University of Dundee

Comparison of machine learning algorithms for the prediction of five-year survival in oral squamous cell carcinoma

Alkhadar, Huda; Macluskey, Michaelina; White, Sharon; Ellis, Ian; Gardner, Alexander

Published in:
Journal of Oral Pathology and Medicine

DOI:
[10.1111/jop.13135](https://doi.org/10.1111/jop.13135)

Publication date:
2021

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Alkhadar, H., Macluskey, M., White, S., Ellis, I., & Gardner, A. (2021). Comparison of machine learning algorithms for the prediction of five-year survival in oral squamous cell carcinoma. *Journal of Oral Pathology and Medicine*, 50(4), 378-384. <https://doi.org/10.1111/jop.13135>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

DR MICHAELINA MACLUSKEY (Orcid ID : 0000-0002-2734-0537)

DR ALEXANDER GARDNER (Orcid ID : 0000-0001-8618-9622)

Article type : Original Article

Comparison of machine learning algorithms for the prediction of five-year survival in oral squamous cell carcinoma.

Running Head: Predicting oral cancer survival

Huda Alkhadar¹, Michaelina Macluskey², Sharon White², Ian Ellis¹, Alexander Gardner^{3*}

1 - Unit of Cell and Molecular Biology, Dundee Dental School, University of Dundee, Dundee, UK

2 - Department of Oral Surgery, Medicine and Pathology, Dundee Dental School, University of Dundee, Dundee, UK

3 - Department of Restorative Dentistry, Dundee Dental School, University of Dundee, Dundee, UK

***Corresponding author:** AGardner001@dundee.ac.uk , Floor 8, Dundee Dental Hospital and School, 2 Park Place, Dundee, DD1 4HR.

Author Contributions:

Huda Alkhadar: Conceptualization, Data curation, Investigation, Software, Validation, Writing-original draft, Writing-review & editing

Michaelina Macluskey: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Supervision, Validation, Writing-original draft, Writing-review & editing

Sharon White: Investigation, Supervision, Validation, Writing-original draft, Writing-review & editing

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/JOP.13135](#). This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

This article is protected by copyright. All rights reserved

Ian Ellis: Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Writing-original draft, Writing-review & editing

Alexander Gardner: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing-original draft, Writing-review & editing

Acknowledgments and funding:

None to declare

Word count: 2648

Abstract

Background/Aim: Machine learning analyses of cancer outcomes for oral cancer remain sparse compared to other types of cancer like breast or lung. The purpose of the present study was to compare the performance of machine learning algorithms in the prediction of global, recurrence-free five-year survival in oral cancer patients based on clinical and histopathological data.

Methods: Data was gathered retrospectively from 416 patients with oral squamous cell carcinoma. The dataset was divided into training and test dataset (75:25 split). Training performance of five machine learning algorithms (Logistic regression, K-nearest neighbours, Naïve Bayes, Decision tree and Random forest classifiers) for prediction was assessed by k-fold cross validation. Variables used in the machine learning models were age, sex, pain symptoms, grade of lesion, lymphovascular invasion, extracapsular

extension, perineural invasion, bone invasion and type of treatment. Variable importance was assessed and model performance on the testing data was assessed using receiver operating characteristic curves, accuracy, sensitivity, specificity and F1 score.

Results: The best performing model was the Decision tree classifier, followed by the Logistic Regression model (accuracy 76% and 60%, respectively). The Naïve Bayes model did not display any predictive value with 0% specificity.

Conclusions: Machine learning presents a promising and accessible toolset for improving prediction of oral cancer outcomes. Our findings add to a growing body of evidence that Decision tree models are useful in models in predicting OSCC outcomes. We would advise that future similar studies explore a variety of machine learning models including Logistic regression to help evaluate model performance.

Introduction

Oral cancer is the sixth most common global malignancy, with an annual incidence of more than 300,000 cases reported worldwide (1). The term oral cancer is synonymous with one histological type, oral squamous cell carcinoma (OSCC), accounting for more than 90% of all head and neck malignancies (2). There were 3,744 people diagnosed with OSCC in 2016 in the United Kingdom, with higher incidence rates in Scotland than other parts of the UK (3). Despite the progress in investigating the pathobiological mechanism and advances in the approach to treatment of OSCC, the prognosis has unfortunately not improved over the last decades, and OSCC is still associated with a high rate of recurrence and a poor five-year survival rate of less than 50% (4). The spread potential of OSCC to surrounding tissues, lymphatics/vascular circulation and nerves is the main reason for the high morbidity and mortality rates of OSCC (5). Surgery is the most well-established mode of initial definitive treatment for a majority of OSCC. However, radiotherapy with or without chemotherapy is employed in conjunction with surgery as post-operative treatment in the patients with advanced-stage cancer (6). Enhanced understating of five-year survival prediction may help guide clinicians and patient on decisions such as choice of treatment, ultimately improving patient outcomes.

Machine learning is an application of artificial intelligence that encompasses a variety of algorithmic techniques whereby a computer can be trained to identify patterns within large complex datasets. The potential applications for machine learning approaches in healthcare to guide clinicians and improve outcomes are increasingly being recognised (7). With respect to cancer, early appreciation of the potential for machine learning to aid risk assessment, diagnosis and prognosis has seen the field grow considerably over the last decade (8). A depiction of the growth in machine learning research in cancer is shown in Figure 1. Despite oral cancer constituting around 2.1% of all cancers globally (9), research outputs applying machine learning to oral cancer appear to consist of only 1% of total outputs combining machine learning and cancer (98 studies out of 9251). It is therefore reasonable to consider machine learning applications with respect to oral cancer as a relatively preliminary field of research.

Kourou *et al.* present a good overview of machine learning application in cancer (10). The current literature can be divided in several ways. Firstly, the outcome being assessed varies from diagnostic prediction, risk assessment (which may include malignant transformation in pre-cancerous lesions), prediction of recurrence and prediction of survival. Kourou *et al.* did not cite any risk-assessment studies for oral cancer, however this is an area that has recently been explored with promising results (11). Cancer machine learning studies may also be divided by the type of input data used, which varies

considerably. Data from clinical, epidemiological, histopathological, imaging, genomic and other molecular pathological analyses are all used as inputs for machine learning models. Furthermore, a range of machine learning algorithms are currently employed in cancer research including Naïve Bayes, Artificial Neural Network, Decision tree, Random forest, Logistic Regression, Linear Discriminant Analysis, Support Vector Machine and K-nearest neighbour models (12).

Due to the heterogeneity of input data and model selection, model performance also varies considerably. With respect to cancer survival, model accuracies between 71% and 93% have been reported for breast cancer (13, 14), whereas accuracies ranging between 71% and 98% have been reported for oral cancer (15, 16). A final caveat is that, until recently, there appears to be a large discrepancy in sample size with breast cancer commonly featuring tens or hundreds of thousands of samples whereas oral cancer studies featured as few as 31 participants (17). The study by Karadaghy *et al.* appears to be the largest machine learning study on oral cancer survival, with 30,000 participants included (18).

The aim of the present study was to evaluate the performance of different machine learning algorithms in global, recurrence-free five-year survival in patients with OSCC. Prediction of local recurrence was also conducted. A secondary aim was to evaluate the clinical and histopathological features deemed most important to the various models.

Materials and methods

Data acquisition and ethical approval

This study was approved by NHS Tayside, Scotland (Caldicott/CSAppHA1350), and Tayside Medical Science Centre Research and Development (14/ES/0015(HIC), allowing retrospective analysis of patient clinical records. Inclusion criteria were a diagnosis of primary OSCC, with follow-up for at least five years or until death. Exclusion criteria were those under 18 years of age or histological findings other than primary OSCC or evidence of oropharyngeal SCC. In total, 430 patient records were deemed suitable for inclusion.

Variable pre-processing

Raw data was compiled in Microsoft excel and read into Python 3.8., (Python Software Foundation, <https://www.python.org>). Survival at 5 years was selected as the target (y) variable and x variables represented histological and patient findings. Non-numeric variables were encoded into categorical variables. Post code were encoded into quintiles of social deprivation based on the Scottish Index of Multiple Deprivation (SIMD) sourced from the Scottish Government website. A summary of the variables is presented in Table 1. Variables with significant numbers of missing data, where information was absent

from patient records were dropped from the final analysis. Dropped variables were smoking history, alcohol consumption, post code, TNM score and depth of invasion. Variables used in the machine learning models were age, sex, pain symptoms (Pain), grade of lesion (Grade), lymphovascular invasion (LVI), extracapsular extension (Extension), perineural invasion (PNI), bone invasion (BI), choice of treatment (Treatment). A small number of cases where data for lesion grade was missing were dropped, giving a total number of participants as 416. A secondary analysis on a subset (n=150) of the data where TNM score and excision margin depth, where available, was performed.

Table 1: A summary of the variables used by the machine learning classification models. * indicates the variable was dropped due to missing data.

Evaluation of machine learning algorithms

Five-year survival prediction was assessed using several machine learning algorithms for classification that are readily available in Python 3.8. These were Logistic regression, K-nearest neighbours, Naïve Bayes, Decision tree and Random forest classifiers. The dataset was first divided into training (75%) and test (25%) data. Imbalance of target variables was then corrected using majority class undersampling via the Python package imbalanced-learn (19). Models were trained using the training dataset, and performances were evaluated by k-fold cross validation where k was set to 5. Model training accuracy was assessed by ANOVA in GraphPad Prism 8 (La Jolla, CA). Trained models were then assessed for their performance on the test dataset. Receiver Operating Characteristic (ROC) curves were plotted for each model and the concordance index (Area Under Curve/AUC) was calculated. Sensitivity (recall), specificity, accuracy and F1 score were also calculated. Where available, relative importance of the variables was assessed for the models.

The same approach was applied to evaluating the models in the prediction of local recurrence. A subset of data was created to include additional variables (TNM and excision margin). TNM score was simply coded as 1-4, whereas excision margin used a 4mm cut-off with a binary coding (<4mm = 0, >4mm = 1).

Results

Model performance on training data

Cross validation of performance on the training data set showed all models displayed relatively similar performance accuracy (Figure 2). Performance accuracy ranged from 56% for the Naïve Bayes classifier to 72% for the Decision tree and Random forest classifier. Comparison of mean performance accuracy

following cross-validation by one-way ANOVA revealed no significant differences between any of the models ($p=0.11$).

Model performance on test data

A summary of model performance measured by ROC-AUC score, accuracy, precision, sensitivity, specificity and F1 score is presented in Table 2.

Table 2: Summary of model performance metrics for the prediction of global, recurrence-free 5-year survival.

The best performing model by all metrics on the test data was the Decision tree model (ROC-AUC score 0.77, accuracy 76%). This was followed by the Logistic Regression model (ROC-AUC score 0.69, accuracy 60%). K-nearest neighbours and Random forest models performed comparatively poorly, and the Naïve Bayes model was essentially a failed classifier with 0% specificity.

Inspection of variable importance

Relative variable importance for the Logistic regression and Decision tree classifiers are presented in Figure 3. The Decision tree classifier displayed very similar variable importance to the Random forest classifier and K-nearest neighbours and Naïve Bayes classifiers do not have inbuilt feature importance parameters. As can be seen in Figure 3, the Logistic regression model assigned high importance to PNI, followed by treatment choice, LVI, extracapsular extension and age. The Decision tree model, by contrast found age as the most important variable by a considerable margin, with site appearing as the next most important variable. Grade, pain PNI and sex showed comparable importance. BI was seen as an unimportant feature, likely reflecting the low proportion of cases featuring evidence of bone invasion.

Model performance in prediction of recurrence

Results for the prediction of recurrence are shown in Supplementary Figure 1 and Supplementary Table 1. ANOVA found no significant differences in the training performance of the models ($p=0.52$). The best performing predictor was the logistic regression model (accuracy = 65%), although model performance was generally poor with low predictive value.

Model performance on subset data including TNM and excision margin

Results for the prediction of global recurrence-free five-year survival are shown in Supplementary Figure 2 and Supplementary Table 2. On the training data, the Naïve Bayes model significantly worse than all other models ($p < 0.05$). Test performance was highest for the Decision tree classifier, followed by Logistic regression and the K-nearest neighbours model.

Discussion

The purpose of the present study was to evaluate the performance of machine learning algorithms and important variables in the prediction of global recurrence-free five-year survival in OSCC patients.

Karadaghy *et al.* recently used machine learning to predict five-year survival in OSCC in 30,000 patients, with reported accuracy of 71% and ROC-AUC score of 0.8 with a Decision Forest model (18). It would seem reasonable to consider these metrics as a benchmark given the robust nature of the study. The best performing model, Decision tree classifier from the present study displayed comparable accuracy (76%) and ROC-AUC score (0.77) to Karadaghy *et al.* Other recent studies have found Decision tree algorithms to perform strongly when applied to predicting OSCC outcomes (20, 21).

Interestingly, we found that Logistic regression was the second-best predictor of global recurrence-free survival, and the strongest predictor of recurrence. Whether Logistic regression can truly be considered a machine learning approach is a topic of some debate in literature, although it is a well-established technique for binary classification. A recent meta-analysis of healthcare literature found that Logistic regression typically performs as well as other machine learning algorithms (22). With respect to cancer outcomes, Logistic regression has been found to perform comparably to other machine learning models in prediction of breast cancer survival (23), as well as in the prediction of nodal metastasis in OSCC (15). Conversely, prediction of five-year survival in OSCC has found that Logistic regression performed considerably poorer than Decision tree and Artificial neural networks (63% accurate compared to ~80% accuracy) (24).

Analysis of variable importance of the machine learning models provides some considerations for improving prediction in future. The Logistic regression model found PNI and treatment choice to be important variables. These would seem logical predictors of OSCC five-year survival and recurrence frequency has been implicated as a critical variable in similar predictive models (24). Age was considered a crucial variable by our Decision tree and Random forest. Alabi *et al.* also found age as the most predictive variable for oral cancer recurrence using a Boosted Decision tree model (21). Bone invasion and lymphovascular invasion were generally found to be unimportant, likely stemming from the fact that these were only present in a small proportion of the dataset (7 and 14%, respectively). Unfortunately, due

to missing data, factors that may be reasonably expected to be significant in predicting survival such as stage (TNM score), smoking and alcohol consumption and depth of invasion were only performed on a subset of our data in the present analyses. The findings were of generally poor performance; however, this may reflect the reduction in the amount of training data. Furthermore, novel ways of analysing histopathology of OSCC may be useful for future prospective studies. For example, margin-to-thickness ratio, which has recently been explored as prognostic histopathological factors in OSCC (25).

The present study used only clinical and histopathological findings as input variables. Our model performance was generally an improvement on those of Chang *et al.*, who reported AUC scores of 0.39 to 0.7 when analysing only clinical data and histopathological. However, the same study noted considerable improvement in model accuracy with the addition of genomic marker variables, such as *p63* expression (17). Similarly, when predicting oral cancer recurrence Exarchos *et al.* reported that the use of imaging data, tissue genomic and blood genomic data significantly improved their model accuracy above clinical/histopathological variables alone (26). Therefore, future work will encompass the inclusion of additional biochemical findings to evaluate whether these variables improve model performance. For example, in the same dataset the presence of nerve growth factor and tyrosine kinase A approached significance in survival analysis by Kaplan-Meier curves, hence could represent important features in a machine learning model (27), as seen for PNI in the present study. Equally, the balance between model performance and ease of attainment for input data must be considered. Data for the present study would be routinely available after clinical examination and a biopsy, whereas the inclusion of genomic data would require microarray or sequencing analyses. Such analyses may not be conducted routinely and have time and cost implications.

A limitation of the present work is the fact that the target variable, global recurrence-free five-year survival, is a binary outcome and does not consider time-to-event. Time-to-event analyses have recently been incorporated into oral cancer survival prediction and have demonstrated promising results, albeit in a relatively small dataset for such an analysis (28). Evidence suggests that in the prediction of survival for different types of cancer, machine learning model performance decreased as survival time is extended (29). Furthermore, our dataset does not consider global versus disease specific mortality. There may some debate as to the effect OSCC would have on all-cause mortality rates, particularly considering the aggressive nature of adjuvant treatments like chemo and radiotherapy. A recent study in a Norwegian cohort found comparable overall survival and disease specific survival rates (62% vs 56%) in patients undergoing treatment for OSCC (30). Our dataset was also characterised by a very low (<2%) proportion of cases being treated with additional chemotherapy or radiotherapy. Another limitation to the present

study is the lack of an external validation dataset. Our approach to using a training-test split seems to be the most widely used in machine learning studies of OSCC. The collection of a prospective validation cohort as done recently by Alabi *et al.* (21), represents an important future direction.

This study found the best performance by all metrics for global, recurrence-free OSCC five-year survival prediction was with the Decision tree model. Logistic regression was the second-best performing model although there was a decrease in all performance metrics. The Naïve Bayes model failed to classify survival in our dataset. Based on our findings we would suggest researchers undertaking similar future prospective studies to consider incorporating Decision tree and Logistic regression alongside a range of other machine learning algorithms. Careful consideration of input variables, making full use of the clinical and histopathological data available, is also likely to help improve overall model performance.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. ETTINGER KS, GANRY L, FERNANDES RP Oral cavity cancer. *Oral and Maxillofacial Surgery Clinics* 2019; 31: 13-29.
2. SPEIGHT P, FARTHING P The pathology of oral cancer. *British Dental Journal* 2018; 225: 841-847.
3. CONWAY D, PURKAYASTHA M, CHESTNUTT I The changing epidemiology of oral cancer: definitions, trends, and risk factors. *British dental journal* 2018; 225: 867-873.
4. RIVERA C, OLIVEIRA AK, COSTA RAP et al. Prognostic biomarkers in oral squamous cell carcinoma: a systematic review. *Oral oncology* 2017; 72: 38-47.
5. AMIT M, NA'ARA S, GIL Z Mechanisms of cancer dissemination along nerves. *Nature reviews Cancer* 2016; 16: 399.
6. D'SILVA N, GUTKIND J Oral cancer: Integration of studies for diagnostic and therapeutic precision. *Advances in dental research* 2019; 30: 45-49.
7. OBERMEYER Z, EMANUEL EJ Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine* 2016; 375: 1216.
8. CRUZ JA, WISHART DS Applications of machine learning in cancer prediction and prognosis. *Cancer informatics* 2006; 2: 117693510600200030.
9. LEVINE R, STILLMAN-LOWE C. Oral Cancers. In. *The Scientific Basis of Oral Health Education*. Springer, 2019; 57-62.

- Accepted Article
10. KOUROU K, EXARCHOS TP, EXARCHOS KP et al. Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal* 2015; 13: 8-17.
 11. WANG X, YANG J, WEI C et al. A personalized computational model predicts cancer risk level of oral potentially malignant disorders and its web application for promotion of non-invasive screening. *Journal of Oral Pathology & Medicine* 2020; 49: 417-426.
 12. NINDREA RD, ARYANDONO T, LAZUARDI L et al. Diagnostic accuracy of different machine learning algorithms for breast cancer risk calculation: a meta-analysis. *Asian Pacific journal of cancer prevention: APJCP* 2018; 19: 1747.
 13. DELEN D, WALKER G, KADAM A Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial intelligence in medicine* 2005; 34: 113-127.
 14. PARK K, ALI A, KIM D et al. Robust predictive model for evaluating breast cancer survivability. *Engineering Applications of Artificial Intelligence* 2013; 26: 2194-2205.
 15. BUR AM, HOLCOMB A, GOODWIN S et al. Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma. *Oral oncology* 2019; 92: 20-25.
 16. ROSADO P, LEQUERICA-FERNÁNDEZ P, VILLALLAÍN L et al. Survival model in oral squamous cell carcinoma based on clinicopathological parameters, molecular markers and support vector machines. *Expert systems with applications* 2013; 40: 4770-4776.
 17. CHANG S-W, ABDUL-KAREEM S, MERICAN AF et al. Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods. *BMC bioinformatics* 2013; 14: 170.
 18. KARADAGHY OA, SHEW M, NEW J et al. Development and assessment of a machine learning model to help predict survival among patients with oral squamous cell carcinoma. *JAMA Otolaryngology–Head & Neck Surgery* 2019; 145: 1115-1120.
 19. LEMAÎTRE G, NOGUEIRA F, ARIDAS CK Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *The Journal of Machine Learning Research* 2017; 18: 559-563.
 20. CHU CS, LEE NP, ADEOYE J et al. Machine learning and treatment outcome prediction for oral cancer. *Journal of Oral Pathology & Medicine* 2020.
 21. 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* 2020; 136: 104068.

22. CHRISTODOULOU E, MA J, COLLINS GS et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of clinical epidemiology* 2019; 110: 12-22.
23. CHAO C-M, YU Y-W, CHENG B-W et al. Construction the model on the breast cancer survival analysis use support vector machine, logistic regression and decision tree. *Journal of medical systems* 2014; 38: 106.
24. TSENG W-T, CHIANG W-F, LIU S-Y et al. The application of data mining techniques to oral cancer prognosis. *Journal of medical systems* 2015; 39: 59.
25. HUANG CY, LIN YS, KANG BH et al. Log margin-to-thickness ratio improves disease-specific survival prediction in oral cancer: A single cancer centre database. *Clinical Otolaryngology* 2019; 44: 63-69.
26. EXARCHOS KP, GOLETIS Y, FOTIADIS DI A multiscale and multiparametric approach for modeling the progression of oral cancer. *BMC medical informatics and decision making* 2012; 12: 136.
27. ALKHADAR H, MACLUSKEY M, WHITE S et al. Perineural invasion in oral squamous cell carcinoma: Incidence, prognostic impact and molecular insight. *Journal of Oral Pathology & Medicine* 2020.
28. KIM DW, LEE S, KWON S et al. Deep learning-based survival prediction of oral cancer patients. *Scientific reports* 2019; 9: 1-10.
29. GUPTA S, TRAN T, LUO W et al. Machine-learning prediction of cancer survival: a retrospective study using electronic administrative records and a cancer registry. *BMJ open* 2014; 4.
30. BJERKLI I-H, JETLUND O, KAREVOLD G et al. Characteristics and prognosis of primary treatment-naïve oral cavity squamous cell carcinoma in Norway, a descriptive retrospective study. *Plos one* 2020; 15: e0227738.

Variable	Number of cases	Coding and distribution (No., %)
Survival (5 years)	416	1 = survived (125, 30%) 0 = deceased (291, 70%)
Age	416	Kept as continuous variable
Sex	416	1 = Male (231, 56%) 0 = Female (185, 44%)
Pain symptoms	416	1 = Present (77, 19%) 0 = Absent (339, 81%)
Site of lesion	416	0 = Tongue (165, 40%) 1 = Floor of mouth (94, 22.5%) 2 = Tongue and floor of mouth (17, 4%) 3 = Other (140, 34.5%)
Grade of tumour	416	0 = Well (60, 14%) 1 = Well to moderate, Moderate (245, 59%) 2 = Poor (111, 27%)
Perineural invasion	416	1 = Present (70, 17%) 0 = Absent (346, 83%)
Lymphovascular invasion	416	1 = Present (60, 14%) 0 = Absent (356, 86%)
Bone invasion	416	1 = Present (30, 7%) 0 = Absent (386, 93%)
Extracapsular extension	416	1 = Present (52, 13%) 0 = Absent (364, 87%)
Treatment	416	0 = None (123, 30%) 1 = Surgery, no neck dissection (108, 26%) 2 = Surgery with neck dissection (185, 44%)
Recurrence	416	0 = Absent (315, 76%) 1 = Present (101, 24%)
Smoking history*	218	NA
Alcohol consumption*	156	NA
TNM score*	288	NA
Depth of lesion*	257	NA

Post code*	328	NA
------------	-----	----

Model	Model performance on test data							
	ROC-AUC score	Accuracy (%)	Sensitivity/True positive rate (%)	Specificity/True negative rate (%)	Precision (%)	F1 Score	False positive rate (%)	False negative rate (%)
Logistic regression	0.69	60	74	50	53	0.62	50	26
K-nearest neighbours	0.65	59	70	50	51	0.59	50	30
Naïve Bayes classifier	0.5	43	100	0	43	0.60	100	0
Decision tree classifier	0.77	76	81	72	69	0.75	28	19
Random forest classifier	0.69	55	74	42	49	0.59	58	26

Growth of publications combining machine learning and cancer

jop_13135_f1.jpg

jop_13135_f2.jpg

jop_13135_f3.jpg