DOI: 10.1590/1415-790X201400010014ENG

ORIGINAL ARTICLE / *ARTIGO ORIGINAL*

Profile of patients with lung cancer assisted at the National Cancer Institute, according to their smoking status, from 2000 to 2007

Perfil dos pacientes com câncer de pulmão atendidos no Instituto Nacional de Câncer, segundo condição tabagística, 2000 a 2007

Mirian Carvalho de Souza[I,II], Ana Glória Godoi Vasconcelos[III], Marise Souto Rebelo[II], Paulo Antonio de Paiva Rebelo[II], Oswaldo Gonçalves Cruz[IV]

ABSTRACT: ***Introduction:*** Tobacco use is directly related to the future incidence of lung cancer. In Brazil, a growing tendency in age-adjusted lung cancer mortality rates was observed in recent years. ***Objective:*** To describe the profile of patients with lung cancer diagnosed and treated at the National Cancer Institute (INCA) in Rio de Janeiro, Brazil, between 2000 and 2007 according to their smoking status. ***Methods:*** An observational study was conducted using INCA's database of cancer cases. To assess whether the observed differences among the categories of sociodemographic variables, characterization of the tumor, and assistance — pertaining to smokers and non-smokers — were statistically significant, a chi-square test was applied. A multiple correspondence analysis was carried out to identify the main characteristics of smokers and non-smokers. ***Results:*** There was a prevalence of smokers (90.5% of 1131 patients included in the study). The first two dimensions of the multivariate analysis explained 72.8% of data variability. Four groups of patients were identified, namely smokers, non-smokers, small-cell tumors, and tumors in early stages. ***Conclusion:*** Smoking cessation must be stimulated in a disseminated manner in the population in order to avoid new cases of lung cancer. The Tumors in Initial Stages Group stood out with greater chances of cure.

Keywords: Lung neoplasia. Neoplasia staging. Multivariate analysis. Biostatistics. Electronic health records. Smoking habit.

[I]Programa de Pós-Graduação em Epidemiologia em Saúde Pública, Escola Nacional de Saúde Pública Sérgio Arouca, Fundação Oswaldo Cruz – Rio de Janeiro (RJ), Brasil.

[II]Instituto Nacional de Câncer José Alencar Gomes da Silva – Rio de Janeiro (RJ), Brasil.

[III]Departamento de Métodos Quantitativos em Saúde; Escola Nacional de Saúde Pública Sérgio Arouca, Fundação Oswaldo Cruz – Rio de Janeiro (RJ), Brasil.

[IV]Programa de Computação Científica; Fundação Oswaldo Cruz – Rio de Janeiro (RJ), Brasil.

Corresponding author: Mirian Carvalho de Souza. Rua Haddock Lobo, 203 ap. 707, Tijuca, CEP: 20260-141, Rio de Janeiro, RJ, Brazil. E-mail: miriancs@inca.gov.br

Conflict of interests: nothing to declare – **Financial support:** none.

RESUMO: ***Introdução:*** O consumo de tabaco está diretamente relacionado à incidência futura de câncer de pulmão. No Brasil foi observada uma tendência de crescimento da taxa de mortalidade ajustada por idade, para esta enfermidade nos últimos anos. ***Objetivo:*** Descrever o perfil dos pacientes com câncer de pulmão diagnosticados e atendidos no Instituto Nacional de Câncer (INCA), no Rio de Janeiro, Brasil, entre 2000 e 2007 segundo condição tabagística. ***Métodos:*** Foi realizado um estudo observacional, utilizando dados do Registro Hospitalar de Câncer do INCA. Para avaliar se as diferenças observadas entre as categorias das variáveis sociodemográficas, de caracterização do tumor e da assistência — para fumantes e não fumantes — são estatisticamente significativas, foi aplicado o teste qui-quadrado. A análise de correspondência múltipla foi utilizada para identificar as características predominantes dos fumantes e não fumantes. ***Resultados:*** Foi observado um predomínio de pacientes fumantes (90,5% dos 1131 incluídos no estudo). As duas primeiras dimensões da análise de correspondência múltipla explicaram 72,8% da variabilidade dos dados. Quatros grupos de pacientes foram identificados: fumantes, não fumantes, tumores de pequenas células e tumores em estádios iniciais. ***Conclusões:*** O estímulo à cessação do tabagismo deve ser realizado de forma disseminada na população para que novos casos de câncer de pulmão sejam evitados. Destaca-se o Grupo Tumores em Estádios Iniciais, que tem maiores chances de cura.

Palavras-chave: Neoplasias pulmonares. Estadiamento de neoplasias. Análise multivariada. Bioestatística. Registros eletrônicos de saúde. Hábito de fumar.

INTRODUCTION

Cancer is a public health problem in developed and developing countries alike. Although it was a practically unknown and rare illness at the beginning of the 20th century, lung cancer has become very frequent over the years[1]. The International Agency for Research on Cancer (IARC) estimated 1.61 million new lung cancer cases in 2008, representing 12.7% of all cases in the world. It was also the most frequent cause of death by cancer worldwide, counting 1.38 million casualties, which is equivalent to 18.2% of the total number of deaths by cancer[2].

In North America, Eastern Asia, and practically all European countries, lung cancer is the most common cause of death by cancer among males[1]. Its incidence and mortality rates are generally lower among women, but in 2008 lung cancer was the 4th most frequent cancer type amidst new cases and the second cause of death[2].

In Brazil, the mortality rates due to lung cancer — adjusted by age — increased between 1980 and 2007[3,4]. The Health Ministry estimated an absolute incidence of 27,310 cases of lung cancer in Brazil for 2012. In terms of incidence rates, malignant lung neoplasia is the 2nd more frequent among men (18/100 thousand) and the 5th most occurring among women (10/100 thousand). States in the southern and southeastern regions of Brazil, known by their elevated urbanization indices and high prevalence of smoking, concentrate the highest incidence rates[5].

The geographical and temporal patterns of lung cancer incidence are greatly determined by tobacco consumption. An increase in tobacco consumption is directly related (20 to 30 years later) to higher incidences of lung cancer. Likewise, decreased consumption leads to less future incidences. In Brazil, 82% of the deaths by lung cancer among men are attributed to smoking; among women, this number reaches 41%[6]. In individuals who quit smoking, the risk of developing lung cancer gradually decreases over 15 years and remains about 2 times higher than among those who never smoked[7].

Although the majority of lung cancer cases is attributed to smoking, this type of neoplasia is also an important problem among individuals who never smoked. There is no epidemiologic evidence of the incidence of lung cancer among non-smokers, but the proportion of non-smokers who develop this disease is increasing, especially in Asian populations. Some explanations for this phenomenon are: better information data on smoking, especially after the arrival of Epidermal Growth Factor Receptors (EGFR); increased life expectancy and, therefore, a longer time of exposure to the risk of falling ill; and improvements in diagnoses of tumors previously classified as carcinomas of unknown origin[8,9].

Over the last years, the etiology of lung cancer among non-smokers has become more well-defined in terms of genetic risk factors and carcinogenesis molecular bases. It is known that this illness occurs more frequently among women and that the predominant histologic type is the adenocarcinoma. A molecular approach has revealed that there are important differences regarding lung cancer between smokers and non-smokers[8,9].

The purpose of this article is to describe the relation between smoking and other variables related to lung cancer among patients with this illness assisted at the Hospital for Stage I Cancer at the *Instituto Nacional de Câncer José Alencar Gomes da Silva* (INCA) between 2000 and 2007. The characterization of individuals who fall ill due to lung cancer can aid in clarifying gaps about the factors that led to the occurrence of this disease.

METHODS

This study is part of a research project approved by the Ethics Committees of INCA and of the Sergio Arouca National School of Public Health, registered on protocols CAAE-012.0.007.031-11 and CAAE-0163.0.031.007-11 in the Research Ethics National System. The authors hereby declare the absence of any conflicts of interest.

DATA SOURCE AND STUDY POPULATION

The data presented in this study were extracted from the database of the Hospital for Stage I Cancer at the INCA by means of an electronic database management tool developed by INCA for keeping hospital registers.

The cases selected to participate in the study were individuals with primary malignant bronchi and lung neoplasia who were diagnosed and assisted at INCA's Hospital for Stage I Cancer between 2000 and 2007. The term "lung cancer" was used to represent malignant bronchi and lung neoplasia whose topography and morphology were classified according to the International Classification of Diseases for Oncology (ICD-O/3)[10]. We considered as smokers individuals who had smoked at some point in life, that is, those who were smokers on the date of the diagnosis as well as former smokers. Non-smokers were the participants who reported never having smoked.

The individuals who were considered eligible to participate in this study were patients older than 29 years of age who had not undergone any previous treatment prior to their arrival at INCA and whose diagnoses were confirmed by histopathological exams that specified the morphology of the tumor.

The data were analyzed in stages by comparing patients who smoked to non-smokers according to variables that characterized the following: sociodemographic profile and risk factors (sex, age range, schooling, marital status, family history of cancer, and alcoholism); the tumor (detailed primary location, histological type and clinical stage according to the TNM-6[11]; treatments used (first treatment received and stage of the disease at the end of the first treatment).

The patients without complete information pertaining to the variables included in the analysis were excluded from the database.

In the first stage, we described the characteristics studied according to each patient's smoking status. In order to evaluate whether the differences observed between smokers and non-smokers were statistically significant, we applied the χ^2 test, considering a significance level of 5%.

In the second stage, the relation between the characteristics studied and smoking was evaluated with the use of a statistical tool known as multiple correspondence analysis followed by a dendrogram to aid the visualization of the similarities. In this stage, we used only the variables that presented a $p < 0.20$ on the χ^2 test.

MULTIPLE CORRESPONDENCE ANALYSIS

Whenever it is necessary to study the relations among a large number of variables simultaneously, tools of multivariate analysis, such as multiple correspondence analyses, can be used, as these techniques allow for synthetic representations of large sets of data. Correspondence analysis is an exploratory and descriptive statistical technique used in analyses of data organized in contingency tables for the purpose of verifying associations or similarities between qualitative or quantitative variables categorized without a probabilistic distribution defined *a priori*[12,13].

A graphic representation of the results obtained through the correspondence analysis displays the entire distribution of the characteristics studied, which can be subjectively interpreted as similarities. Each category of each variable is represented by a point and the distance from one point to the other represents the relations among the categories of the variables[13,14].

Our starting point to conduct the multiple correspondence analysis was an *(n x p)* matrix in which each *(n)* line corresponded to one patient and each *(p)* column referred to one characteristic studied. Each patient presents a (p_i, $i = 1,...,n$) profile defined by his/her characteristics; likewise, a (p_j, $j = 1,...,p$) profile can be drawn for each variable based on the patient's answers[12].

Considering the *(n x p)* matrix as a set of *n* points within a space of *p* dimension, the center of gravity of the mass of data corresponds to the mean value of all profiles, and can be therefore denominated the "profile expected value". The distances between each point and the center of gravity are distances between observed and expected values, which, for this reason, are called χ^2 distances[12,15].

The average of the χ^2 distances corresponds to a measure of similarity called *inertia*; it takes on a zero (0) value when all points of the data matrix are superimposed to the center of gravity. The total inertia can be decomposed in relative inertias pertaining to each one of the evaluated dimensions[15,16].

The square root of the inertia corresponds to a measurement called *eigenvalue*, which indicates how much of the total data variability is being explained by that dimension[15].

The analysis of the absolute contribution of each category — obtained through the inertia — along with the observation of the points on the graph of the correspondence analysis allow for the conceptual characterization of a graph's axis, also known as "dimensions". The relative contribution of a category, in turn, measures how much of the variability of a given category is explained by the analyzed dimension[14].

In the present study, it was expected that the graphic representation of the dimensions would display grouping areas of the categories of the variables included in the analysis on the categories of smoking, so that we could identify the predominant characteristics of the patients who smoked as well as of those who did not smoke.

With the purpose of complementing the interpretation of the results yielded by the multiple correspondence analysis, we devised a dendrogram that divides the data in similar groups based on the average of the coordinates obtained through the correspondence analysis[17].

The statistical procedures were carried out on the free R software, version 2.11 (The R Foundation for Statistical Computing, Vienna, Austria; http://www.r-project.org/) with the aid of the program ca version 0.33[18] and the statistical package Stata 9.0.

RESULTS

Out of the 2265 patients who met the study's inclusion criteria, 1131 had complete information about all the variables and were thus considered in the analysis of the data.

The highest percentages of lack of information were observed in relation to the variables tumor stage (29.5%), family history of cancer (20.0%) and alcoholism (17.6%).

SOCIODEMOGRAPHIC PROFILE OF THE PATIENTS

Overall, we observed a predominance of male patients (male/female ratio 2,3:1) who smoked (90.5%). Between 2000 and 2007, the prevalence of smoking increased about 1,5% per year on average.

Statistically significant differences were observed between smokers and non-smokers when the data were analyzed by sex, age range, marital status, and alcoholism (Table 1). The male/female sex ratio (2,6:1) among patients who smoked was almost 3 times higher when compared to the same ratio among non-smokers (0,9:1). We verified that the disease manifested itself in more advanced ages (64 years on average) among the non-smokers than amidst the smokers (61 years on average). The percentage of widowed individuals among the non-smokers was two times higher in relation to the smokers. Approximately two thirds of the patients who smoked also consumed alcohol; on the other hand, alcoholism was registered in less than one quarter of the individuals who did not smoke.

We did not find statistically significant differences regarding schooling and family history of cancer. Concerning years of study, a high percentage of patients with only a few years of schooling was observed, both smokers and non-smokers. The occurrence of cases of cancer in relatives up to the second degree was reported by about half of the patients, regardless of smoking (Table 1).

PROFILE OF TUMOR-RELATED CHARACTERISTICS

In regards to tumor-related characteristics, the differences observed between smokers and non-smokers were statistically significant (Table 1).

The predominant location of the tumor was the upper lobe of the lung, and it was 35.9% more frequent among smokers than among non-smokers. Although more rare, tumors located in the lower lobe of the lung were twice more frequent among the non-smokers in relation to the smokers. Among the non-smokers, cases of tumors located in the main bronchus were not registered.

Lung adenocarcinoma was the predominant histological type among the non-smokers, and it corresponded to more than half of the tumors in this group. Among the smokers, adenocarcinomas and squamous-cell carcinomas were the most frequent. Small-cell carcinomas occurred more frequently among the smokers.

Concerning the clinical stages, 85.9% of the patients arrived at INCA when the disease was advanced (stages III and IV). Among the patients who did not smoke, we observed an increase of 44.1% of patients in stages I and II in relation to those who smoked.

Table 1. Distribution of the study population according to characteristics related to the patients, tumors, and treatment by smoking status, *Hospital do Câncer/Instituto Nacional de Câncer*, 2000 – 2007.

Characteristics studied	Smokers		Non-smokers		Total		p-value
	n	%	n	%	n	%	
All cases	1023	90.5	108	9.5	1131	100	
Sex							
Male	740	72.3	51	47.2	791	69.9	< 0.001
Female	283	27.7	57	52.8	340	30.1	
Age range							
30 to 49 years	136	13.3	12	11.1	148	13.1	0.007
50 to 59 years	295	28.8	22	20.4	317	28.0	
60 to 69 years	371	36.3	35	32.4	406	35.9	
70 to 89 years	221	21.6	39	36.1	260	23.0	
Schooling							
Up to elementary school – not finished	619	60.5	63	58.3	682	60.3	0.737
Finished elementary school or more	404	39.5	45	41.7	449	39.7	
Marital status							
Married	674	65.9	65	60.2	739	65.3	0.008
Separated	86	8.4	8	7.4	94	8.3	
Single	163	15.9	13	12.0	176	15.6	
Widow(er)	100	9.8	22	20.4	122	10.8	
Family history of cancer							
Present	514	50.2	60	55.6	574	50.8	0.343
Absent	509	49.8	48	44.4	557	49.2	
Alcoholism							
Present	616	60.2	25	23.1	641	56.7	< 0.001
Ausent	407	39.8	83	76.9	490	43.3	
Detailed primary location							
Main bronchus	10	1.0	0	0.0	10	0.9	< 0.001
Upper lobe of the lung	515	50.3	40	37.0	555	49.1	
Middle lobe of the lung	12	1.2	6	5.6	18	1.6	
Lower lobe of the lung	119	11.6	21	19.4	140	12.4	
Overlapping lesion or unspecified lung	367	35.9	41	38.0	408	36.1	
Histological type							
Adenocarcinomas	427	41.7	62	57.4	489	43.2	0.002
Squamous carcinomas	339	33.1	20	18.5	359	31.7	
Other carcinomas	187	18.3	23	21.3	210	18.6	
Small-cell carcinomas	70	6.8	3	2.8	73	6.5	
Stages							
Stage I	75	7.3	12	11.1	87	7.7	0.022
Stage II	63	6.2	9	8.3	72	6.4	
Stage III	527	51.5	39	36.1	566	50.0	
Stage IV	358	35.0	48	44.4	406	35.9	
First treatment received							
Surgery	40	3.9	3	2.8	43	3.8	0.193
Radiotherapy	368	36.0	41	38.0	409	36.2	
Chemotherapy	217	21.2	32	29.6	249	22.0	
Initiated by surgery	29	2.8	4	3.7	33	2.9	
Initiated by radiotherapy	139	13.6	8	7.4	147	13.0	
Initiated by chemotherapy	230	22.5	20	18.5	250	22.1	
State of the disease at the end of the first treatment							
Complete remission	55	5.4	4	3.7	59	5.2	0.558
Partial remission	38	3.7	2	1.9	40	3.5	
Stable disease	50	4.9	3	2.8	53	4.7	
Disease in progression	222	21.7	29	26.9	251	22.2	
No possibility of treatment	211	20.6	27	25.0	238	21.0	
Death	187	18.3	19	17.6	206	18.2	
No follow-up	260	25.4	24	22.2	284	25.1	

PROFILE OF THE TREATMENT AND DISEASE EVOLUTION

Upon assessing the patients' situation in regards to the conduction of their first antineoplastic treatment (Table 1), we found that approximately half of the smokers were treated with radiotherapy (either isolated or as the initial step when more than one therapy was employed). Among the non-smokers, the predominant types of treatment were chemotherapy and radiotherapy, carried out separately. Surgeries, which are generally resorted to in the initial stages of lung cancer, were utilized by a small percentage of patients, regardless of their smoking condition. The differences in the proportions observed between smokers and non-smokers concerning the type of treatment used were not statistically significant at a significance level of 0.05, but this variable was included in the correspondence analysis whenever the p-value yielded by the χ^2 test was lower than 0.20.

Despite the high percentage of patients who were not followed up, their vital state at the end of the first treatment was included in this study to illustrate the prognostic of lung cancer patients diagnosed and assisted at INCA between 2000 and 2007. Around 40.0% of the patients either died or were considered unfit to undergo therapy during the course of the first treatment; for 26.9% the disease was either stable or in progression, and only 8.8% presented complete or partial remission at the end of the first treatment. These statistics were homogeneous for smokers and non-smokers alike.

RESULTS OF THE CORRESPONDENCE ANALYSIS

The first three dimensions explained 40.2, 32.6, and 14.1% of the total variability of the data, respectively. In the analysis that follows, only the first two dimensions — which together explained 72.8% of the data's variability — were considered.

As displayed on Table 2, it is possible to verify that the following categories of variables had an absolute contribution higher than 10% over dimension 1: stage I, isolated surgical treatment, and female sex. In dimension 2, the categories that stood out were isolated chemotherapy treatment, and stage IV.

In dimension 1, the categories that presented relative contributions above 70% were alcoholism, stage I, and isolated surgical treatment. In dimension 2, the categories that presented relative contributions above 70% were stage IV, isolated chemotherapy treatment, and squamous carcinomas.

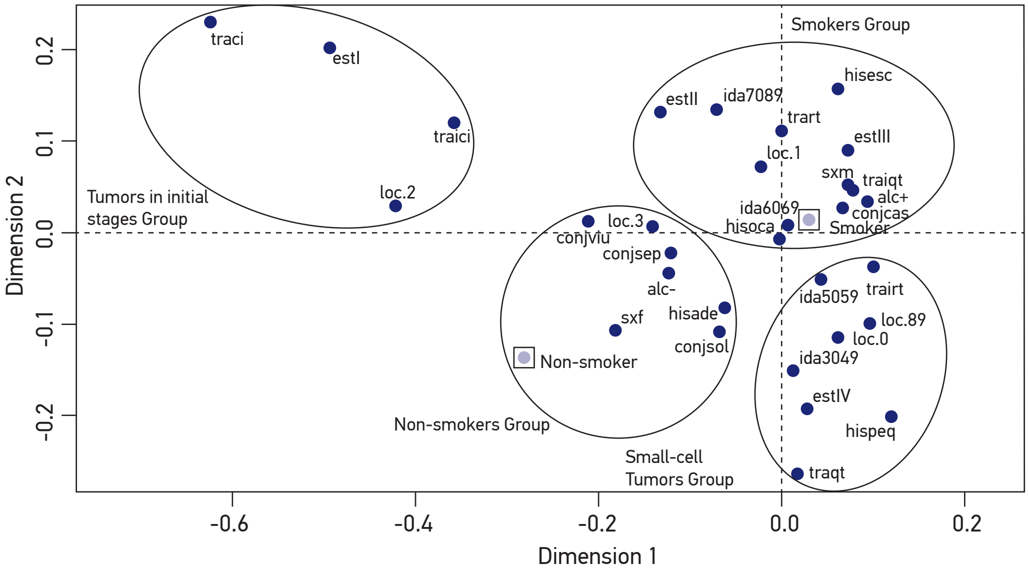
IDENTIFICATION OF THE GROUPS

Upon visual inspection of the joint distribution of the first two dimensions, obtained through the correspondence analysis (Figures 1 and 2), it was possible to identify the four groups described as follows:

Table 2. Absolute and relative contributions of the first and second dimensions of the correspondence analysis according to the characteristics studied.

Characteristics of the patients	Dimension 1		Dimension 2	
	Absolute	Relative	Absolute	Relative
Smoking habits				
Smoker	0.008	0.606	0.002	0.142
Non-smoker	0.074	0.606	0.021	0.142
Sex				
Male	0.042	0.628	0.018	0.219
Female	0.097	0.628	0.042	0.219
Age range				
30 to 49 years	0.000	0.003	0.036	0.427
50 to59 years	0.005	0.167	0.009	0.236
60 to 69 years	0.000	0.014	0.000	0.019
70 to 89 years	0.011	0.101	0.050	0.366
Marital status				
Married	0.028	0.678	0.006	0.109
Separated	0.012	0.473	0.001	0.016
Single	0.007	0.122	0.022	0.307
Widow(er)	0.047	0.430	0.000	0.002
Alcoholism				
Present	0.049	0.710	0.008	0.092
Absent	0.064	0.710	0.010	0.092
Detailed primary location				
Main bronchus	0.000	0.051	0.001	0.175
Upper lobe of the lung	0.003	0.064	0.031	0.645
Middle lobe of the lung	0.028	0.694	0.000	0.003
Lower lobe of the lung	0.024	0.648	0.000	0.002
Overlapping lesion or unspecified lung	0.033	0.427	0.043	0.452
Histological type				
Adenocarcinomas	0.016	0.267	0.035	0.470
Squamous carcinomas	0.012	0.126	0.095	0.821
Other carcinomas	0.000	0.001	0.000	0.005
Small-cell carcinomas	0.009	0.163	0.032	0.461
Stage				
Stage I	0.183	0.714	0.038	0.120
Stage II	0.011	0.392	0.013	0.389
Stage III	0.026	0.310	0.049	0.480
Stage IV	0.003	0.019	0.161	0.919
First treatment received				
Surgery	0.145	0.711	0.024	0.097
Radiotherapy	0.000	0.000	0.054	0.614
Chemotherapy	0.001	0.004	0.185	0.886
Initiated by surgery	0.037	0.527	0.005	0.060
Initiated by radiotherapy	0.013	0.453	0.002	0.062
Initiated by chemotherapy	0.013	0.298	0.006	0.109

- *Tumors in Initial Stages Group*: formed by patients who presented tumors in early stages during diagnosis (*estI*), located in the middle lobe of the lung (*loc.2*), and whose treatment was isolated or combined surgery (*traci* and *traici*). It clearly stands out among the other groups regardless of sociodemographic characteristics.
- *Small-cell Tumors Group*: among the patients with small-cell tumors (*hispeq*), we identified the patients who were under 60 years of age (*ida3049* and *ida5059*), whose tumor was on stage IV (*estIV*), located in the main bronchus or with no specific location (*loc.0* and *loc.89*). They were either treated with isolated chemotherapy (*traqt*) or their treatments were initiated with radiotherapy (*trairt*).
- *Non-Smokers Group*: among the non-smokers (*NON SMOKER*), we identified a group of female patients (*sxf*), with no partners (*conjsol, conjsep, conjviu*), who did not consume alcohol (*-alc*), and whose tumors were adenocarcinomas (*hisade*) located in the lower lobe of the lung (*loc.3*).
- *Smokers Group*: this group was composed of smokers (*SMOKER*) of the male sex (*sxm*), who were elderly (*ida6069* and *ida7089*), married (*conjcas*) and consumed alcohol (*+alc*). Their tumors – squamous carcinomas (*hisesc*) or other carcinomas (*hisoca*) - were on stage II or III (*stII* and *stIII*) and located in the upper lobe of the lung (*loc.1*). The treatments that stood out in this group were isolated radiotherapy (*trart*) and those initiated with chemotherapy (*traiqt*).



Fumante: smoker
Não fumante: non-smoker

sxm: male
sxf: female

ida3049: age group 30 to 49 years
ida5059: age group 50 to 59 years
ida6069: age group 60 to 69 years
ida7089: age group 70 to 89 years

conjcas: married
conjsep: divorced
conjsol: single
conjviu: widow

alc+: alcoholic
alc-: non-alcoholic

loc.0: main bronchus
loc.1: upper lobe
loc.2: middle lobe
loc.3: lower lobe
loc.89: overlapping or unspecified lesion

hisade: adenocarcinoma
hisesc: squamous carcinoma
hispeq: small cell carcinomas
hisoca: other carcinomas

estI: stage I
estII: stage II
estIII: stage III
estIV: stage IV

traci: surgery alone
trart: radiotherapy alone
traqt: chemotherapy alone
traici: treatment initiated by surgery
trairt: treatment initiated by radiotherapy
traiqt: treatment initiated by chemotherapy

Figure 1. Joint distribution of the correspondence analysis dimensions.

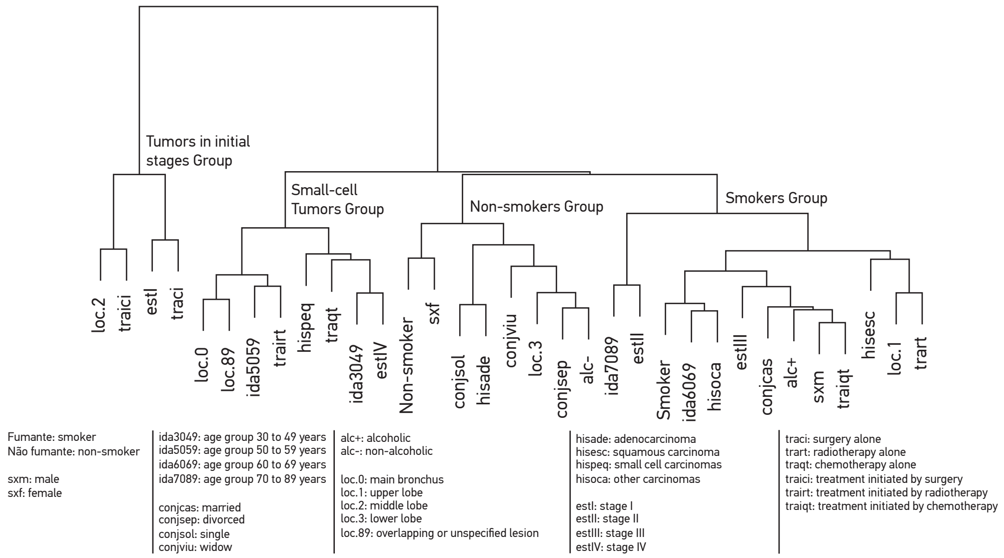


Figure 2. Dendrogram of the coordinates of the first two dimensions of the correspondence analysis.

DISCUSSION

The results of the present study are in agreement with previous findings on lung cancer with respect to the predominance of male patients who smoke[19-23].

In a review about the occurrence of lung cancer in the United States, the authors reported that around 10% of the patients were non-smokers and that non-smoking women are more frequently afflicted by this illness in comparison to men who do not smoke[24]. These results are also similar to those found in the present study.

Concerning the presence of smoking habits, the age range gradient — observed in the descriptive analysis — is coherent. Data from the Special Survey on Smoking, conducted in Brazil in 2008, reveal that the prevalence of smokers increases according to age, but that it is lower among more elderly individuals[25].

We observed a high proportion (43.2%) of patients with adenocarcinomas in the present study. In other national articles, this proportion has varied between 25.0 and 47.4%[19-23]. A review of 12 published studies revealed that the frequency of adenocarcinomas among non-smokers varied between 47.0 and 76.0%, and that the frequency of squamous carcinomas ranged between 3.0 and 27% within the same group[24]. Another review of studies showed that squamous-cell carcinomas were more common among smokers (35.7%) than amidst non-smokers (5.9%)[26]. The predominance of squamous carcinomas among smokers and of adenocarcinomas among non-smokers found in the present study is in accordance with the two reviews in question. Some authors have pointed out that although smoking increases the risk of lung cancer development, this factor has less influence over adenocarcinomas than over squamous carcinomas[8,9].

Overall, the distribution of cases according to the stage of the tumors found in this study was similar to that of other studies carried out with lung cancer patients in which stratifications were not conducted based on smoking status[19-21]. For some authors, it is not clear whether the clinical stage on the date of the diagnosis is different for non-smokers and smokers with lung cancer[24].

The multiple correspondence analysis allowed us to characterize four groups of patients. As displayed on Figure 1, the formation of the group of patients with tumors in initial stages stands out, regardless of smoking or other sociodemographic characteristics. According to the literature on the topic, surgery is the treatment that offers better prognoses for patients in initial stages[27]. The results of the present study indicate that this treatment is the most used in this group.

Generally, systemic chemotherapy is an important treatment component for patients with small-cell carcinomas, as this histological type is usually in advanced stages in the majority of the cases. For those in less advanced stages of the disease, radiotherapy is used together with chemotherapy[28]. This description is compatible with the characteristics of the group of patients with small-cell carcinomas observed on Figures 1 and 2.

We identified a group of smokers formed by male, elderly and married male patients who consumed alcohol. They had squamous carcinomas or other carcinomas in stages II or III and received isolated radiotherapy or treatments preceded by chemotherapy, which is generally employed in cases of worse prognoses in which therapy is still a possibility[29].

The group of non-smokers was composed of female, single patients who did not consume alcohol. They had adenocarcinomas located in the lower lobe of the lung.

LIMITATIONS

The main limitation of this study was the loss of a large number of cases due to a lack of important information, such as disease staging. This limitation can be circumvented by stimulating individuals in charge to completely fill in medical charts and information systems. Another limitation concerns the methodology, which, although useful to outline the patients' profile and thus point out priority groups that must be addressed, is not conducive to inferences based on what was found. In other words, the results herein described refer only to the population of this study.

CONCLUSION

Smoking is still the main cause of lung cancer, and every effort must be made to diminish the prevalence of the use of tobacco-derived products. In regards to primary prevention, smoking cessation must be stimulated in a disseminated manner in the population, without restrictions to subgroups. Nevertheless, the identification of characteristics that are common to patients with lung cancer can aid in planning more specific early diagnosis strategies and developing new targeted therapies.

Although lung cancer is a silent disease, generally detected in advanced stages, the Tumors in Initial Stages Group had a better prognosis and clearly stood out in relation to the other groups. The data used in the present study do not allow us to understand the reasons that led to the formation of this group, but given the higher chances of curing lung cancer in its initial stages, a detailed study of these patients' characteristics can contribute to the development of intervention measures that increase the proportion of cases detected early.

REFERENCES

1. Boyle P, Smans M, editors. Atlas of cancer mortality in the European Union and the European economic area 1993-1997. Lyon: International Agency for Research on Cancer Scientific Publication n. 159; 2008.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008 v1.2, cancer incidence and mortality worldwide. Lyon: International Agency for Research on Cancer CancerBase n.10; 2010. Disponível em http://globocan.iarc.fr. 9. (Acessado em 3 de fevereiro de 2012).
3. Schmidt MI, Duncan BB, Azevedo e Silva G, Menezes AM, Monteiro CA, Barreto SM, et al. Chronic non-communicable diseases in Brazil: burden and current challenges. Lancet 2011; 377(9781): 1949-61.
4. Souza MC, Vasconcelos AGG, Cruz OG. Trends in lung cancer mortality in Brazil from the 1980s into the early 21st century: age-period-cohort analysis. Cad Saúde Pública 2012; 28(1): 21-30.
5. Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA). Estimativa 2012: incidência de câncer no Brasil. Rio de Janeiro: INCA; 2011.
6. Organização Mundial da Saúde. WHO global report: mortality attributable to tobacco. Genebra: WHO; 2012.
7. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004; 328(7455): 1519.
8. Hadoux J, Besse B, Planchard D. Lung cancer in never smoker: epidemiology, molecular profiles and treatment. Presse Med 2011; 40(4 Pt 1): 371-8.
9. Uehara C, Jamnik S, Santoro I. Câncer de pulmão. Medicina (Ribeirão Preto) 1998; 31: 266-76.
10. Organização Mundial da Saúde. CID-O: Classificação Internacional de Doenças para Oncologia. 3a ed. São Paulo: EDUSP; 2005.
11. Instituto Nacional de Câncer (Brasil). TNM: classificação de tumores malignos. 6ª ed. Rio de Janeiro: INCA; 2004.
12. Greenacre MJ, Blasius J (eds.). Multiple correspondence analysis and related methods. Boca Raton: Chapman & Hall-CRC; 2006.
13. Carvalho MS, Struchiner CJ. Análise de correspondência: uma aplicação do método à avaliação de serviços de vacinação. Cad Saúde Pública 1992; 8(3): 287-301.
14. Mota JC, Vasconcelos AGG, Assis SG. Análise de correspondência como estratégia para descrição do perfil da mulher vítima do parceiro atendida em serviço especializado. Ciênc Saúde Coletiva 2007; 12(3): 799-809.
15. Pereira JCR. Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais. 3ª ed. São Paulo: EDUSP; 2004.
16. Paula FL, Fonseca MJM, Oliveira RVC, Rozenfeld S. Perfil de idosos com internação por quedas nos hospitais públicos de Niterói (RJ). Rev Bras Epidemiol 2010; 13(4): 587-95.
17. Maechler M, Rousseeuw P, Struyf A, Hubert M. Cluster analysis basics and extensions. Disponível em http://CRAN.R-project.org/package=cluster. (Acessado em 25 de julho de 2011).
18. Greenacre M, Nenadic O. ca: simple, multiple and joint correspondence analysis. Disponível em http://CRAN.R-project.org/package=ca. (Acessado em 25 de julho de 2011).
19. Franceschini J, Santos AA, El Mouallem I, Jamnik S, Uehara C, Fernandes ALG, et al. Avaliação da qualidade de vida em pacientes com câncer de pulmão através da aplicação do questionário Medical Outcomes Study 36-item Short-Form Health Survey. J Bras Pneumol 2008; 34(6): 387-93.
20. Barros JA, Valladares G, Faria AR, Fugita EM, Ruiz AP, Vianna AGD, et al. Diagnóstico precoce do câncer de pulmão: o grande desafio. Variáveis epidemiológicas e clínicas, estadiamento e tratamento. J Bras Pneumol 2006; 32(3): 221-7.

21. Mora PAR. Análise de sobrevida de pacientes com câncer de pulmão [dissertação de mestrado]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2004.
22. Uehara C, Santoro IL, Jamnik S. Câncer de pulmão: comparação entre os sexos. J pneumologia 2000; 26(6): 286-90.
23. Xavier F, Henn LA, Oliveira M, Orlandine L. Smoking and its relation to the histological type, survival, and prognosis among patients with primary lung cancer. São Paulo Med J 1996; 114(6): 1298-302.
24. Subramanian J, Govindan R. Lung cancer in 'Never-smokers': a unique entity. Oncology (Williston Park) 2010; 24(1): 29-35.
25. Instituto Nacional de Câncer (Brasil). Global adults tobacco survey Brazil 2008. Rio de Janeiro: INCA; 2010.
26. Toh CK, Lim WT. Lung cancer in never-smokers. J Clin Pathol 2007; 60(4): 337-40.
27. Raz DJ, Zell JA, Ou SH, Gandara DR, Anton-Culver H, Jablons DM. Natural history of stage I non-small cell lung cancer: implications for early detection. Chest 2007; 132(1): 193-9.
28. Puglisi M, Dolly S, Faria A, Myerson JS, Popat S, O'Brien ME. Treatment options for small cell lung cancer - do we have more choice? Br J Cancer 2010; 102(4): 629-38.
29. Bareschino MA, Schettino C, Rossi A, Maione P, Sacco PC, Zeppa R, et al. Treatment of advanced non small cell lung cancer. J Thorac Dis 2011; 3(2): 122-33.

Received on: 02/18/2012
Final version presented on: 09/16/2012
Accepted on: 02/13/2013