ORAL CARCINOMA

Oral Squamous Cell Carcinomas: A Statistical Study to Determine the Potential Prognostics Factors that Affect the Tumor Size

Wan Muhamad Amir W Ahmad[1), Masitah Hayati Harun[1), Nor Azlida Aleng[2), Farah Muna Mohamad Ghazali[1), Rabiatul Adawiyah Abdul Rohim[1), Nurul Asyikin Nizam Akbar[3)

ABSTRACT

Introduction: **The increase in cancer patients is very downright scary. In 2012, it has been reported about 14.1 million new cancer cases and 8.2 million cancer death because of Oral squamous cell carcinoma (OSCC). Because of that, many researches had been conducted to investigate the factors that have direct and indirect or both associated with the OSCC, including their survival time**

Objective: **The aim of this research paper is to determine the potential clinic pathological prognostic factors in patients who attended Hospital Universiti Sains Malaysia (USM) from 2005 to 2015(based on secondary data)**

Method: **Three methods of statistical analysis will be used which were ordinal regression analysis, correlation path analysis, and decision tree analysis. All the factors that contributed to the growth of tumor size will be used to educate people or stakeholder of how important this factor toward patient's management.**

Results: **From the ordinal regression point of view, age factor, betel quid, and tumor ploidy is the main factor that contributing to the growth of a tumor. Correlation path analysis shows a there is an association between gender, tumor size and tumor size. Using decision tree analysis, three-factor was in the listed as a potential factor towards tumor growths, which are gender, tumor size, and nerve invasion.**

Conclusion: **From all the analysis, it can be concluded that betel quid, age, gender, tumor ploidy, tumor size, and nerve invasion is the potential factor toward the cell carcinoma growth.**

KEY WORDS

oral squamous cell carcinoma (OSCC), clinic pathological and ordinal regression, correlation and decision tree analysis

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is the most common oral cancer, making up 80-90% of all malignant neoplasms of the oral cavity. Although the incidence of oral cancer varies widely the world, it is accepted that oral cavity ranges from the 6th to the 9th most common anatomical location for cancer, mostly dependent on the country (and even certain areas in some countries) and gender of the patients. although this means incidence, it can represent the most common location for cancer in certain regions, especially in southeastern Asia (Johnson, N. W., *et. al*. 2011). According to the latest World Health Organization (WHO) data published in May 2014, oral cancer deaths in Malaysia reached 1,060 or 0.83% of total deaths (WHO, 2014). The main factors of etiological and predisposing of OSCC include smoking, drinking habits, and ultraviolet radiation (specifically for lip cancer), but several other factors such as human papillomavirus (HPV) and *Candida* infections, nutritional deficiencies, and genetic predisposition have been also associated (Marur, S., *et. al*. 2010). OSCC is an adult and elderly disease and its most common clinical aspect is an ulcerated lesion with a necrotic central area surrounded by elevated rolled borders (Neville, B. W., and Day, T. A., 2002). Understanding the factors that cause cancer will contribute to the prevention of disease. Age is frequently named as a risk factor for oral cancer, as historically it occurs in those over the age of 40. The diagnostic age of the patients may indicate a component of time in the biochemical or biophysical processes of aging cells that allows malignant transformation, or perhaps, immune system competence diminishes with age. The very recent data (through 2008-2011)

Received on December 24, 2019 and accepted on March 11, 2020
1) School of Dental Sciences, Health Campus, Universiti Sains Malaysia (USM)
16150 Kubang Kerian, Kelantan, Malaysia
2) Faculty of Ocean Engineering Technology and Informations, Universiti Malaysia Terengganu (UMT)
Terengganu, Malaysia
3) Department of Hematology &Transfusion Medicine Unit, Health Campus, Universiti Sains Malaysia (USM)
16150 Kubang Kerian, Kota Bharu, Kelantan, Malaysia
Corresponding to: Wan Muhamad Amir W Ahmad
(e-mail: wmamir@usm.my)

Table 1: Description of data among oral squamous cell carcinoma patients

Variable	Code	Description
Tumour_size	Y	Size of tumor 0 = t > 4 cm, 1 = 2 cm < t > 4 cm, t <= 2 cm
Age	x_1	Age of oral squamous cell carcinoma patients
Betel_quid	x_2	Betel quid chewing habits 0 = never, 1 = stop, 2 = current
Tumor_ploidy	x_3	

Table 2: Model Fitting Information and Goodness-of-Fit

Model	-2 Log Likelihood	Chi-Square	d.f	Sig.
Intercept Only	49.038			
Final	32.526	16.512	5	0.006

	Chi-Square	d.f	Sig.
Pearson	31.027	37	0.744
Deviance	32.526	37	0.679

Link function: Logit.

Table 3: Parameter Estimates of Oral Squamous Cell Carcinoma (OSCC)

							95% Confidence Interval	
		Estimate	Std. Error	Wald	d.f	Sig.	Lower Bound	Upper Bound
Threshold	[Tumour_size = 0]	-29.092	4.404	43.637	1	0.000	-37.724	-20.460
	[Tumour_size = 1]	-26.123	4.087	40.860	1	0.000	-34.133	-18.113
Location	Age	-0.178	0.077	5.339	1	0.021	-0.330	-0.027
	[Betel_Quid = 0]	2.444	1.123	4.733	1	0.030	0.242	4.646
	[Betel_Quid = 1]	4.975	2.150	5.354	1	0.021	0.761	9.190
	[Betel_Quid = 2]	0a	.	. 0	.	.	.	
	[Tumour_ploidy = 0]	-18.538	1.093	287.732	1	0.000	-20.680	-16.396
	[Tumour_ploidy = 1]	-18.807	0.000	. 1	.	-18.807	-18.807	
	[Tumour_ploidy = 2]	0a	.	. 0	.	.	.	

Link function: Logit.

a. This parameter is set to zero because it is redundant.

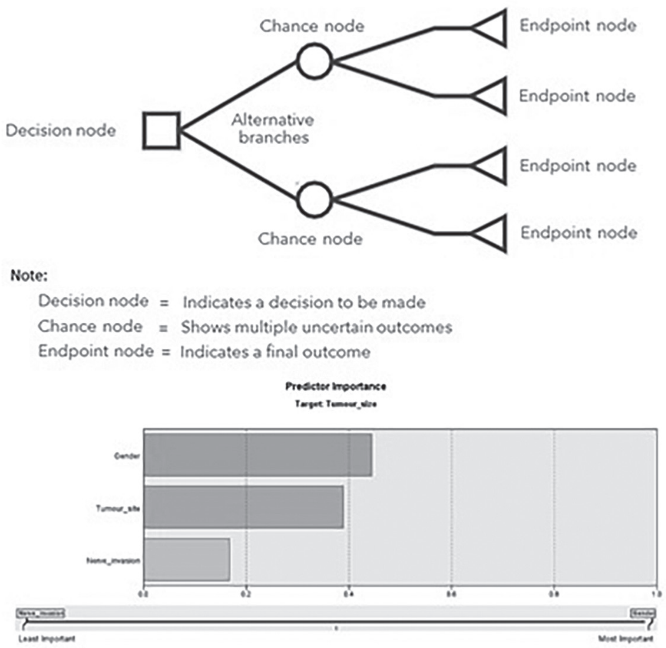

Figure 1: Decision trees are graphical models for describing sequential decision problems and Predictor Important (using decision tree analysis)

prompted us to believe that the fastest growing segment of oral cancer patients were non-smokers under the age of fifty years, which would show a paradigm shift in the cause of the disease, and at locations where it most often occurs in the oral environment. The interior of the mouth, tobacco, and alcohol-related cancers have declined along with the same reduction in smoking, and posterior of the oral cavity sites associated with the HPV16 viral cause are increasing. Tumors of the tongue and the floor of the mouth may be associated with excessive alcohol consumption and cigarette-smoking habits, whereas lesions on the buccal mucosa and gingiva may be related to tobacco/betel quid chewing habits (De Camargo Cancela, M., *et. al*. 2012). Taken together, these factors can explain the high percentages of tumors on the buccal mucosa since OSCC population in this study have betel quid chewing habits. In addition, betel quid chewing, chronic exposure to sunlight and human papillomavirus are among well-known risk factor for oral squamous cell carcinoma (OSCC) (Kaminagakura, E., *et. al*. 2012). Approximately, 90% of epithelial malignancy found within oral mucosa is squamous cell carcinoma (SCC) (Paul, D. L. Vol., 2012). It is one of the most difficult malignancies to control and has been associated with poor prognosis that may be explained by frequent lymph node metastases and local invasion characteristic (Noguti, J., *et. al*. 2012).

Ploidy is an abnormal DNA content has been associated with advanced stage OSCC and other markers of poor prognoses, such as lower degree differentials and lymph node metastasis. It appears to be an independent prognostic factor for relapse and death; it was found useful also as a valuable differential diagnosis marker for nondysplastic oral white patches or as a predictor of occult nodal metastasis. However, the debate has been maintained over this issue owing to the reported intratumoral heterogeneity of DNA ploidy, with some authors defending a homogeneous distribution of ploidy in the tumor maintained even in the metastasis (although ploidy has not been correlated with prognosis) and others reporting heterogeneity and thus limited application to predict prognosis. The study of DNA content of cells in the tumor invasive front, considered important to measure tumor aggressiveness (and therefore predict outcome), suggested an influence on disease-specific survival, especially if in conjunction with clinical findings (Diwakar, N., *et. al*. 2005).

DATA AND METHODS

Ordinal Regression

Ordinal regression models are the most popular embedded under different link functions in the framework of generalized linear models. Many variations of regression models to analyze ordinal response variables have been developed and described in previous years. Compared to frequently used methods for binary and nominal data, ordinal regression models have the advantage that they make full use of ranked data. Therefore, the analysis of epidemiological data on risk factors depends on regression models. The choice of a model is largely determined by

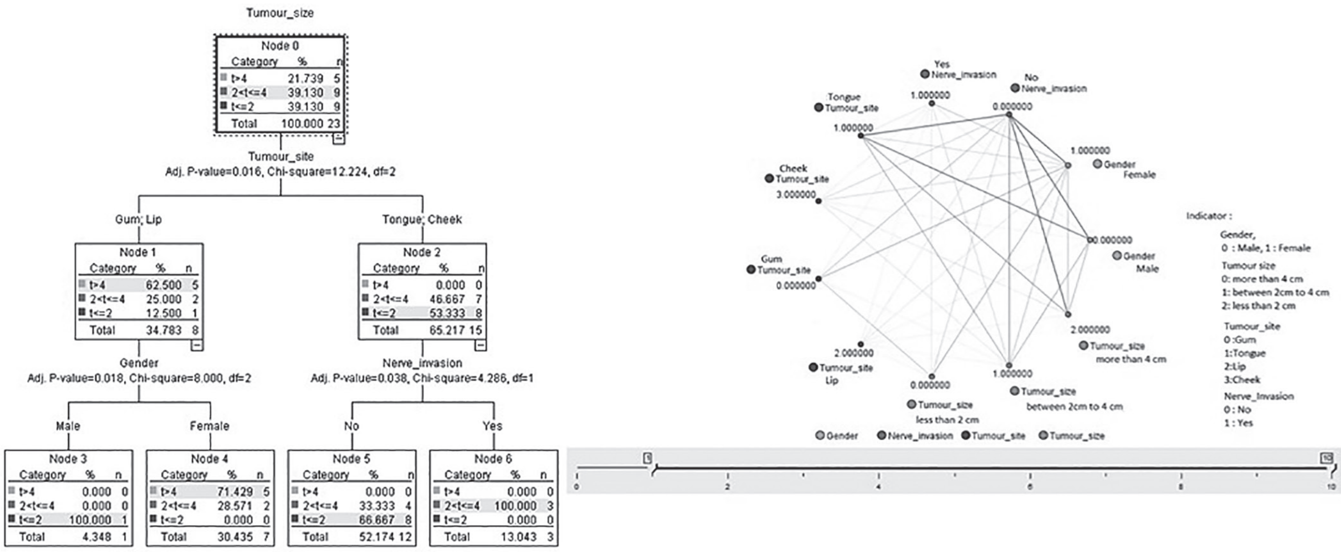

Figure 2: Decision tree for Oral Squamous Cell Carcinoma (OSCC) classification and Co-occurrence frequency graph between categorical values

the scale of measurement of the response variable. In ordinal regression analysis, there are link functions, e.g., logit and clog log links, are used to build specific models. There is no clear cut method to distinguish the preference for using different link functions. However, the logit link is generally suitable for analyzing the ordered categorical data evenly distributed among all categories. The clog log link may be used to analyze the ordered categorical data when higher categories are more probable (Bender, R., and Benner, A., 2000). In this study, the ordinal regression model was used to model the relationship between the ordinal outcome (i.e., different levels of tumor size) and independent variables. The framework of the ordinal regression model is described with the data set in the following section. Table 1 shows the description of data among oral squamous cell carcinoma patients. The association between tumor size among oral squamous cell carcinoma patients and independent variables was analyzed using an ordinal regression model, adjusted for age, betel quid and tumor ploidy. The result was considered statistically significant for a p-value ≤ 0.05. All statistical analyses in this study were performed using the statistical software Statistical Package for Social Sciences version 21.0.

Fitting Decision Tree Analysis

A decision tree is a graphical model describing decisions and their possible outcomes. A decision tree is a great and efficient method for classification, prediction and for facilitating decision making in sequential decision problems. Decision trees are indispensable graphical tools in such settings, its allow for an intuitive understanding of the problem and aid for the optimal decision making. A decision tree consists of three types of nodes (a) decision node (b) chance node (c) Endpoint node/Terminal node (Mesarić J and Šebalj D, 2016). Figure 1 shows the concept of a decision tree.

RESULTS AND DISCUSSION

The first section, the ordinal regression result will be discussed and the second section decision tree analysis result.

RESULTS

Ordinal Regression

Model fitting information provides results of ordinal regression versus reduced model (intercept) with a complementary log-log link function. The presence of a relationship between the dependent variable and a combination of independent variables is based on the statistical significance of the final model. From Table 2, the -2LL of the model with the only intercept is 49.038 while the -2LL of the model with intercept and independent variables is 32.526. The difference (Chi-square statistics) is 16.512-32.526 = -16.014 which is significant at $\alpha = 0.05$, $p < 0.01$. The conclusion is that there is an association between the dependent and independent variable(s) in complimentary Log-log link function. Goodness-of-fit Pearson is widely used in statistics to measure the degree of the relationship between the linear related variables. Deviance is a likelihood-ratio test used under full maximum likelihood. The deviance can be regarded as a measure of lack of fit between model and data. Generally, the larger the deviance, the poorer the fit to the data. The deviance is usually compared to deviances from other models fitted to the same data. The null hypothesis states that the observed data are consistent with the fitted model. The null hypothesis is accepted and one concludes that the observed data were consistent with the estimated values in the fitted model since the p was not significant, $p = 0.679 > 0.05$. Table 2 shows that the goodness of fit. From the Table 3, Three factors were found to influence the tumor size of oral squamous cell carcinoma patients namely; age, $p = 0.021 < 0.05$ but the negative sign indicates that younger age group have a more aggressive disease with a higher incidence of local recurrence or regional lymph node involvement after treatment compared with older patients. Betel quid chewing habits, $p = 0.030 < 0.05$ and tumour ploidy, $p = 0.0001 < 0.05$. Since the estimated $p < 0.05$ the null hypothesis is rejected and conclude that the regression coefficient for the three predictor variables was found to be statistically different from zero in estimating the general question in the presence of other independent variables.

Decision Tree Analysis

According to Figure. 1, the top two predictors are ranking to their contribution is the gender and tumor site. Using the CHAID method, these two are the best predictors of tumor size among oral squamous cell carcinoma (OSCC) patients. In total there were three predictors that the model deemed important.

According to the predictor important of decision tree analysis in Figure. 2, the prevalence of tumor size among oral squamous cell carcinoma (OSCC) patients is about 39.13% (below 4 cm). SPSS modeler had suggested that first split on the decision tree analysis is determined by the tumor site (which is the most important toward the stage of tumor size). There is only one group had been suggesting by decision trees analysis. This group is focusing on the tumor site which the most common site observed in our study for oral squamous cell carcinoma was the site at the gum and lip for the first split and tough and cheeks for the second split. The first split had found that the case of tumor size of oral squamous cell carcinoma is lower (34.783%) compared to the second split (65.217%). The second split from group one is the last split which emphasized on the gender (which having gum and lip tumor). From the analysis, it is clearly stated that female have contributed about 30.435%

of the case from the tumor size than male. The second split from group two is the last split which emphasized the nerve invention (which having tongue and cheek tumor). From the analysis, its show that the patients who don't have nerve invention most contribute to tumor size (52.174%).

Figure 2 shows the co-occurrence frequency graph between categorical values. Based on the Figure 2, mostly tongue squamous cell carcinoma normally affects male patients. According to the literature which conducted by (Jae-Ho Jeon, *et. al*. 2017), it stated that from May 2001 to August 2011, 15 patients were male and 8 patients were female with ages of 40 years or younger and 51 patients was male and 43 patients were female with ages of over 40 years were treated underdiagnosis of tongue squamous cell carcinoma at the National Cancer Center (Goyang, South Korea). Besides that, tumor cells spread easily through the nerve and the surrounding tissues and compartment. Without nerves invasion, the tumor size will be increased more and more. According to the finding, a tumor which is less than 4 cm has a strong association with the nerve invasion.

DISCUSSION

From this study show that age, betel quid, and tumor ploidy were significant and have an association with tumor size of oral squamous cell carcinoma. From the previous study have shown that patients in the younger age group have a more aggressive disease with a higher incidence of local recurrence or regional lymph node involvement after treatment compared with older patients. Our comparison of the local recurrence-free survival for the group < 30 years and the group above 30 years yielded no significant result ($p = 0.940$). In the present study, the overall survival of < 30 years of age group was less as compared to more than 30 years group but was not significant. This was consistent with the study by (Hilly, O., *et. al*. 2013), who demonstrated that there was an insignificant difference between the overall survival and disease-free survival in the younger group as compared to the older one. This has been found in the study that in spite of aggressive treatment in younger age group, the overall survival and disease-free survival were not improved in younger age group (Udeabor, S. E., *et. al*. 2012). In this study, 57% of OSCC patients were at age ≤ 57 years old. Our results are quite high showing significant value ($p < 0.005$) compare to other reports (Komolmalai, N., *et. al*. 2015; Udeabor, S. E., *et. al*. 2012). Previous studies found evidence of the synergistic effects of betel quid chewing on the risk of developing oral cavity cancer. This might be explained by the fact that betel quid chewers are proportionately heavier smokers, which was also true in the current study. Furthermore, alcohol might alter the intracellular metabolism of the epithelial cells at the target site. As a result, the oral mucosa was more vulnerable to carcinogens brought by smoking and betel quid chewing. Betel quid chewing are well-known risk factors associated with oral cavity cancer (Castellsague, X., *et. al*. 2004).

A study by (Razak, A. A., *et. al*. 2010) showed that betel quid, a traditional stimulant was common in Malaysia associated in these ethnic groups. Betel-quid is a masticatory mixture combining the areca nut, betel leaf, slaked lime, and locally varied flavorings. It is the fourth most frequently consumed psychoactive substance worldwide. Moreover, our results were also supported as in the study by (Razak, A. A., *et. al*. 2010) which reported that in Malaysia, habits of chewing betel leaf are a risk factor associated with OSCC (Razak, A. A., *et. al*. 2010). Epidemiological studies have shown that the sites of occurrence for oral cancer differ widely. It is stated that the tumor site is related to specific risk factors. Tumors of the tongue and the floor of the mouth may be associated with excessive alcohol consumption and cigarette-smoking habits, whereas lesions on the buccal mucosa and gingiva may be related to tobacco/betel quid chewing habits. Taken together, these factors could explain the high percentages of tumors on the buccal mucosa since OSCC population in this study have betel quid chewing habits. Second affected site was on the tongue followed by gingiva (Kerawala, C., *et. al*. 2016).

Interestingly, there was a relationship between the number of chromosomal aberrations and DNA ploidy and tumor size. The present study shows that DNA aneuploidy tumors have more chromosomal alterations than diploid tumors ($p < 0.05$). DNA aneuploidy is a hallmark of cancer. Although many potential causes of DNA aneuploidy have been suggested, a persistent defect in chromosomal segregation has been reported (Donadini, A., *et. al*. 2010). Recently, (Lengauer C, *et. al*. 1997) reported that DNA aneuploidy in cancers reflects an underlying chromosomal instability. (Harada, K., *et. al*. 1998) also suggested that DNA aneuploidy predisposes tumor cells to subsequent genetic alterations, which results in intra tumour cytogenetic heterogeneity. Genetic aberrations detected by CGH were more frequent in tumors with intratumor heterogeneity than in cancers without it. The present study suggests that DNA aneuploidy may be linked to genetic instability that is detected as genetic aberrations by CGH. The relationship of DNA copy number abnormalities to tumor size in adrenocortical tumors was described by (Kjellman M, *et. al*. 1996) using CGH. They reported that gains and losses were found only in tumors with diameters of 5 cm or more and that the number of alterations per a tumor was proportional to the tumor size. In this study, we found that the tumor ploidy is the risk factor of tumor size in OSCC ($p < 0.05$). Along with another factor, genetic also contribute to OSCC. The previous study from (Fan Y, *et. al*. 2014), reports that 18% of patients with a family history of cancer. It is now established that up to 10% of all cancers have a strong hereditary component. There was a study showing that clustering of oral cancer has been seen in certain ethnic groups, like Askenazi group in Israel which the incidence is double as compared to other Jewish population in that country.

CONCLUSIONS

In this prospective cohort study, we found a strong relationship between ages, betel quid chewing and tumor ploidy in tumor size of oral squamous cell carcinoma (OSCC). Synergistic effects endured patients with all the above habits had a higher risk of developing tumor size of oral squamous cell carcinoma than patients who abstained. Therefore, we recommend those aged < 57 years who are habitual betel quid chewers and who had tumor ploidy undergo oral mucosa examinations frequently so that potential oral cancer can be identified as early as possible.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Universiti Sains Malaysia (USM) for providing the research funding (Grant no.1001/PPSG/8123051, School of Dental Sciences, (PPSG), Health Campus, Universiti Sains Malaysia).

REFERENCES

Bender, R., and Benner, A. 2000. Calculating Ordinal Regression Models in SAS and S-Plus. Biometrical Journal, 42(6), pp. 677-699.

Castellsague, X., Quintana, M. J., Martınez, M. C., Nieto, A., Sanchez, M. J., *et al*. (2004). The role of type of tobacco and type of alcoholic beverage in oral carcinogenesis. International Journal of Cancer, 108, pp. 741-749.

De Camargo Cancela, M., de Souza, D. L., and Curado, M. P. (2012). International incidence of oropharyngeal cancer: a population-based study. Oral Oncol, 48(6), pp. 484-90.

Diwakar, N., Sperandio, M., Sherriff, M., Brown, A., and Odell, E. W. (2005). Heterogeneity, histological features and DNA ploidy in oral carcinoma by image-based analysis. Oral Oncol: 41, pp. 416-22.

Donadini, A., Maffei, M., Cavallero, A., Pentenero, M., Malacarne, D., *et al*. (2010). Oral cancer genesis and progression: DNA near-diploid aneuploidization and endoreduplication by high resolution flow cytometry. Cell Oncol, 32(5-6), pp. 373-383.

Fan Y, Zheng L, Mao M-H, Huang M-W, Liu S-M, and Zhang J. (2014). Survival analysis of oral squamous cell carcinoma in a subgroup of young patients. Asian Pac J Cancer Prev, 15, pp. 8887-8891.

Harada, K., Nishizaki, T., Ozaki, S., *et al*. (1998). Intratumoral cytogenetic heterogeneity detected by comparative genomic hybridization and laser scanning cytometry in human gliomas. Cancer Res 1998, 58, pp. 4694-700.

Hilly, O., Shkedy, Y., Hod, R., Soudry, E., Mizrachi, A., and Hamzany, Y. (2013). Carcinoma of the oral tongue in patients younger than 30 years: Comparison with patients older than 60 years. Oral Oncol, 49, pp. 987-90.

Hizlinda, T., Noriah, M. I., Noor Azimah, M., Farah Naaz, M. A., Anis Ezdiana, A. A., *et al*. (2012). Perceived Effects of the Malaysian National Tobacco Control Programme on Adolescent Smoking Cessation: A Qualitative Study. The Malaysian Journal of Medical Sciences. MJMS, 19(2), pp. 35-47.

Jae-Ho Jeon, Min Gyun Kim, Joo Yong Park, Jong Ho Lee, Myung Jin Kim, Hoon Myoung, and Sung Weon Choi. (2017). Analysis of the outcome of young age tongue squamous cell carcinoma. Maxillofac Plast Reconstr Surg., 39(1): 41.

Johnson, N. W., Jayasekara, P., and Amarasinghe, A. A. (2011). Squamous cell carcinoma and precursor lesions of the oral cavity: epidemiology and etiology. Periodontol 2000., 57, pp. 19-37.

Kaminagakura, E., Villa, L. L., Andreoli, M. A., Sobrinho, J. S., Vartanian, J. G., *et al*. (2012). High-risk human papillomavirus in oral squamous cell carcinoma of young patients. International Journal of Cancer, 130(8), pp. 1726-1732.

Kerawala, C., Roques, T., Jeannon, J. P., and Bisase, B. (2016). Oral cavity and lip cancer: United Kingdom National Multidisciplinary Guidelines. The Journal of Laryngology and Otology, 130 (Suppl 2), S83-S89.

Kjellman M, Kallioniemi OP, Karhu R, *et al*. (1996). Genetic aberrations in adrenocortical tumors detected using comparative genomic hybridization correlate with tumor size and malignancy. Cancer Res 1996, 56, pp. 4219-23.

Komolmalai, N., Chuachamsai, S., Tantiwipawin, S., Dejsuvan, S., Buhngamongkol, P., *et al*. (2015). Ten-year analysis of oral cancer focusing on young people in northern Thailand. J Oral Sci, 57(4), pp. 327-34.

Lengauer C, Kinzler KW, and Vogelstein B. (1997). Genetic instability in colorectal cancers. Nature 1997; 386: 623-7.

Marur, S., D'Souza, G., Westra, W. H., and Forastiere, A. A. (2010). HPV-associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol., 11, pp. 781-789.

Mesarić J and Šebalj D (2016). Decision trees for predicting the academic success of students. Croatian Operational Research Review, 7(2), pp. 367-388.

Neville, B. W., and Day, T. A. (2002). Oral cancer and precancerous lesions. CA Cancer J Clin., 52, pp. 195-215.

Noguti, J., De Moura, C. F. G., De Jesus, G. P. P., Da Silva, V. H. P., Hossaka, T. A., *et al*. (2012). Metastasis from Oral Cancer: An Overview. Cancer Genomics - Proteomics, 9(5), pp. 329-335.

Paul, D. L. Vol. (2012). The PhD University of Medicine and Pharmacy Craiova, University of Medicine and Pharmacy Craiova.

Razak, A. A., Saddki, N., Naing, N. and Abdullah, N. (2010). Oral Cancer Survival among Malay Patients in Hospital Universiti Sains Malaysia, Kelantan. Asian Pacific J Cancer Prev, 11(1), pp. 187-191.

Udeabor, S. E., Rana, M., Wegener, G., Gellrich, N. C., and Eckardt, A. M. (2012). Squamous cell carcinoma of the oral cavity and the oropharynx in patients less than 40 years of age: A 20-year analysis. Head Neck Oncol, 4: 28.

World Health Organization (2014). World health statistics 2014 [Online]. [Assessed Available from World Wide Web: World Health Organization (2016). World Health Statistics 2016.