Prognosis prediction in head and neck cancer

Conor Whitley

The University of Liverpool, Department of Physics

Abstract

1   Introduction

Oral squamous cell carcinoma (OSCC) is the sixth most common form of cancer worldwide with a recent increase in incidence reported [1]. The majority of head and neck cancers are squamous cell carcinomas that originate in the upper aerodigestive epithelium and are typically associated with exposure to carcinogens present in tobacco and alcohol [2]. The regions in which head and neck tumours typically develop are anatomically complex and play a vital physiological role in the patient; early diagnosis and selection of appropriate treatment is needed to ensure patient longevity and retention of vital organ function.

Many previous studies [3–9] have investigated the viability of a range of prognositic biomarkers for head and neck/ oropharyngeal cancer (is oropharyngeal too different?) with varying degrees of success. The discovery of effective prognostic biomarkers has been difficult, and thus far has largely focused on immunohistochemistry techniques. Magnetic Resonance Imaging (MRI) has also been utilised [10–12] to estimate various clinical biomarkers in a non-invasive manner. However MRI based techniques often either overestimate or underestimate when attempting to quantify biomarkers when they are evaluated using direct measurements using pathological staging sections [13, 14].

In cases where lower biological aggression is demonstrated, a de-escalation of therapy is possible [15]. Identification of these cases is paramount to minimising the adverse effects of treatment, and improving patient outcomes. Previous work [16–18] has hypothesized that tumours which may be responsive to novel therapeutic treatment may carry a distinct molecular fingerprint; the identification of which may allow for screening of patients towards appropriate treatment.

This work explores the potential efficacy of fourier transform infra-red (FTIR) microscopy in combination with a known prognositic biomarker: α-Smooth Muscle Actin (ASMA) expression, as a method of identifying cases which may be appropriate for therapeutic treatment. Previous work [4–6] has explored the efficacy of ASMA and SERPINE1 [6] as predictive variables for

extra capsular spread (ECS), an important prognostic biomarker for OSCC. FTIR microscopy images taken of the same sample set will be used in order to perform a comparison to established biomarkers.

FTIR microscopy is a well established technique and has been utilised to investigate a range of biomedical applications in recent years [19–21]. Due to its ability to access chemical information present within the sample; analysis of FTIR microscopy data is able to lend insight into a range of label-free discrimination tasks. Zawlik et.al [22] investigated FTIR coupled with principal component analysis (PCA) to investigate the efficacy of chemotherapy in triple-negative breast cancer. They determined that it was possible to monitor changes in the biochemical composition of the tissue in order to monitor the effectiveness of received treatment.

Butler et.al [23] have undertaken development of a high-throughput ATR-FTIR based instrument for use in biofluid assays. Their work concluded that it was possible to utilise IR spectroscopy to triage brain cancer using biofluid samples with a sensitivity and specificity of 93.2% and 92.8% respectively.

2 Materials and Methods

Tissue preparation

The dataset is comprised of FTIR spectra taken from primary tumour sites of 28 patients. All samples were formalin-fixed-paraffin-embedded (FFPE) tissue biopsies collated into 1mm diameter core tissue micro array (TMA). Four adjacent sections were taken from the TMA, the first and last sections were used to ensure reasonable sample consistency by inspection using a Haematoxylin and Eosin (H&E) staining procedure, and later for tissue annotation. Middle TMA sections were divided, and floated onto CaF_2 disks using a warm water bath. Images of stained sections were scanned using an Aperio CS2 scanner (Leica Biosystems). notation was performed by a maxillofacial pathologist using adjacent H&E stained sections to ensure accurate labelling of tumour regions.
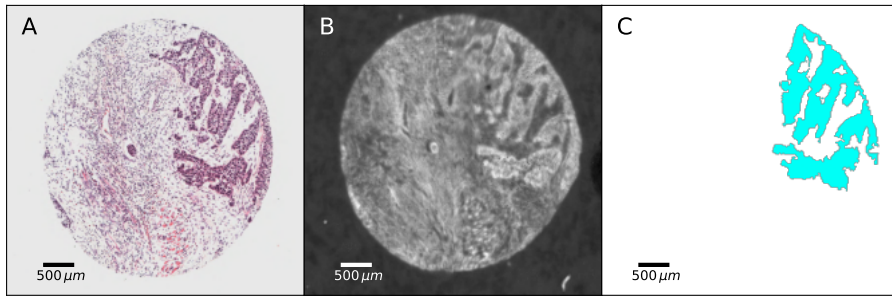

Figure 1: Annotation of FTIR images

Figure 1.A shows an example of an H&E image used by a pathologist to

~~annotate a measured FTIR image (Figure 1.B). The annotation Figure 1.C was then used to extract relevant tumour areas from Figure 1.B.~~

FTIR Microspectroscopy

FTIR micoscopy is a technique which allows for imaging of sample specimens at a multitude of infra-red wavelengths simulataneously. This is achieved through the use of a broadband light source and Michelson interferometer set-up. Individual chemical spectra are then obtained by performing a Fourier transform on the resulting interferogram. Marginal spectral differences in biochemical compounds of interest are typically located in a region known as the fingerprint region ($1000cm^{-1}$-$1800cm^{-1}$). It is differences in these absorption bands which contain information which can be utilised to discriminate between samples of interest.

FTIR measurements of TMA cores were taken at room temperature using a Varian Cary 670-FTIR spectrometer with an attached Varian Cary 620-FTIR microscope produced by Varian (now Agilent Technologies, Santa Clara CA, USA); with a liquid nitrogen cooled 128×128 pixel mercury-cadmium-telluride (MCT) focal plane array with an effective field of view for each pixel of 5.5µm. The sample stage was enclosed in a perspex box and pumped with dry air until a humidity of 1% was achieved into order to mitigate the effects of water contributions on measured IR spectra. Images were acquired at a resolution of $6cm^{-1}$ over a spectral range of $990cm^{-1}$

to $3800 cm^{-1}$ using a co-addition of 128 scans. Attenuator and integration time of the FPA were chosen to gain the maximum signal-to-noise ratio. Background scans were acquired using a blank CaF_2 disk situated within the perspex box before each session of measurements.

ASMA Expression

Previous work [4–6] has shown ASMA expression to be a viable prognostic biomarker. ASMA expression is closely associated with the presence of activated fibroblasts, also known as *myofibroblasts* in tumour associated stroma. The degree of ASMA expression can be interrogated through the use of appropriate chemical stains, and evaluated using an optical microscope.

Data Preprocessing

Annotated areas selected by an oral and maxillofacial pathologist AT with comparison to adjacent H&E slices were used to extract areas of tumour tissue from each FTIR image. Extracted spectra were pre-processed using an open-source extended multiplicative scattering correction (EMSC) code provided by Kohler et.al [24]. Distribution of computing tasks was achieved using the HTCondor computing software package [25].

A number of preprocessing steps were carried out on the dataset before a final analysis was performed using a Logistic Regression (LR) classifier. An

unsupervised quality control check ~~was performed on~~ all data to eliminate anomalous spectra through the use of the multivariate Hotelling's T^2 statistic [26, 27].

Vector normalisation step ~~was carried out on the dataset in order~~ to account for sample thickness. Wavenumber absorbance features ~~were zero centered~~ and variance ~~was scaled~~ to one ~~in order to facilitate a final~~ Principal Component Analysis (PCA). Five principal components were taken ~~to consider only those of greatest variance and to assist convergence of LR classifier parameters.~~

Data Analyisis

Data was analysed using a number of scientific Python packages. The classification power of FTIR ~~as~~ a biomarker was estimated using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and a precision-recall (PR) curve. A number of classification statistics were calculated over a 100 sample subsets drawn with replacement, and stratified by case number in order to obtain an estimate of the variability.

Survival analysis was conducted using a Kaplan Meier (KM) estimator ~~in order to compare the predictive power in a way more appropriate for prognosis prediction.~~ Mean survival functions across all sample subsets were computed to illustrate the prognostic prediction power of the classification.

Spectra were labelled as belonging either to a patient who survived a year beyond the ~~last review date~~ or died within this time frame. A supervised classification training procedure was performed by training a LR classifier on a subset of patient spectra and then validated on a ~~separate set of~~ patients set ~~aside.~~ A sampling without replacement was used to ~~repeatedly~~ draw training and testing sets ~~twenty times to allow for a~~ distribution of statistics ~~to be obtained~~ across the entire dataset.

3   Results

When stratifying the dataset by survival beyond one year 205,701 spectra (68%) belonging to 24 patients died within one year; compared to 96,402 spectra (32%) belonging to 10 patients who survived. Stratifying for survival beyond two years; 151,408 (50%) spectra belonged to 19 patients who died within two years, with 150,695 (50%) spectra belonging to 15 patients who survived beyond two years. Table 1 lists the clinical and pathological characteristics of the patient cohort for both strata.

Table 1: Clinical and pathological characteristic of patient cohort (n=34)

	1 year death	True	False	2 year death	True	False	All
Age (years)							
Mean		70.4	56.3		64.5	57.2	
Range		59-85	29-72		48-85	29-72	
Median		70.5	56.5		64	59	
Gender							
F		0	9		2	7	9
M		9	15		12	12	24
T Stage							
1		0	1		0	1	1
2		4	14		8	10	18
3		0	2		2	0	2
4		4	5		3	6	9
4a		2	2		2	2	4
N Stage							
0		2	5		1	6	7
1		2	5		2	5	7
2a		1	0		1	0	1
2b		4	12		8	8	16
2c		1	2		3	0	3
Pathological Site							
Floor of mouth		2	7		4	5	9
Oropharynx		0	3		0	3	3
Other		5	8		5	8	13
Tongue		3	6		6	3	9

Table 1 shows that males were more prevalent than females in this cohort. The commonest TNM stage characteristics were T2 and T4; with N2b being the most prevalent nodal metastasis status.

Prediction of death within one year

Figure 2 shows the mean ROC and PR curves for models using FTIR data (shown in orange), and FTIR with ASMA data (shown in blue).
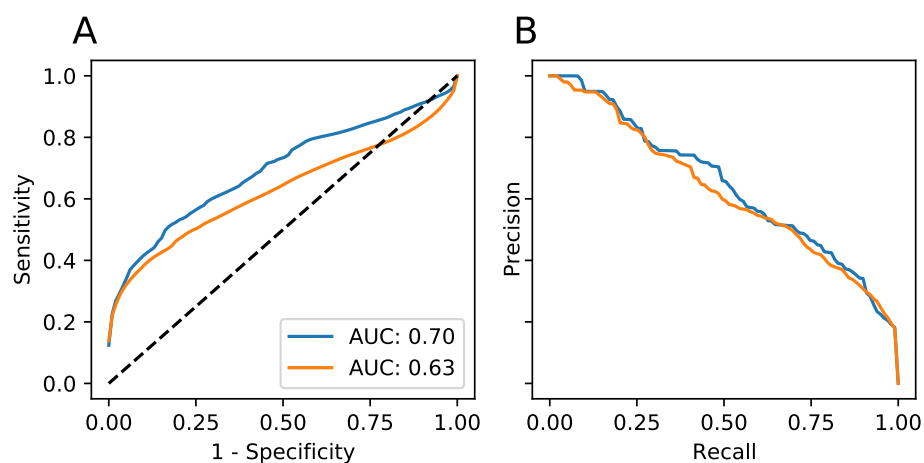

Figure 2: Mean ROC and PR curves for 1 year prediction results. AUC scores are also shown for each model in fig. 2A. FTIR - Orange, FTIR+ASMA - Blue

There is a considerable difference between the mean ROC AUC performance of each model, with the model using both FTIR and ASMA data showing a marked improvement over the model using FTIR data.
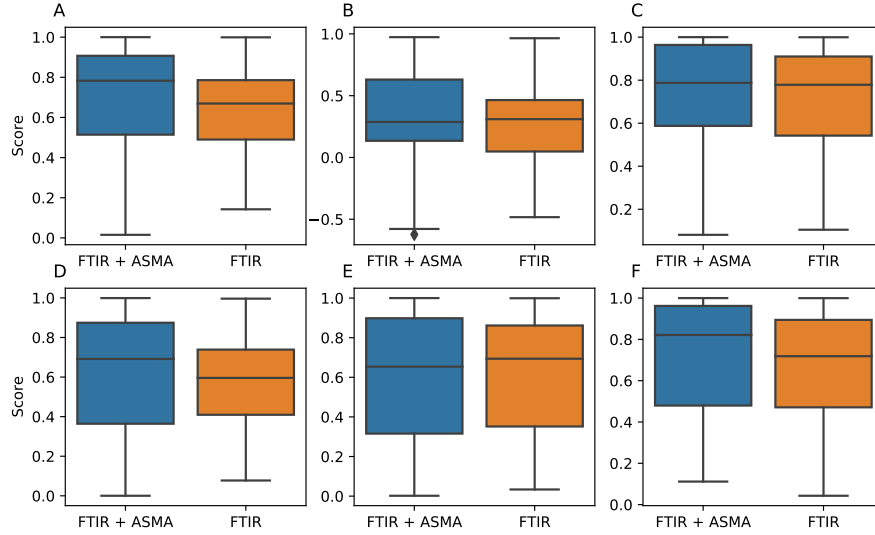
Figure 3: Box plots of classification statistics calculated across all sample subsets. AUC (A), Matthews Correlation Coefficient (MCC) (B), Specificity (C), Sensitivity (D), Positive Predictive Value (PPV) (E), Negative Predictive Value (NPV) (F)

Figure 3 shows boxplots of classification metrics calculated across sample subsets for each model. High classification scores can be achieved across all subsets with a consistently high AUC score demonstrating that good classifier performance is achievable across a wide range of decision thresholds. All other statistics were calculated using an acceptance threshold of 0.5 on the value of the logistic function to determine prediction outcome.
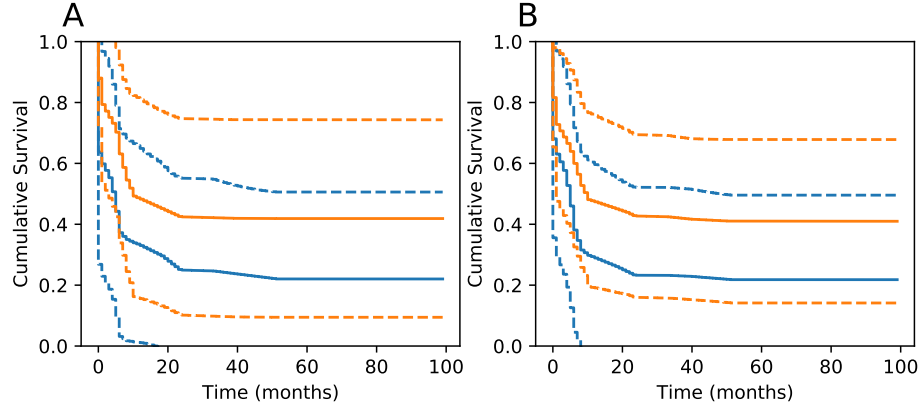
Figure 4: Kaplan-Meier survival curves showing mean survival curves predicted across all sample subsets. FTIR + ASMA (A), FTIR(B). Survival curves shown in orange (blue) indicate patients predicted to have survived beyond (within) one year

Figure 4 shows mean survival curves predicted across each subset with standard deviations depicted with dashed lines in their respective colours. A clear difference in the predicted survival curves demonstrates that the classifier has prognostic value. Each pair of survival curves for each sample split was subject to a log-rank (Mantel-Cox) test to compare overall survival between predicted outcomes. All sample splits were found to have $p < 0.001$ implying that a systematic difference between the two survival curves exists.

Prediction of death within two years

In order to compare with previous results the same process was repeated but with the model attempting to predict the death of a patient within two years. Whilst potentially more clinically useful; any models would likely have to detect biomarkers which may be less well-progressed and thus have a weaker IR signature.

Figure 5 shows the mean ROC and PR curves for models using FTIR data (shown in orange), and FTIR with ASMA data (shown in blue).
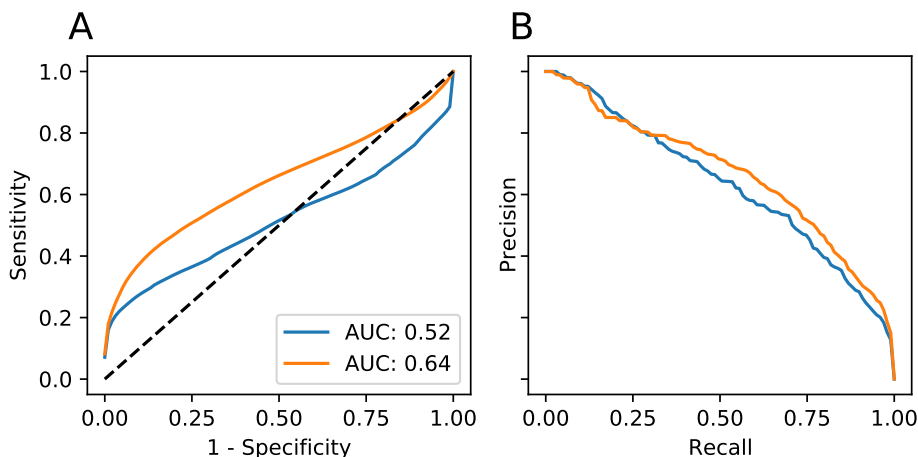

Figure 5: Mean ROC and PR curves for 2 year prediction results.AUC scores are also shown for each model in fig. 5A.FTIR-Orange,FTIR+ASMA - Blue

The classification performance for the prediction of two year survival is substantially less than for a single year. Interestingly the addition of ASMA data seems to be detrimental to classifier performance.
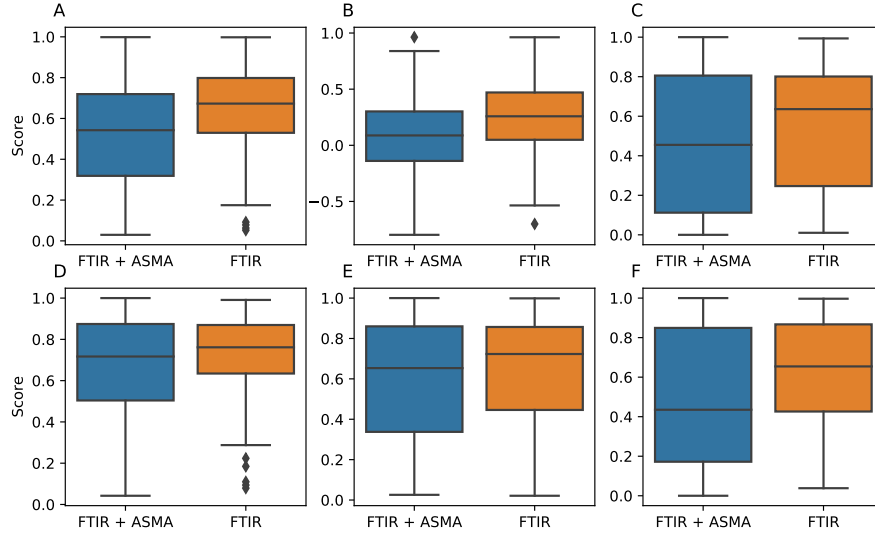
Figure 6: Box plots of classification statistics calculated across all sample subsets. AUC (A), MCC (B), Specificity (C), Sensitivity (D), PPV (E), NPV (F)

Figure 6 shows boxplots of classification metrics calculated across sample subsets. Classification statistics are more modest in comparison to a prediction of death within one year. This is expected due to the lesser performance as indicated by fig. 5. Predictive models using only FTIR information are on average better and vary less than those also using both FTIR and ASMA data as predictive variables.
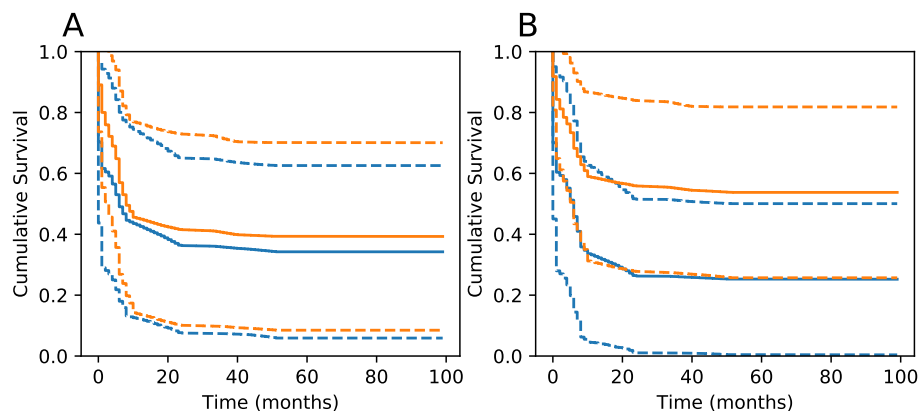
Figure 7: Kaplan-Meier survival curves showing mean survival curves predicted across all sample subsets. FTIR + ASMA (A), FTIR (B). Survival curves shown in orange (blue) indicate patients predicted to have survived beyond (within) two years.

Figure 7A indicates that predictive models using both FTIR and ASMA are on average not effective prognostic biomarkers due to the large degree of overlap between predicted outcomes. In contrast Figure 7B shows a much more useful prognostic prediction as the distribution of survival curves overlap much less. Each pair of survival curves for each sample split was subject to a log-rank (Mantel-Cox) test to compare overall survival between predicted outcomes. All sample splits were found to have $p < 0.001$ implying that a systematic difference between the two survival curves exists.

4 Discussion

The use of FTIR in clinical diagnostics is growing quickly but a limited amount has been aimed towards prognostic predictions. The combination of both FTIR microscopy with techniques more familiar to oncologists such as Immuno Histo Chemical (IHC) staining has the potential to improve prognostic predictive capabilities significantly as shown in this work. The addition of ASMA data increased classification performance considerably for the prediction of death within one year. Although the benefit to prognostic prediction in this case was lesser, the importance of classifying a patient as its relevant prognostic group cannot be understated. Survival curves of predicted risk groups have shown that FTIR and ASMA can be used to construct a prognostically useful tool.

Many other prognostic biomarkers exist [28–30] but a large proportion are still in the "discovery phase" - requiring further study to ascertain prognostic benefit [31]. The use of FTIR within clinical diagnostics likely fits into this category, as many potential barriers facing other potential biomarkers are still present. A relatively small sample set was a key issue facing this study due to the difficulty in acquiring and imaging large numbers of samples. Despite attempts to determine the feasibility of FTIR as a prognostic tool through multiple sampling of the dataset; a larger study would likely need to be conducted in the future to ascertain the true ability to translate to a wider clinical cohort.

5   Conclusion

The use of FTIR in a clinical setting is still in its infancy, however the work covered here shows that it has the potential to be of significant benefit as a prognostic tool. The addition of ASMA information was shown to be beneficial in one case, and demonstrates that additional information from other modalities could lead to the creation of a novel and informative prognostic tool. In order to progress the work covered here a larger patient sample set would be needed to affirm the conclusions reached here. The addition of informaiton from a broader range IHC stains or imaging modalities could further increase the prognostic benefit outlined here and yield powerful tools for clinicians.

References

[1]   Carla Conti et al. "Microimaging FTIR of head and neck tumors. IV". In: *Microscopy Research and Technique* 72.2 (2009), pp. 67–75. ISSN: 1059910X. DOI: `10.1002/jemt.20644`.

[2]   Athanassios Argiris et al. "Head and neck cancer Athanassios". In: *Lancet* 371.9625 (2008), pp. 1695–1709. DOI: `10.1016/S0140-6736(08) 60728-X.Head`.

[3] David Chin et al. "Novel markers for poor prognosis in head and neck cancer". In: *International Journal of Cancer* 113.5 (2005), pp. 789–797. ISSN: 00207136. DOI: 10.1002/ijc.20608.

[4] M. G. Kellermann et al. "Myofibroblasts in the stroma of oral squamous cell carcinoma are associated with poor prognosis [3]". In: *Histopathology* 51.6 (2007), pp. 849–853. ISSN: 03090167. DOI: 10.1111/j.1365-2559.2007.02873.x.

[5] Marilena Vered et al. "Tumor-host histopathologic variables, stromal myofibroblasts and risk score, are significantly associated with recurrent disease in tongue cancer". In: *Cancer Science* 101.1 (2010), pp. 274–280. ISSN: 13479032. DOI: 10.1111/j.1349-7006.2009.01357.x.

[6] J Dhanda et al. "SERPINE1 and SMA expression at the invasive front predict extracapsular spread and survival in oral squamous cell carcinoma". In: August (2014), pp. 2114–2121. ISSN: 0007-0920. DOI: 10.1038/bjc.2014.500.

[7] Chun Ta Liao et al. "Pathological risk factors stratification in pN3b oral cavity squamous cell carcinoma: Focus on the number of positive nodes and extranodal extension". In: *Oral Oncology* 86.September (2018), pp. 188–194. ISSN: 18790593. DOI: 10.1016/j.oraloncology.2018.09.021. URL: https://doi.org/10.1016/j.oraloncology.2018.09.021.

[8]   Corey C. Foster et al. "Dose and Volume De-Escalation for Human Papillomavirus–Positive Oropharyngeal Cancer is Associated with Favorable Posttreatment Functional Outcomes". In: *International Journal of Radiation Oncology Biology Physics* 107.4 (2020), pp. 662–671. ISSN: 1879355X. DOI: `10.1016/j.ijrobp.2020.04.014`. URL: `https://doi.org/10.1016/j.ijrobp.2020.04.014`.

[9]   Gerard Mamelle et al. "Lymph node prognostic factors in head and neck squamous cell carcinomas". In: *The American Journal of Surgery* 168.5 (1994), pp. 494–498. ISSN: 00029610. DOI: `10.1016/S0002-9610(05)80109-6`.

[10]   Christine T. Lwin et al. "Accuracy of MRI in prediction of tumour thickness and nodal stage in oral squamous cell carcinoma". In: *Oral Oncology* 48.2 (2012), pp. 149–154. ISSN: 13688375. DOI: `10.1016/j.oraloncology.2011.11.002`. URL: `http://dx.doi.org/10.1016/j.oraloncology.2011.11.002`.

[11]   Peng Wang et al. "An approach to identify, from DCE MRI, significant subvolumes of tumors related to outcomes in advanced head-and-neck cancer". In: *Medical Physics* 39.8 (2012), pp. 5277–5285. ISSN: 00942405. DOI: `10.1118/1.4737022`.

[12]   Tobias Waech et al. "Measurement variations of MRI and CT in the assessment of tumor depth of invasion in oral cancer: A retrospective

study". In: *European Journal of Radiology* 135 (2021). ISSN: 18727727. DOI: 10.1016/j.ejrad.2020.109480.

[13]   Jenny K. Hoang et al. "Evaluation of cervical lymph nodes in head and neck cancer with CT and MRI: Tips, traps, and a systematic approach". In: *American Journal of Roentgenology* 200.1 (2013), pp. 17–25. ISSN: 0361803X. DOI: 10.2214/AJR.12.8960.

[14]   Ming Hui Mao et al. "Accuracy of magnetic resonance imaging in evaluating the depth of invasion of tongue cancer. A prospective cohort study". In: *Oral Oncology* 91.October 2018 (2019), pp. 79–84. ISSN: 18790593. DOI: 10.1016/j.oraloncology.2019.01.021. URL: https://doi.org/10.1016/j.oraloncology.2019.01.021.

[15]   Conor P. Barry et al. "De-escalation of surgery for early oral cancer-is it oncologically safe?" In: *British Journal of Oral and Maxillofacial Surgery* 51.1 (2013), pp. 30–36. ISSN: 15321940. DOI: 10.1016/j.bjoms.2012.02.014. URL: http://dx.doi.org/10.1016/j.bjoms.2012.02.014.

[16]   Rebekah K. O'Donnell et al. "Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity". In: *Oncogene* 24.7 (2005), pp. 1244–1251. ISSN: 09509232. DOI: 10.1038/sj.onc.1208285.

[17]   Paul Roepman et al. "An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas".

In: *Nature Genetics* 37.2 (2005), pp. 182–186. ISSN: 10614036. DOI: `10.1038/ng1502`.

[18]   D. S. Rickman et al. "Prediction of future metastasis and molecular characterization of head and neck squamous-cell carcinoma based on transcriptome and genome analysis by microarrays". In: *Oncogene* 27.51 (2008), pp. 6607–6622. ISSN: 09509232. DOI: `10.1038/onc.2008.251`.

[19]   Alexandra Sala et al. "Biofluid diagnostics by FTIR spectroscopy: A platform technology for cancer detection". In: *Cancer Letters* 477.December 2019 (2020), pp. 122–130. ISSN: 18727980. DOI: `10.1016/j.canlet.2020.02.020`. URL: `https://doi.org/10.1016/j.canlet.2020.02.020`.

[20]   Vera E. Sitnikova et al. "Breast cancer detection by ATR-FTIR spectroscopy of blood serum and multivariate data-analysis". In: *Talanta* 214.October 2019 (2020), p. 120857. ISSN: 00399140. DOI: `10.1016/j.talanta.2020.120857`. URL: `https://doi.org/10.1016/j.talanta.2020.120857`.

[21]   Daniela Lazaro-Pacheco et al. "Deciphering the structural and chemical composition of breast cancer using FTIR spectroscopy". In: *Applied Spectroscopy Reviews* 0.0 (2020), pp. 1–15. ISSN: 0570-4928. DOI: `10.1080/05704928.2020.1843471`. URL: `https://doi.org/10.1080/05704928.2020.1843471`.

[22] Izabela Zawlik et al. "FPA-FTIR Microspectroscopy for Monitoring Chemotherapy Efficacy in Triple-Negative Breast Cancer". In: *Scientific Reports* 6 (2016), pp. 1–8. ISSN: 20452322. DOI: `10.1038/srep37333`.

[23] Holly J. Butler et al. "Development of high-throughput ATR-FTIR technology for rapid triage of brain cancer". In: *Nature Communications* 10.1 (2019), pp. 1–9. ISSN: 20411723. DOI: `10.1038/s41467-019-12527-5`. URL: `http://dx.doi.org/10.1038/s41467-019-12527-5`.

[24] A. Köhler et al. "Estimating and correcting Mie scattering in synchrotron-based microscopic fourier transform infrared spectra by extended multiplicative signal correction". In: *Applied Spectroscopy* 62.3 (2008), pp. 259–266. ISSN: 00037028. DOI: `10.1366/000370208783759669`.

[25] HTCondor Team, Computer Sciences Department, University of Wisconsin Madison, WI. *HTCondor*. Accessed on 11.05.2021. URL: `https://research.cs.wisc.edu/htcondor/`.

[26] Pitard F.F. "An Introduction to the Theory of Sampling: An Essential Part of Total Quality Management". In: *Comprehensive Chemometrics* (2009), pp. 1–16. URL: `http://www.sciencedirect.com/science/article/pii/B9780444527011000892`.

[27] Wolfgang Karl Härdle and Léopold Simar. *Applied multivariate statistical analysis, fourth edition*. 2015, pp. 1–580. ISBN: 9783662451717. DOI: `10.1007/978-3-662-45171-7`.

[28] J Zapała et al. "Proliferative index activity in oral squamous cell carcinoma : indication for postoperative radiotherapy ?" In: (2014), pp. 1189–1194. DOI: 10.1016/j.ijom.2014.03.013.

[29] Thomas Scholzen and Johannes Gerdes. "The Ki-67 Protein : From the Known and". In: 322.August 1999 (2000), pp. 311–322.

[30] Neelam G Shah et al. "Prognostic significance of molecular markers in oral squamous cell carcinoma : a multivariate analysis". In: December (2009), pp. 1544–1556. DOI: 10.1002/hed.

[31] César Rivera et al. "Prognostic biomarkers in oral squamous cell carcinoma: A systematic review". In: *Oral Oncology* 72 (2017), pp. 38–47. ISSN: 18790593. DOI: 10.1016/j.oraloncology.2017.07.003.