Epidemiological analysis of Lung cancer in Erbil province of Iraqi Kurdistan: Incidence, Survival, Relative Risk Ratio, and Treatment Regimes in males and females

Kurdistan Ibrahim Mawlood  
Salahaddin University-Erbil

Dler Hussein Kadir (dhkadir2@gmail.com)  
Salahaddin University-Erbil

Waleed Al-Murrani  
University of Plymouth

Research Article

Keywords: Lung cancer, Epidemiological analysis, Relative Risk Ratio, Treatment Regimes

Posted Date: February 2nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1280286/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

This study compares survival function, relative risk, incident rates and treatment regime between genders. A total of 590 cases of Lung Cancer admitted to Nanakali hospital, Erbil province of Iraqi Kurdistan, were collected for 5 years period between 1st January 2013 to 31st of December 2017. The follow-up of the cases continued till the 1st of April 2018 to complete the record. Chi-square, correlation, relative risks and basic exploratory data analysis were carried out. Simple linear regression was carries out for number of lung cancer among males and females. A multivariate Cox-regression model was used to determine the prognostic factors for lung cancer patients. Pearson's Correlation Coefficient (r) for total cases of Lung cancer (Males and Females) was equal to (0.875 ± 0.033 with P= 0.044) and R-square (Precision) of 0.766. The Prediction Regression equation that Female Lung Cancer (F) = 11.79 + 0.6714 Age group. This means any age group can be selected to predict for expected incidence. The prediction equation is that Male Lung Cancer cases = - 2.857 + 0.7690 Age group. The Regression coefficient is + 0.769 per 10 years of age and it is highly significant (P=0001).

The result of multivariable cox regression model indicates that gender had no influence on survival outcome (HR ~= 0.81, 95%CI: 0.56 to 1.16, p=0.0.247). However, taking surgery and immune system are statistically significant prognostic factors for lung cancer patients. The model indicated that the risk of mortality increases by 92% if lung cancer patients do not take surgery (HR ~= 1.92, 95%CI: 0.31 to 0.97, p=0.039). Furthermore, the risk of mortality is reduced by 44% among those patients who took immune system. The study concludes that female patients survive longer than males and median survival probability in female is greater than male lung cancer patients. Taking surgery and immune system are statistically significant prognostic factors for lung cancer patients.

1. Introduction

Cancer is a wide term that encompasses more than 270 distinct kinds of cancer illness. Several stages of cancer have been discovered, suggesting that many gene mutations are involved in cancer ethology. The aberrant cell growth is caused by these gene alterations. The increase of cell proliferation is aided by genetic abnormalities induced by heritance or hereditary factors. Extra data has been acquired with the help of technical advancements in bioinformatics and molecular methods, which may be helpful for early diagnosis and appropriate therapy, [Fisher et. al. 2016, Aizawa et al. 2016, Poon et al. 2014]. The effects of medicines on cancer patients may be predicted and even managed in certain cases. Molecular genetic research has discovered cancer pathways in recent years. The findings of these research contributed to a better understanding of the role of genetic abnormalities in the development of cancer. The goal of this research was to look at the molecular features of cancer, [Antwi et al. 2015, Shtivelman et al. 1985, Cigudosa et al. 1999].

Cancer is a broad term for a variety of illnesses defined by the uncontrolled division of aberrant cells with the potential to infiltrate and destroy normal body tissue. Cancer has a high proclivity for spreading throughout the body. Changes (mutations) in the DNA of cells are the cause of cancer. Each gene in a cell
includes a collection of instructions that teach the cell what tasks to execute as well as how to grow and divide, [Wood et al. 2001, Alvarez-Buylla et al. 2008, Goelz et al. 1985]. Errors in the instructions may cause a cell to cease functioning normally and even lead it to become malignant. The vast majority of malignancies strike individuals with no recognized risk factors. Age, lifestyle, family history, health problems, and the environment are all known to raise your cancer risk, [Fraga et al. 2005, King et al. 1985].

Cancer is a leading cause of illness and death globally, affecting both men and women of all ages. According to the World Health Organization (WHO), an estimated 10 million fatalities would occur worldwide in 2020, with cancer accounting for one out of every six deaths (excluding deaths caused by conflicts). Also, it was estimated that about 70 percent of Cancers occur in low-and middle-income countries. Tobacco is the most significant risk factor for cancer, and lung cancer is the most frequent disease in males (16.7 percent), while breast cancer is the most common cancer in women (34.2 per 100,000), [Heinrich et al. 2002, Thomas et al 2007, Roninson et al. 2002].

The lungs are two sponge-like organs in the human body. The lobes of your right lung are divided into three parts. There are two lobes in your left lung. Because the heart occupies more space on the left side of the body, the left lung is smaller. When you breathe in, air enters via your mouth or nose and travels down the trachea to your lungs (windpipe). The trachea is divided into bronchi tubes, which enter the lungs and split into smaller bronchi tubes. These split into bronchioles, which are smaller branches. Alveoli are small air sacs that sit at the end of the bronchioles, [Muller et al. 2014, Esteller et al. 2007, Doi et al. 2009].

When you inhale, the alveoli take oxygen into your blood and through the exhalation they eliminate CO₂. The primary tasks of your lungs are to take in oxygen and expel carbon dioxide. Lung malignancies begin in the cells that line the bronchi and other regions of the lung, such as the bronchioles and alveoli. The pleura is a thin lining layer that surrounds the lungs. Your lungs are protected by the pleura, which allows them to move back and forth against the lung wall as they expand and relax throughout respiration, [Espada et al. 2007, Dalgliesh et al. 2010, Noonan et al. 2009, Mulero-Navarro et al. 2008].

the diaphragm divides the Lung from the abdomen and it travels up and down as you breathe and pushes air into and out of the lungs, Figure 1.

The cells in our chest as well as other areas of our bodies normally undergo a growth and death cycle that keeps the number of cells in control. Any kind of cancer arises when a series of particular changes, known as mutations, occur in a previously healthy cell. Uncontrolled cell division may result in an excessive number of cells when a collection of mutations alters genes in ways that disrupt the normal development and death cycles of cells. When the gas pedal becomes stuck or the brakes fail, the cells continue to divide with nothing to stop them. A tumour, neoplasm, or lesion is a mass formed by mutated and abnormally growing cells. On a Lung X-ray or CT scan, this tumour may be identified as a nodule that is either malignant (cancer) or benign (non-cancerous), [Sporn et al. 2009, Thun et al. 2008, Proctor et al. 2012]. The mass may be benign or cancerous, as seen in Figure 2.
The tumour is deemed malignant when the tumour cells are able to infiltrate normal tissues. Lung cancer is defined as a tumour in which the malignant cells originated in the lungs. Metastasis is the spread of cancer from one area of the body to another, and metastases are the tumours produced by the cancer cells that have spread. Lung cancer metastases may spread to lymph nodes surrounding the lungs, as well as to other organs such as the bones, adrenal glands, and the brain, through the circulation. Cancer may begin in other areas of the body and then move to the lungs, [Osmani et al. 2018, Horn et al. 2008, Saunders et al. 1997, Miller et al. 1969]. This is referred to as metastasis of the initial malignancy, not lung cancer. Lung cancer is defined as cancer that begins in the lungs, Figure 3.

Beside the wide distribution of cancer incidence overall the world, Kurdistan Region in Iraq has been exposed to several carcinogenic hazards. Although, the few reports about the increased risk of cancer in different cities in Iraq were present during the recent years, but these reports did not cover Kurdistan region. This is while that cancer incidence rate and possible risks of cancer in this region is considerably increasing. Among the cities of Kurdistan region (KR), Erbil province covers most of the KR population and cancer incidence at diverse types. Through this study, we focused on the epidemiological analysis of Lung cancer as one of the most type of cancer incidence in Erbil province according the reports of hospitals and Erbil cancer society and investigated its incidence, survival, relative risk ratio, and treatment regimes in males and females through the study of a wide spectrum of the people with lung cancer, [Al-Janabi. Et al. 2017, Alhalfi. Et al. 2016].

Rashid et al 2017 analysed 611 of males and female cases of Lung Cancer recorded during 2011-2015 in Sulaimani province. They found that the ASR equals 6.3 per 100000 of the population per year and also lung cancer incidence rate is highest in over 70 age group and is lowest in age group 20-29 for males and females. Lung cancer constituted 78.15% among smokers. Mawlood et al (2019) used two Survival models (Cox-Proportional Hazard and Accelerated Failure Time Models) and two survival tests (Kaplan Meier and Log Rank test) to detect the significant factors effecting on Lung cancer. They showed significant difference between levels of treatments, namely Surgery, Radio, Age and Gender. They also found significant difference in survival time, between patients having surgery and radiotherapy. Khoshnaw et al (2015) reported that the annual incidence rate at all age groups in Sulaimani rose from 38 to 61.7 cases per 100000 of the population, [Rashid. Et al. 2018, Mawlood. et al. 2019, Khoshnaw. et al, 2015].

Habib et al (2016), reported on lung cancer in Basrah, south of Iraq for 2005-2012 and demonstrated that lung cancer is more prevalent in males (75.9%) than in females (24.1%). incidence of lung cancer increased exponentially with age. Mathatra et al (2016) reported that Lung cancer accounts for an estimated 27% of the total cancer deaths in USA (2015) and 20% in EU in 2016, [Habib. Et al. 2016, Mathatra. et al. 2016].

Through the current study, the results revealing some more useful indicators and statistical diagnostic tools in the fight against cancer. It is also a retrospective study which analysed cancer registry data in Erbil from 2013 to 2017.
2. Materials And Methods

2.1 Data collection

The data used in the manuscript as indicated in supporting information were taken from the Department of Lung cancer, Nanakali hospital, Erbil-Iraq, (https://www.facebook.com/NANAKALI-HOSPITAL-144316382578149/). These data include 590 cases of Lung Cancer admitted to Nanakali hospital collected for 5 years period between 1st January 2013 to 31st of December 2017.

2.2 Method

A total of 590 cases of Lung Cancer admitted to Nanakali hospital, mostly from Erbil province, Kurdistan, were collected for 5 years period between 1st January 2013 to 31st of December 2017. The record also contains few lungs cancer cases from Baghdad, Diyala and Anbar living in Erbil. The follow-up of the cases continued till the 1st of April 2018 to complete the record. 462 cases died (D) and 128 cases remained alive (A) up to the end of the follow-up date bringing death rate to 85% of all lung cancer cases in this study. Ethical approvals are granted by the hospital, to use the record, without names of patients.

Age at diagnosis of positive lung cancer is recorded in years, and survival period is calculated in months. The analysed record also includes treatment regime in each case, in males and females, and the overall outcome, dead or alive, is also recorded for both sexes.

Treatment regime followed the main worldwide factors to treat cancer cases, which are Surgery (S), Chemotherapy (C), Radiotherapy (R), Immunotherapy (I), Hormone (H) and combinations of two or more of mentioned factors, Scheme 1.

All statistical analysis is carried using STATA software version 12 and include: Descriptive statistics of all parameters, age group distribution of Erbil population (2017) for every 10 years (10, 20, 30, 40, 50, 60, 70, 80+) as in Table 1 with percentages for males and females, age distribution of lung cancer patients (Table 2), Anderson-Darling test of normality and of each age group (every ten years) and Percentiles and Probability distribution of Survival data with Anderson-Darling test of normality.

Measures of association of occurrence of Lung cancer cases with age groups of patients, for males, females and the total, were calculated and presented in terms of Pearson’s Correlation Coefficient (r) and Regression analysis and Regression Coefficients (b). The Prediction equations for the occurrence of Lung cancer at any age was also calculated. This will help health planners in predicting Lung cancer frequency at any age.

Relative Risk Ratio (RRR) ratio is also calculated, for females/males, at different age groups as probability of occurrence of Lung cancer in females over the probability of occurrence in males.

The Chi-Square test of significance is performed to compare the relationship between study variables in males and females. The chi-square compares Observed (O) with Expected (E) frequencies and a
3. Results And Discussion

3.1. Age groups distribution of Erbil population in 2017

The distribution percentage of Erbil population of 2,113,391 in 2017 by age group of tens and gender was calculated and shown in Table 1. The figures in this table are used for the calculation of measures of association (correlations and regressions). The age of about 25% of the population of females and males falls within the age group up to 20 years and that less than 7% falls over 50 years. This relation partly explains why the relative higher numbers of Lung Cancer cases occurred in the over 60 age groups.

| Age groups 10 Y | Female | Male |
|-----------------|--------|------|
| %               | Number | %    | Number |
| 10              | 11.8   | 127587 | 12.3 | 132993 |
| 20              | 10.7   | 115694 | 11.6 | 125425 |
| 30              | 8.6    | 92097  | 8.9  | 96231  |
| 40              | 7.1    | 76769  | 6.8  | 73526  |
| 50              | 5      | 54062  | 4.8  | 27572  |
| 60              | 3.3    | 35682  | 2.5  | 27031  |
| 70+             | 3.5    | 37844  | 3.3  | 38925  |
| Total           | 50%    | 1056695 | 50% | 1056695 |

3.2. Age distribution and descriptive statistics of the 590 Lung cancer cases

Age distribution of the 590 Lung Cancer cases, from 2013 to 2017 and the Percentiles is given in Figure 1. Mean age at diagnosis is 51.17 ± 0.823 with a Standard Deviation (SD) which describes variability, of 19.98 year and the Q1 (first Quartile) is 40.00 year. The Median is 54 years with minimum age of 1.0 year and maximum of 99.0 years. Age at diagnosis in years is tested for normality using Anderson-Darling test of normality and shown in Figure 1. Anderson Darling test value is 4.454 with a P ≤ 0.005. This indicates age in years of the 590 Lung cancer cases is not normally distributed. This is the reason we adopted the Chi-Square non-parametric test of significance in the analysis of results.

3.3. Confidence interval and limits for age of the 590 Lung cancer cases

Mean age at diagnosis equals 51.17 years with a SD of 19.98. To reflect on the spread of the data of age at diagnosis, the 95% Confidence Interval and Limits of Lung Cancer mean are calculated and found that
the Upper Limit = 90.33 and the Lower Limit = 12.00 Year. This Confidence Interval and Limits includes 95% of the cases, the 99% Confidence Interval is wider and effectively covers all age groups. Falling within a certain Confidence Interval and Limits help in reasoning when taking the history from patients. The overall Crude Incidence Rate Standardized per 100000 of the population in 2017 is 27.28 per 100000, it was 24.03 in males and 31.79 in females. Gender Percentage Ratio in the 590 Lung Cancer cases: Males: Females = 43.05 %: 56.95% (254: 336 cases). Percentile 50 is about 50 years which nearly equals the mean, Figure 4.

3.4. Measures of Association between occurrence of Lung cancer and age groups:

For the purpose of calculating Pearson's Correlation Coefficient (r) and the Regression Coefficient (b) as measures of association between age and incidence of Lung Cancer cases, in males and females, data are sorted according to age groups, Table 2.

Table 2: Age groups in 10’s and the number of Lung cancer cases in females and males.

| Age group | Male | Female | Total |
|-----------|------|--------|-------|
| 10 years  | 13   | 14     | 27    |
| 20 years  | 12   | 13     | 25    |
| 30 years  | 16   | 20     | 36    |
| 40 years  | 22   | 48     | 70    |
| 50 years  | 28   | 74     | 102   |
| 60 years  | 46   | 69     | 115   |
| 70 years  | 59   | 66     | 125   |
| 80+ years | 58   | 32     | 90    |
| Total     | 254  | 336    | 590   |

3.5. Correlation Coefficients (r)

Correlation Coefficients are qualitative measures of association and runs from (+1 to -1) according to the nature and direction of association. Data in Table (2) are used to calculate Pearson's Correlation Coefficient (r) for total cases of Lung cancer (Males and Females) and found to equal (0.875 ± 0.033 with P= 0.044) and R-square (Precision) of 0.766 (Any below 80% should be taken with caution in scientific work). Female's Correlation (0.643 ±0.052 with P=0.085) pulled down by the number of Lung cases in the age group 80+, yet, it is intermediate. Pearson's Correlation Coefficient (r) between Age in years and occurrence of Cancer in males is very high (r = 0.953 ± 0.129) and significant at P = 0.0002. It should be mentioned that the association with the increase in age group is very clear up to age of 80 years and over,
where the number of recorded cases fell to only 90 cases in total. This might be an error of reporting particularly on the female side where the over 80 female lung cancer cases dropped to only 32 cases compared to 66 cases in age group 70. Correlations Coefficient, in both males and females are positive and confirm the fact that cancer incidence grows with age and covers all ages from birth to death.

3.6. Regression coefficient of number of lung cancer cases (Y) on age groups (X)

To quantify the relation between Lung cancer cases and age groups, Table 2, a Regression analysis and calculation of the Regression Coefficients (b) of Y (number of cases) and X (age groups) for both males and females separately and for the total cases was conducted. The Prediction Equation was also calculated which enables planners to estimate the expected incidence of Lung cancer at any age. This, of course, may be added to the jigsaw of diagnosis process.

3.7. Regression of female lung cancer cases on age

Regression Analysis of Female cases on Age is conducted and is shown below and in Figure 5. The Prediction Regression equation that Female Lung Cancer (F) = 11.79 + 0.6714 Age group. This means any age group can be selected to predict for expected incidence.

**Table 3**: Analysis of Variance of female lung cancer

| Source     | DF | SS      | MS    | F    | P     |
|------------|----|---------|-------|------|-------|
| Model      | 1  | 1893.4  | 1893.4| 4.24 | 0.085 |
| Residual   | 6  | 2680.6  | 446.76|      |       |
| Total      | 7  | 4574    |       |      |       |

Female Regression Coefficient of number of Lung cancer cases and age in years equals + 0.6714 case/10 years which is marginally significant at P= 0.085. The reason is as already mentioned when calculating the Correlation coefficient for females. This means that Lung cancer in females is on rise with years, Table 3.

3.8. Regression Analysis of Males Lung cancer cases on age

Association as Regression coefficient and the Prediction Equation for males Lung cancer on age is calculated and shown in the ANOVA table (Table 4) and Figure 6. The prediction equation is that Male Lung Cancer cases = - 2.857 + 0.7690 Age group. The Regression coefficient is + 0.769 per 10 years of age and it is highly significant (P=0001).

**Table 4**: Analysis of Variance for male cancer patients.
This means that Lung Cancer cases linearly increased with age. The Prediction equation, which assumes linearity, can be used to predict Lung Cancer occurrence at any age. This regression analysis and results are trustworthy with an R-Square of about 90%.

3.9. The Regression of Total Lung Cancer Cases (Males and Females) on Age Groups:

Regression analysis for data in Table (2) and Figure (4) showed that the Regression equation is that Total lung cancer cases = 8.93 + 1.440 Age group. Regression Coefficient (b=1.44) is highly significant (P = 0.004) and estimated with precision (R-square = 76.50%). This means, with this linear relationship between the occurrence of Lung Cancer and age we can predict the level of incidence will increase by 1.44 every ten years. Figure 7 shows Linear regression line of total Lung cancer cases on age groups and the prediction equation.

3.10. Exploratory analysis

Basic exploratory of the data was performed to evaluate the covariate distribution in uncensored/ censored patient and in each ethnic group. It can be noticed that 42.4% of censored patients were male while 57.6% of censored were female patients. Table 5 indicates that more female patients were died than male lung cancer patient.

Table 5: Distribution of covariates in the study population by gender
| Variables          | Male          | Female         | P-value* |
|-------------------|---------------|----------------|----------|
|                   | n (%)         | n (%)          |          |
| Surgery           |               |                |          |
| Made surgery      | 172 (67.7)    | 52 (76.47)     | 0.07     |
| Does not make surgery | 82 (32.3)   | 14 (20.59)     |          |
| Radio             |               |                |          |
| Took Radiotherapy | 73 (28.7)     | 135 (40.2)     | 0.004    |
| Does not take Radiotherapy | 181 (71.3) | 201 (59.8)     |          |
| Chemotherapy      |               |                |          |
| Injected Chemotherapy | 220 (86.6) | 304 (90.5)     | 0.141    |
| Does not inject Chemotherapy | 34 (13.4) | 32 (9.5)       |          |
| Hormone           |               |                |          |
| Used hormone      | 38 (15.0)     | 91 (27.1)      | <0.001   |
| Does not use hormone | 216 (85.0) | 216 (72.9)     |          |
| Immune system     |               |                |          |
| Took immune system| 23 (9.1)      | 26 (7.4)       | 0.566    |
| Does not take immune system | 231 (90.9) | 310 (92.3)     |          |
| Statue            |               |                |          |
| Dead              | 58 (22.8)     | 70 (20.8)      | 0.559    |
| Alive             | 196 (77.2)    | 266 (79.2)     |          |
| Age at diagnose   |               |                |          |
| Less than 20 years| 25 (9.8)      | 27 (8.0)       | <0.001   |
| 20 to 39          | 38 (15.0)     | 68 (20.2)      |          |
| 40-59             | 74 (29.1)     | 143 (42.6)     |          |
| More than 60 years| 117 (46.1)    | 98 (29.2)      |          |

*Chi-square tests were performed for categorical variables

It our interest to investigate the relationship between variables of interest and gender. Chi-square test revealed that each variable of radiotherapy, taking hormone and age at diagnose has a statistically significant relationship with gender since their p-values are less than 0.05. The above table illustrates that there were association between gender and radiotherapy (p= 0.004), mode of gender and Hormone (p < 0.001) and gender and Age at diagnose (p < 0.001). The other remained categorical covariates were not statistically significant.

Figure 8 shows that female patients survive longer than male. However, the confidence intervals illustrate that the uncertainty is greater in survival curve for Female patients. It also indicates that the median survival probability is greater than male lung cancer patients.

**Table 6**: Multivariate Cox regression modelling for each covariate.
Table 6 shows the result of multivariable cox regression model which indicates gender had no influence on survival outcome (HR = 0.81, 95% CI: 0.56 to 1.16, p=0.247). However, taking surgery and taking immune system are statistically significant prognostic factors for lung cancer patients. The model indicated that the risk of mortality increases by 92% if lung cancer patients do not take surgery (HR = 1.92, 95% CI: 0.31 to 0.97, p=0.039). Furthermore, the risk of mortality is reduced by 44% among those patients who took immune system. The proportional hazard assumptions were checked to investigate whether the hazard ratio is approximately proportional using both Log-log plot and a plot of log hazard ratio over time (i.e., the ratio of PH is approximately constant over time or whether it is time-dependent).

Log-log plot of left-hand side in figure 9 for surgery shows that the hazards are approximately proportional since the difference between two lines are approximately constant. The plot of log hazard ratio with time shows the approximate straight lines illustrating that hazard ratio remains constant with time.

As for immune system, the hazards are not proportional since the difference between two lines are not constant. Also, the plot of log hazard is not constant over time. Thus, the proportionality hazard assumption for immune system is not met.

3.11. The outcome of the most common treatment regime: Surgery, Chemotherapy and Radiotherapy alone and in combination

Data are sorted for repetitions of main three treatment regime alone and in combination and put in Table (4). This is to show, what was the treatment regime (s) preferred to treat Lung cancer in the 590 cases of
Lung cancer in Erbil.

**Table 7.** The repeated usage of the 3 treatment regimens and combinations for the 590 Lung cancer cases from 2013 to 2017.

| Treatment     | Male | Female | Total | Percent |
|---------------|------|--------|-------|---------|
|               | Dead | Alive  | Dead  | Alive   |         |
| Surgery plus  | 22   | 156    | 39    | 218     | 435     | 73.7   |
| Chemo plus    | 36   | 189    | 51    | 256     | 532     | 90.16  |
| Radio plus    | 12   | 60     | 18    | 116     | 206     | 34.9   |
| Total treatment | 70  | 405    | 108   | 590     | 1173    |        |
| Percentage    | 5.97 | 34.52  | 9.2   | 50.3    |         |

Table 7 shows that the total treatments offered to the 590 patients is 1173 with an average of 2 treatments per patient. The table also shows that the treatment of choice is Chemotherapy, in combination with other regimes. Its usage constituted 90.16% of all treatments conducted on the 590 cases followed by surgery alone and in combination with 73.70% and the least used is Radiotherapy alone and in combination with 34.90%.

The results also indicated that Surgery and combination succeeded to keep alive 35.80% males and 49.43% females, Chemotherapy and combination helped to keep alive 35.53% males and 48.12% females while Radiotherapy and combination protected 29.13% males and 56.31% females. Total percentage of (success) in keeping life is about the same (85%).

**3.12. Relative Risk (RR) Ratio of Lung Cancer for Females versus Males for 10 years age groups (10 to 70 years and over):**

Relative Risk Ratio (RRR) measures the strength of association or relation. In this analysis, the concept of calculation of RR slightly modified to express the Probability of occurrence of Lung cancer in females to that in males. The RR ratio shown in Table (5) expresses which gender is at a higher risk of having lung cancer and at what age group?.

**Table 8.** Relative Risk (RR) ratio of female Lung cancer incidence versus male Lung Cancer
| Age Groups | Female | Male | Relative Risks F / M |
|------------|--------|------|---------------------|
| 10 years   | 14     | 13   | 0.81                |
| 20         | 13     | 12   | 0.82                |
| 30         | 20     | 16   | 0.95                |
| 40         | 48     | 22   | 1.64                |
| 50         | 74     | 28   | 1.99                |
| 60         | 69     | 46   | 1.30                |
| 70+        | 98     | 117  | 1.14                |
| RR Total   | 336    | 254  | 1.32                |

It is clear that Males are at higher risk than Females to have Lung cancer up to the age of 40 then the trend changed direction where Females show higher risk than Males from the age 40 onwards, it doubled at the age group of 50-60.

4. Conclusions

During this study, it was found that the epidemiological data are not different from similar data on lung cancer reviewed. This is expected on light of similarity in environmental influences, socio-economic and genetic background of the populations studied. Also, the treatment of choice is chemotherapy, alone and in combination with other regimes constituted 90.16%, surgery alone and in combination with 73.70% and the least used is radiotherapy alone and in combination with 34.90%. Further, female patients survive longer than males and median survival probability in female is greater than male lung cancer patients. Besides the above results, we showed that surgery and immune system are statistically significant prognostic factors for lung cancer patients. The Relative Risk Ratio (RRR) revealed that among the gender and age groups, males are at higher risk than females to have Lung cancer up to the age of 40 then the trend changed direction where females show higher risk than males from the age 40 and it doubled at the age group of 50-60. This might be attributed to early smoking in males and age effect in females. Lung cancer incidence grows with age and covers all ages and both sexes from birth to death.

Moreover, it was confirmed that, the prediction regression equation of the total lung cancer cases (males and females) (Y) on age groups (years) (X) is $8.93 + 1.440 \text{ age group}$, female Lung Cancer (F) = $11.79 + 0.6714 \text{ age group}$, and male lung cancer cases = $-2.857 + 0.7690 \text{ age group}$ estimated with precisions ($R^2 = 76.50\%$, ($R^2 = ? \%$) and ($R^2 = 90\%$) respectively. This means that Lung Cancer cases linearly increased with age. Also, using only chemotherapy led to significantly higher deaths in males ($P \leq 0.05$) and lower alive (A) in females though not significant ($P > 0.05$). The other significant results are in the use of the combination of the four treatment regimes. This also led to significantly less alive (A) in males ($P \leq 0.01$) but more (A) in females indicating a higher success rate in saving lives in females ($P \leq 0.01$).
Declarations

- Ethical Approval and Consent to participate

Not applicable

- Consent for publication

Not applicable

- Availability of supporting data

Not applicable

- Competing interests

Not applicable

- Funding

Not applicable

- Authors’ contributions

All authors collaborated in all steps of manuscript preparations including data collection, data analysis, data interpretation and manuscript writing

- Acknowledgements

We appreciate Nanakali Hospital-Erbil and Salahaddin University for partial support of this work

- Authors' information

Dr. Kurdistan Ibrahim Mawlood: Professor of Statistics at department of Statistics, College of Administration and Economics, Salahaddin University-Erbil, Kurdistan Region - F.R. Iraq., Dr. Dler Hussein Kadir is lecturer of Statistics at department of Statistics, College of Administration and Economics, Salahaddin University-Erbil, Kurdistan Region - F.R. Iraq and also College of Administration and Economics, Cihan University-Erbil, Iraq., Dr. Waleed Al-Murrani is professor of biology at School of biological sciences, University of Plymouth/UK.

-Human Ethics statements

During this study we studied a large spectrum of 590 cases of Lung Cancer admitted to Nanakali hospital, Erbil province of Iraqi Kurdistan. All of our investigations and assessments were along with satisfaction of the mentioned ills with respect to the ethical rules and under supervision of Nanakali hospital.
-Animal Ethics statements

Not applicable because no animal investigated in our study.

-Data availability statement

The data used in the manuscript as indicated in supporting information were taken from the Department of Lung cancer, Nanakali hospital, Erbil-Iraq, (https://www.facebook.com/NANAKALI-HOSPITAL-1443163382588149/). These data include 590 cases of Lung Cancer admitted to Nanakali hospital collected for 5 years period between 1st January 2013 to 31st of December 2017.

References

1. Aizawa K, Liu C, Tang S, et al. Tobacco carcinogen induces both lung cancer and non-alcoholic steatohepatitis and hepatocellular carcinomas in ferrets which can be attenuated by lycopene supplementation. International journal of cancer. 2016, 139, 1171-81.
2. Alhalfi. HS. New Epidemiological Data on Lung Cancer in Iraq people. Pulmonology Research and Respiratory Care. 2016, 1(1), 2575-9776.
3. Al-Janabi. AH, Naseer. ZI, Hamady. T. Epidemiological Study of Cancer in Iraq-Kerbala province from 2008-2015. International J of Medical Research and Health Sciences 2017: 6(1), 79-86.
4. Alvarez-Buylla ER, Chaos Á, Aldana M, et al. Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. Plos one. 2008, 3, e3626.
5. Antwi SO, Eckert EC, Sabaque CV, et al. Exposure to environmental chemicals and heavy metals, and risk of pancreatic cancer. Cancer Causes & Control. 2015, 26, 1583-91.
6. Cigudosa JC, Parsa NZ, Louie DC, et al. Cytogenetic analysis of 363 consecutively ascertained diffuse large B-cell lymphomas. Genes, Chromosomes and Cancer. 1999, 25, 123-33.
7. Dalglish GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature. 2010, 463, 360-3.
8. Doi A, Park I-H, Wen B, et al. Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nature genetics. 2009, 41, 1350-3.
9. Espada J, Ballestar E, Santoro R, et al. Epigenetic disruption of ribosomal RNA genes and nucleolar architecture in DNA methyltransferase 1 (Dnmt1) deficient cells. Nucleic acids research. 2007, 35, 2191-8.
10. Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. Human molecular genetics. 2007, 16, R50-R9.
11. Fisher R, Pusztai L, Swanton C. Cancer heterogeneity: implications for targeted therapeutics. British journal of cancer. 2013, 108, 479-85.
12. Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nature genetics. 2005, 37, 391-400.

13. Goelz SE, Vogelstein B, Hamilton SR, Feinberg AP. Hypomethylation of DNA from benign and malignant human colon neoplasms. Science. 1985, 228, 187-91.

14. Habib. OS, Al-Sadi. JN, Hussein. OG. Lung Cancer in Basrah during 2005-2012. Saudi Med J. 2016, 37(11), 1214-1219.

15. Heinrich MC, Blanke CD, Druker BJ, Corless CL. Inhibition of KIT tyrosine kinase activity: a novel molecular approach to the treatment of KIT-positive malignancies. Journal of clinical oncology. 2002, 20, 1692-703.

16. Horn L, Johnson DH. Evarts A Graham and the first pneumonectomy for lung cancer. J Clin Oncol. 2008, 26 (19), 3268-3275

17. Khoshnaw. N, Mohamed. HA, Abdulla. DA. Asian Pac. J. Cancer Prev. 2015, 10 (18) 8625-31.

18. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science. 1985, 229, 974-7.

19. Mathatra. J, Malvezzi. M, Nrgri. E, Lavecci. C, Bofella. P. Risk Factors for Lung Cancer Worldwide. European Respiratory Journal. 2016, 48, 889-902.

20. Mawlood. KI, Obed. SA. Study and Analysis of the Chest Cancer Data Using Survival Models. QALAAI ZANIST SCIENTIFIC JOURNAL. 2019, 4 (2), 1-22.

21. Miller AB, Fox W, Tall R. Five-year follow-up of the Medical Council comparative trial of surgery and radiotherapy for the primary treatment of small-celled or oat-celled carcinoma of the bronchus. Lancet. 1969, 2 (7619), 501-5.

22. Mulero-Navarro S, Esteller M. Chromatin remodeling factor CHD5 is silenced by promoter CpG island hypermethylation in human cancer. Epigenetics. 2008, 3, 210-5.

23. Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. Cancer cell. 2014, 25, 304-17.

24. Noonan E, Place R, Pookot D, et al. miR-449a targets HDAC-1 and induces growth arrest in prostate cancer. Oncogene. 2009, 28, 1714-24.

25. Osmani L, Askin F, Gabrielson E, Li QK. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): Moving from targeted therapy to immunotherapy. Semin Cancer Biol. 2018, 52(Pt1), 103–109.

26. Poon SL, McPherson JR, Tan P, Teh BT, Rozen SG. Mutation signatures of carcinogen exposure: genome-wide detection and new opportunities for cancer prevention. Genome medicine. 2014, 6, 24.

27. Proctor RN. The history of the discovery of the cigarette-lung cancer risk: evidentiary traditions, corporate denial, global toll. Tobacco control.2012,21(2):87-91.

28. Rashid. NG, Mohamed. S, Abdulla. RM, Hama. SH, HamaSalih. THM, Abdullah. KE. Lung Cancer Incidence in Sulaimani Province during 2011-2015. Journal of Cancer Policy 14, September 2018: 11-15.
29. Roninson IB. Oncogenic functions of tumour suppressor p21 Waf1/Cip1/Sdi1: association with cell senescence and tumour-promoting activities of stromal fibroblasts. Cancer letters. 2002, 179, 1-14

30. Saunders M, Dische S, Barrett A, et. al. Continuous hyper fractionated accelerated radiotherapy (CHART) versus conventional radiotherapy in non-small cell lung cancer: a randomized multicentral trial. Lancet. 1997, 350 (9072), 161-65.

31. Shtivelman E, Lifshitz B, Gale RP, Canaani E. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature.1985, 315, 550-554

32. Sporn JC, Kustatscher G, Hothorn T, et al. Histone macroH2A isoforms predict the risk of lung cancer recurrence. Oncogene. 2009, 28, 3423-8.

33. Thomas RK, Baker AC, DeBiasi RM, et al. High-throughput oncogene mutation profiling in human cancer. Nature genetics. 2007, 39, 347-51.

34. Thun MJ, Hannan LM, Adams Cambell LL, et. al. Lung cancer occurrence in never smokers: an analysis of 13 cohorts and 22 cancer registry studies. PLoS Med. 2008, 5(9), 1185.

35. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. Science. 2001, 291, 1284-9.

**Scheme 1**

Scheme 1 is available in Supplemental Files section.

**Figures**

![Diagram](image_url)

**Figure 1**

schematic comparison of healthy lung and lung cancer
Figure 2

spreading of lung cancer to the other organs
Figure 3

distribution map of studied lung cancer cases admitted to Nanakali hospital in Erbil, KR, Iraq
Figure 4

Age in years at diagnosis of the 590 Lung cancer cases and the Percentiles at any specific age.
Figure 5

Regression of number of female Lung (Lung) cancer cases on age groups of 10's and the dotted lines represent the 95% Confidence Intervals, Upper and Lower
Figure 6

Regression of number of male Lung (Lung) cancer cases on age groups of 10's and the dotted lines represent the 95% Confidence Intervals, Upper and Lower.
Figure 7

Regression line of number of Lung cancer cases on age group and the Regression prediction equation.

Figure 8

KM survival curves for lung cancer patients based on gender.
Figure 9

Checking Proportional hazard assumptions for immune system and surgery.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- Supportinginformation.docx
- GraphicalAbstract.docx
- scheme1.png