Prevalence and diagnostic of head and neck cancer in Pakistan

Azeem Akhtar[1], Ijaz Hussain[2], Muhammad Talha[3], Muhammad Shakeel[4] Muhammad Faisal[5], Muhammad Ameen[6] and Tajammal Hussain[7]

[1]College of Statistical and Actuarial Sciences, University of Punjab, Lahore, Pakistan
[2]Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
[3]Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia
[4]Department of Statistics, Government College of Science Wahdat Road, Lahore, Pakistan
[5]Faculty of Health Studies, University of Bradford, Bradford, UK; Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK
[6]Bureau of statistics Punjab, Pakistan
[7]Department of Statistics, COMSATS Institute of Information Technology, Lahore, Pakistan

Abstract: Head and Neck (HN) cancer is the most frequently identified cancer across the globe and is considered as 8-10 % of all cancers in Southeast Asia. The social and cultural habits are the main cause of the remarkable increase in HN cancer in Pakistan. HN cancers are more prevalent in men as compared to women. The HN cancer in Pakistan is mainly attributed to discrete demographic profile, risk elements, eating patterns and family history. The foremost factors of risk are linked with cigarette smoking, alcoholic drinks and tobacco chewing like paan, gutka, etc. The actual burden of HN cancer in Pakistan is 18.74% of all new cancer cases recorded during 2004 -2014. The data is collected from HN cancer patients who are diagnosed from September 2011 till May 2012 in the Institute of Nuclear medicine and Oncology, Lahore Pakistan. In order to diagnose risk factors of HN cancer, the Logistic Regression (LR) and Linear Discriminant Analysis (LDA) are used as analytical tools. The findings of using these both tools concluded that gender, age, occupation, socioeconomic status, family history, chewing habits, smoking habit, worked in the chemical or dust place, alcohol use are main causes for HN cancer. The main aim of this study is to develop the awareness about HN cancer, its causes and early diagnosis.

Keywords: Logistic Regression, Linear Discriminant Analysis, HN Cancer, Estimation, Risk Factors, early diagnosis.

INTRODUCTION

Billions of cells join together to assemble human body and each cell undergoes the process of birth and death. The uncontrolled division and growth of cells may lead to cancer. The tumor (benign or malignant) is occurring when the extra tissues swell together without inflammation. If a tumor enlarges, it moves to another place, which is known as metastasizing (Stadler, 2014). In most of Asian countries, the cancer is a severe health risk and is the main source of mortality in Asian Pacific countries (Park *et al*., 2008). Early diagnosis is the only way to survive. The new cancer incidences in Asia may elevate to 7.1 million till 2020, if no changes are introduced in prevention and management policies (Mackay *et al*., 2006).

HN cancer consists of a various group of cancer (tumor) kinds occurring from the upper aero digestive tract, which includes the lip, mouth, nose, tongue, throat and oral cavities, sinuses, pharynx, vocal cords (larynx), and some other places in this anatomical part as shown in the fig. 1 (Razak *et al*., 2010). The cancer (tumor) can occur at any of these sites with the possibility of being benign or malignant. The HN cancer challenges connected with both its manifold anatomical sub-regions and related unique biological features so it can be said that HN cancer is a composite of the large number of malignancies (Waldron *et al*., 2008). A Computed Tomography (CT) of the HN cancer (benign) is presented in fig. 2 whereas fig. 2 demonstrates the malignant HN cancer (Liu, 2007 and Cummings *et al*., 2007).

HN cancer is similar to other cancers such as, lip, oral cavity, cavity of nasal, paranasal sinuses, pharynx, and larynx. The 90% of HN cancers are squamous cell carcinomas, which are the most frequently appeared malignancies of the HN site. It is also frequently occurred tumor among men in India, and among females it is the fifth most common cancer. The burden of HN cancer is 21% of the cancers among men and 11% among women across Pakistan (Joshi *et al.*, 2014). The HN cancer ranked six among all cancers in the globe with about 630,000 new patients identified per year causing above 350,000 deaths annually. The 5 years rate of survival of patients with HN squamous cell carcinoma is approximately 45-55% (Vigneswaran & Williams, 2014).

The smoking is one of the major risk elements in growing squamous cell carcinoma of HN cancer (Herrero *et al*., 2011). The oral and orpharynx cancers are caused by tobacco and alcohol. Additionally in Asia, placing a

**Corresponding author:* e-mail: ijaz@qau.edu.pk

fraction of tobacco in the mouth, chewing of betel quid and areca nut is another factor of HN cancer risk (Lambert *et al*., 2011). The regular use of betel nut, betel quid and paan masala increases the chances of growing HN cancer (Channa & Khan, 2014; Mazumder, *et al*., 2014).

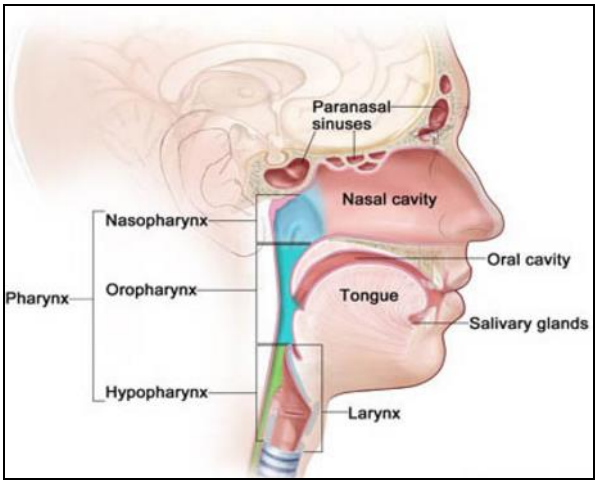

Fig. 1: Head and neck cancer regions

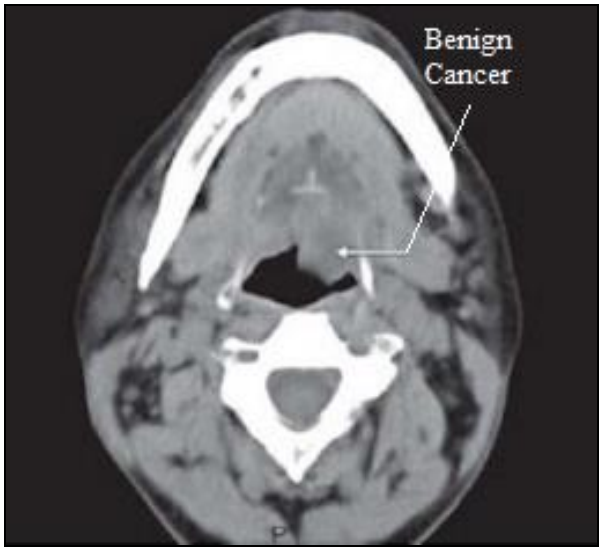

Fig. 2: A Computed Tomography image of head and neck cancer (benign)

Conventionally, 80-90% of HN squamous cell carcinoma (HNSCC) occurrence reasons are connected with the use of alcohol and tobacco. The consistent tobacco smoking is the main source of HN cancer. The unwilling or second-hand smoking can also increase the risk of HN cancer. Alcohol also causes the drastic increase in HN cancer jeopardy, which is caused by regular use of tobacco, food habits, viruses, and working places. The average 5-year rate of survival from HN cancer is below 50% (King & Agulnik, 2010).

There are two major risk factors (smoking and heavy alcohol consumption) for oral cavity and pharynx cancers. In worldwide, there are 42% death due to smoking and 16% because of heavy alcohol consumption. In high-income countries, there are fatalities of about 70% for smoking and 30% for drinking alcohol (Jemal *et al*., 2011). The oropharyngeal cancers have been associated with smoking, ionizing radiations, too much dental work, and unhygienic foods (Mishra & Meherotra, 2014). Oral cancer mortality rate is very low (2 in 100,000) in the Middle East, that is much lesser than India and US (Anis & Gaballah, 2013). HNSCC can be cured through surgical treatments, radiation therapy, systemic therapy, or combinations of these procedures (Hinni *et al*., 2013).

Fig. 3: A Computed Tomography image of head and neck cancer (Malignant)

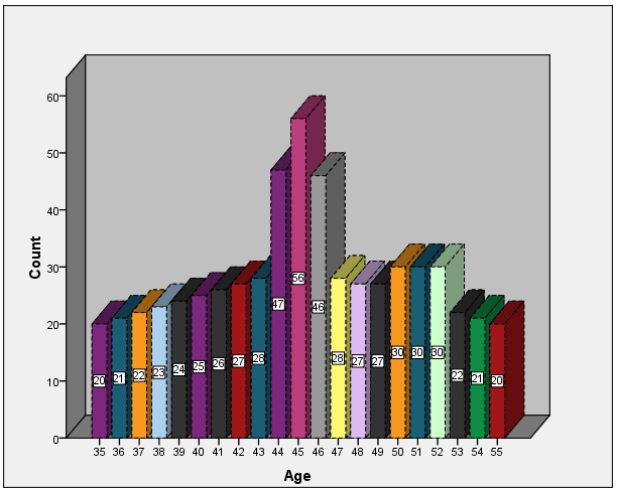

Fig. 4: Age wise distribution of head and neck cancer patients.

A small number of studies of epidemiology have explored the danger of laryngeal cancer associated with a family history of HN cancer (Garavello *et al*. 2012). Chewing of

betel nut is considered as a fashion in South Asia, while it is a rigorous risk element for HN cancer causing a major life threat. It may cause to develop squamous cell carcinoma cancer. The People who chewed betel nut 5 times per day and continue it for about 15 years have larger chances to get HN cancer comparatively (Shah *et al.,* 2012). The HN cancer is more common in males than females because 66% to 95% males are main victims of this fatal disease. Conversely, the fig. is changing step-by-step with the smoking trends among females. The proportion of the pharyngeal and oral cavity cancers among males and females is almost 3:1. The Age factor is also an aspect of HN cancer. From 40-50 years age, cancer occurrence threat augmented as compared to the younger people. The chain smokers have comparatively high probability of getting this disease. The probability of occurrence of HN cancer is relatively more in aged people than the younger ones (Foulkes *et al.*, 2006).

Generally in Pakistan and especially in Karachi, the probability of getting HN cancer is higher than any other city on the globe, which is a real burden on Pakistan. It is because of the huge amount of people are habitual of smoking, chewing chalia and gutka. The Government must create the awareness about HN cancer and show the sincere efforts in order to stop this chronic disease. The HN cancer incidences in Pakistan and its nearby countries like India and Sri Lanka is 10 out of 100,000. The five-year survival rate of HN cancer is below 40% (Mazahir *et al.,* 2006). The overall age of individuals (cases and controls) in the current study was 45 years, ranging from 35-55 years.

The earlier studies on HN cancer were aimed at addressing the various objectives. Some of them have paid attention to find out the association between human Papiloma virus and HN cancer. Others aimed at finding the most important risk elements, which produce HN cancers. In the current study, the primary objective is to probe the possible link between risk factors and HN cancers. LR and LDA tools were implemented to investigate the association and the importance of risk factors regarding HN cancers.

MATERIALS AND METHODS

Description of study area

The present study is based on case-control where HN cancer patients are considered as cases. The data for these cases is obtained from INMOL Hosptial, which is one of the leading cancer hospitals in the Punjab province. The relatives of the patients who accompanied them while visiting hospital are considered as control.

Analytical procedure

The cases were collected during the nine-month time period, i.e from Spetmber 2011 to May 2012. During this period, a total of 300 cases and also an equal number of controls were included in this study. A well structured questionnaire was designed and used in order to collect the required data from the cases and controls.

The maximum questions related to possible risk factors contributing to HN cancer were included in the questionnaire. It was observed from literature that age, gender, occupation, smoking, family history of HN cancer patients, drinking alcohol, are the major risk factors in HN cancer. Here some additional risk elements are launched by considering the regional and social effects. The risk elements considered in the present paper are age, gender, occupation, residing area, socioeconomic status, smoking habits, alcohol, chewing habits, family history of HN cancer, used on meat, soft drink used, worked in the chemical or dust place, fast food intake, to have any gum infection previously. The risk elements are regarded as predictors or independent variables, whereas the response variable i.e HN cancer (dichotomous variable) is considered as the dependent variable.

Linear Discriminant Analysis (LDA) is frequently used for classification problems, it was proposed by R. Fisher, also famous as the Fisher's Linear Discriminant Analysis (Guo *et al*., 2007). Logistic Regression (LR) analysis is an efficient tool to classify dichotomous variables (Liu, *et al*., 2014). LR model and LDA functions are used to observe the relationship between HN cancer and the risk factors. Nagelkerke R square is used to observe how much variation of the response variable is explained through the risk factors in logistic regressions. Fisher's linear discriminant function applied to observe the significant disrciminator of HN cancer.

STATISTICAL ANALYSIS

Logistic regression model

LR model or logit is the best suitable tool in binary situations, where the response variable contains two categorical values. The LR model is primarily used to find out the relationship of a dependent variable with one or more independent variables X_i. The LR model has been used to predict and determine the most powerful explanatory variables on the dependent variable (Agresti and Kateri 2011. Collett, 1991. Cox and Snell, 1989. Hosmer and Lemeshow, 1989).

The logistic regression model for response variable P_i (probability of success for response variable) related with *k* independent variables $x_1, x_2, x_3, \ldots\ldots\ldots\ldots\ldots x_k$ can be described as follows (Collett, 2003):

$$\log it(p_i) = \log(\frac{p_i}{1-p_i}) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \ldots\ldots\ldots\ldots\ldots + \beta_k x_{ki}$$

which is linear in terms of coefficients and its parameters can be estimated like multiple linear regression models. Here $\frac{p_i}{1-p_i}$ is called the odd ratio. β_o, β_1 are parameters to

be estimated and P_i is the probability of success for response variable.

Fisher linear discriminant analysis

LDA comprises of the determination of a linear equation like multiple linear regression. These linear equations are used to predict the belonging of an object to specified classes. The linear equation of Fisher's linear discriminant is presented as follows;

$$F_{LDF} = a + t_1 X_1 + t_2 X_2 + t_3 X_3 + \ldots\ldots\ldots\ldots + t_n X_n$$

For details about linear discriminant analyses, see (Guo, *et al*., 2007).

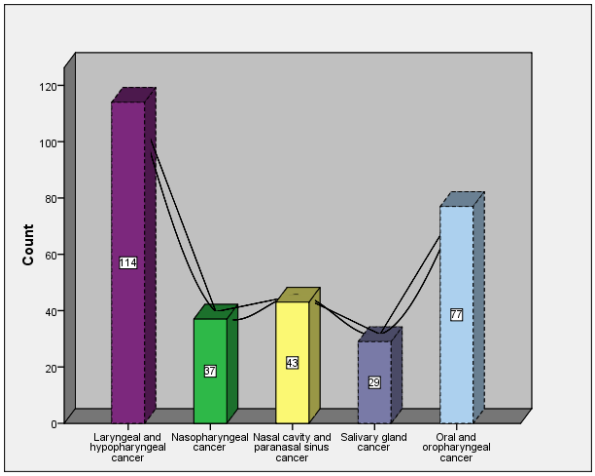

Fig. 5: The distribution of H & N cancer patients.

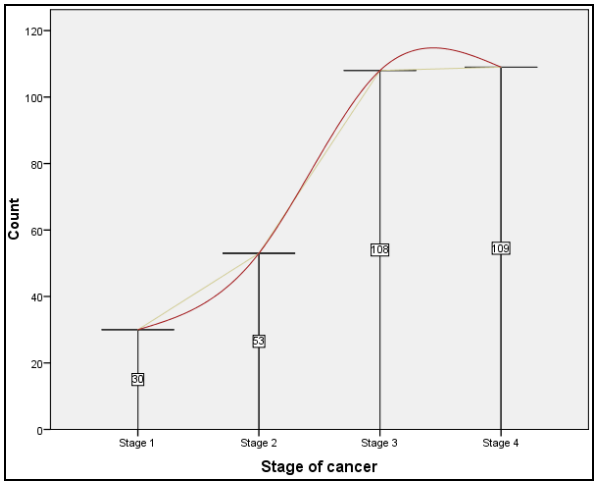

Fig. 6: The number of patients diagnosed at different stages.

Data analysis

Graphical method is used for exploratory analysis and two supervised learning techniques are used to identify the risk factors for HN cancer**.** fig. 4 shows that most of the patients are older and their disease was not diagnosed and treated in early stages. The different types of HN cancer diagnosed from the observed data are presented in fig. 5. Laryngeal and hypo pharyngeal cancer type are common because most of the patients were smokers. The main risk factors for laryngeal and hypo pharyngeal cancer are alcohol and tobacco consumption. The numbers of cases reported at different stages of HN cancer are displayed in fig. 6. The number of patients diagnosed at stage 3 and stage 4 are higher as compared with the number of patients diagnosed in the first two stages. fig. 7 shows the distribution treatment methods that were adopted as a remdial. The number of treated by surgery are higher as compared to treated with radiation therapy, and chemotherapy. The summary statistics and percentage distribution of different cases for different variables are presented in table 1.

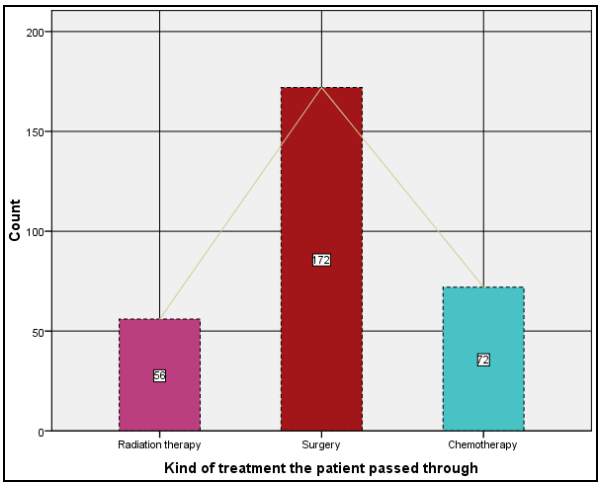

Fig. 7: The number of patients treated under radiation therapy, surgery and chemotherapy treatment.

RESULTS

In the present study, the age, gender, occupation, residing area, socioeconomic status, smoking habit, family history of HN cancer, used meat, alcohol use, worked on chemical or dust place, chewing habits, fast food intake, and soft drink used are assumed as significant contributors toward HN cancer. From the results of logistic regression it is inferred that age, gender, socioeconomic status, occupation, family history of HN cancer, smoking habit, alcohol use, worked on chemical or dust place and chewing habit are significantly contributing towards HN cancer. The estimated logistic regression coefficients along with odds ratios are presented in table 2.

The value of a Nagelkerke R Square test is used for the fitted logistic regression which is 0.789, it means 78.90% variability of HN cancer is explained by independent variables used in the present study. Omnibus test is also used to observe the goodness of fit for logistic regression coefficients. The estimated value of Chi-square test statistic under state hypothesis is 213.59, which validates that there is a significant association between HN cancer and selected predictors. The estimated logistic regression model is presented as follows;

Table 1: Frequency and Percentage Distribution of Cases and Controls

Variables	Categories	N (%)	Cases (%)	Controls (%)
Sex	Male	461 (76.8)	216 (46.9)	245 (53.2)
	Female	139 (23.2)	84 (60.4)	55 (39.6)
Type of residential area	Rural	369 (61.5)	151 (40.9)	218 (59.1)
	Urban	231 (38.5)	149 (64.5)	82 (35.5)
Age	< 45	262 (43.7)	108 (41.2)	154 (58.8)
	≥ 45	338(56.3)	192 (56.8)	146 (43.2)
Occupation	Labour	396 (66.0)	210 (53.0)	186 (47.0)
	Others	204 (34.0)	116 (56.9)	88 (43.1)
Socioeconomic status	High	121 (20.2)	73(60.3)	48(39.7)
	Middle	167 (27.8)	90(53.9)	77(46.1)
	Low	312 (52.0)	145(46.5)	167(53.5)
H& N cancer family history	Yes	125 (20.8)	73 (58.4)	52 (41.6)
	No	475 (79.2)	227 (47.8)	248 (52.2)
Smoking habit	Yes	401 (66.8)	172 (42.9)	229 (57.1)
	No	199 (33.2)	128 (64.3)	71 (35.7)
Alcohol use	Yes	257 (42.8)	166 (64.6)	91 (35.4)
	No	343 (57.2)	134 (39.1)	209 (60.9)
Chewing habit	Yes	450 (75)	244 (54.2)	206 (45.8)
	No	150 (25)	56 (37.3)	94 (62.7)
Used meat regularly	Yes	364 (60.7)	202 (55.5)	162 (44.5)
	No	236 (39.3)	98 (41.5)	138 (58.5)
Fast food intake	Yes	386 (64.3)	229 (59.3)	157 (40.7)
	No	214 (35.7)	71 (33.2)	143 (66.8)
Used soft drink	Yes	442 (73.7)	247 (55.9)	195 (44.1)
	No	158 (26.3)	53 (33.5)	105 (66.5)
Worked on chemical or dust	Yes	401 (66.8)	223 (55.6)	177 (44.1)
	No	199 (33.2)	77 (38.6)	122 (61.3)
Gum disease earlier	Yes	338 (56.3)	165 (48.8)	173 (51.2)
	No	262 (43.7)	135 (51.5)	127 (48.5)

Table 2: Estimated parmeters and odd ratios logistic regression

Variables	β	S.E	Wald	d.f	Sig.	Exp (β)	95% C.I for Exp (β)	
							Lower	Upper
Sex	1.48	1.18	14.59	1	0.00*	4.43	0.44	7.65
Type of residential area	-0.36	0.55	0.42	1	0.52	0.70	0.24	2.07
Age	0.05	0.03	9.96	1	0.02*	1.67	0.60	4.68
Occupation	-0.36	0.17	7.69	1	0.03*	1.99	0.79	3.97
Socioeconomic status	-1.74	0.72	5.85	1	0.02*	1.18	0.44	3.72
H& N cancer family history	-0.73	0.57	12.69	1	0.01*	3.43	0.55	6.76
Sex	1.92	0.63	9.42	1	0.00*	2.55	1.04	4.50
Smoking	0.92	0.33	23.91	1	0.00*	2.45	1.21	4.76
Alcohol	1.87	0.66	18.08	1	0.00*	2.18	1.09	4.50
Chewing	0.23	0.26	20.81	1	0.00*	1.79	0.48	5.31
Used meat	-2.73	0.79	1.86	1	0.60	0.29	0.04	0.57
Fast food	-1.07	0.43	1.25	1	0.09	0.72	0.26	1.98
Soft drink	-0.92	0.47	3.62	1	0.06	0.57	0.10	1.27
chemical or dust	3.07	0.84	13.53	1	0.00*	3.60	1.50	8.03
Gum disease earlier	-0.59	0.406	2.632	1	0.105	0.649	0.246	1.458

Table 3: Estimated Classification Function and Discriminant Coefficients of linear disciminant Analysis.

Variables	Classification Function		Discriminant Coefficient
	Case	Control	
Sex	27.58	26.08	1.50
Residentail area	8.68	9.05	-0.37
Age	6.37	6.33	0.05
Occupation	2.14	2.51	-0.38
Socioeconomic status (1)	11.45	13.16	-1.71
Socioeconomic status (2)	2.35	3.13	-0.77
H&N cancer Family history	6.33	4.34	1.99
Smoking habit	3.12	2.21	0.91
Alcohol use	11.86	10.02	1.84
Chewing habit	0.96	0.66	0.30
Used meat	6.16	8.89	-2.73
Fast food intake	2.54	3.61	-1.07
Soft drink used	4.08	5.00	-0.92
Worked at chemical or dust place	5.97	2.82	3.15
Having any gum disease earlier	3.19	3.85	-0.66
Constant	-91.42	-84.53	-6.89

$$\hat{Y} = P(HNC) = \frac{1}{1+e^{-Z}}$$

Where

$$Z = -6.659 + 1.487X_1 - 0.360X_2 + 0.053X_3 - 0.375X_4 - 1.742X_{51} - 0.730X_{52} + 1.925X_6 + 0.917X_7 + 1.868X_8 + 0.233X_9 - 2.727X_{10} - 1.073X_{11} - 0.924X_{12} + 3.073X_{13} - 0..594X_{14}$$

Linear discriminant analysis is performed by using SPSS, it can be observed that age, gender, occupation, socioeconomic status (1), socioeconomic status (2), family history of HN cancer, smoking habit, alcohol use, chewing habits, worked in the chemical or dust place, are significant discriminator. The output of linear discriminant coefficients is presented in table 3.

Here, the Chi-square test is used to observe the significant association of predictors and HN cancer. The estimated value of Chi-square statistics is $X^2=223.32$, which is significant at $p=0.00$ with 15 degrees of freedom. This test statistic shows that there exists a multivariate association in HN cancer and risk factors (predictors). The estimated Eigen value 0.825 discriminant analysis shows that 82.5% variation of HN cancer is explained by the risk factors.

In comparison to both techniques it could be inferred that in the current study it is inferred that age, gender, family history of HN cancer, alcohol use, smoking habit, chewing habits and worked in chemical or dust place had the greater effect on the HN cancer.

$$Z=-6.89+1.50X_1-0.37X_2+0.05X_3-0.38X_4-1.71X_{51}-0.77X_{52}+1.99X_6+0.91X_7+1.84X_8+0.30X_9-2.73X_{10}-1.07X_{11}-0.92X_{12}+3.15X_{13}-0.66X_{14}$$

The estimated classification rate in logistic regression and Linear Discriminant Analysis is 87%, which means that both of the techniques are 87% correctly classifying to the actual population.

DISCUSSION

Both, the LR and LDA, are used to identify the most important risk factors associated with HN cancer. Ten significant factors, i.e. occupation (-0.38), gender (1.49), socioeconomic status (1) (-1.742), socioeconomic status (2) (-0.73), age (0.05), family history of head and neck cancer (1.93), smoking habit (0.92), alcohol use (1.868), chewing habit (0.23), worked in the chemical or dust place (3.07) is identified by logistic regression. In Fisher's linear discriminant the family history of Head and Neck cancer (1.99), age (0.05), smoking habit (0.91), gender (1.50), use of alcohol (1.8), habits of chewing (0.30), worked in the chemical or dust place (3.15) are found as significant discriminators. The accuracy rate of correct classification is assessed for both techniques. It is concluded that the rate of correctly assignation is 87% for logistic regression and linear discriminant analysis.

It can be observed from the results of logistic regression and discriminant analysis that smoking habit is contributing towards HN cancer, which supports to findings of Herrero *et al.*, (2011). The findings of the present study also indicate that tobacco, continuous chewing of betel nut, betel quid, and pan masala are significant contributors towards HN cancer, the similar results are drawn by (Lambert *et al.*, 2011, Channa and Khan, 2014; Mazumder, *et al.*, 2014). In addition to indicated factor in earlier studies, it is detected that the

age, occupation, socioeconomic status (1), socioeconomic status (2), family history, and worked at chemical or dust place are significant contributors towards HN cancer. The limitation of this could be a small sample size. In this study, only one year data has been used and from one hospital. A large sample size study is recommended from more than one hospital to validate these risk factors of HN cancer.

CONCLUSION

The relation between the risk factors and HN cancer was established in the present study. The relationship and significance of risk factors upon HN cancer were observed based on logistic regression and linear discriminant analysis. It is concluded that patients suffering from HN cancer have an age between 34 and 55. Since the patients diagnosed as HN cancer were smokers therefore their type of cancer was hypo pharyngeal and laryngeal cancer. Most of the patients diagnosed as HN cancer were on the fourth stage, reflecting a lack of awareness. Moreover, it is inferred that the use of paan, chalia, worked on chemical and dust place, family history are also the significant contributors toward HN cancer.

ACKNOWLEDGEMENT

We would like to thank Dr. Kamran (Head & Neck cancer specialist), INMOL Hospital Lahore for his kind cooperation in the way of collecting data from the hospital. He also gave more understandings about head and neck cancer from medical point of view. The authors are thankful to the Deanship of Scientific Research, King Saud University Riyadh Saudi Arabia for funding through the Research Group Project no. RG-1437-037.

REFERENCES

Agresti A and Kateri M (2011). *Categorical data analysis* . Springer Berlin Heidelberg, pp.206-208.

Anis R and Gaballah K (2013). Oral cancer in the UAE: A multicentre, retrospective study. *Libyan J. Med.*, 2013.

Channa NA and Khan A (2014). Hazards of mainpuri chewing in Hyderabad and adjoining areas, Pakistan *Rawal Medical Journal*, **39**(2): 216-219.

Collett D (2003). Modeling binary data, London: Chapman & Hall. 45-101.

Cox DR and Snell EJ (1989). The Analysis of Binary Data. Second Edition, London: Chapman & Hall. 01-102.

Cummings BJ, Keane T, Pintilie M, Warde P, Waldron J, Payne D, Liu FF, Bissett R, McLean M, Gullane P and O'Sullivan B (2005). Five year results of a randomized trial comparing hyperfractionated to conventional radiotherapy over four weeks in locally advanced head and neck cancer. *Rad. Oncol.*, **85**: 7-16.

Draper NR and Smith H (1998). *Applied Regression Analysis*, (3rded.). New York: John Wiley & Sons. 401-408.

Foulkes WD, Brunet JS, Kowalski LP, Narod SA and Franco EL (2006). Family history of cancer is a risk factor for squamous cell carcinoma of the head and neck in Brazil: A case-control study. *International Journal of Cancer*, **63**(6): 769-773.

Garavello W, Turati F, Bosetti C, Talamini R, Levi F, Lucenteforte E, Chiesa F and Franceschi S (2012). Family history of cancer and risk of laryngeal cancer: A case control study from Italy and Switzerland. *International Journal of Cancer*, **130**: 665-670.

Guo Y, Hastie T and Tibshirani R (2007). Regularized linear discriminant analysis and its application in microarrays. *Biostatistics*, **8**(1): 86-100.

Herrero JI, Pardo F, Avola, DD, Alegre F, Rotellar F, Inarrairaequi M, Marti P, Sangro B and Quiroga J (2011). Risk factors of lung, head and neck, esophageal and kidney and urinary tract carcinomas after liver transplantation: The effect of smoking withdrawal. *Liver Transplantation*, **17**(4): 402-408.

Hinni ML, Ferlito A, Silver CE, Westra WH, Seethala RR, Rodrigo JP, Orry J, Bradford CR, Hunt JL, Strojan P and Devaney KO (2013). Surgical margins in head and neck cancer: A contemporary review. *Journal of the Sciences and Specialties of the Head And Neck*, **35**(9): 1362-1370.

Hosmer DW and Lemeshow S (1989). *Applied Logistic Regression*, New York: John Wiley and Sons, Inc. 243-268.

Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D (2011). Global cancer statistics. *CA Cancer J. Clin.*, **61**: 69-90.

Johnson RA, Wichern DW (1992). *Applied multivariate statistical analysis*, 6th edition, Englewood Cliffs, NJ: Prentice hall, 575-649.

Joshi P, Dutta S, Chaturvedi P and Nair S (2014). Head and neck cancers in developing countries. *Rambam Maimonides Medical Journal*, **5**(2): e0009, 1-6.

King T and Agulnik M (2010). Head and neck cancer: Changing epidemiology and public health implications, *Oncology*, 915-919.

Kleinbaum DG and Klein M (2002). *Logistic Regression: A Self Learning Text*, (3rd ed.) Springer Publications. 1-77.

Lambert R, Sauvaget C, de Camargo Cancela M and Sankaranarayanan R (2011). Epidemiology of cancer from the oral cavity and oropharynx. *European Journal of Gastroenterology & Hepatology*, **23**(8): 633-641.

Liu D, Li T and Liang D (2014). Incorporating logistic regression to decision-theoretic rough sets for classifications. *International Journal of Approximate Reasoning*, **55**(1): 197-210.

Liu FF (2007). Nasopharyngeal cancer: New frontiers from the laboratory to the clinic. *Int. J. Radia. Oncol. Biol. Phys.*, **69**: S122-124.

Mackay J, Jemal A, Lee N and Parkin D Eds. (2006). The Cancer Atlas. Atlanta, American Cancer Society.

Mazahir S, Malik R, Maqsood M, Malik F, Majeed A, Fatmi Z, Khawaja MR and Ghaffar S (2006). Socio-demographic correlates of betel, areca and smokeless tobacco use as a high-risk behavior for head and neck cancers in a squatter settlement of Karachi, Pakistan. *Bio Medical Central*, **1**: 1-10.

Mazumder TH, Nath S, Nath N and Kumar M (2014). Head and neck squamous cell carcinoma: Prognosis using molecular approach. *Cent. Eur. J. Biol.*, **9**(6): 593-613.

Mishra A and Meherotra R (2014). Head and neck cancer: Global burden and regional trends in India. *Asian Pac. J. Cancer Prev.*, **15**(2): 537-550.

Park S, Bae J, Nam HB and Yoo KY (2008). Cancer etiology in Asia. *Asian Pac. J. Cancer Prev.*, **9**: 371-80.

Razak AR, Siu LL, Liu FF, Ito E, O'Sullivan B and Chan K (2010). Nasopharyngeal carcinoma: the next challenges. *Eur. J. Cancer*, **46**(11):1967-1978.

Stadler WM (2014). Cancer Biology Review, a case-based approach, Demos Medical. 117-122.

Shah J, Patel S and Singh B (2012). *Head and Neck Surgery and Oncology*. (4thed.). 15-16.

Vigneswaran N and Williams MD (2014). Epidemiologic trends in head and neck cancer and aids in diagnosis. *Oral Maxillofacial Surg. Clin. N. Am.*, **26**: 123-141.

Waldron J, Warde P, Irish J, Pintilie M, Sellman S, Bayley A,Cummings BJ, Kim J, Liu FF, Payne D, Ringash J, Brown D, Freeman J, GilbertR, Gullane P, Witterick I and O'Sullivan B (2008). A dose escalation study of hyper fractionated accelerated radiation delivered with integrated necksurgery (HARDWINS) for the management of advanced head and neck cancer. *Radiother. Oncol.*, **87**: 173-180.