published version of this article in the publication archive of the University of Vaasa. It might differ from the original.

Advanced-stage tongue squamous cell carcinoma: a machine learning model for risk stratification and treatment planning

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

Title: Advanced-stage tongue squamous cell carcinoma: a machine learning model for risk stratification and treatment planning

Year: 2023

Version: Accepted Manuscript

Copyright ©2023 Taylor & Francis. This is an Accepted Manuscript of an article published by Taylor & Francis in Acta Oto-laryngologica on 15 Feb 2023, available online: http://www.tandfonline.com/10.1080/00016489.2023.2172208

Please cite the original version:

Alabi, R. O., Elmusrati, M., Leivo, I., Almangush, A. & Mäkitie, A. A. (2023). Advanced-stage tongue squamous cell carcinoma: a machine learning model for risk stratification and treatment planning. *Acta Oto-laryngologica* 143(3), 206-214. https://doi.org/10.1080/00016489.2023.2172208

Advanced-stage tongue squamous cell carcinoma: A machine learning model for risk stratification and treatment planning

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: 3782

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

DRAFT COPY OF THE MANUSCRIPT

Abstract

Background: A significant number of tongue squamous cell carcinoma (TSCC) patients are diagnosed at a late stage. **Objectives:** We primarily aimed to develop a machine learning (ML) model based on ensemble ML paradigm to stratify advanced-stage TSCC patients into the likelihood of overall survival (OS) for evidence-based treatment. We compared the survival outcome of patients who received either surgical treatment only (Sx) or surgery combined with postoperative radiotherapy (Sx+RT) or postoperative chemoradiotherapy (Sx+CRT). **Material and Methods:** A total of 428 patients from Surveillance, Epidemiology, and End Results (SEER) database were reviewed. Kaplan-Meier and Cox proportional hazards models examine OS. In addition, a ML model was developed for OS likelihood stratification. **Results:** Age, marital status, N stage, Sx, and Sx+CRT were considered significant. Patients with Sx+RT showed better OS than Sx+CRT or Sx alone. A similar result was obtained for T3N0 subgroup. For T3N1 subgroup, Sx+CRT appeared more favorable for 5-year OS. In T3N2 and T3N3 subgroups, the numbers of patients were small to make insightful conclusions. The OS predictive ML model showed an accuracy of 86.3% for OS likelihood prediction. **Conclusions and Significance:** Patients stratified as having high likelihood of OS may be managed with Sx+RT. Further external validation studies are needed to confirm these results.

KEYWORDS: Tongue cancer; Overall survival; Machine learning; Chemoradiotherapy; Radiotherapy; Radiation; Surgery; SEER

DRAFT COPY OF THE MANUSCRIPT

Introduction

Tongue squamous cell carcinoma (TSCC) represents the most frequent site for oral cavity cancer [1]. The rates for incidence, recurrence, and mortality of TSCC have shown a marked increase in recent decades in the Western world due to the aggressive nature of this type of cancer in terms of its rapid local invasion and early lymph node metastasis [2]. Therefore, a concise effort is needed to predict TSCC tumor behavior, but the lack of specific prognostic indicators still constitutes a major challenge [2]. In addition, the decision-making regarding the best treatment approach is somewhat challenging for many cases with TSCC despite the generally improved overall survival (OS) in this patient population.

Traditionally, the main tumor-related parameters for TSCC management decision are based on the American Joint Committee on Cancer (AJCC) tumor-nodal-metastasis (TNM) staging scheme. However, there are valid concerns regarding the sufficiency of TSCC stage for prognostication [2]. Additionally, the ideal treatment for patients with advanced-stage TSCC at diagnosis remains controversial. Therefore, accurate and efficient risk-stratification of patients at advanced stage TSCC is necessary for their treatment recommendations.

TSCC patients at advanced stage usually experience higher rates of recurrence and poor survival despite the recent improvements in diagnostic and management approaches [3,4]. Additionally, the treatment options for these patients may contribute to significant morbidity and psychosocial concerns. Advanced-stage TSCC is characterized by significant burden on the patient's physical appearance (i.e. disfiguration), proper functioning (mastication and deglutition), major senses, airway, upper gastrointestinal tract, and low self-esteem in terms of social interactions and normal daily activities [3]. Thus, it is important to properly examine the patients to plan a targeted individualized treatment option by stratifying the advanced stage patients into risk groups i.e. – either low likelihood or high likelihood of 5-year overall survival (OS).

DRAFT COPY OF THE MANUSCRIPT

Of note, some advanced-stage TSCC patients may be deemed at diagnosis unsuitable for curatively intended therapy [5]. Therefore, a non-curative approach aiming at palliative care and enhancing quality of life has been proposed [3]. The clinicians need to make informed decisions regarding curative or non-curative treatment intent for advanced-stage TSCC. Based on this, it is important to create an OS likelihood stratification system that enables the survival outcomes of patients with advanced-stage TSCC at initial diagnosis to be predicted. This is poised to offer a targeted therapeutic intervention for the patient.

In recent years, machine learning (ML), a subfield of artificial intelligence (AI) has gained increased interest in clinical research due to its feasibility and ability to analyze the complex relationships between clinical variables for prognostication [6]. These relationships can be modeled by ML techniques for outcome prediction. Thus, several studies have been published regarding the prognostication of outcomes in various sites of the head and neck region [7,8]. Due to the promising results obtained from these ML algorithms and advancements in technology, several modifications to these algorithms have been developed. For example, some of these modifications include understanding the dependent and independent relationships between these variables for risk of bias to enhance the effective estimation of various outcomes in oncology. These modifications are aimed at improving the performance of these models for effective prognostication of outcomes to enhance patient-centered care (personalized medicine) and proper management of cancer.

In this study, we examined the treatment options regarding the 5-year OS in advanced stage TSCC. Furthermore, we aimed to explore the potential of two distinct ML algorithms – k-nearest neighbor oracle (KNORA), which is derived from dynamic ensemble selection techniques and extreme random trees for OS likelihood prediction of advanced-stage TSCC patients. This prognostication can assist in the personalized risk stratification, to allow tailoring an adequate treatment alternative for each patient.

DRAFT COPY OF THE MANUSCRIPT

2. Material and Methods

2.1. Study population

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 (alabir) for research plus data for November 2020 Submission. From this selection, the SEER database was queried for tongue cancer patients from 2010 - 2015. A total of 428 cases was extracted and considered advanced-stage (T3N0 and T3N1) oral tongue cancer patients. Owing to the limitations of the SEER database in terms of the recorded cases, there were no extracted cases for T3N2 and extremely limited cases for T3N3. Therefore, these were excluded from the analyses. A detailed description of data extraction from the SEER database is given in Figure 1.

2.2. Data process and selection of attributes

2.2.1 Inclusion and exclusion criteria

The inclusion criteria comprised patients with pathologically confirmed tongue cancer with clinicopathological parameters available including age at diagnosis, ethnicity, gender, marital status, grade, and stage classification according to the American Joint Committee on Cancer (AJCC) tumor-nodal-metastasis (TNM) 7^{th} edition (Table 1). Overall survival was the primary endpoint and target variable. The extracted clinicopathological parameters were differentiated into categorical and quantitative variables (Table 1). The categorical parameters were normalized for the ML training phase (Section 2.4). In the SEER database, overall survival is considered as the time until death as a result of any cause and follow-up period as the time from diagnosis until the last follow-up or death of any cause. The exclusion criteria include all unknown cases (i.e., cases with missing information) in each of the considered clinicopathological parameters. In addition, all other tumor sites than oral tongue were excluded.

DRAFT COPY OF THE MANUSCRIPT

2.2.2 Treatment endpoints

In terms of the treatment parameters (surgery, radiotherapy, and chemotherapy), surgical treatment is the primary option for tongue cancer patients (Table 1). For advanced-stage tongue cancer patients, surgical treatment only (Sx), surgery followed by either postoperative radiotherapy (Sx+RT) or postoperative chemoradiotherapy (Sx+CRT) were the used treatment options (Table 1).

2.3. Statistical analysis

The statistical analyses were conducted using Python version 3.10.0 in the Jupyter notebook. The Kaplan–Meier method was used to calculate the OS. Additionally, Cox proportional hazards models were used to assess the association of variables with the survival and unadjusted hazard ratio (HR) and confidence interval (CI). Statistical significance was defined as $p < 0.05$.

2.4. Machine learning model training

We compared an extreme random tree algorithm with a dynamic ensemble selection (DES) to stratify advanced-stage TSCC patients into the likelihood of overall survival. The DES was achieved using k-nearest neighbor oracle (KNORA). Following the data preprocessing (sub-section 2.2.1), the processed data were loaded into Jupyter notebook using Python version 3.10.0. Then, the advanced-stage tongue cancer was extracted from the main data loaded to the Jupyter notebook. A 10-fold cross-validation was used with various hyperparameters tunning to maximize the performance of the model (Figure 1). The hyperparameters that give the best accuracy for each model were selected and used in the KNORA model training (k = 7 for the selection of the local neighborhood for making predictions). We used bagged decision trees as the pool of classifier model (repeated k-fold cross-validation with 3 repeats and 10 folds). To ensure that a vibrant model was developed, we considered two variants of this algorithm – KNORA-eliminate (KNORA-E) and KNORA-Union (KNORA-U) for the likelihood of OS of

DRAFT COPY OF THE MANUSCRIPT

advanced-stage TSCC. The performances of each of these variants were evaluated in terms of mean accuracy. The schematic of the ML training process with regards to KNORA methodology is given in Figure 1. The mean accuracy of the KNORA methodology was compared with an extreme gradient boosting algorithm that was trained using a 10-fold cross validation. The training of the extreme boosting algorithm was done using Microsoft Azure Machine Learning Studio (Azure ML 2021) to build the predictive model. The performance of the extreme random tree was evaluated using weighted accuracy.

2.5. Model prediction: High and low likelihood of overall survival

The trained ML model predicts the individual patients with advanced-tongue cancer as either having a high likelihood of OS or low likelihood of OS. A high likelihood of OS means that the trained model predicts the patient to be alive from the date of diagnosis to the last follow-up. Conversely, a low likelihood of survival indicates that the trained ML model predicts the patient to die during the follow-up and consider all causes of death.

3. Results

3.1. Patient characteristics

The patient-, tumor- and treatment-related characteristics of the main group (entire patient population) are summarized in Table 1. The median age of patients at diagnosis was 63 years (range, 18–85; average age of 63.0; SD $\pm$ 14.5) and stratifying the age group into young and old using 40 years as the cut-off point indicated that 28 (6.5%) were younger patients while 400 (93.5%) can be regarded as older patients with advanced-stage TSCC. Furthermore, the male-to-female ratio was 1.9:1 (278 male, 150 female). Other important parameters such as ethnicity showed that, 336 (78.5%) were from the white origin, 52 (12.2%) were black, and 40 (9.4%) were from other origins including American Indian/AK Native, Asian/Pacific Islander. Considering marital status, 195 (45.6%) were married while 233 (54.4%) were considered unmarried (single, divorced, widowed, and separated) at the time of diagnosis (Table 1).

DRAFT COPY OF THE MANUSCRIPT

According to the SEER classification, the most common primary site was the mobile tongue (58.4%, n = 250) including the anterior two-third while the remaining sites (41.6%, n = 178) which included dorsal surface of tongue, border of tongue, overlapping lesion of tongue, and ventral surface of the tongue. All of the extracted cases were T3 (100.0%, n = 428) [Table 1]. The most prominent pathologic N-stage was N0 which covered a total of 273 (63.8%) patients, while 142 (33.2%) had N1 and 13 (3.0%) had N3 in accordance with the AJCC staging scheme. Of note, there were no extracted cases with N2 (Table 1). Regarding M stage, 418 (97.7%) had M0 and 10 (2.3%) had M1. Regarding tumor grading, in 83 (19.4%) out of the 428 cases, the grade was well differentiated, 255 (59.6%) moderately differentiated, 90 (21.0%) poorly differentiated, and there were no undifferentiated cases. The histopathologic characteristics are briefly summarized in Table 1. Summarily, the sub-groups considered for the advanced stage of the disease include a total of 273 (63.8%) having T3N0, 142 (33.2%) had T3N1, and 13 (3.0%) had T3N3. Thus, the T3N3 subgroup were excluded in this analysis due to few numbers of cases. There were no T3N2 cases.

3.1.4 Treatment options for the subgroups

The treatment options were considered from three different perspectives – standard treatment of surgery followed by either postoperative radiotherapy (Sx+RT) or postoperative concurrent chemoradiotherapy (Sx+CRT), and surgical treatment only (Sx). A total of 194 (45.3%) patients had Sx+RT, 72 (16.8%) had Sx+CRT, and 138 (32.2%) had Sx only, while 24 (5.6%) patients received none of the available treatment options (Table 1). The follow-up time ranged between 0 to 107 months with the majority (61.4%) of the patients being alive at the last follow-up.

3.2. Survival analysis

In the entire TSCC cohort, the following parameters showed to be significant with $p < 0.005$ and hazard ratio (HR) of 1.03 for the age of the patient, HR of 0.74 for the marital status, HR

DRAFT COPY OF THE MANUSCRIPT

of 1.49 for N stage, HR of 0.25 for surgery, and HR of 2.23 for Sx+CRT (Figure 2). All the considered cases were of advanced stage (T3) with majority of them presenting as M stage (M0). Therefore, these two parameters (i.e., T stage and M stage) were kept constant in the analysis. In terms of other treatment approach, the Kaplan-Meier analysis demonstrated that the treatment option of surgery followed by postoperative radiotherapy (Sx+RT) showed better OS than either surgery with chemoradiotherapy (Sx+CRT) or surgery only (Sx) [Figure 3]. Similarly, in the T3N0 subgroup, the same result was obtained (Figure 4).

However, in the T3N1 subgroup, the Kaplan-Meier plot for the treatment option revealed that Sx+CRT appeared promising for 5-year OS (Figure 5). There were no cases extracted with T3N2 from the dataset and too few cases with T3N3, and thus no insightful conclusion regarding the 5-year OS could be made from these subgroups. Therefore, the conclusions from this study are based on the treatment outcome of stages T3N0 and T3N1.

3.3. Model performance

The hyperparameter tunings showed the performance of various values for the k-nearest neighbour for a vibrant KNORA model (k = 7 and bagging). In terms of the performances of the KNORA variants, the KNORA-Union (KNORA-U) showed a mean accuracy of 84.6% and 0.050 as the standard deviation of the accuracy of the model while KNORA-Eliminate (KNORA-E) gave a mean accuracy of 83.9% and 0.048 as the standard deviation of the accuracy of the model across all repeats and folds in the OS likelihood estimation of advanced-stage TSCC. Likewise, an extreme random tree showed a weighted accuracy of 86.3% and 0.91 for the weighted area under curve (Figure 6). Other performance metrics of the extreme random tree is presented in Table 2. Hence, these ML models showed extremely similar performance ability in the OS likelihood stratification of advanced-stage TSCC. This further emphasized that ensemble ML algorithms have great potential in predictive oncology.

DRAFT COPY OF THE MANUSCRIPT

4. Discussion

We examined the 5-year overall survival (OS) of patients with advanced-stage tongue cancer in the SEER database with respect to their given treatment options. We developed a machine learning model that can stratify the patients into a high likelihood or low likelihood of OS. Additionally, we examined the Kaplan-Meier plot for the available treatment options within the extracted cohort, which revealed that patients that had received surgery followed by postoperative radiotherapy (Sx+RT) showed a better 5-year OS than their counterparts that had received either surgery only (Sx) or surgery with postoperative chemoradiotherapy (Sx+CRT). A similar result was obtained in the T3N0 subgroup. However, for patients with T3N1, Sx+CRT showed a more promising 5-year OS. Therefore, it is evident that the survival results are not uniform in all the advanced-stage TSCC, which remains a challenging group to manage. This makes the decision-making regarding the design of the best treatment approach challenging for advanced-stage TSCC. To the best of our knowledge, our study represents the first attempt to examine the potential of ML model for the likelihood of overall survival in advanced-stage tongue cancer.

The examined ML models showed promising results in the likelihood of OS prediction for this patient population to enhance targeted treatment. These models were built based on ensemble machine learning paradigm. For example, an extreme random tree used in this study randomly samples the entire dataset. During training phase, that is, in the model development, it constructs trees over every observation in the dataset but with different subsets of features [9]. This makes it faster since the node splits are randomized [9]. Thus, producing a low variance which makes it suitable for likelihood of OS estimation. Similarly, the k-nearest Neighbor Oracle (KNORA) is based on the dynamic ensemble selection (DES) paradigm where the oracle appellation in the KNORA describes the selected ensemble of models used for the prediction of outcome [10]. It dynamically selects one or more models (i.e., an ensemble

DRAFT COPY OF THE MANUSCRIPT

of models) to make a prediction. The final prediction is derived based on voting on the predictions made by these ensembles of models. Therefore, the final prediction made by using the DES paradigm is based on the consideration of predictions of multiple models instead of a single best model [10]. In effect, the DES approach is poised to provide a highly accurate outcome prediction that is of utmost importance in cancer management such as in the likelihood of OS prediction for advanced-stage TSCC.

Remarkably, advanced-stage tongue cancer is characterized by a worse prognosis at diagnosis. By standard, the recommended treatment for late-stage tongue cancer is surgery followed by postoperative radiotherapy [11]. This recommendation was corroborated in this study as it was observed that patients that received Sx+RT showed a better chance of OS (Figure 3 – 4). Similarly, the European Organization for Research and Treatment of Cancer (EORTC) and Radiation Therapy Oncology Group (RTOG) has suggested the use of concurrent chemotherapy [11]. This latter suggestion was also supported by the finding in this study specifically for advanced-stage patients with T3N1 disease where having received Sx+CRT showed promising results in terms of 5-year OS [11] (Figure 5). With the current radiotherapy (RT) techniques such as the intensity-modulated radiotherapy treatment (IMRT), this treatment modality may be poised to be a vibrant treatment approach for advanced-stage tongue cancer. Post-operative chemoradiotherapy is typically indicated for patients with multiple metastases, extra nodal spread, or positive surgical margin.

Therefore, based on the afore-mentioned treatment options, patients with a high likelihood of OS may be targeted with either Sx+RT or Sx+CRT as their primary treatment approach (Figure 4). Similarly, patients with more advanced stage (≥ T3N1) with a high chance of disease recurrences, Sx+RT or Sx+CRT may be the treatment option of choice. Conversely, in cases where the ML model stratifies the patients as having low likelihood of survival (i.e., low chance of survival category), curatively intended treatment modalities may be

DRAFT COPY OF THE MANUSCRIPT

unwarranted. In such a scenario, non-curative treatment aimed at life-prolonging treatment or palliative care may be considered portion of patients are reported to experience early death after head and cancer treatment. Several studies have been published emphasizing the importance of palliative care with or without traditional life-prolonging treatment [5,12]. Notably, the study by Mroueh et al. emphasized the importance of non-curative treatment as it comprises both life-prolonging treatments including Sx+RT, Sx+CRT, and RT and those aimed at alleviating tumor-related symptoms [3]. With this likelihood of survival prediction by a ML model and individualization of the treatment approach, the quality of life of advanced-stage TSCC patients can be improved and survival time may be prolonged [13].

Certainly, the stage of tongue cancer at diagnosis remains the most important prognostic factor [14]. Therefore, advanced stages are usually associated with high disease recurrence and mortality rates. Despite the concerted efforts towards early-stage detection of tongue cancer, late-stage presentation is still common. In this study, age of the patient was found to be associated with OS in late-stage OTSCC as has been reported previously [15]. However, some conflicting results have been published regarding the prognosis between younger and older adults [15]. Recent studies have improved the methodology and an insightful cut-off point regarding age stratification as young or old. As a result, it was considered that younger patients were associated with better OS due to their age [15]. Several reasons have been attributed to this. Notably, younger patients may be able to better tolerate aggressive approaches [15]. Similarly, ethnicity of the patient was deemed significant to the OS in advanced TSCC. This may be largely attributable to socioeconomic status, stage of presentation at diagnosis, level of awareness regarding TSCC, disposition towards the health care system, and attitude towards oral screening examination [16]. Likewise, marital status was found to be significant to the OS in TSCC. This is in accordance with several of previous studies considering the social support

DRAFT COPY OF THE MANUSCRIPT

from the family members and the consequent emotional state having a decreasing effect on the psychological distress [17].

Focusing on lymph nodal status and the presence of distant metastases are directly related to the oncological stage and, consequently, to the prognosis [15]. Thus, as the patients are characterized by advanced stage of the disease, it is expected that these two parameters (lymph nodal status and distant metastasis) are significant as reported in our study (Table 1). The surgical treatment was considered significant for OS and this corroborates with the study by Boonpoapichart et al. where the survival improved in patients who were presented with a late disease stage but had surgery [18]. Therefore, the primary treatment strategy for TSCC is upfront surgery [19]. In advanced stages, surgical treatment is followed by radiotherapy or chemoradiotherapy based on final histopathological findings, according to international guidelines [15,19].

Having a ML model that can stratify the patients into respective likelihood of OS will ensure that targeted treatment intervention is offered to the patients even at presentation. This approach is poised to improve the quality of life of the patients, most importantly, incorporate the possibility of palliative care from the onset, especially for low likelihood of survival patients. Despite this stratification possibility by the ML model, it is important to monitor the treatment intervention for advanced and progressive disease. Thus, a swift shift in the treatment objective becomes pertinent where progression of the disease is noticed during the treatment phase or follow-up or the patient becomes unresponsive to the initially planned treatment intervention [3]. However, the final decision regarding the most appropriate treatment lies with the clinicians after considering many parameters such as growth pattern of the tumor, margins and the histopathological examination of the specimen, performance status, and co-morbidities of the patients.

DRAFT COPY OF THE MANUSCRIPT

Despite the potentials of AI for the proper management of cancer, there are certain ethical and legal issues relating to the use of ML in daily clinical practice. Alabi et al. mentioned that the data privacy and confidentiality, bias or discrimination, peer disagreement (contradictory diagnostic or prognostic opinion between the model and the clinician), and patient's liberty to decide the type of treatment may be violated were some of the ethical and legal issues [20]. Indeed, there are currently no well-defined regulations in place to address the legal and ethical issues that may arise because of the potential usage of AI in daily clinical practice [20]. Therefore, there is a need for AI/ML algorithmic transparency, data privacy and confidentiality protection, and an overall ethical and legal framework for the use of AI in oncology [20].

This study has some limitations. First, there are certain limitations that are inherent to the SEER database. For example, confounding factors such as comorbidity, tumor recurrence, duration of therapy, dose of radiation, and chemotherapy agents that could affect survival could not be directly extracted from the SEER database. Thus, this constitutes a limitation in terms of the dataset used in this study. Second, the retrospective nature of the dataset means that the developed model may be poised to be biased. In some of the cases, the number of cases used in the analyses were relatively small. Due to this small sample size, the authors could not provide the clinical characteristics of the subgroups. As a result of the limitations of the SEER database, further studies in the form of external independent validation should be performed to verify the results and recommendations in terms of the treatment.

In conclusion, our registry-based study showed that for advanced stage TSCC, the patients predicted to have a high likelihood of OS can benefit from any of the treatment options (Sx+RT or Sx+CRT) while the patients with low likelihood of survival should be offered the option for palliative care in addition to the traditional treatment option. This approach is poised to ensure that the quality of life of advanced-stage tongue cancer patients can be improved, and

DRAFT COPY OF THE MANUSCRIPT

the survival of these patients can be prolonged. Remarkably, the onus is on the clinicians to make insightful conclusions regarding the best treatment for the patients. For example, the administration of RT techniques if necessary. Due to these limitations, external validation studies may be warranted to confirm these results.

Acknowledgment

Minerva Foundation: Selma and Maja-Lisa Selander’s Fund for Research.
Finksa Läkaresällskapet, The Sigrid Jusélius Foundation. The Helsinki University Hospital Research Fund.

DRAFT COPY OF THE MANUSCRIPT

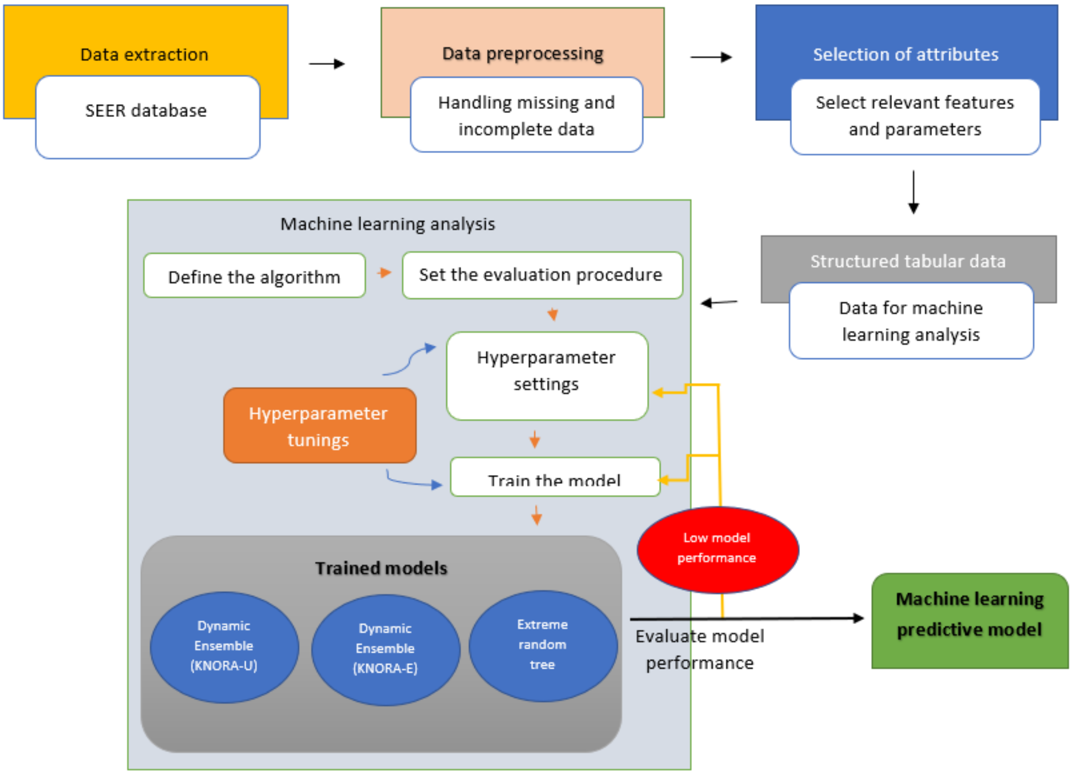

Figure 1. Schematic of the machine learning workflow for model training.

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	z	p	-log2(p)
Age	0.03	1.03	0.01	0.02	0.04	1.02	1.04	5.40	<0.005	23.88
Race	0.13	1.14	0.10	-0.06	0.31	0.94	1.37	1.34	0.18	2.47
Gender	-0.12	0.89	0.13	-0.38	0.15	0.68	1.16	-0.86	0.39	1.35
Marital	-0.30	0.74	0.13	-0.55	-0.04	0.58	0.96	-2.25	0.02	5.36
Grade	0.03	1.03	0.11	-0.18	0.24	0.84	1.28	0.32	0.75	0.42
N-Stage	0.40	1.49	0.10	0.21	0.59	1.23	1.81	4.11	<0.005	14.65
Surgery	-1.38	0.25	0.22	-1.81	-0.94	0.16	0.39	-6.20	<0.005	30.72
Radiation	-0.57	0.56	0.33	-1.23	0.08	0.29	1.08	-1.72	0.09	3.55
Sx+RT	0.11	1.11	0.37	-0.62	0.84	0.54	2.31	0.29	0.77	0.38
Chemotherapy	-0.01	0.99	0.37	-0.74	0.73	0.48	2.07	-0.02	0.99	0.02
CRT	-0.70	0.50	0.49	-1.65	0.26	0.19	1.29	-1.43	0.15	2.72
Sx+CRT	0.80	2.23	0.39	0.04	1.57	1.04	4.79	2.05	0.04	4.64

Figure 2. The hazard ratios of the input features.

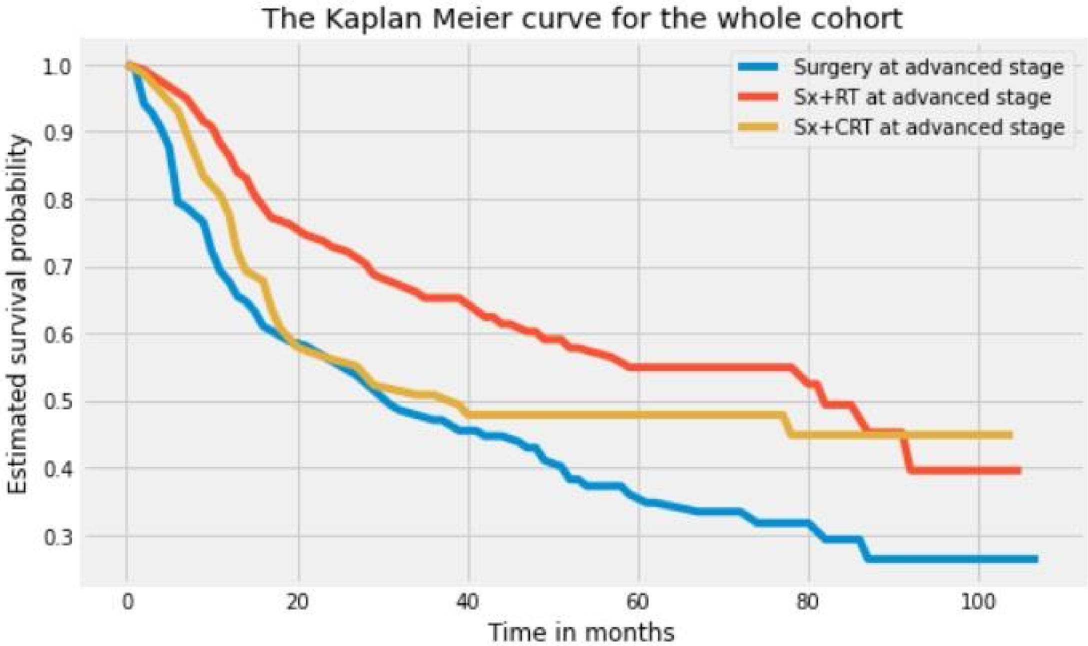

Figure 3. The Kaplan Meier curve for the advanced stage.

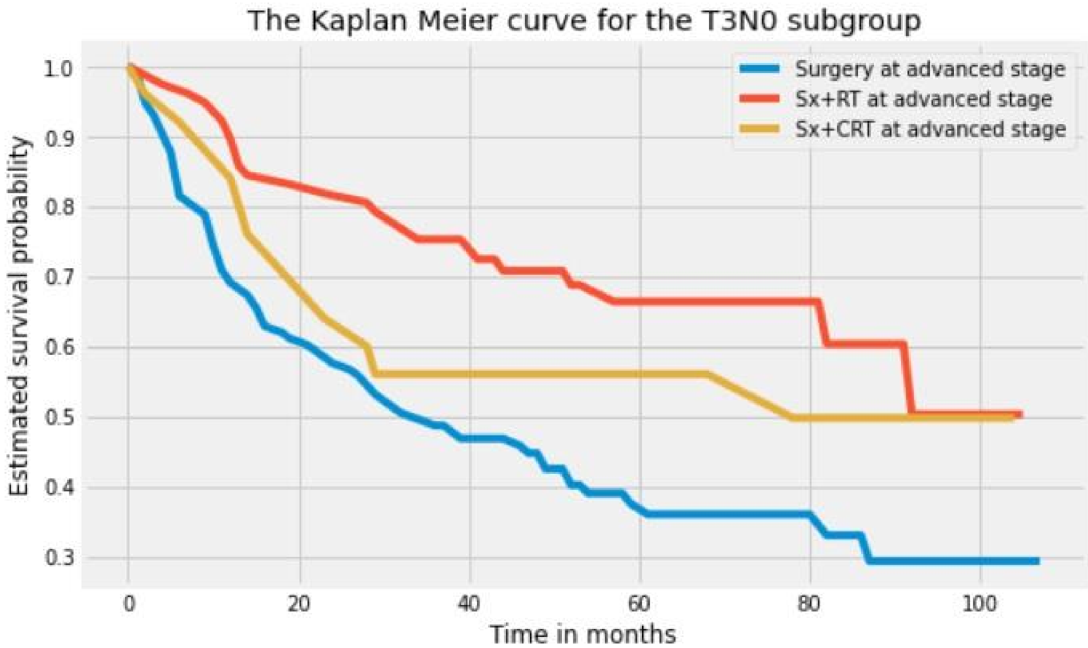

Figure 4. The Kaplan Meier curve for the T3N0 advanced stage.

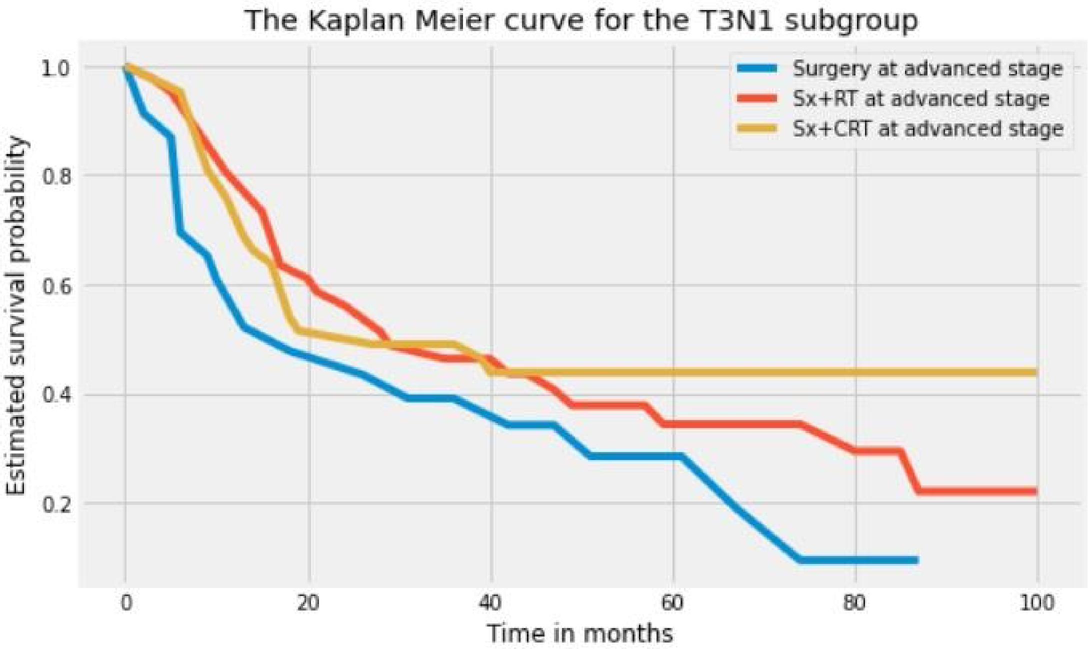

Figure 5. The Kaplan Meier curve for the T3N1 advanced stage.

Disclosure statement

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

Figure Legend

Figure 1. Schematic of the machine learning workflow for model training.

Figure 2. The hazard ratio plot of the included variables.

Figure 3. The Kaplan Meier curve for the advanced stage.

Figure 4. The Kaplan Meier curve for the T3N0 advanced stage.

Figure 5. The Kaplan Meier curve for the T3N1 advanced stage.

Figure 6. Area under curve for extreme random tree model training.

DRAFT COPY OF THE M

Table 1. Baseline demographic and tumor characteristics of advanced stage tongue cancer patients in the SEER database.

Variables (Definition)	**Total (N = 428) (%)**	**Categorization for machine learning analysis**	**P value and HR Significance**
Year of TSCC diagnosis			
2010	62 (14.5%)	No categorization (Not used in training)	*Not evaluated*
2011	69 (16.1%)		
2012	74 (17.3%)		
2013	70 (16.4%)		
2014	71 (16.6%)		
2015	82 (19.2%)		
Ethnicity			
White	336 (78.5%)	0 = White	$p < 0.18$ HR: 1.14 *Not significant*
Black	52 (12.1%)	1 = Black	
Others	40 (9.4%)	2 = Others (American Indian /AK Native, Asian pacific	
Age at diagnosis			
<40 years old (young)	28 (6.5%)	No categorization (Not used in training)	$p < 0.005$ HR: 1.03 *Significant*
>=40 years old (old)	400 (93.5%)		
Gender			
Male	278 (64.9%)	0 = Male	$p < 0.39$ HR: 0.89 *Not significant*
Female	150 (35.1%)	1 = Female.	
Marital status at the time of diagnosis			
Single (never married)	106 (24.8%)	0 = Single (never married)	$p < 0.02$ HR: 0.74 *Significant*
Married	195 (45.6%)	1 = Married	
Divorced	58 (13.6%)	0 = Divorced	
Widowed	60 (14.0%)	0 = Widowed	
Separated	9 (2.1%)	0 = Separated	
Unmarried/Domestic Partner	0 (0.0%)	0 = Unmarried/ Domestic Partner	
AJCC 7th edition, T stage			
AJCC T3	428 (100.0%)	T3 = 3	Not evaluated
AJCC 7th edition, N stage			
AJCC N0; No regional lymph node metastasis	273 (63.8%)	N0 = 0	$p < 0.005$ HR: 1.49 *Significant*
AJCC N1; Single regional lymph node metastasis	142 (33.2%)	N1 = 1	
AJCC N3; Advanced neck metastases	13 (3.0%)	N3 = 3	
AJCC 7th edition, M stage			
AJCC M0; No distant metastasis	418 (97.7%)	M0 = 0	*Not evaluated*
AJCC M1; Presence of distant metastasis	10 (2.3%)	M1 = 1	
Grade Differentiation of tumor			
Grade I: Well differentiated	83 (19.4%)	Grade I = 1	$p < 0.75$ HR: 1.03 *Not significant*
Grade II: Moderately differentiated	255 (59.6%)	Grade II = 2	
Grade III: Poorly differentiated	90 (21.0%)	Grade III = 3	
Grade IV: Undifferentiated	0 (0.0%)	Grade IV = 4	
Treatment options			
Only surgery (**Sx**)	138 (32.2%)	1 = Only Sx	*Significant* ($p < 0.005$)
Surgery followed by postoperative radiotherapy (**Sx+RT**)	194 (45.3%)	1 = Sx + RT	*Not significant*
Surgery followed by postoperative chemoradiotherapy (**Sx+CRT**)	72 (16.8%)	1 = Sx + CRT	*Significant* ($p < 0.005$)
None of the available treatments	24 (5.6%)	0 = None	*Not significant*
Status at the end of follow-up			
Alive	165 (38.6%)	0 = Alive	*Target parameter*
Dead	263 (61.4%)	1 = Dead	

DRAFT COPY OF THE MANUSCRIPT

Table 2. Performance metrics of extreme random tree ML algorithm

Metrics	Values			
Accuracy	Accuracy 86.0%	Balanced_accuracy 82.1%	Weighted_accuracy 85.3%	
F1 score	F1 score 0.866	F1 score_weighted 0.864	F1 score_micro 0.859	F1 score_macro 0.818
AUC	AUC_weighted 0.914	AUC_micro 0.923	AUC_macro 0.914	
Other metrics	Sensitivity coefficient 0.80	Specificity 0.90	Precision 0.84	Mathew's correlation 0.64

Tables

Table 1. Baseline demographic and tumor characteristics of advanced stage tongue cancer patients in SEER database.

Table 2. Performance metrics of extreme random tree machine learning model

DRAFT COPY OF THE MANUSCRIPT

References

[1] Marur S, Forastiere AA. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. Mayo Clin Proc. 2016;91:386–396.

[2] Kantola S, Parikka M, Jokinen K, et al. Prognostic factors in tongue cancer – relative importance of demographic, clinical and histopathological factors. Br J Cancer. 2000;83:614–619.

[3] Mroueh R, Haapaniemi A, Saarto T, et al. Non-curative treatment of patients with oral tongue squamous-cell carcinoma. European Archives of Oto-Rhino-Laryngology. 2019;276:2039–2045.

[4] 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.

[5] Mroueh R, Haapaniemi A, Grénman R, et al. Improved outcomes with oral tongue squamous cell carcinoma in Finland: Oral tongue carcinoma in Finland. Head & Neck. 2017;39:1306–1312.

[6] 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.

[7] 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.

[8] 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.

[9] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn. 2006;63:3–42.

[10] Brownlee J. Dynamic Ensemble Selection (DES) Classification [Internet]. 2021 [cited 2022 Aug 1]. Available from: https://machinelearningmastery.com/dynamic-ensemble-selection-in-python/.

[11] Kim TH, Cha I-H, Choi EC, et al. Postoperative Concurrent Chemoradiotherapy Versus Radiotherapy Alone for Advanced Oral Cavity Cancer in the Era of Modern Radiation Techniques. Front Oncol. 2021;11:619372.

[12] Heinonen T, Loimu V, Saarilahti K, et al. End-of-life care pathway of head and neck cancer patients: single-institution experience. Eur Arch Otorhinolaryngol. 2018;275:545–551.

[13] Rodrigues G, Warde P, Pickles T, et al. Pre-treatment risk stratification of prostate cancer patients: A critical review. cuaj. 2012;6:121–127.

DRAFT COPY OF THE MANUSCRIPT

[14] Seoane-Romero Jm, Vazquez-Mahia I, Seoane J, et al. Factors related to late stage diagnosis of oral squamous cell carcinoma. Med Oral. 2012;e35–e40.

[15] Ansarin M, De Berardinis R, Corso F, et al. Survival Outcomes in Oral Tongue Cancer: A Mono-Institutional Experience Focusing on Age. Front Oncol. 2021;11:616653.

[16] 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.

[17] Sun W, Qiu Z, Tan W, et al. The influence of marital status on survival in patients with oral tongue squamous cell carcinoma. Oncotarget. 2017;8:82092–82102.

[18] Boonpoapichart S, Punyavong P, Jenwitheesuk K, et al. Significant Prognostic Factors Influencing the Survival Difference of Oral Tongue Squamous Cell Carcinoma. Plastic and Reconstructive Surgery - Global Open. 2021;9:e3889.

[19] Chinn SB, Myers JN. Oral Cavity Carcinoma: Current Management, Controversies, and Future Directions. JCO. 2015;33:3269–3276.

[20] Alabi RO, Tero V, Mohammed E. Machine learning for prognosis of oral cancer: What are the ethical challenges? CEUR-Workshop Proceedings. 2020;

DRAFT COPY OF THE MANUSCRIPT