

Lung Cancer: A Nationwide Study to Characterize Sex Differences, Incidence, and Spatial Patterns in Portugal

TERESA GUERREIRO¹, LUIS ANTUNES^{2,3}, JOANA BASTOS⁴, ALEXANDRA MAYER⁵,
GONCALO FORJAZ^{6,7}, ANTÓNIO ARAUJO^{8,9} and CARLA NUNES^{1,10}

¹NOVA National School of Public Health, Universidade NOVA de Lisboa, Lisbon, Portugal;

²Portuguese Institute of Oncology Francisco Gentil, Porto, Portugal;

³Cancer Epidemiology Group, IPO Porto Research (CI-IPOP),

Portuguese Institute of Oncology Francisco Gentil, Porto, Portugal;

⁴Portuguese Institute of Oncology Francisco Gentil, Coimbra, Portugal;

⁵Portuguese Institute of Oncology Francisco Gentil, Lisbon, Portugal;

⁶Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD, U.S.A.;

⁷Azores Oncological Centre, Azores, Portugal;

⁸Universitary Hospital Center of Porto, Porto, Portugal;

⁹Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal;

¹⁰Public Health Research Center, Universidade NOVA de Lisboa, Lisbon, Portugal

Abstract. *Background/Aim:* This study characterized lung cancer (LC) in Portugal, to capture sex differences, regional variation, and spatial distribution. *Patients and Methods:* Variables: age, sex, vital status, region of residence, degree of urbanization, histologic types and stage at diagnosis. Spatial analyses conducted to identify high (HIC) and low incidence (LIC) clusters. *Results:* In total, 11,642 cases were diagnosed (76.6% male, 23.4% female), with a similar mean age at diagnosis. There were 23.6% locally advanced and 58.4% metastatic disease cases, with 13.4% alive five years after diagnosis. Non-small-cell LC accounted for 77.3% of which 40.8% was adenocarcinoma and 22.7% was squamous cell carcinoma. Standardized incident rate (IR) was $37.5/10^5$, $65.2/10^5$ in males, and $15.7/10^5$ in females, with four HIC (urban) clusters and four LIC (non-urban). *Conclusion:* This study highlighted the sex differences in incidence, mortality, histology, and geographic distribution of LC in Portugal. Considering the advanced stages and the poor overall

survival, understanding sex and geographic differences is important for public health interventions.

Worldwide, lung cancer (LC) remains the leading cause of cancer incidence and mortality, with 2.1 million new LC cases (11.6% of the total) and 1.8 million deaths reported in 2018, representing 18.4% of all cancer deaths (1). In males, LC is the most frequent cancer and the leading cause of cancer-related death, while amongst females it is the third most frequent and second cause of cancer-related death (1).

Incidence rate (IR) and patterns of LC vary according to the geographical area, ethnicity, differences in tobacco exposure (including intensity and duration of smoking), type of cigarettes, and degree of inhalation, as well as, environmental factors (1, 2). Globally, there is a reported higher LC incidence in males than females with a sex ratio of 2.1:1 (2, 3).

In 2017, within the European Union, LC had the highest predicted age-standardized cancer mortality rate in both sexes, corresponding to about 20% of total cancer deaths (4). Since 2012, LC mortality in females has risen by 5.1% whereas it has decreased by 10.7% in males. This decrease in male LC mortality correlates with the decreased prevalence of smoking in males in recent decades (4). In Portugal, LC is the fourth most common cancer representing 9.8% of all cases and the main cause of cancer-related mortality (17.3%) in 2018 (5). Regarding sex, LC is the third most frequently occurring cancer in males and the fourth in females with a male to female IR of 3.1:1 (5).

According to the last report in 2017, published by Direção Geral da Saúde (DGS) (Portuguese General Health Directorate),

This article is freely accessible online.

Correspondence to: Teresa Guerreiro, National School of Public Health, Universidade NOVA de Lisboa, Avenida Padre Cruz, 1600-560 Lisbon, Portugal. Tel: +35 1917236613, e-mail: mt.guerreiro@ensp.unl.pt

Key Words: Lung cancer, incidence, Portugal, sex differences, spatial cluster analysis.

mortality in males slightly decreased in the last five years by 1.2%, but conversely increased by 16.3% in females. Mortality is expected to continuously rise over the next few years due to the changes in tobacco consumption in females (6).

More than 80% of lung cancers in Western populations are attributed to smoking and the disease can be largely prevented through tobacco control (1). Other factors than active tobacco use, account for about 15% of cases in females and 10% in males (7). Studies also reported an increasing trend in the incidence of LC among non-smokers, especially in Asia (8). Despite the limited data available on never-smokers, studies show that females are more likely to have non-smoking associated LC than males (2).

Globally, LC imposes a major disease burden and little improvement in LC survival was made despite considerable advances in patient survival over the last several decades for other cancer types (*e.g.*, breast and prostate). The lack of improvement in LC survival is largely attributed to late diagnosis of advanced stage disease, limited treatment options, and poor survival prognosis (9).

The use of spatial statistical analysis is an effective method to understand geographic patterns of disease and to detect spatial clusters based on smaller units such as municipalities (10). This type of geographic information is important for planning cancer control and prevention activities that need to be adjusted to the different profiles identified.

The last report from the Portuguese National Cancer Registry published in 2010 provided a national overview on LC incidence and mortality per sex and district of residence (11). Our study, characterized LC in Portugal by adding to the current knowledge regarding histology, stage at diagnosis, and place of residence. To our knowledge, this is the first study applying spatial statistical analysis using population-based lung cancer data, which may be critical for future health interventions.

To provide a national overview of LC in Portugal, this study aimed to: 1. characterize LC patients registered in the regional oncologic registries between 2009 and 2011, and followed up until 2016, in terms of age, sex, vital status at 1, 3, and 5 years, region of residence, degree of urbanization, and tumor histology and stage at diagnosis; 2. analyze incidence rates per sex and region; and 3. identify mainland's spatial distribution of high (HIC) and low incidence clusters (LIC) and describe their characteristics in terms of age, degree of urbanization, histology, and stage at diagnosis. Understanding LC patient characteristics, identifying critical incidence zones and potentially related factors, may be helpful for the development of future health strategies to support early detection and a better control of LC in Portugal.

Patients and Methods

Data and sources. We identified invasive lung cancer cases (ICD-10 code C34) (12) registered on the years 2009, 2010, and 2011

from all population-based cancer registries in Portugal: North (RORENO), Central (ROR-CENTRO), South and Madeira (ROR-SUL), and Azores (RORA) to have national coverage. We submitted data requests to each registry coordinator and obtained ethics committee approvals. Informed consent was not required as data received were already anonymized.

The following variables across all registries were analyzed: age at diagnosis (years and according to age group: <40, 40-49, 50-59, 60-69, 70-79 and 80+), sex (male/female), tumor stage (I, II, III, IV), and vital status (alive/dead at 1, 3, and 5 years following diagnosis). The regions of residence were North, Central Region, Lisbon Region, Alentejo, Algarve, Madeira islands, and Azores islands. Histologic types and respective morphology codes were aggregated by non-small cell lung cancer (NSCLC): NSCLC not otherwise specified (NOS) (ICD-O-3 code: 8046), adenocarcinoma (ADC) (8050, 8140-41, 8143, 8190, 8211, 8250-55, 8260, 8310, 8320, 8323, 8480-81, 8490, 8550, 8572, 8574), squamous cell carcinoma (SQCC) (8052, 8070-76, 8078, 8083-84, 8120, 8123), and others (8004, 8012-13, 8021-22, 8031-33, 8035, 8082, 8200, 8240-41, 8244, 8246 8249, 8430, 8560, 8562, 8575). Small cell lung carcinoma (SCLC) (8002, 8041-45) and other unspecified (OU) (8000, 8001, 8010, 8020, 8230) were also aggregated and reported (13).

The 278 mainland and the 11 Madeira island municipalities were classified as densely, intermediate, and thinly populated according to the degree of urbanization. The information was not available for the 19 Azores islands municipalities' (14, 15). The population per year, sex, age group, and municipality was obtained from Statistics Portugal and Pordata (16, 17). Smoking habits data were not available at the individual or municipality levels. Smoking habits at a regional level, obtained from the last Portugal National Health Inquiry published in 2016, were considered in the context of the discussion (18).

Methods. We performed a descriptive analysis of the total number of LC cases from 2009 to 2011 and patient characteristics according to sex. Additionally, we characterized the tumor histology per region and degree of urbanization. A chi-squared test was used for the comparison of proportions of variables between sexes. Differences were statistically significant at $p < 0.05$. Data were analyzed using Statistical Package for Social Services (IBM SPSS Statistics for Windows, version 24.0, Armonk, NY, USA).

The annual incident rates (IR) according to sex were defined as the number of new LC cases among males and females, in a given area, in a specific year, divided by the number of person-years at risk and corresponding to the sum of the mid-year population of males and females that lived in that same area during the same year (19).

The same approach was followed in the IR analysis per region. We analyzed the crude IR per region to characterize the real incidence. For the IR analysis per age group, patients were divided into six groups. We calculated both crude and age-adjusted incidence rates. Age-adjusted IR was computed by the direct standardization method using two different European standard populations: EU1976 and EU2010 (20). The former allows comparison with previous studies whereas the latter is based on an age distribution closer to the current demographic pattern (21, 22).

The identification of spatial clusters was carried out using Software for the Spatial and Space-Time Scan Statistics (SaTScan™ version 9.6, Martin Kulldorff, National Cancer Institute, Boston, MA, USA) (23). Circular windows with a maximum of 20% of the studied population were applied (24). A purely spatial statistical analysis was conducted using a discrete Poisson model to detect the spatial distribution of high

and low incidence clusters of LC on the Portuguese mainland. This analysis was not conducted for the islands. Statistical significance of clustering was based on Monte Carlo hypothesis testing by comparing the likelihood ratio statistic test from the observed dataset with the statistic test from 999 random datasets generated under the null hypothesis of no clustering (10). We used the crude number of cases per municipality because the objective was to identify real critical areas independently of modifiable or non-modifiable (age-sex populations distribution) risk factors. The results were mapped using the open source geographic information system software QGIS™ (version 3.4) (25).

Results

During the time period 2009-2011, 11,642 LC cases were diagnosed in Portugal, of which 8,914 (76.6%) were male and 2,728 (23.4%) female, with an incidence ratio per sex of 3.3:1. Mean age at diagnosis was 65.9 years old (standard deviation=11.536) and median age (Quartile: Q1-Q3) was 66 years old (range=58-75 years). Mean and median ages were similar when compared by sex. The number of LC cases increased with age, with the age groups of 60-69 and 70-79 years old presenting the highest number of cases, in both males and females (Table I).

From the total number of cases, 6,183 (53.1%) had information on tumor stage; most presented with locally advanced (N=1,460, 23.6%) or metastatic disease (N=3,612, 58.4%) at diagnosis. However, considering the lack of disease stage information for a high proportion of cases, these data must be interpreted with caution. At years 1, 3, and 5 after diagnosis, 41.3%, 18.8%, and 13.4% were alive, respectively, which reflects the high mortality associated with LC. Mortality was higher in males than in females at all time points post-diagnosis. The North and Lisbon Region presented the highest number of cases, respectively 4,426 (38.0%) and 3,387 (29.1%). Azores islands and Alentejo had the highest male to female ratios of 7.8:1 and 4.2:1, respectively.

Degree of urbanization analysis excluded data from the Azores Islands as no data were available per municipality. For mainland Portugal and Madeira islands, Densely and Intermediate populated areas accounted for more than 85% of cases. NSCLC accounted for 77.3% of cases of which the most common histological type was ADC (40.8%), SQCC (22.7%), and NSCLC NOS (7.5%). ADC was more frequent in females (57.0%) than in males (35.8%), but SQCC was more common in males (26.8%) than females (9.3%). SCLC represented 10.1% of all cases and was more frequently seen in males (10.8%) than in females (7.6%). Other unspecified (OU) histological types accounted for 12.6%.

When analyzing the frequency of histological types per degree of urbanization and region, we observed that ADC was the most common reported type of cancer in all areas of the mainland regardless of the degree of urbanization. In the Madeira islands, OU types, and in the Azores islands, SQCC and OU types, were more common than ADC (Table II).

Incidence. Lung cancer cases increased by 9.5% from 2009 to 2011, ranging from 3,667 to 4,015 cases. Crude IR was 36.7 per 100,000 inhabitants with 58.8/100,000 and 16.5/100,000 respectively for male and female. Standardized IR for EU2010 were similar with an overall IR of 37.5/100,000, 66.2/100,000 in males, and 15.7/100,000 in females, reflecting the better matching of the revised population per age group compared to the standardized IR for EU1976 with substantially lower IR (27.1/100,000, 46.2/100,000 in males and 11.5/100,000 in females). Crude IR increased more markedly after the age of 50 and peaked in the age group 70-79 years old (122.8/100,000), with major differences between male (221.1/100,000) and female (49.6/100,000). Marked differences were also observed within regions with the highest crude IR – in Azores islands with 56.8/100,000, which was nearly 40% more than Algarve, Lisbon Region, and the North (42.2, 40.0, and 39.9/100,000 respectively) and more than double of the Central Region (25.6/100,000) (Table III).

Spatial analysis. A pure spatial scan statistical analysis identified four significant high incidence clusters (HIC) and four significant low incidence clusters (LIC) (Table IV; Figure 1). The primary cluster was a LIC and included 90 municipalities that covered most of the Central Region and a substantial part of northeast Alentejo, with a relative risk (RR) of 0.62, matching the results of IR per region. The secondary LIC comprised one in the inner North Region, and two in the Lisbon Region, with RR of 0.78, 0.63, and 0.74, respectively.

The most relevant HIC were in the North and Lisbon Region, specifically in the areas of the country's major cities, Oporto and Lisbon, with RR of 1.54 and 1.72, respectively. We identified one additional HIC in the Lisbon Region and the other in the south regions of Alentejo and Algarve, with a RR of 1.24 and 1.17 respectively.

Characterization of the specific clusters in terms of sex, age groups, histological type, and stage at diagnosis is included in Table IV.

Discussion

This analysis showed a higher male to female ratio of new LC cases in Portugal than what was reported worldwide (3.3:1 vs. 2.1:1) (3).

The incidence of LC correlates with age for both sexes; incidence below the age of 40 is very similar among males and females and the gap between both sexes widens from the age of 50 onwards. Traditionally, lung cancer incidence reflects smoking patterns with a latency period of two to three decades, and the prevalence of smoking (male:female) (2, 26). LC data in never smokers is limited, but there is some evidence that non-smoking females are more affected by LC than males (2).

Table I. Characteristics of lung cancer patients and tumor (N=11,642) in Portugal 2009-2011, according to sex.

	Total N=11,642	Sex		p-Value
		Male N=8,914 (76.6)	Female N=2,728 (23.4)	
Age group, N (%)		N=8,914 (76.6)	N=2,728 (23.4)	<0.0001
<40 years, N (%)	152 (1.3)	80 (0.9)	72 (2.6)	
40-49 years, N (%)	857 (7.4)	596 (6.7)	261 (9.6)	
50-59 years, N (%)	2,373 (20.4)	1,840 (20.6)	533 (19.5)	
60-69 years, N (%)	3,471 (29.8)	2,819 (31.7)	652 (23.9)	
70-79 years, N (%)	3,387 (29.1)	2,603 (29.2)	784 (28.8)	
≥80 years, N (%)	1,402 (12.0)	976 (10.9)	426 (15.6)	
Age (years)		N=8,914 (76.6)	N=2,728 (23.4)	
Mean (SD)	65.9 (11.536)	65.93 (11.037)	65.87 (13.035)	
Median (Q1-Q3)	66 (58-75)	66 (58-74)	67 (57-76)	
Vital status ¹ @1YR, N (%)	N=11,560	N=8,854 (76.6)	N=2,706 (23.4)	<0.0001
Alive	4,771 (41.3)	3,335 (37.7)	1,436 (53.1)	
Dead	6,789 (58.7)	5,519 (62.3)	1,270 (46.9)	
Vital status ¹ @3YR, N (%)	N=11,541	N=8,843 (76.6)	N=2,698 (23.4)	<0.0001
Alive	2,171 (18.8)	1,380 (15.6)	791 (29.3)	
Dead	9,370 (81.2)	7,463 (84.4)	1,907 (70.7)	
Vital status ¹ @5YR, N (%)	N=11,525	N=8,836 (76.7)	N=2,689 (23.3)	<0.0001
Alive	1,548 (13.4)	940 (10.6)	608 (22.6)	
Dead	9,977 (86.6)	7,896 (89.4)	2,081 (77.4)	
Region of residence, N (%)	N=11,642	N=8,914 (76.6)	N=2,728 (23.4)	<0.0001
North	4,426 (38.0)	3,414 (38.3)	1,012 (37.0)	
Central region	1,789 (15.4)	1,334 (15.0)	455 (16.7)	
Lisbon region	3,387 (29.1)	2,501 (28.0)	886 (32.5)	
Alentejo	802 (6.9)	649 (7.3)	153 (5.6)	
Algarve	567 (4.9)	448 (5.0)	119 (4.4)	
Madeira (islands)	250 (2.1)	195 (2.2)	55 (2.0)	
Azores (islands)	421 (3.6)	373 (4.2)	48 (1.8)	
Degree of urbanization, N (%)	N=11,221	N=8,541 (76.1)	N=2,680 (23.9)	0.014
Densely populated	5,679 (50.6)	4,275 (50.0)	1,404 (52.4)	
Intermediate populated	3,987 (35.5)	3,040 (35.6)	947 (35.3)	
Thinly populated	1,555 (13.9)	1,226 (14.4)	329 (12.3)	
Histology, N (%)	N=11,642	N=8,914 (76.6)	N=2,728 (23.4)	<0.0001
NSCLC	9,002 (77.3)	6,791 (76.2)	2,211 (81.0)	
NSCLC NOS	872 (7.5)	726 (8.1)	146 (5.4)	
ADC	4,751 (40.8)	3,194 (35.8)	1,557 (57.0)	
SQCC	2,641 (22.7)	2,388 (26.8)	253 (9.3)	
Others ²	738 (6.3)	483 (5.4)	255 (9.3)	
SCLC	1,173 (10.1)	966 (10.8)	207 (7.6)	
OU	1,467 (12.6)	1,157 (13.0)	310 (11.4)	
Stage ³ , N (%)	N=6,183 (53.1)	N=4,703 (76.1)	N=1,480 (23.9)	<0.0001
I	816 (13.2)	525 (11.2)	291 (19.7)	
II	295 (4.8)	225 (4.8)	70 (4.7)	
III	1,460 (23.6)	1,183 (25.2)	277 (18.7)	
IV	3,612 (58.4)	2,770 (58.8)	842 (56.9)	

¹Information on vital status not available for 117 cases (1%): 82 missing at 1YR, 101 missing at 3YR, and 117 missing at 5YR post diagnosis.

²Others: include large cell carcinoma, mixed carcinoma, neuroendocrine lung cancer, and others. ³46.9% (N=5,459) of cases without disease stage information. NSCLC NOS: Non-small cell lung cancer not otherwise specified; ADC: adenocarcinoma; SQCC: Squamous cell carcinoma; SCLC: small cell lung cancer; OU: other unspecified.

Our data showed that mortality was higher in males than in females at all time points (from 1 to 5 years after diagnosis). This finding concurs with the literature, that females have a lower mortality when compared to males (27).

During the 3-year time frame in our analysis (2009-2011), the IR increased more in females than in males (17.9% vs. 7.7%), which is also reflected in the literature (26, 28). In Portugal, the 5-year lag time existing between

Table II. *Histological characterization of lung cancer tumors of patients (N=11,642) in Portugal in the period 2009-2011, according to region and degree of urbanization.*

	Total	NSCLC NOS	ADC	SQCC	Others ¹	SCLC	OU
Number of cases, N (%)	N=11,642	872 (7.5)	4,751 (40.8)	2,641 (22.7)	738 (6.3)	1,173 (10.1)	1,467 (12.6)
Region of residence, N (%) [*]							
North	4,426 (38.0)	316 (7.1)	1 844 (41.7)	1,046 (23.6)	277 (6.3)	464 (10.5)	479 (10.8)
Central region	1,789 (15.4)	55 (3.1)	702 (39.2)	404 (22.6)	156 (8.7)	146 (8.2)	326 (18.2)
Lisbon region	3,387 (29.1)	277 (8.2)	1,506 (44.4)	758 (22.4)	200 (5.9)	369 (10.9)	277 (8.2)
Alentejo	802 (6.9)	105 (13.1)	288 (35.9)	162 (20.2)	53 (6.6)	90 (11.2)	104 (13.0)
Algarve	567 (4.9)	68 (12.0)	229 (40.4)	120 (21.2)	33 (5.8)	47 (8.3)	70 (12.3)
Madeira (islands)	250 (2.1)	9 (3.6)	67 (26.8)	45 (18.0)	11 (4.4)	13 (5.2)	105 (42.0)
Azores (islands)	421 (3.6)	42 (10.0)	115 (27.3)	106 (25.2)	8 (1.9)	44 (10.4)	106 (25.2)
Degree of urbanization ^{2*} , N (%)	N=11,221	830 (7.4)	4,636 (41.3)	2,535 (22.6)	730 (6.5)	1,129 (10.1)	1,361 (12.1)
Densely populated	5,679 (50.6)	466 (8.2)	2,484 (43.7)	1,271 (22.4)	343 (6.0)	595 (10.5)	520 (9.2)
Intermediate populated	3,987 (35.5)	246 (6.2)	1,594 (40.0)	903 (22.6)	286 (7.2)	392(9.8)	566 (14.2)
Thinly populated	1,555 (13.9)	118 (7.6)	558 (35.9)	361 (23.2)	101 (6.5)	142 (9.1)	275 (17.7)

¹Others: include large cell carcinoma, mixed carcinoma, neuroendocrine lung cancer, and others. ²No degree of urbanization was available for Azores islands. NSCLC NOS: Non-small cell lung cancer not otherwise specified; ADC: adenocarcinoma; SQCC: squamous cell carcinoma; SCLC: small cell lung cancer; OU: other unspecified. **p*<0.0001.

Table III. *New lung cancer cases, crude incidence rate (IR), and standardized incidence rate (SIR) for the European population, in Portugal per 10⁵ inhabitants, and crude IR by age group and region of residence, for the years 2009-2011, according to sex.*

	Aggregate 2009-2011			2009			2010			2011		
	Total	M	F	Total	M	F	Total	M	F	Total	M	F
Number of new cases (N)	11,642	8,914	2,728	3,667	2,834	833	3,960	3,046	914	4,015	3,034	981
Crude IR (%)	36.7	58.8	16.5	34.7	56.0	15.1	37.5	60.3	16.6	38.1	60.3	17.8
SIR EU1976 (%)	27.1	46.2	11.5	26.2	45.0	10.7	27.6	47.2	11.5	27.7	46.3	12.4
SIR EU2010 (%)	37.5	65.2	15.7	36.0	63.1	14.6	38.3	66.8	15.8	38.3	65.8	16.8
Age group (IR)												
<40 years (%)	1.0	1.1	1.0	1.2	1.2	1.2	1.0	1.4	0.6	0.9	0.6	1.1
40-49 years (%)	18.4	26.3	10.9	19.2	27.5	11.3	18.1	26.5	10.1	17.8	24.7	11.2
50-59 years (%)	56.6	91.6	24.4	54.3	90.3	21.2	56.6	92.6	23.5	58.8	91.7	28.5
60-69 years (%)	99.9	176.1	34.8	93.5	167.8	30.2	102.4	179.3	36.7	103.7	180.9	37.4
70-79 years (%)	122.8	221.1	49.6	118.8	214.0	48.2	128.2	231.0	51.5	121.3	218.3	49.0
≥80 years (%)	90.5	179.4	42.4	82.9	165.0	38.4	90.7	178.6	43.3	97.3	193.1	45.1
Region of residence (IR)												
North (%)	39.9	64.3	17.5	34.1	55.6	14.2	41.9	67.7	18.2	43.8	69.6	20.1
Central region (%)	25.6	40.0	12.5	25.6	40.4	12.1	25.2	40.0	11.7	26.0	39.6	13.6
Lisbon region (%)	40.0	62.5	19.9	40.6	64.0	19.6	40.3	63.2	19.7	39.2	60.2	20.4
Alentejo (%)	35.3	58.9	13.0	33.1	57.6	9.9	36.5	58.5	15.8	36.2	60.5	13.4
Algarve (%)	42.2	68.5	17.3	41.3	63.8	19.7	41.9	69.5	15.6	43.5	72.2	16.5
Madeira (islands) (%)	31.3	51.8	13.0	31.1	51.6	12.8	28.4	46.7	12.0	34.4	57.4	14.2
Azores (islands) (%)	56.8	102.2	12.8	58.7	101.8	16.8	58.8	109.4	9.6	53.0	95.3	12.0

M: Male; F: female; IR: incidence rate; SIR EU1976: standardized incidence rate by European population of 1976; SIR EU2010: standardized incidence rate by European population of 2010.

the year of diagnosis and the publication of incidence data by DGS, does not allow a timely understanding of the situation at the present. Nevertheless, considering the increase in tobacco consumption among females in the last

decade, we expect that the IR trend is currently maintained (18, 29).

Globally, the most common histological type of LC is ADC in both sexes, followed by SQCC and SCLC (8). The

Table IV. Spatial analysis of lung cancer incidence at municipality level and characterization of high and low incidence clusters in mainland Portugal for the years 2009-2011.

	High incidence clusters					Low incidence clusters				
	All	Cluster A/H	Cluster B/H	Cluster C/H	Cluster D/H	All	Cluster A/L	Cluster B/L	Cluster C/L	Cluster D/L
Number of cases ¹ , N (%)	5,017 (43.1)	2,025 (17.4)	983 (8.4)	965 (8.3)	1,044 (9.0)	3,256 (27.9)	1,417 (12.2)	1,298 (11.1)	250 (2.1)	291 (2.5)
Population (n)	3382,790	1,284,562	545,260	726,939	826,029	4134,896	1941,548	1479,125	359,346	354,877
Number of municipalities ² , N (%)	47 (16.9)	8	1	4	34	150 (54.0)	90	50	6	4
Cases		2,025	983	965	1,044		1,417	1,298	250	291
Expected cases		1,402	595	794	902		2,120	1,615	392	387
Annual cases/100,000		52.6	60.1	44.3	42.2		24.3	29.3	23.2	27.4
Relative risk (RR)		1.54	1.72	1.24	1.17		0.62	0.78	0.63	0.74
Log likelihood ratio (LLR)		142.09	112.68	18.76	11.68		159.16	38.63	30.61	13.58
p-Value		<0.0001	<0.0001	<0.0001	0.0037		<0.0001	<0.0001	<0.0001	0.00064
Sex, N (%)										
Male, N (%)	3,847 (76.7)	1,590 (78.5)	704 (71.6)	719 (74.5)	834 (79.9)	2,448 (75.2)	1,052 (74.2)	999 (77.0)	175 (70.0)	222 (76.3)
Female, N (%)	1,170 (23.3)	435 (21.5)	279 (28.4)	246 (25.5)	210 (20.1)	808 (24.8)	365 (25.8)	299 (23.0)	75 (30.0)	69 (23.7)
Age group, N (%)										
<40 years, N (%)	56 (1.1)	23 (1.1)	12 (1.2)	7 (0.7)	14 (1.3)	56 (1.7)	23 (1.6)	23 (1.8)	6 (2.4)	4 (1.4)
40-49 years, N (%)	352 (7.0)	143 (7.1)	69 (7.0)	59 (6.1)	81 (7.8)	266 (8.2)	106 (7.5)	114 (8.8)	26 (10.4)	20 (6.9)
50-59 years, N (%)	1,004 (20.0)	444 (21.9)	190 (19.3)	181 (18.8)	189 (18.1)	658 (20.2)	243 (17.1)	298 (23.0)	50 (20.0)	67 (23.0)
60-69 years, N (%)	1,476 (29.4)	564 (27.9)	286 (29.1)	309 (32.0)	317 (30.4)	973 (29.9)	404 (28.5)	383 (29.4)	83 (33.2)	103 (35.4)
70-79 years, N (%)	1,503 (30.0)	585 (28.9)	298 (30.4)	292 (30.3)	328 (31.4)	939 (28.8)	466 (32.9)	336 (25.9)	59 (23.6)	78 (26.8)
≥80 years, N (%)	626 (12.5)	266 (13.1)	128 (13.0)	117 (12.1)	115 (11.0)	364 (11.2)	175 (12.4)	144 (11.1)	26 (10.4)	19 (6.5)
Degree of urbanization, N (%)										
Densely populated	3,792 (75.6)	1,608 (79.4)	983 (100)	965 (100)	236 (22.6)	870 (26.7)	247 (17.4)	304 (23.4)	108 (43.2)	211 (72.5)
Intermediate populated	908 (18.1)	417 (20.6)	0 (0)	0 (0)	491 (47.0)	1,695 (52.1)	806 (56.9)	682 (52.6)	127 (50.8)	80 (27.5)
Thinly populated	317 (6.3)	0 (0)	0 (0)	0 (0)	317 (30.4)	691 (21.2)	364 (25.7)	312 (24.0)	15 (6.0)	0 (0)
Histology, N (%)										
NSCLC, N (%)	4,049 (80.7)	1,658 (81.9)	814 (82.8)	761 (78.9)	816 (78.1)	2,420 (74.3)	1,009 (71.2)	963 (74.1)	204 (81.6)	244 (83.8)
NSCLC NOS, N (%)	399 (8.0)	149 (7.4)	77 (7.8)	60 (6.2)	113 (10.8)	181 (5.6)	40 (2.8)	106 (8.2)	19 (7.6)	16 (5.5)
ADC, N (%)	2,213 (44.1)	910 (44.9)	479 (48.8)	433 (44.9)	391 (37.4)	1,223 (37.5)	516 (36.4)	461 (35.4)	113 (45.2)	133 (45.7)
SQCC, N (%)	1,111 (22.1)	460 (22.7)	192 (19.5)	221 (22.9)	238 (22.8)	773 (23.7)	327 (23.1)	320 (24.6)	48 (19.2)	78 (26.8)
Others ³ , N (%)	326 (6.5)	139 (6.9)	66 (6.7)	47 (4.9)	74 (7.1)	243 (7.5)	126 (8.9)	76 (5.9)	24 (9.6)	17 (5.8)
SCLC, N (%)	541 (10.8)	225 (11.1)	104 (10.6)	110 (11.4)	102 (9.8)	315 (9.7)	120 (8.5)	137 (10.6)	27 (10.8)	31 (10.7)
OU, N (%)	427 (8.5)	142 (7.0)	65 (6.6)	94 (9.7)	126 (12.1)	521 (16.0)	288 (20.3)	198 (15.3)	19 (7.6)	16 (5.5)
Stage ⁴ , N (%)										
I	3,176 (63.3)	1,254 (61.9)	513 (52.2)	588 (60.9)	821 (78.6)	1,052 (32.3)	173 (12.2)	631 (48.6)	144 (57.6)	104 (35.7)
II	439 (13.8)	168 (13.4)	96 (18.7)	99 (16.8)	76 (9.3)	145 (13.8)	16 (9.2)	77 (12.2)	29 (20.1)	23 (22.1)
III	162 (5.1)	69 (5.5)	26 (5.1)	27 (4.6)	40 (4.9)	52 (4.9)	9 (5.2)	27 (4.3)	7 (4.9)	9 (8.7)
IV	764 (24.1)	280 (22.3)	115 (22.4)	132 (22.4)	237 (28.9)	244 (23.2)	34 (19.7)	149 (23.6)	36 (25.0)	25 (24.0)
	1,811 (57.0)	737 (58.8)	276 (53.8)	330 (56.2)	468 (56.9)	611 (58.1)	114 (65.9)	378 (59.9)	72 (50.0)	47 (45.2)

Cluster A/H: North (Oporto); Cluster B/H: Lisbon City; Cluster C/H: Lisbon Region (Oeiras; Cascais; Amadora, Almada); Cluster D/H: South (Alentejo and Algarve); Cluster A/L: Central Region (and most municipalities of Northeast Alentejo); Cluster B/L: North Region (Interior); Cluster C/L: Lisbon Region (Mafra, Vila Franca Xira) and Central Region (West); Cluster D/L: Lisbon Region (left side River Tagus). ¹Percentage calculated from the total number of cases in 2009-2011 (N=11,642); ²Percentage calculated from the total number of mainland municipalities (N=278); ³Others: include large cell carcinoma, mixed carcinoma, neuroendocrine lung cancer, and others; ⁴Number of cases in the cluster with disease stage information. NSCLC NOS: Non-small cell lung cancer not otherwise specified; ADC: adenocarcinoma; SQCC: squamous cell carcinoma; SCLC: small cell lung cancer; OU: other unspecified.

same situation was observed in Portugal; however, the proportion of ADC was substantially higher in females compared to males (57.0% *vs.* 35.8%). However, the opposite was observed for SQCC, which was much higher in males (26.8% *vs.* 9.3%). In more developed countries, there is a trend of increasing rates of ADC and a decrease in SQCC and SCLC, which is likely linked with reduced tobacco consumption, changes in cigarette composition, and the use of filters in cigarettes (26) (30).

The ADC was also the most common histological type in all mainland regions, but not in the Madeira islands where the OU type was the highest and Azores islands where the ADC presented at proportions similar to SQCC and OU type. The reason for the high percentages of unspecified types in both islands are not known; however, an unspecified diagnosis can be linked to the nature of the biopsy, which can lack distinctive features to allow accuracy, or to the poor differentiation of certain tumor cells (31).

As in most countries, the vast majority of lung cancer patients were diagnosed with locally advanced or metastatic disease, which accounted for more than 75% of cases regardless of sex, and which contributes to the low survival rates of LC (32). A slightly higher proportion of stage I cancer at diagnosis was observed in female patients, most likely linked with more frequent detection in routine screening or diagnostic imaging. However, these results must be interpreted with caution due to the high proportion of cases (46.9%) without disease stage information. In general, the prognosis of patients with LC is poor, but patients with operable disease do have a better prognosis (33). A strong effort should be made to improve early diagnosis of LC as there is evidence that if the condition is diagnosed early, survival for both sexes can be improved.

Regarding geography, there are substantial regional differences in IR; Azores islands had the highest incidence rate (56.8/100,000), as well as the highest male to female ratio (7.8:1). These results are strongly aligned with a high tobacco consumption in this region, which is much more widespread in males than in females (18). Similar results, although not as strong, were observed in Alentejo with a male to female ratio of 4.2:1. On the mainland, we observed the lowest IR (25.6/100,000) in the Central Region, that also had the higher proportion of non-smokers in both sexes (18). This finding also reflected the results obtained in the spatial cluster analysis where the most relevant cluster was a LIC that covered most of the Central Region.

Comparing high (HIC) and low incidence clusters (LIC), the main differences observed were related to the degree of urbanization. In HIC, 75.6% of the cases occurred in densely populated areas and only 6.3% in rural municipalities, whereas in the LIC densely populated areas accounted only for 26.7% and the remaining 52.1% and 21.2% in intermediate and thinly populated areas, respectively.

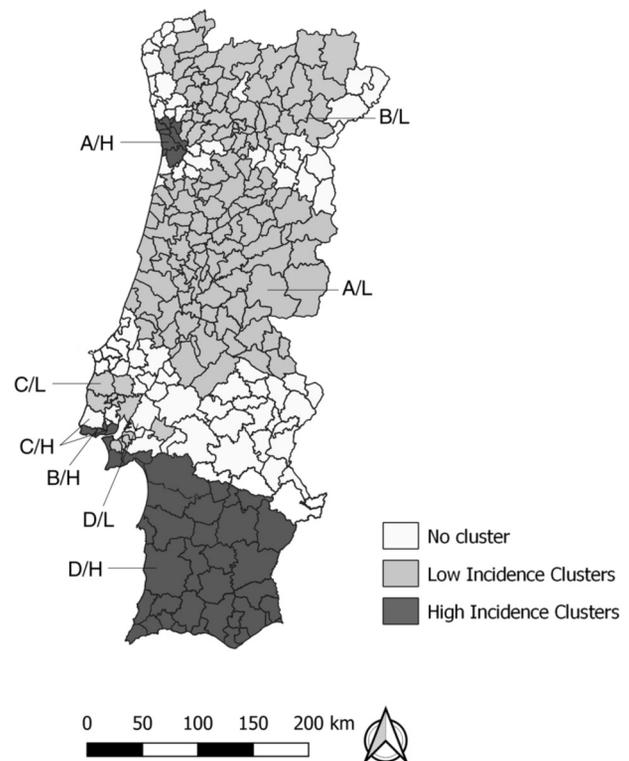


Figure 1. Geographic distribution of high and low incidence clusters in mainland Portugal from 2009-2011.

Information about disease stages was much higher in HIC, but when comparing the existing cases, the split was similar with more than 80% of patients presenting in stages III and IV. Within HIC, we observed that in purely urban clusters more patients were diagnosed with stage I disease (in a range of 13.4 to 18.7%) as compared with more rural clusters (9.3%). However, the opposite occurred with stage III disease where rural clusters had a higher proportion (28.9%) compared to urban clusters (23.4%). These differences might be linked with a timely access to differentiated care. Nevertheless, from an overall perspective, the differences identified between HIC and LIC cannot explain the disparities in IR leading to the conclusion that other risk factors not present in the analysis might have played a critical role.

The main strength of our study was the use of population-based data which originated from cancer registries that combined provided nationwide coverage. Additionally, cluster identification and spatial analysis were based on smaller administrative units that can help to outline regional disparities that could be missed or disregarded if a simple urban *versus* rural approach was used to observe incidence rates spatially in Portugal (10).

However, the study has several limitations. Incomplete information or potential data entry errors in the official registry

datasets is possible. Individual-level information on smoking behavior and other known risk factors for LC, namely exposure to occupational or environmental pollutants, second hand smoke, presence of comorbidities, and history of previous LC, are not routinely captured in cancer registries thus, they cannot be used to directly measure the contribution of these factors to the differences in IR (27, 28, 34). Additionally, given the short time period used for the analyses, we only analyzed the spatial distribution of LC incidence and not temporal trends that could provide valuable insight and information on the changing incidence rates in Portugal. Moreover, we did not assess potential influencing factors associated with clustering since they were absent from our analysis dataset. Finally, the time lag from diagnosis to data availability for evaluation purposes was long and can negatively impact the timely implementation of public health policies.

Further studies are required to evaluate spatial and temporal changes in the pattern of LC incidence in Portugal using data from a longer surveillance period. Additionally, more detailed individual and contextual level information is needed to evaluate potential incidence determinants in the identified clusters (implementation of local studies). Evidence-based knowledge is important to adequately plan health interventions; thus, it is important to consider including in the National Cancer Registry additional information related to key risk factors and tumor characteristics, in order to enable a deeper analysis of the root causes for lung cancer incidence and mortality differences.

Conclusion

This study highlighted the sex differences observed in terms of incidence, mortality, histologic type, and geographic distribution of LC in Portugal. The spatial analysis identified four high and four low incidence clusters on the mainland, complementing the differences observed between regions. The HIC were mainly urban clusters and LIC more non-urban. As most LC cases were diagnosed at an advanced stage and thus have poor survival expectations, identification of high incidence areas can be very valuable for the prioritization of public health interventions. Identifying specific subgroups at highest risk and a clear description of the most critical causal factors can help to tailor strategies that are essential to reduce the burden of lung cancer in Portugal.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

Teresa Guerreiro is an employee of Novartis Farma Portugal. This study was prepared in the context of the first author's PhD thesis, a personal project. There was no interference from the employer in

the research project and resulting publications. Luís Antunes, Joana Bastos, Alexandra Mayer, Gonçalo Forjaz, António Araújo and Carla Nunes declare no conflicts of interest related to this study.

Authors' Contributions

Teresa Guerreiro was responsible for the concept and writing of the manuscript, data aggregation, analysis and interpretation. Luís Antunes, Joana Bastos, Alexandra Mayer and Gonçalo Forjaz for data collection and manuscript revision, António Araújo for the concept and revision of the manuscript and Carla Nunes for concept of the manuscript, data analysis and interpretation and manuscript revision.

Acknowledgements

The Authors thank the four population-based cancer registries for allowing access to the data that enabled this study. The Authors would also like to thank Alicia A. Livinski, National Institutes of Health Library for manuscript editing assistance.

References

- 1 Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68(6): 394-424, 2018. PMID: 30207593. DOI: 10.3322/caac.21492
- 2 Malhotra RK, Manoharan N, Nair O, Deo S and Rath GK: Trends in lung cancer incidence in Delhi, India 1988-2012: Age-period-cohort and joinpoint analyses. *Asian Pac J Cancer Prev* 19(6): 1647-1654, 2018. PMID: 29937537. DOI: 10.22034/APJCP.2018.19.6.1647
- 3 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F: Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 136: E359-86, 2015. PMID: 25220842. DOI: 10.1002/ijc.29210
- 4 Malvezzi M, Carioli G, Bertuccio P, Boffetta P, Levi F, La Vecchia C and Negri E: European cancer mortality predictions for the year 2017, with focus on lung cancer. *Ann Oncol Off J Eur Soc Med Oncol* 28(5): 1117-1123, 2017. PMID: 28327906. DOI: 10.1093/annonc/mdx033
- 5 IARC. Global Cancer Observatory: Population Fact Sheets: Portugal. Vol. 2019. Available at: <http://gco.iarc.fr/today/data/factsheets/populations/620-portugal-fact-sheets.pdf> [Last accessed on May 2, 2020]
- 6 Gonçalves MB, Miranda N and da Saúde D-G: Programa Nacional para as Doenças Oncológicas 2017. Direção-Geral da Saúde. Direção-Geral da Saúde 1-24, 2017. Available at: <https://comum.rcaap.pt/bitstream/10400.26/22531/1/Programa%20Nacional%20para%20as%20Doen%C3%A7as%20Oncol%C3%B3gicas%202017.pdf> [Last accessed on May 2, 2020]
- 7 Jemal A, Miller KD, Ma J, Siegel RL, Fedewa SA, Islami F, Devesa SS and Thun MJ: Higher lung cancer incidence in young women than young men in the United States. *N Engl J Med* 378(21): 1999-2009, 2018. PMID: 29791813. DOI: 10.1056/NEJMoal715907
- 8 Zhang X, Wu L, Xu Y, Zhang B, Wu X, Wang Y and Pang Z: Trends in the incidence rate of lung cancer by histological type

- and gender in Sichuan, China, 1995-2015: A single-center retrospective study. *Thorac Cancer* 9(5): 532-541, 2018. PMID: 29504256. DOI: 10.1111/1759-7714.12601
- 9 Schabath MB, Cress WD and Muñoz-Antonia T: Racial and ethnic differences in the epidemiology of lung cancer and the lung cancer genome. *Cancer Control* 23(4): 338-346, 2016. PMID: 27842323. DOI: 10.1177/107327481602300405
 - 10 Fu Z, Li Y, Chu J, Sun J, Lu Z, Zhang J, Chen X, Zhang G, Xue F, Guo X and Xu A: Lung cancer mortality clusters in urban and rural areas of Shandong Province, China: A spatial scan statistical analysis. *Precis Radiat Oncol* 3(1): 15-22, 2019. DOI: 10.1002/pro6.62
 - 11 RORENO. Registo Oncológico Nacional 2010 Elaborado pelo Registo Oncológico Regional do Norte Editado pelo Instituto Português de Oncologia do Porto Francisco Gentil – EPE. Regist Oncológico Nac 2010. 2016. Available at: http://www.roreno.com.pt/images/stories/pdfs/ro_nacional_2010.pdf [Last accessed on May 2, 2020]
 - 12 World Health Organization W. ICD-10 International Statistical Classification of Diseases and Related Health Problems. World Heal Organ (2010 Edition), 2010. Available at: https://www.who.int/classifications/icd/ICD10Volume2_en_2010.pdf?ua=1 [Last accessed on May 2, 2020]
 - 13 Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM and Whelan S: WHO. International Classification of Diseases for Oncology (ICD-O). International Agency for Research on Cancer. ICD-O-3 online <https://codes.iarc.fr/codegroup/2>. 2011. Available at: <https://codes.iarc.fr/codegroup/2> [Last accessed on May 2, 2020]
 - 14 Eurostat. RAMON - Reference And Management of Nomenclatures Degree of Urbanisation (DEGURBA) - Local Administrative Units. 1-2, 2016. Available at: http://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA [Last accessed on May 2, 2020]
 - 15 Dijkstra L and Poelman H: Regional Working Paper 2014. A harmonised definition of cities and rural areas: the new degree of urbanisation. *Eur Comm* 28, 2014. Available at: https://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf [Last accessed on May 4, 2020]
 - 16 INE – Instituto Nacional de Estatística Portugal. Available at: <http://www.ine.pt> [Last accessed on 13 Jan, 2019]
 - 17 FMS Foundation. PORDATA: Base de dados Portugal Contemporâneo. [Database of Contemporary Portugal] Fundação Francisco Manuel dos Santos, 2018. Available at: <http://https://www.pordata.pt> [Last accessed on 16 Nov, 2018]
 - 18 INE, INSA. Inquérito Nacional de Saúde 2014. [National Health Survey 2014]. Lisbon, Portugal: Instituto Nacional de Estatística Portugal; 2016. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=263714091&PUBLICACOESmodo=2 [Last accessed on May 2, 2020]
 - 19 Ellis L, Woods LM, Estève J, Eloranta S, Coleman MP and Rachet B: Cancer incidence, survival and mortality: Explaining the concepts. *Int J Cancer* 135(8): 1774-1782, 2014. PMID: 24945976. DOI: 10.1002/ijc.28990.
 - 20 Naing NN: Easy way to learn standardization : direct and indirect methods. *Malays J Med Sci* 7(1): 10-15, 2000. PMID: 22844209.
 - 21 Eurostat. Revision of the European Standard Population: Report of Eurostat's task force, 2013. Available at: <http://ec.europa.eu/eurostat/documents/3859598/5926869/KS-RA-13-028-EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f> [Last accessed on May 2, 2020]
 - 22 Commission E: European Cancer Information System (ECIS) Glossary. Vol. 2019. Available at: <https://ecis.jrc.ec.europa.eu/info/glossary.html> [Last accessed on May 2, 2020]
 - 23 Kulldorf M: Information Management Services I. SaTScan TM v8.0: Software for the spatial and space-time scan statistics. 2009. Available at: <http://www.satscan.org/> [Last accessed on May 2, 2020]
 - 24 Kulldorf M: A Spatial scan statistic. *Commun Stat-Theor M* 26(6): 1481-1496, 2017. DOI: 10.1080/03610929708831995
 - 25 QGIS. QGIS Geographic Information System. Open Source Geospatial Foundation Project Vol. 2019. Available at: <http://qgis.osgeo.org> [Last accessed on May 2, 2020]
 - 26 Egleston BL, Meireles SI, Flieder DB and Clapper ML: Population-based trends in lung cancer incidence in women. *Semin Oncol* 36: 506-515, 2009. PMID: 19995642. DOI: 10.1053/j.seminoncol.2009.09.003
 - 27 Thakur M, Ruterbusch J, Schwartz A, Gadgeel S, Beebe-Dimmer J and Wozniak A: Risk of second lung cancer in patients with previously treated lung cancer: Analysis of Surveillance, Epidemiology, and End Results (SEER) Data. *J Thorac Oncol* 13(1): 46-53, 2018. PMID: 28989038. DOI: 10.1016/j.jtho.2017.09.1964
 - 28 Kligerman S and White C: Epidemiology of lung cancer in women: Risk factors, survival, and screening. *Am J Roentgenol* 196(2): 287-295, 2011. PMID: 21257878. DOI: 10.2214/AJR.10.5412
 - 29 INE, INSA. Inquérito Nacional de Saúde 2005/2006 [National Health Survey 2005/2006]. Instituto Nacional de Estatística Portugal, 2009. Available at: <http://www2.insa.pt/sites/INSA/Portugues/Publicacoes/Outros/Paginas/INS2005-2006.aspx> [Last accessed on May 2, 2020]
 - 30 Lewis D, Check D, Caporaso N, Travis W and Devesa S: U.S. lung cancer trends by histologic type. *Cancer* 120(18): 2883-2892, 2014. PMID: 25113306. DOI: 10.1002/cncr.28749
 - 31 Tane S, Nishio W, Ogawa H, Hokka D, Tane K, Tanaka Y, Tauchi S, Uchino K, Sakai Y, Ohbayashi C, Yoshimura M and Maniwa Y: Clinical significance of the “not otherwise specified” subtype in candidates for resectable non-small cell lung cancer. *Oncol Lett* 8(3): 1017-1024, 2014. PMID: 25120650. DOI: 10.3892/ol.2014.2302
 - 32 Cheng TYD, Cramb SM, Baade PD, Youlten DR, Nwogu C and Reid ME: The international epidemiology of lung cancer: Latest trends, disparities, and tumor characteristics. *J Thorac Oncol* 11(10): 1653-1671, 2016. PMID: 27364315. DOI: 10.1016/j.jtho.2016.05.021
 - 33 Rosen JE, Keshava HB, Yao X, Kim AW, Detterbeck FC and Boffa DJ: The natural history of operable non-small cell lung cancer in the national cancer database. *Ann Thorac Surg* 101(5): 1850-1855, 2016. PMID: 27041452. DOI: 10.1016/j.athoracsur.2016.01.077
 - 34 Malhotra J, Malvezzi M, Negri E, La Vecchia C and Boffetta P: Risk factors for lung cancer worldwide. *Eur Respir J* 48(3): 889-902, 2016. PMID: 27174888. DOI: 10.1183/13993003.00359-2016

Received April 10, 2020

Revised May 1, 2020

Accepted May 4, 2020