

RESEARCH ARTICLE

Associations of Most Prevalent Risk Factors with Lung Cancer and Their Impact on Survival Length

Mohammad Haroon Khan^{1*}, Shahid Hussain¹, Raisa Bano¹, Hafiz Jamshed-ul-Hassan¹, Muhammad Aadil ur Rehman²

Abstract

Lung cancer is one of the most common malignancies in the world. Its incidence and mortality rates are on the rise in Pakistan. However, epidemiological studies to identify common lung cancer determinants in the Pakistani population have been limited. In this study, data of 440 cases and 323 controls were collected from different hospitals in Peshawar and Islamabad, along with information about socio-demographic factors including age, sex and smoking. Univariate and multi-factorial analyses of socio-demographic factors in association with each other were also performed. Overall survival analysis highlighted that, out of 440 patients in the lung cancer dataset, 204 people were uncensored with a median survival time of 13 months (95% CI=12-18). There were 41 female and 399 male patients. Differences were observed between length of survival in the males and females ($\chi^2 = 6.1$; p -value = 0.01). Gender was observed to be significantly related to survival (p -value < 0.01), with better survival in females (hazard ratio=2). Cox regression was extended to adjust for the covariate age ($z = 2.5$; p -value = 0.02). Survival analysis was also performed on the basis of smoking groups (current smokers, former smokers and never smoked individuals) and smoking duration (smoking duration >10 years, <10 years and never smoked). Smoking duration was significantly associated with survival (p -value < 0.01), with better survival in never smokers in comparison to both smoking for greater or less than 10 years. Strong associations were observed for smoking group with duration greater than 10 years, OR=6.1(3.9-9.5) on univariate and multi-factorial analysis OR=11.3(CI=6.8-19.3).

Keywords: Lung cancer - smoking - risk factors - occupational factors - dietary factors - Pakistani population

Asian Pac J Cancer Prev, 17, Cancer Control in Western Asia Special Issue, 65-70

Introduction

Lung cancer is globally one of the most leading and prevailing malignancies (Tas et al., 2008). It is also one among the most fatal diseases known to mankind (Wahbah et al., 2007). Relative prevalence of its various histological patterns has changed over a period of time, which may be attributed to changes in various risk factors (Rennert, et al 1982). However, the prognosis is still poor as most of them have usually metastasized distantly before being diagnosed. Therefore, almost all types have a poor 5 year survival rate.

Several major risk factors exist in our society including cigarette smoking, tobacco chewing and occupation which contributes to lung carcinoma. A limited data of lung cancer is available related to Pakistani population due to lack of awareness and insufficient provision of healthcare. In developing countries, lung cancer incidence and mortality are continuously increasing (Lam, 2005), while the figures are declining in the developed countries (Kwong et al., 2005). High mortality rates are attributed majorly to the genetic and

non-genetic risk factors along with late detection of the disease (Ahmed et al., 2013).

Smoking is a well-established risk factor of lung cancer but according to literature, it also affects non-smokers in high frequency (Sun et al., 2007) which may be due to the multi-factorial nature of the disease. Multiple risk factors including genetic, familial and social and lifestyle factors contribute towards the disease development in non-smokers. Smoking trend in Pakistan is on the rise and tobacco is being used in different forms like hukka, cigarettes, cigar, pan shisha etc. (Alam, 1998). Like other developing countries, lung cancer is the most frequent type of malignancies in Pakistani population (Hussain et al., 2009).

Increasing rates of lung cancer in non-smokers suggests that individuals are exposed to multiple risk factors including occupational exposures, lifestyle and dietary factors. Only very limited epidemiological studies in Pakistani population has been reported to evaluate these risk factors. This study was conducted to investigate relative prevalence of various histological patterns of lung carcinoma in our region and to highlight

¹Molecular Biology and Bioinformatics Research Group, Department of Bioinformatics and Biosciences, ²Department of Computer Sciences, Capital University of Science and Technology, Islamabad, Pakistan, *For correspondence: haroon.khan@cust.edu.pk

the association of common socio-demographic factors with increased risk of lung cancer and their impact on survival length in Pakistani Population.

Materials and Methods

Data was recorded on a specially designed questionnaire along with informed and signed consents from lung cancer patients along with comparative controls at NORI hospital, Islamabad, IRNUM hospital, Peshawar and private clinics. They have specified facilities for cancer treatments and also have state of the art oncology technologies. This study was approved by the research and ethics committee of the department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad, Pakistan. Data was recorded from July 2013 to June 2014 including a total of 440 cases, with 323 controls from the same ethnic groups. Histologically confirmed lung cancer cases from different geographic distributions of Pakistan were included in this case-control study. The Performa includes questions related to age, sex and smoking. A total of 323 controls which were free of any type of cancer and chronic respiratory diseases. All variables recorded were used in this study. ORs and their correspondent 95% confidence intervals were estimated using logistic regression, Risk Ratios and 95% confidence intervals were estimated through Cox regression. ORs and RRs were obtained in both univariate and multi-variate models including all independent variables. Analysis of socio-demographic factors in association with each other was also performed. Kaplan-Meier analysis was used to access the impact of risk factors on length of survival. All analysis was conducted using R 3.1.

Results & Discussion

Study participants belonged to diverse regions of Pakistan. All the cases were collected after

appropriate confirmation of lung cancer. It was our effort to collect controls from the same ethnic group and age category throughout the country. It was observed that most of the study participants were educated and so were comparatively easy to convince. Majority of the cases (58.6%) belonged to the smokers group having smoking duration more than 10 years while current smokers were recorded for 53.4% of the cases. It was also observed that majority of the cases (49.2%) belongs to the age category of 39-54 years while the lowest group (7.2%) are less than 34 years of age. Age group ranging from 39-54 years are at higher risk (OR=2.3 (1.5-3.6)) of lung cancer. Similar pattern of gender and age differences among cases and controls were observed in Kerala, India (Bhaskarapillai et al., 2012) and Nepal (Hashibe et al., 2010). Positive association was observed for the majority of variables but with varying degrees of strength (Table 1).

In Uni-factorial analysis, smoking showed the most significant association with lung cancer (OR=6.1 (3.9-9.5), which suggests that smokers in Pakistani population are at much higher risk of lung cancer. Our results are in line with the previous results of global and regional research on lung cancer (IARC, 1994), Stellman et al., (2001), Hussain et al., (2009), Ganesh et al., (2011) and Matteis, (2013). Elevated association of smoking is representing an increasing lung cancer incidence in Pakistani population, where smoking habit is increasing specially with the aid of sheesha. It is also an enigma that Pakistani cigarettes have comparatively higher tar contents (Alam, 1998). Gender (male) was another variable showing positive association (OR=2.9 (2-4.5) with lung cancer. In the multifactorial analysis of socio-demographic analysis, smokers with more than 10 years smoking duration showed a very strong association (OR= 11.3, 95%CI =6.9-19.3, Z=9.2 and P<0.01) with lung cancer. Males (OR=2.6, 95%CI= 1.6-4.1, Z=4.1, P<0.01) and the age group of 39-54 years (OR=1.9, 95%CI =1.2-3, Z=2.7, P<0.01) also showed positive association.

Similarly, multifactorial analysis of socio-demographic

Table 1. Uni-Factorial Analysis and Distribution of Socio-Demographic Factors

Age	Cases (441)	Controls (323)	OR(95%CI)	Z	P
a. >70 Y	48.0 (10.9%)	57.0 (17.6%)	1.0		
b. 55-69 Y	143.0 (32.5)	94.0 (29.1)	1.8(1.1-2.9)	2.5	0.01
c. 39-54 Y	218.0 (49.2)	113.0 (35.0)	2.3(1.5-3.6)	3.6	<0.01
d. <34	32.0 (7.3)	59.0 (18.3)	0.6(0.4-1.1)	-1.5	0.13
Sex					
a. Male			2.9(1.9-4.5)	5.1	<0.01
b. Female			1.0		
Smoking					
a. Current smoker	236.0 (53.4)	66.0 (20.4)	1.0		
b. Never smoked	63.0 (14.3)	185.0 (57.3)	0.1(0.1- 0.1)	-11.7	<0.01
c. Former smoker	142.0 (32.3)	72.0 (22.3)	0.5(0.4- 0.8)	-2.9	<0.01
Smoking Duration					
a. <10 years	121.0 (27.5)	102.0 (31.6)	1.0		
b. ≥ 10 years	259.0 (58.6)	36.0 (11.1)	6.1(3.9-9.5)	8.1	<0.01
c. Never smoked	61.0 (13.9)	185.0 (57.3)	0.3(0.2-0.4)	-6.4	<0.01

Table 2. Multi-Factorial Analysis of Socio-Demographic Factors

Age	OR(95%CI)	Z	P
a. >70 Y	1.0		
b. 55-69 Y	1.6(1-2.6)	1.9	0.06
c. 39-54 Y	1.9(1.2-3)	2.7	<0.01
d. <34	0.5(0.3-0.9)	-2.5	0.01
Sex			
c. Male	2.6(1.6-4.1)	4.0	<0.01
d. Female	1.0		
Smoking			
a. Current smoker	1.0		
b. Never smoked	0.1(0.1-0.1)	-11.3	<0.01
c. Former smoker	0.6(0.4-0.9)	-2.7	<0.01
Smoking Duration			
a. <10 years	1.0		
b. ≥ 10 years	11.3(6.9-19.3)	9.2	< 0.01
c. Never smoked	0.3(0.2-0.4)	-5.3	<0.01

factors in association with each other was also performed e.g. Age : sex, age: smoking, age: sex: smoking and age: smoking: smoking duration. In the results the category age: sex: smoking (35-54 Y: male: former smokers) were at high risk (OR=5.4 (2.8-10.3). after it in age: smoking: smoking duration category 35-54Y: former smoked: >10 years were also found highly associated with each other (OR=4.3 (3-6.2). in the age: smoking analysis the age group 35-53Y with former smokers were also in association with each other (OR=2.5 (1.5-4.2). the category age: sex (35-54 Y: male) were also found associated (OR=2.1 (1.6-2.9). In the multifactorial analysis to observe association among various factors the age group 35-54 years is highly associated with lung cancer.

Current study has documented smoking, smoker with duration more than 10 years, sex male and 39-54 Y group are at high risk factor in Pakistani population.

Survival Analysis

Overall survival analysis tells us that for the 440 people in the lung cancer dataset, 204 people were uncensored (followed for the entire time, until occurrence of event) and among these 204 people there was a median survival time of 13 months (the median is used because of the skewed distribution of the data). The 95% confidence interval for the median survival time for the 204 uncensored individuals is (12, 18). (Table 4 and Figure 1, Figure 2a)

Association between sex and length of survival

We observed that there were a total of 41 females and 399 male patients. In females, 13 died while in males 191 died. The median follow up was 25 for males with 95% CI=25-NA, and 13 for females with 95% CI=12-17. NA in this case means infinity which is due to the fact that the data is skewed. Statistically significant differences

were observed between length of survival for the males and females ($\chi^2 = 6.1$; p-value = 0.01 on 1 degree of freedom). We also tested to confirm if there is a difference in survival functions between the two groups after adjusting for a potential confounder through a proportional hazard model. Cox regression was applied and the results were more or less similar to the log rank test ($\chi^2 = 7.2$; p-value <0.01). Gender was observed significantly related to survival (p-value <0.01), with better survival in females in comparison to males (HR= 2.0). Cox regression was extended to adjust for the covariate age and was thus tested for any difference when comparing males to females after adjusting for age. Again difference was observed (z=2.5; p-value = 0.02). After adjusting for age, females have significantly better survival in comparison to males. Males have 2.1 times the hazard of dying in comparison to females, adjusting for age (HR<1).

Association between smoking and length of survival.

The data was also analyzed on the basis of smoking groups (current smokers, former smokers and never smoked individuals). The median follow up was 12, 20 and 31 for current smokers, former smokers and non-smokers respectively. According to the results, there were significant variations among the smoking groups $\chi^2 = 15.9$ with a p-value <0.01. The results were also tested after adjusting for a potential confounder, a proportional hazard model was applied. The results are very similar to the log rank test ($\chi^2 = 17.9$ with p-value <0.01). The Cox regression estimates the hazard ratio of dying when comparing smoking groups, smoking either former or current is significantly related to survival (p-value <0.01), with better survival in never smokers in comparison to both current and former smokers (HR= 0.5). Former smokers have 0.6 times the hazard of dying in comparison to never smoked and current smokers have even higher hazard ratios. To extend the cox regression to adjust for age, again comparison among the smoking group was performed after adjusting for age. Differences were recorded again (for former smokers, z = -3.3, p-value = 0.0011; never smoked,

z=-2.3 and p=0.02). After adjusting for age, never smoked have significantly better survival in comparison to both current and former smokers. Never smoked individuals have 0.5 times the hazard of dying in comparison to current

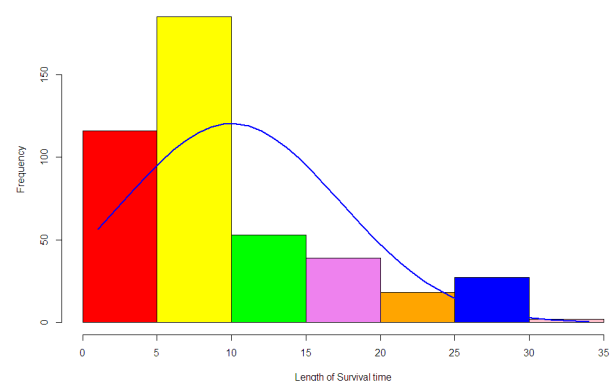


Figure 1: Histogram of Survival Time with Normal Curve in Lung Cancer Patients

Table 3. Multi-Factorial Analysis of Socio-Demographic Factors in Association with Each Other

		OR(95% CI)	Z	P
Age : sex				
1	55-69yrs:Male	1.7(1.2-2.4)	3.0	<0.01
2	35-54yrs:Male	2.1(1.6-2.9)	4.9	0.00
3	<34yrs:Male	0.4(0.2-0.6)	-4.2	0.00
age : smoking				
1	55-69ssyrs: Never Smoked	0.5(0.3-0.9)	-2.6	0.01
2	55-69yrs:Former Smoker	0.7(0.4-1.1)	-1.6	0.11
3	35-54yrs:Never Smoker	0.8(0.5-1.2)	-1.1	0.26
4	35-54yrs:Former Smoker	2.5(1.5-4.2)	3.7	0.00
5	<34yrs:Never Smoked	0.1(0.0-0.2)	-5.1	0.00
6	<34yrs:Former Smoker	1.3(0.5-3.1)	0.6	0.57
age : sex : smoking				
1	55-69yrs:Male:Former Smoker	0.8(0.5-1.4)	-0.7	0.47
2	35-54yrs:Male:Former Smoker	5.3(2.8-10.3)	5.0	0.00
3	<34yrs:Male:Former Smoker	1.3(0.5-3.1)	0.6	0.57
Age : smoking : smoking duration				
1	55-69yrs:former Smoker:>10Y	1.7(1.2-2.4)	2.8	<0.01
2	35-54yrs:former Smoked:>10y	4.3(3.0-6.2)	8.1	0.00
3	<34yrs:former Smoked:>10y	0.9(0.5-1.6)	-0.3	0.73

and former smokers, adjusting for age (HR<1) (Table 4 and Figure 2b).

Association between smoking duration and length of survival.

The data was also analyzed on the basis of smoking duration (smoking duration >10 years, <10 years and never smoked). The median follow up was 15, 12 and 30 respectively. According to log-rank test, there were significant variations among the group ($\chi^2 = 11$ with a p-value <0.01 on 2 degree of freedom). The same was confirmed after adjusting for a potential confounder, a proportional hazard model was applied. When using the Cox regression to perform the test, the results are very similar to the log rank test ($\chi^2 = 13.3$ with p-value <0.01). Smoking duration is significantly associated to survival (p-value < 0.01), with better survival in never smokers in comparison to both either smoking for greater or less than

10 years (hazard ratio of dying for never smoked= 0.4). Smokers of less than 10 years have 0.9 times the hazard of dying in comparison to never smoked and those who have smoked for more than 10 years have even higher hazard ratios. To extend the cox regression to adjust for age, again comparison among the smoking group was performed after adjusting for age. Statistically significant differences were observed (<10 years smokers, z = -0.6, p-value = 0.6; never smoked, z=-3.2 and p=0.0016). After adjusting for age, never smoked have significantly better survival in comparison to the other smokers. Never smoked individuals have 0.4 times the hazard of dying in comparison to the other smoking duration groups, adjusting for age (HR<1) (Table 4 and Figure 2b).

Smoking is the leading determinant of lung cancer in Pakistan. Males are at higher risk and lung cancer risk further increases as the period of exposure to variable increases.

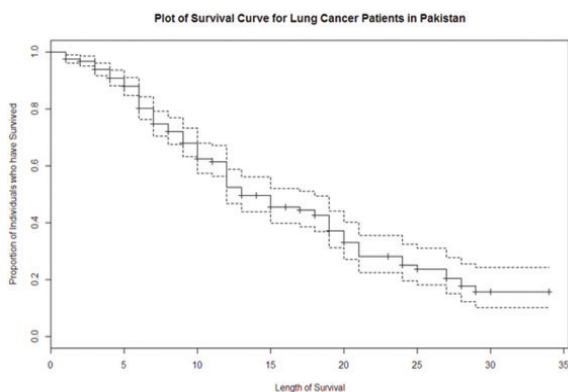


Figure 2 . Kaplan - Meier Plot of Overall Survival of Lung Cancer Patient in Pakistan

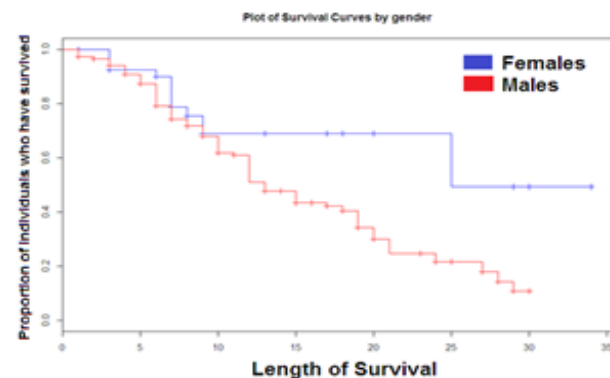


Figure 3 . Kaplan - Meier Plot by Sex of Lung Cancer Patient in Pakistan

Table 4. Survival Analysis on the Basis of Different Risk Factors

S. No	Analysis Type	Records	Events	Median	95% LCI	95% UCI
Overall survival						
1.		440.0	204.0	13.0	12.0	18.0
Survival analysis on the basis of sex.						
2.	Males	399.0	191.0	13.0	12.0	17.0
3.	Females	41.0	13.0	25.0	25.0	NA
Survival analysis on the basis of smoking.						
4.	Current Smokers	236.0	139.0	12.0	12.0	13.0
5.	Former Smokers	142.0	50.0	20.0	18.0	24.0
6.	Never Smoked	62.0	15.0	31.0	29.0	NA
Survival analysis on the basis of smoking duration.						
7.	<10 Years	121	56	15	12.0	21.0
8.	>10 Years	257.0	134.0	12.0	12.0	17.0
9.	Never Smoked	62.0	14.0	30.0	29.0	NA

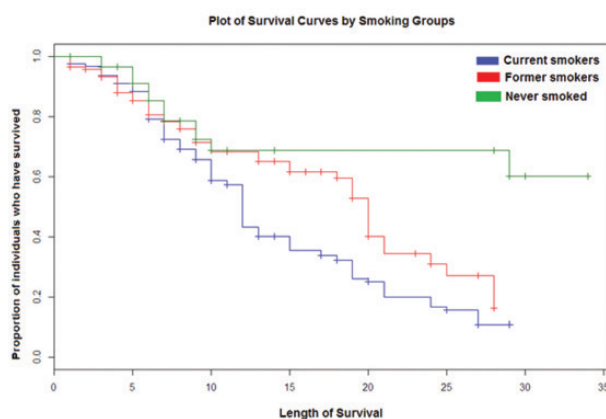


Figure 4 . Kaplan - Meier Plot of by Smoking of Lung Cancer Patient in Pakistan

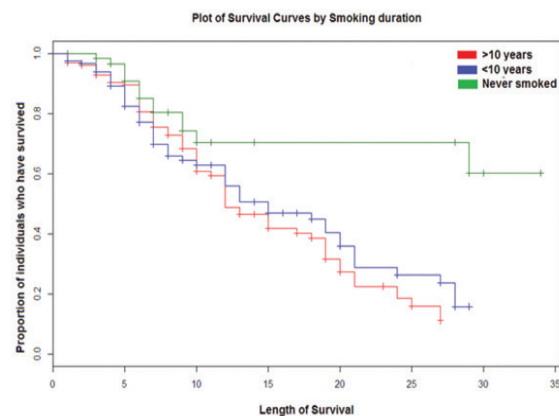


Figure 5 . Kaplan - Meier Plot of by Smoking of Lung Cancer Patient in Pakistan

References

- Ahmed K, Abdullah AIE, Jesmin T, et al (2013). Early detection of lung cancer risk using data mining. *Asian Pac J Cancer Prev*, **14**, 595-8.
- Alam, SE (1998). Prevalence and pattern of smoking in Pakistan. *Journal of Pakistan. J Pak Med Assoc*, **48**, 64-6.
- Bailey WJE, Amos CI, Pinney SM, et al, (2004). A major lung cancer susceptibility locus maps to chromosome 6q23-25. *Am J Hum Genet*, **75**, 460-74.
- Bhaskarapillai B, Kumar SS, Balasubramanian S (2013). Lung cancer in malabar cancer center in kerala- a descriptive analysis. *Asian Pac J Cancer Prev*, **13**, 4639-43.
- Bruske HI, Mohner M, Pohlabein H, et al (2000). Occupational lung cancer risk for men in Germany, results from a pooled case-control study. *Am J Epidemiol*, **151**, 384-95.
- Ferreccio C, Gonzalez C, Milosavljevic V, et al (2000). Lung cancer and arsenic concentrations in drinking water in Chile. *Epidemiology*, **11**, 673-9.
- Ganesh B, Sushama S, Monika S, et al (2011). A case-control study of risk factors for lung cancer in Mumbai, India. *Asian Pac J Cancer Prev*, **12**, 357-62.
- Hashibe M, Siwakoti B, Wei M, et al (2010). Socioeconomic status and lung cancer risk in Nepal. *Asian Pac J Cancer Prev*, **12**, 1083-8.
- Hussain MR, Khan SA, Bukhari MH (2009). Role of smoking in primary pulmonary malignancies in central punjab. *Annals*, **17**, 18-21.
- International Agency for Research on Cancer, IARC (2012). diesel engine exhaust carcinogenic (press release no 213), Lyon, France.
- Kwong SL, Chen MS, Snipes KP, et al (2005). Asian subgroups and cancer incidence and mortality rates in California. *Cancer*, **12**, 2598-975.
- Lam WK (2005). Lung cancer in Asian women, the environment and genes. *Respirology*, **4**, 408-17.
- Matteis DS, Consonni D, Pesatori AC, et al (2013). Are women who smoke at higher risk for lung cancer than men who smoke. *Am J Epidemiol*, **7**, 601-12.
- Rennert G, Rennert HS, Epstein L (1982). Lung cancer histology in Jews and Arabs in Israel. *Am Rev Respir Dis*, **143**, 721-6.
- Stellman SD, Takezaki T, Wang L, et al (2001). Smoking and lung cancer risk in American and Japanese men: an international case-control study. *Cancer Epidemiol Biomarkers Prev*, **11**, 1193-9.
- Sun S, Schiller JH, Gazdar AF (2007). Lung cancer in never smokers, a different disease. *Nature Rev*, **7**, 778-90.
- Tas D, Okutan O, Kaya H, et al (2008). Analysis of 138 cases of lung cancer in a training hospital compared to the data of lung cancer cases diagnosed ten years previously. *Marmara*

Med J, **213**, 231-7.

Wahab M, Boroumand N, Castro C, et al (2007). Changing trends in the distribution of the histologic types of lung cancer, a review of 4,439 cases. *Ann Diagn Pathol*, **112**, 89-96.

Wahbah M, Boroumand N, Castro C, et al (2007). Changing trends in the distribution of the histologic types of lung cancer: a review of 4,439 cases. *Ann Diagn Pathol*, **11**, 89-96.