Published in final edited form as:

Lung Cancer. 2011 December ; 74(3): 378–383. doi:10.1016/j.lungcan.2011.04.020.

Association of smoking with tumor size at diagnosis in non-small cell lung cancer

Xing Chen^{1,*}, Ivan P. Gorlov², Kelly W. Merriman¹, Shih-Feng Weng¹, Millennia Foy^{1,3}, Gwendolyn Keener¹, Christopher I. Amos¹, Margaret R. Spitz¹, Marek Kimmel³, and Olga Y. Gorlova¹

¹Departments of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas

²Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas

³Department of Statistics, Rice University, Houston, Texas

Abstract

Tumor size at diagnosis (TSD) indirectly reflects tumor growth rate. The relationship between TSD and smoking is poorly understood. The aim of the study was to determine the relationship between smoking and TSD. We reviewed 1712 newly diagnosed and previously untreated non-small cell lung cancer (NSCLC) patients' electronic medical records and collected tumor characteristics. Demographic and epidemiologic characteristics were derived from questionnaires administered during personal interviews. Univariate and multivariate linear regression models were used to evaluate the relationship between TSD and smoking controlling for demographic and clinical factors. We also investigated the relationship between the rs1051730 SNP in an intron of the *CHRNA3* gene (the polymorphism most significantly associated with lung cancer risk and smoking behavior) and TSD. We found a strong dose dependent relationship between TSD and smoking. Current smokers had largest and never smokers smallest TSD with former smokers having intermediate TSD. In the multivariate linear regression model, smoking status (never, former, and current), histological type (adenocarcinoma vs SqCC), and gender were significant predictors of TSD. Smoking duration and intensity may explain the gender effect in predicting TSD. We found that the variant allele of rs1051730 in *CHRNA3* gene was associated with larger TSD of squamous cell carcinoma. In the multivariate linear regression model, both rs1051730 and smoking were significant predictors for the size of squamous carcinomas. We conclude that smoking is positively associated with lung tumor size at the moment of diagnosis.

Keywords

Lung cancer; tumor size; epidemiologic characteristics; risk factors; *CHRNA3*

*Corresponding author: Department of Epidemiology, Unit 1340, 1155 Pressler Street, Houston, TX 77030, Tel. (713) 792-7754, Fax. (713) 792-8261, cnxingchen@gmail.com.

CONFLICT OF INTEREST STATEMENT

None declared. There is no conflict of interest for all authors of manuscript "Association of smoking with tumor size at diagnosis in non-small cell lung cancer".

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

One of the most crucial biological characteristics of lung cancer is tumor growth rate or doubling time. Tumor growth rate is associated with cancer progression and survival^{1, 2}. Tumor growth rate is one of the major parameters in mathematical modeling of lung cancer progression³⁻⁵.

Observational studies strongly suggest a positive association between tumor size at diagnosis (TSD) and tumor growth rate. The current belief is that there is a marked variation in growth rates, which is reflected in variation of tumor size at detection. In particular, in a chest X-ray screening study, Usuda et al² noted a significantly longer doubling time in tumors smaller than 3 cm in diameter compared to those greater than 3 cm. In a CT screening study, Hasegawa et al.⁶ observed monotonously decreasing doubling times as the tumor size increased, within a group of small tumors (<2 cm), although based on a small case series. Therefore TSD can serve as an indirect, surrogate measure of the tumor growth rate. Unfortunately, direct estimates of the lung tumor growth rate are difficult to obtain and they exclusively come from lung cancer screening programs⁷, and therefore are likely to be biased toward slower-growing tumors⁸. Since the data on the TSD are easier to obtain than data on growth rate, we can use the TSD to identify epidemiological and clinical factors potentially associated with lung tumor growth rate, that might be indicators of more aggressive disease.

In this paper we report the results of the analysis of the association between TSD on one hand, and demographic, clinical, and epidemiologic variables such as sex, age, ethnicity, histologic cell type, smoking history, carcinogen exposure, and respiratory conditions, on the other. We also analyzed the effect of SNP rs1051730 in CHRNA3 gene on TSD. We chose this SNP because this is the most significant SNP associated with lung cancer risk independently in three genome wide studies⁹⁻¹¹.

PATIENTS AND METHODS

Patients

We reviewed the medical records of 1712 patients diagnosed with non-small cell lung cancer (NSCLC) between 1994 and 2009 at The University of Texas MD Anderson Cancer Center (Houston, Texas) and for whom tumor size measurements were available (Table 1). A subset of 1190 patients selected for the analysis was part of an ongoing lung cancer case-control study that had no age, sex, ethnicity, or stage restrictions. The reason for choosing these patients was the wide array of epidemiological characteristics available for them from the personal interviews based on a structured questionnaire¹². This study was approved by the MD Anderson Cancer Center's Office of Human Research Protections (OHRP) with institutional review board IRB00005015. Additionally, to expand the sample size for the earlier stages, we identified 522 stage I and II patients through the MDACC Tumor Registry, a database capturing clinical information including stage, cell type, and degree of differentiation, along with basic demographic information. Smoking histories for these additional patients were obtained from the Patient History Database that stores information on demographics and select exposures of all MDACC patients, including detailed tobacco use history. The study of these additional patents was also approved by the MDACC IRB. None of the patients was enrolled from a screening trial.

The following clinical information from patients' records was collected: tumor size (from radiology or pathology reports; see below); presence of chronic obstructive pulmonary disease or emphysema; and tumor histology/cell type and the degree of differentiation provided by the Tumor Registry and verified by chart review. Only non-small cell lung

cancer patients were included in this study. From the interviewer-administered questionnaires or Patient History Database, the following demographic and epidemiology characteristics were obtained: detailed smoking histories; age; sex; and ethnicity. Additionally, self-reported height and weight; self-reported dust, X-ray, and asbestos exposure; self-reported physician-diagnosed respiratory conditions (bronchitis, asthma, and hay fever); and hormone replacement therapy in women were available for the subset of patients enrolled in the case-control study. Never smokers were defined as those who reported having smoked less than 100 cigarettes in their lifetime. Former smokers were those who had quit more than 1 year prior to diagnosis. More recent quitters were grouped together with current smokers; henceforth this group is collectively referred to as current smokers.

Assessment of the tumor size

Radiologic tumor size was obtained from computed tomography scan reports; surgical tumor size was available from surgical reports through the electronic medical records system. In 1017 (59.4%) cases, the tumor size had been assessed radiologically, and in 1201 (70.2%), surgically; 506 (29.6%) patients had undergone both evaluations. We used the maximum measurement (if more than one dimension were available) for both radiologic size and surgical size. The maximum surgical size was used as the primary measure when radiological size was not available. The correlation between the radiologic and surgical maximum tumor dimensions based on the subset of patients with complete information was 0.71 ($N=506$; $p<0.001$).

Data on SNP rs1051730

We used genotyping data on the most significant SNP rs1051730 identified as the top candidate in three independent genome wide studies⁹⁻¹¹. CHRNA3 SNP rs1051730 has been implicated in both nicotine dependence and lung cancer risk^{12, 13}. Heterozygous individuals have 1.3-fold lung cancer risk and variant homozygotes 1.8-fold risk compared to wild type homozygotes^{9, 10}. Both radiologic tumor size and genotyping data were available for 532 patients.

Estimation the frequency of medical care utilization

The frequency of medical care may influence diagnosis, particularly through an incidental finding of an asymptomatic disease such as early lung cancer.¹⁴ The more exposure a patient has to medical care, the higher the chance a smaller tumor will be detected. Because data on patients' health care visits prior to NSCLC diagnosis were not available, we used data on medical service use from the National Health and Nutrition Examination Survey (NHANES). We stratified NHANES participants by age (25-39, 40-64, or ≥ 65 years), sex (male or female), race/ethnicity (non-Hispanic white, African-American, or Hispanic), and smoking status (never, former, or current). The combination of the categories resulted in 54 subgroups. For each subgroup, we calculated the mean number of health care visits, and these numbers were assigned as proxy values to each patient in our study according to his/her demographic characteristics.

Statistical Analysis

The *t*-test, Kruskal-Wallis test, and one-way analysis of variance were used to compare TSDs. Conservative Bonferroni correction was used to adjust for multiple testing, resulting in the nominal $p<0.003$ being significant. A multivariate linear regression analysis was performed to determine the effect of smoking history and gender on tumor size. We also conducted statistical analysis stratified by major histological types of lung cancer: adenocarcinoma and squamous cell carcinoma (SqCC). We also performed analyses

stratified by AJCC stages and by nodal status ($N=0$ to 3), excluding patients with distant metastases from the latter. All analyses were performed using SPSS software, version 16.0, for Windows.

RESULTS

The complete results of the analysis of associations between epidemiologic and clinical variables used in this study and TSD are shown in Supplementary Table 1. Out of 16 studied variables, three - histological type, smoking status, and gender - showed an association with TSD that remained significant after correction for multiple testing (Table 2). We found that adenocarcinoma tumors were significantly smaller than SqCC tumors (median, 2.80 versus 4.00 cm, respectively; $p<0.001$). Men presented with larger tumors (median, 3.4 cm) than women (median, 3.00 cm) ($p<0.001$). Never smokers had significantly smaller tumors (median, 2.50 cm) than did smokers (former smokers, 3.00 cm, $p=0.001$; current smokers, 3.50 cm, $p<0.001$). Among ever smokers, former smokers had significantly smaller tumors than did current smokers ($p=0.014$). There is a monotonic relationship between smoking status and TSD.

We next determined whether differences in stage explained the observed associations between tumor size and demographic and clinical variables. To distinguish between the effects of stage and demographic characteristics, we analyzed those variables that remained significant after the Bonferroni correction ($p<0.003$), as well as age (because of its high clinical significance), after stratifying by stage.

We found significant differences in stage distribution by age, sex, ethnicity, smoking status and histological cell type (Supplementary Table 2). Nevertheless, after stage stratification, the associations of tumor size with histological cell type, gender, and smoking status remained the same within each stage, although they did not always reach statistical significance. In all stages (I - IV), SqCC tumors were always larger than adenocarcinoma. (Figure 1)

Women presented with stage I disease more often than did men (Supplementary Table 2) but the stage difference did not completely explain the size difference because men tended to have larger tumors than did women at all stages (Figure 2).

In the stage-stratified analysis, the difference in tumor size remained significant only for stage I, IIIA and IIIB between groups with different smoking status (Figure 3). Ever smokers always presented with larger tumor size at diagnosis compared with never smokers, with an exception of stage II patients, likely due to the small number of cases in this group.

We also compared the tumor size after the stratification by N factor of TNM staging, limiting the analysis to those without distant metastases (M0). The results were consistent with the results of analysis by stage (Supplementary figures 1-3)

Multivariate general linear model with TSD as the outcome and histology, gender and smoking status as predictors was used to estimate the effects of these three TSD-associated variables in a single model. All three variables remained significant (smoking status, $p=0.015$; gender, $p=0.018$; histology, $p<10^{-6}$). After adjustment for stage status, the results remained virtually the same (smoking status, $p=0.017$; gender, $p=0.019$; histology, $p=2.2\times 10^{-5}$). In cell type-stratified analyses, neither smoking nor gender was significant in adenocarcinoma patients. In SqCC patients smoking status was a significant predictor of TSD ($p = 0.028$) while gender did not reach significance ($p = 0.06$).

To estimate the effect of health care use on tumor size at presentation, we used information on medical services use from the NHANES dataset (1999-2006). Patients with a high number of visits tended to present with smaller tumors, but this trend did not reach statistical significance ($p=0.338$). When we adjusted our previous comparisons for the estimated number of health care visits, the results remained virtually unchanged.

We also analyzed the association between rs1051730 and TSD. In the univariate analysis we found that 3 genotypes differed by their radiological TSD, with smallest TSD in normal homozygotes and largest in variant homozygotes in SqCC patients (Figure 4). In this group the difference between wild type homozygous and the heterozygous genotypes was borderline significant ($p=0.07$) and the difference between wild type and variant homozygotes was significant ($p=0.02$). After adjustment for smoking status, gender and stage status the effect of the SNP became even more significant ($p=0.001$). No difference in TSD between 3 genotypes was detected in adenocarcinoma patients.

DISCUSSION

Though all major histological types of lung cancer are associated with smoking, the association is stronger for squamous cell carcinoma than for adenocarcinoma¹⁵. By comparison, adenocarcinoma is the most common histological type of lung cancer in never smokers¹⁶. We found that SqCC patients have larger TSD than adenocarcinoma patients. We also found that smoking is associated with TSD in SqCC but not in adenocarcinoma patients. These two observations raise the possibility that smoking may be a primary driving force for larger TSD in smoking SqCC patients. Notably, when we analyzed never smokers only, there was no difference in TSD between SqCC and adenocarcinoma patients. In fact, in never smokers, TSD was non-significantly larger for adenocarcinoma than for SqCC patients (3.13 ± 0.20 and 2.75 ± 0.63 , correspondingly), although based on a very limited sample (only 5 squamous and 79 adenocarcinoma patients).

A number of studies^{17, 18} suggest that nicotine stimulates cell growth via activation of nicotinic cholinergic receptors (e.g. *CHRNA3*). Recent paper by Lam et al.¹⁹ found that nicotinic acetylcholine receptors change their expression in response to nicotine exposure. Our reanalysis of the lung tissue gene expression data from the study by Gruber et al.²⁰ shows that the expression of *CHRNA3* is higher in smokers (former plus current) compared to never smokers ($p=0.008$). These data suggest that tobacco smoke may accelerate tumor growth through up-regulation of *CHRNA3* and other cholinergic nicotinic receptors. This is consistent with our observation that the SNP rs1051730 has a significant effect on TSD.

Another significant predictor of TSD was gender. The gender difference in tumor size might reflect difference in smoking behavior between men and women. To differentiate between the effects of smoking and sex, we compared tumor sizes by sex, stratified by smoking status (Supplementary Table 3). We did not observe a sex difference in tumor size among never smokers. However, among former smokers (borderline significant) and current smokers (significant), men presented with larger TSD than women did. This suggests that larger TSD in men might be explained by men's heavier smoking (Supplementary Table 4) rather than by the gender difference. We therefore performed a multivariable regression analysis including gender as well as detailed smoking characteristics in the model. In former and current smokers, only the effect of smoking variables (age at smoking initiation ($p<0.001$), years of smoking ($p=0.004$), and number of cigarettes per day ($p<0.001$)) but not gender or age at diagnosis were significant in predicting the TSD (Supplementary Table 5). Therefore, smoking variables may explain the gender difference in tumor size.

One of the potential biases that can affect our results is the difference in number of doctor visits between smokers and non-smokers. According to our analysis of NHANES data, smokers use health care more frequently than non-smokers ($p = 0.032$) (Supplementary figure 4). However, this difference does not explain the observed differences in TSD. Indeed, after adjustment for the estimated frequency of health care visits, based on projections derived from NHANES data, we found the number of health care visits did not explain the observed difference in tumor size by sex, age, ethnicity, or smoking history.

The biological mechanism by which tobacco smoke might influence TSD is not completely understood. Yoshino et al.²¹ and Hecht et al.²² indicated that carcinogens in tobacco smoke may act not only as genetic inducers but also as promoters of the disease progression. A recent review study by Parsons²³ have reviewed 10 observational studies, showing that people who continue to smoke after a diagnosis of early stage lung cancer have higher risk of recurrence, second primary tumor compared with those who stop smoking at that time. These results together with our observation that smoking history is associated with larger tumors (possibly through faster progression) suggest that smoking may promote tumor growth and that cessation may be effective in slowing down the disease.

The difference between groups in tumor size observed in this study may seem moderate. However, this is the difference in linear size, which translates into a more considerable difference in terms of tumor volume. For example, a change from 3 cm to 3.7 cm in diameter corresponds to almost two-fold volume difference; for a typical volume doubling time of 207 days^{7, 24}, it would correspond to about half a year delay in diagnosis. A number of studies suggest that the tumor size may affect survival independently from stage. Wisnivesky et al.²⁵ studied an association between survival by tumor size in stage I in non-small cell lung cancer. They found that cure rates decreased with increasing tumor size among 7,620 stage I patients ($p < 0.05$). In a series of 246 stage IA lung cancer Gajra et al.²⁶ found that patients with tumors ≤ 1.5 cm in diameter had an improved outcome compared to those with tumors ≥ 1.5 cm ($p = 0.03$). The results of our analysis (not shown) also suggest a negative association between TSD and survival.

Although longitudinal tumor size measurements resulting from lung cancer screening^{2, 6} strongly suggest a positive association between TSD and tumor growth rate, factors independent of tumor biology TSD such as access to care, presence of symptoms, comorbidities etc also may influence TSD. Though the effect of such factors on TSD is difficult to estimate the available data suggest that this effect is likely to be relatively small. We found that the differences in the number of health care visits between smokers and non-smokers can not explain the difference in TSD between these groups. A recent study by Smith et al.²⁷ evaluated the association of demographic factors with the time between self-recognized symptoms and diagnosis. The authors found that patients with comorbidities were diagnosed ~ 12 days earlier compared to the patients without comorbidities. The authors also found that patients with knowledge of lung cancer symptoms were diagnosed, paradoxically, about a month later compared to those lacking such knowledge. Based on the existing estimates of a typical lung tumor doubling time of 207 days, a month-long delay will result in about 11% change of tumor volume, a noticeable but moderate effect on TSD.

Among the limitations of our study is a limited sample size for certain patient demographic, clinical, and epidemiologic groups. Carcinogen exposure and respiratory conditions were reported (except for emphysema) by patients, and the accuracy of these data has not been verified. By necessity, our sample was not representative of the overall patient population in terms of stage because little or no information was available on the size of the primary tumor in most stage IV patients. However, the relative proportions of stage I, II, IIIA, and IIIB cases in our study population did not significantly differ from those in the overall MD

Anderson lung cancer patient population. A further comparison of the stage distribution (excluding stage IV) by demographic subgroup revealed no significant differences between our patients and the overall MD Anderson lung cancer patient population (p -values not shown) for sex, age, ethnicity, or emphysema. The only subgroup that differed significantly was former smokers ($p = 0.011$): our dataset included more stage I cases than did the overall population (data not shown but available from the authors). We have addressed any possible bias related to stage distribution by performing stage-specific analyses.

Another limitation is that the MD Anderson patient population may be different from the U.S. lung cancer patient population; therefore, our findings should be verified in an independent study.

In conclusion, the results of this study suggest that lung tumor size at diagnosis is influenced by smoking. We found that the rs1051730 SNP is associated with lung tumor size for squamous cell carcinoma patients. The results of this study suggest that lung tumorigenesis may be driven not only by somatic mutations caused by tobacco carcinogens but also by pre-existing polymorphisms that may bring about physiological changes. Our findings are also important for development of mathematical models of the natural history of lung cancer. We and others have developed models to predict the lung cancer progression and estimate the screening efficiency.^{24, 28-30} With the very recent report from the NLST (National Lung Screening Trial) that screening with a spiral CT can reduce mortality among high risk subjects by 20 percent³¹, our findings may help to identify subsets of individuals for whom spiral CT screening may be most effective. Differences in tumor growth rate may require different screening schedule with early screening initiation and/or more frequent examinations for subgroups with higher growth rate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by the following grants: CISNET 5 U01 CA097431 to M.K. and O.G., FAMRI Young Clinical Scientist Award and Prevent Cancer Foundation grant to O.G., R03CA133885 and R03128025 to I.G., and CA55769 to M.R.S. We also thank the Department of Scientific Publications at MD Anderson for editorial assistance.

REFERENCES

1. Stein WD, Figg WD, Dahut W, Stein AD, Hoshen MB, Price D, et al. Tumor growth rates derived from data for patients in a clinical trial correlate strongly with patient survival: a novel strategy for evaluation of clinical trial data. *Oncologist*. 2008; 13(10):1046–54. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18838440. [PubMed: 18838440]
2. Usuda K, Saito Y, Sagawa M, Sato M, Kanma K, Takahashi S, et al. Tumor doubling time and prognostic assessment of patients with primary lung cancer. *Cancer*. 1994; 74(8):2239–44. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7922975. [PubMed: 7922975]
3. Weiss W. Implications of tumor growth rate for the natural history of lung cancer. *J Occup Med*. 1984; 26(5):345–52. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6327961. [PubMed: 6327961]
4. Geddes DM. The natural history of lung cancer: a review based on rates of tumour growth. *Br J Dis Chest*. 1979; 73(1):1–17. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=435370. [PubMed: 435370]

5. Bastogne T, Samson A, Vallois P, Wantz-Mezieres S, Pinel S, Bechet D, et al. Phenomenological modeling of tumor diameter growth based on a mixed effects model. *J Theor Biol.* 262(3):544–52. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19835891. [PubMed: 19835891]
6. Hasegawa M, Sone S, Takashima S, Li F, Yang ZG, Maruyama Y, et al. Growth rate of small lung cancers detected on mass CT screening. *Br J Radiol.* 2000; 73(876):1252–9. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11205667. [PubMed: 11205667]
7. Winer-Muram HT, Jennings SG, Tarver RD, Aisen AM, Tann M, Conces DJ, et al. Volumetric growth rate of stage I lung cancer prior to treatment: serial CT scanning. *Radiology.* 2002; 223(3): 798–805. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12034952. [PubMed: 12034952]
8. Patz EF Jr, Goodman PC, Bepler G. Screening for lung cancer. *N Engl J Med.* 2000; 343(22):1627–33. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11096172. [PubMed: 11096172]
9. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet.* 2008; 40(5):616–22. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18385676. [PubMed: 18385676]
10. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* 2008; 452(7187):633–7. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18385738. [PubMed: 18385738]
11. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008; 452(7187): 638–42. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18385739. [PubMed: 18385739]
12. Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *J Natl Cancer Inst.* 2008; 100(21):1552–6. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18957677. [PubMed: 18957677]
13. Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PA, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet.* 2007; 16(1):36–49. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17135278. [PubMed: 17135278]
14. Villamil Cajoto I, Garcia Prim JM, Gonzalez Barcala FJ, Pose Reino A, Paredes Vila S, Calvo Alvarez U, et al. Reason for a medical visit in the diagnosis of lung cancer. A review of 481 cases. *Rev Clin Esp.* 2009; 209(2):67–72. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19798842. [PubMed: 19798842]
15. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung Cancer.* 2001; 31(2-3):139–48. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11165392. [PubMed: 11165392]
16. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. *Nat Rev Cancer.* 2007; 7(10):778–90. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17882278. [PubMed: 17882278]
17. Zhang S, Togo S, Minakata K, Gu T, Ohashi R, Tajima K, et al. Distinct roles of cholinergic receptors in small cell lung cancer cells. *Anticancer Res.* 30(1):97–106. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20150622. [PubMed: 20150622]
18. Spindel ER. Is nicotine the estrogen of lung cancer? *Am J Respir Crit Care Med.* 2009; 179(12): 1081–2. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19498063. [PubMed: 19498063]

19. Lam DC, Girard L, Ramirez R, Chau WS, Suen WS, Sheridan S, et al. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. *Cancer Res.* 2007; 67(10):4638–47. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17510389. [PubMed: 17510389]
20. Gruber MP, Coldren CD, Woolum MD, Cosgrove GP, Zeng C, Baron AE, et al. Human lung project: evaluating variance of gene expression in the human lung. *Am J Respir Cell Mol Biol.* 2006; 35(1):65–71. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16498083. [PubMed: 16498083]
21. Yoshino I, Maehara Y. Impact of smoking status on the biological behavior of lung cancer. *Surg Today.* 2007; 37(9):725–34. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17713724. [PubMed: 17713724]
22. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst.* 1999; 91(14):1194–210. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10413421. [PubMed: 10413421]
23. Parsons A, Daley A, Begh R, Aveyard P. Influence of smoking cessation after diagnosis of early stage lung cancer on prognosis: systematic review of observational studies with meta-analysis. *BMJ.* 340:b5569. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20093278. [PubMed: 20093278]
24. Gorlova O, Peng B, Yankelevitz D, Henschke C, Kimmel M. Estimating the growth rates of primary lung tumours from samples with missing measurements. *Stat Med.* 2005; 24(7):1117–34. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15568189. [PubMed: 15568189]
25. Wisnivesky JP, Yankelevitz D, Henschke CI. The effect of tumor size on curability of stage I non-small cell lung cancers. *Chest.* 2004; 126(3):761–5. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15364754. [PubMed: 15364754]
26. Gajra A, Newman N, Gamble GP, Abraham NZ, Kohman LJ, Graziano SL. Impact of tumor size on survival in stage IA non-small cell lung cancer: a case for subdividing stage IA disease. *Lung Cancer.* 2003; 42(1):51–7. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14512187. [PubMed: 14512187]
27. Smith SM, Campbell NC, MacLeod U, Lee AJ, Raja A, Wyke S, et al. Factors contributing to the time taken to consult with symptoms of lung cancer: a cross-sectional study. *Thorax.* 2009; 64(6): 523–31. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19052045. [PubMed: 19052045]
28. Bartoszynski R, Edler L, Hanin L, Kopp-Schneider A, Pavlova L, Tsodikov A, et al. Modeling cancer detection: tumor size as a source of information on unobservable stages of carcinogenesis. *Mathematical Biosciences.* 2001; 171(2):113–42. Available from <Go to ISI>://000169363500001. [PubMed: 11395047]
29. Gorlova OY, Kimmel M, Henschke C. Modeling of long-term screening for lung carcinoma. *Cancer.* 2001; 92(6):1531–40. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11745232. [PubMed: 11745232]
30. Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA.* 2007; 297(9):953–61. Available from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17341709. [PubMed: 17341709]
31. HARRIS G. CT Scans Cut Lung Cancer Deaths, Study Finds. *The New York Times.* 2010 Available from http://www.nytimes.com/2010/11/05/health/research/05cancer.html?ref=national_cancer_institute.

Figure 1. Mean tumor size at diagnosis (TSD) in adenocarcinoma and SqCC stratified by stage. Except stage IIIB, SqCC patients always had significant larger tumors than did adenocarcinoma patients (*** $p < 0.001$, ** $p < 0.01$).

Figure 2. Mean tumor size at diagnosis (TSD) in men and women stratified by stage (*** $p < 0.001$, ** $p < 0.01$).

Figure 3.
Mean tumor size at diagnosis (TSD) in smoking stratified by stage (** $p < 0.01$, * $p < 0.05$)

Figure 4.
Mean radiological SqCC TSD for 3 genotypes in SNP rs1051730.

Table 1

Demographic characteristics of the study population

Characteristics	Tumor size dataset
Gender, n(%)	1704
Male	891 (52.3)
Female	813 (47.7)
Stage, n (%)	1712
I	846 (49.4)
II	289 (16.9)
IIIA	237 (13.8)
IIIB	233 (13.6)
IV*	107 (6.3)
Age, mean (SD)	64.25(10.66)
Race/Ethnicity, n (%)	1684
White	1451 (86.2)
African-American	169 (10.0)
Hispanic	64 (3.8)
Smoking, n (%)	1701
Never	238 (14.0)
Former	808 (47.5)
Current/Recent quitter	655 (38.5)

* Stage IV is underrepresented in our study population because little or no information was available on the size of the primary tumor in most stage IV patients

Table 2

Tumor size at presentation (maximum dimension) by demographic and clinical subgroups

Characteristic	<i>n</i>	Maximum size		
		Median/ mean, cm	IQR/SD*	<i>P</i> median/ <i>P</i> mean
Sex				
Male	891	3.40/4.08	2.2-5.3/2.61	<0.001/<0.001
Female	813	3.00/3.41	2.0-4.2/2.16	
Cell type				
Adenocarcinoma	472	2.80/3.33	2.00-4.00/2.02	<0.001/<0.001
Squamous cell carcinoma	241	4.00/4.55	2.50-6.00/2.66	
Smoking history				Pairwise <i>P</i> value
Never	238	2.50/3.15	1.60-4.00/2.43	N-F: <0.001/0.001
Former	808	3.00/3.72	2.00-5.00/2.27	N-C: <0.001/<0.001
Current/Recent quitter	655	3.50/4.03	2.00-5.00/2.58	F-C: 0.050/0.014

* IQR means interquartile range, SD means standard deviation