Published in final edited form as:

Lung Cancer. 2009 December ; 66(3): 386–392. doi:10.1016/j.lungcan.2009.02.012.

Impact and interactions between smoking and traditional prognostic factors in lung cancer progression

Nancy L. Guo^{*}, Kursad Tosun, and Kimberly Horn

Mary Babb Randolph Cancer Center/Department of Community Medicine, West Virginia University, Morgantown, WV 26506-9300, United States

Summary

Background—Cigarette smoking is a well-known risk factor of lung carcinogenesis. The clinical impact of smoking on lung cancer metastases and survival remains unclear. We sought to investigate the effect of smoking intensity on lung cancer treatment failure (represented by overall survival), and the interactions between smoking and clinicopathological factors in lung cancer progression.

Methods—Clinical information was obtained from four non-small cell lung cancer patient cohorts ($n = 347$). Twenty patients were excluded from the analysis because their smoking history was not available. The distribution of smoking intensity on patient age (≥ 60 y or < 60 y), gender, tumor differentiation (poor, moderate, and well differentiated), and clinical stage (1, 2, or 3) was assessed with Kruskal-Wallis rank sum tests. The effect of smoking on cause-specific lung cancer mortality was estimated by using Cox proportional hazard models and Kaplan-Meier analysis. The interactions between clinicopathological factors and smoking intensity with regard to lung cancer overall survival were evaluated with Analysis of Variance (ANOVA) for Cox modeling.

Results—Greater smoking intensity at diagnosis was found in older patients (≥ 60 y; $p = 0.022$), male ($p = 1.35e-7$), poorly differentiated tumors ($p = 8.51e-5$), patients with tumor stage 2 ($p = 0.031$), and squamous cell lung cancer patients ($p = 2.2e-16$). Patients who smoked more than 61 packs/year had an increased risk for lung cancer recurrence (hazard ratio = 1.41, 95% CI: [1.03, 1.94], log-rank $p = 0.032$) and shorter overall survival period (log-rank $p = 0.033$, Kaplan-Meier analysis) than those who smoked less than 61 packs/year. ANOVA analysis showed that smoking intensity ($p = 0.03$) and tumor stage ($p = 1.2e-6$) are the only significant prognostic factors of lung cancer, whereas patient age, gender, and tumor differentiation were not significant in lung cancer prognostication. There were significant interactions between smoking and clinical stage ($p = 0.02$) as well as patient age and tumor differentiation ($p = 0.03$) in lung cancer progression.

Conclusion—Smoking intensity at diagnosis is an independent, significant prognostic factor of non-small cell lung cancer. This factor could be used in patient selection for chemoprevention of tumor metastases and relapse. Additionally, the information may be used for clinically relevant tobacco prevention and intervention messages.

*Corresponding author: Nancy L. Guo, Ph.D., Assistant Professor, Mary Babb Randolph Cancer Center, Morgantown, WV 26506-9300, Tel: (304) 293-6455, Fax: (304) 293-4667, Email: lguo@hsc.wvu.edu.

Conflict of interest statement

The authors do not have any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations that could inappropriately influence (bias) this work.

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.

Keywords

cigarette smoking; non-small cell lung cancer; prognosis; tumor stage; tumor differentiation; adjuvant chemotherapy; overall survival

Introduction

Lung cancer remains the leading cause of cancer mortality, and its incidence is increasing worldwide (1). Non-small cell lung cancer (NSCLC) accounts for almost 80% of deaths from lung cancer (2;3). Most patients with NSCLC are diagnosed at advanced disease stage, and current multimodality therapy is of limited efficacy, with an overall 5-year survival rate of about 15% (4). A minority (~25–30%) of patients with non-small cell lung cancer have stage I disease and receive surgical resection as the major treatment option (5). Nevertheless, 35–50% of stage I NSCLC patients will relapse within five years (5), indicating that a subgroup of these patients might benefit from adjuvant chemotherapy. Similarly, patients with clinical stage IB, IIA or IIB, or IIIA NSCLC receive adjuvant chemotherapy, but some may unnecessarily receive potentially toxic chemotherapy (6). While tumor recurrence remains the major treatment failure for lung cancer, it is not currently possible to identify specific high-risk patients for chemoprevention.

Lung cancer is associated strongly with exposures to environmental carcinogens, with the highest population-attributable risk from cigarette smoking (1;7;8). Nevertheless, only approximately 10% of smokers develop lung cancer, and the disease also occurs in the absence of exposure to cigarette smoke (9). Recent studies identified genetic susceptibility locus for lung cancer carcinogenesis (4) and prognosis (9;10). In a Japanese population-based study, stage I lung adenocarcinoma patients with smoking intensity of less than 20 packs-year-index (PYI) showed a more favorable prognosis than those with a PYI of 20 or more (11). Similar results were observed in Cancer Prevention Studies conducted by American Cancer Society, indicating that cigarette smoking was more strongly associated with death from lung adenocarcinoma (12;13). Cigarette smoking causes a shortened life expectancy in general, and lung cancer mortality for ex-smokers is related to the number of years since cessation (14).

In this study, we hypothesize that smoking intensity by the time of diagnosis is an independent prognostic factor of non-small cell lung cancer, and this factor could be used in patient selection for chemoprevention. We sought to identify a clinically relevant indicator of smoking intensity for lung cancer prognostication in patients with resectable disease (stage I, II, or III). Furthermore, the association and interactions between smoking and clinicopathological factors, including patient age, gender, tumor differentiation, and clinical stage, were assessed in lung cancer progression.

Materials and Methods

Acquisition of Patient Cohorts and Clinical Information

Four non-small cell lung cancer patient cohorts were obtained from previous publications, including Beer et al. (5) ($n = 86$), Bhattacharjee et al. (15) ($n = 84$), Larsen et al. (16) ($n = 48$), and Raponi et al. (17) ($n = 129$). Patients from Beer et al. (5), Bhattacharjee et al. (15), and Larsen et al. (16) were diagnosed as lung adenocarcinoma, whereas patients from Raponi et al. (17) were squamous cell lung cancers. These data contain patient smoking information by the time of diagnosis, and clinical information including patient age, gender, tumor differentiation, stage, and overall survival status during the follow-up period. All patients were with resectable disease (stage I, II, or III). These data also contain genome-scale gene

expression profiles for each patient, which were used in a separate gene-environment interaction study to assess smoking-related gene expression patterns in lung cancer prognosis.

Association of Smoking Intensity and Clinicopathological Factors

This study analyzed the distribution of smoking intensity on five prognostic factors of non-small cell lung cancer, including patient age (≥ 60 y or < 60 y), gender, tumor differentiation (poor, moderate, or well differentiated), AJCC stage (1, 2, or 3) (18), and histological classification (lung adenocarcinoma vs. squamous cell lung cancer) using Kruskal-Wallis rank sum tests. The Kruskal-Wallis one-way analysis is a non-parametric method for testing equality of population medians among groups, which does not assume a normal population. However, the test does assume an identically-shaped distribution for each group, except for any difference in medians (19).

Smoking intensity indicated by the packs-year-index (PYI) was fitted in a Cox proportional hazards model. A clinically relevant cutoff in terms of PYI was identified as a significant predictor of the disease outcome. Survival probabilities in two different PYI-defined patient groups were assessed by Kaplan-Meier analysis and log-rank tests. The association between the smoking intensity-defined patient groups and clinicopathological parameters was evaluated by χ -square tests (two-sided). $P < 0.05$ indicates a significant association.

A Cox proportional hazard model defines the relationship between the survival of a patient and a set of predictors (e.g., smoking intensity). Let T be a variable representing the survival time with cumulative distribution function $P(t) = \Pr(T \leq t)$. The complement of the distribution function is the *survival function* $S(t) = \Pr(T > t) = 1 - P(t)$. The *hazard function* assesses the instantaneous risk of death at time t , conditional on survival to that time point:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

The Cox model (shown below) defines the hazard at time t for an individual with a given set of predictors denoted by $\mathbf{X}$:

$$h(t, \mathbf{X}) = h_0(t) e^{\sum_{i=1}^n \beta_i X_i} \quad \mathbf{X} = (X_1, X_2, \dots, X_n)$$

A hazard ratio is defined as the hazard for one individual divided by the hazard for a different individual:

$$\text{hazard ratio} = \frac{\widehat{h}(t, \mathbf{X}^*)}{\widehat{h}(t, \mathbf{X})} = e^{\sum_{i=1}^n \beta_i (X_i^* - X_i)}$$

where $\mathbf{X}^*$ denotes the set of predictors for one individual and $\mathbf{X}$ denotes the set of predictors for the other individual (20).

Interactions of Smoking Intensity and Clinicopathological Factors in Lung Cancer Progression

Analysis of Variance (ANOVA) for Cox modeling was used to evaluate the significance of smoking intensity and traditional prognostic factors (age, gender, tumor differentiation, and stage) on cause-specific lung cancer mortality. Furthermore, ANOVA for Cox modeling was

employed to assess all possible two-way interactions among these five studied factors in lung cancer progression.

All computational analyses were carried out with statistical packages in *R* (21).¹

Results

Distribution of smoking intensity on clinicopathological factors

In the studied cohort ($n = 327$), patients older than 60 smoked more intensely than patients younger than 60 ($p = 0.022$, Kruskal-Wallis rank sum test). The average smoking intensity was 163.8 packs/year in older patients (≥ 60 y; $n = 229$; median = 60 packs/year; SE = 16.6 packs/year) and 129 packs/year in younger patients (< 60 y; $n = 98$; median = 45 packs/year; SE = 23.4 packs/year). Male patients smoked more intensely than female patients ($p = 1.35e-07$, Kruskal-Wallis rank sum test). The average smoking intensity was 188.9 packs/year in male patients ($n = 175$; median = 75 packs/year; SE = 21.3 packs/year) and 112.4 packs/year in female patients ($n = 152$; median = 40 packs/year; SE = 15.3 packs/year). Patients who had poorly differentiated tumors smoked more than patients who had moderately or well differentiated tumors ($p = 8.51e-05$, Kruskal-Wallis rank sum test). The average smoking intensity was 193.2 packs/year in patients with poorly differentiated tumors ($n = 105$; median = 75 packs/year; SE = 24.4 packs/year), 172.5 packs/year in patients with moderately differentiated tumors ($n = 147$; median = 54 packs/year; SE = 21.4 packs/year), and 90.4 packs/year in patients with well differentiated tumors ($n = 75$; median = 40 packs/year; SE = 22.5 packs/year). Patients in stage 2 smoked more than those in stage 1 and stage 3 by the time of diagnosis ($p = 0.031$, Kruskal-Wallis rank sum test). The average smoking intensity was 129.7 packs/year in patients with stage 1 tumors ($n = 234$; median = 50 packs/year; SE = 14.1 packs/year), 229.6 packs/year in patients with stage 2 tumors ($n = 44$; median = 80 packs/year; SE = 43.6 packs/year), and 198 packs/year in patients with stage 3 tumors ($n = 49$; median = 50 packs/year; SE = 44.4 packs/year). Squamous cell lung cancer patients smoked more than lung adenocarcinoma patients ($p = 2.2e-16$, Kruskal-Wallis rank sum test). The average smoking intensity was 351.5 packs/year in squamous cell lung cancer patients ($n = 115$; median = 365 packs/year; SE = 30.8 packs/year) and 45.9 packs/year in lung adenocarcinoma patients ($n = 212$; median = 40 packs/year; SE = 2.24 packs/year). These results (Figure 1) showed that smoking intensity is significantly associated with classic clinicopathological factors of non-small cell lung cancer, including age, gender, tumor differentiation, and clinical stage. Furthermore, smoking intensity was found to be strongly associated with squamous cell lung cancer, whereas the association is significantly weaker for lung adenocarcinoma. These observations are consistent with the editorial report by Gazdar and Minna (12).

Smoking intensity in lung cancer prognostication

Smoking intensity indicated by packs-year-index (PYI) was fitted in a Cox hazard function to investigate its impact on lung cancer overall survival, independently of other clinicopathological factors. Patients were stratified into two groups by different smoking intensity cutoff values in the analyses. A series of smoking intensity cutoffs ranging from 61–84 packs/year were found to be significant predictors of lung cancer overall survival. For clinical prevention purposes, the smallest cutoff value of smoking intensity, 61 packs/year, was modeled in the Cox hazard function. Computed from the Cox model, the hazard ratio between lung cancer patients who smoked more than 61 packs/year and those who smoked less than 61 packs/year is 1.41 (coefficient $\beta = 0.345$, SE(β) = 0.161, log-rank $p = 0.031$; 95% CI: [1.03, 1.94]). These results indicate that lung cancer patients who smoked more than 61 packs/year at diagnosis had an increased risk of treatment failure (indicated by death from lung

¹<http://www.r-project.org/>

cancer) of 41% than those who smoked less than 61 packs/year. Kaplan-Meier analysis further confirmed the results (Figure 2A), showing that two smoking intensity-defined patient groups (stratified by 61 PYI) had significantly different postoperative survival periods ($p = 0.033$, log-rank test). To assess if smoking intensity-defined prognosis is dependent on histological classification of cell types or tumor stage, Kaplan-Meier analysis was performed on subgroups of patients with lung adenocarcinoma, squamous cell lung cancer, and stage 1, 2, and 3, respectively. The results showed that neither lung adenocarcinomas (Figure 2B) nor squamous cell lung cancers (Figure 2C) stratified by smoking intensity of 61 packs/year had significantly different postoperative survival periods. Stage 1 non-small cell lung cancer patients had significantly ($p = 0.046$) better prognosis in those who smoked less than 61 packs/year (Figure 2D), whereas the prognostic effect of less smoking was favorable but not significant in patients with stage 2 ($p = 0.083$; Figure 2E) or stage 3 ($p = 0.067$; Figure 2F). These results indicate that lower smoking intensity is associated with favorable postoperative overall survival of non-small cell lung cancer. This effect is most significant in early stage patients and is not specific to either lung adenocarcinoma or squamous cell lung cancer.

Furthermore, the association between the smoking intensity-stratified patient groups and lung cancer clinicopathological factors was evaluated with χ -square tests (two-sided). The results indicate that smoking intensity has significant association with patient age ($p < 0.011$), gender ($p < 9.1\text{e-}6$), tumor differentiation ($p < 0.0013$), tumor stage ($p < 0.0056$), and histology ($p < 3.7\text{e-}16$; Table 1). Together, smoking intensity indicated by a PYI of 61 at diagnosis is clinically relevant for prognostic prediction of lung cancer treatment outcome.

Interactions between smoking and clinicopathological factors in lung cancer progression

The above univariate analysis showed that smoking intensity indicated by a PYI of 61 is a significant prognostic factor of non-small cell lung cancer. Next, we sought to explore whether smoking intensity, together with other clinicopathological factors, are independent and significant predictors of lung cancer intervention outcome. Analysis of Variance (ANOVA) for Cox modeling was used to assess the significance of smoking intensity and traditional prognostic factors (age, gender, tumor differentiation and stage) on cause-specific lung cancer mortality. The results indicate that smoking intensity indicated by 61 PYI ($p = 0.03$) and tumor stage ($p = 1.2\text{e-}06$) are the only significant prognostic factors of lung cancer treatment outcome, whereas patient age (stratified by 60 y), gender, and tumor differentiation are not significant in the analysis (Table 2).

To investigate the interactions between cigarette smoking and clinicopathological factors in lung cancer progression, all possible two-way interactions between smoking intensity (stratified by 61 PYI), age (partitioned by 60 y), gender, tumor stage (1, 2, or 3), and tumor differentiation (poor, moderate, or well) were modeled in ANOVA analysis for Cox modeling of overall survival. There were significant interactions between smoking intensity and tumor stage ($p = 0.02$) as well as patient age and tumor differentiation ($p = 0.03$) in lung cancer prognostication. Other interactions were not significant in the analysis. These results are included in Table 3.

Quantification of cigarette smoking induced risk for lung cancer recurrence

To model the impact of cigarette smoking on lung cancer treatment outcome, smoking intensity was fitted in Cox modeling of patient overall survival as a continuous variable. In the studied cohort ($n = 327$), smoking intensity ranged from 0 to 1277.5 packs/year as recorded at diagnosis in the clinical data. Different smoking intensities indicated by 1/5 packs/day (≤ 4 cigarettes/day), 1/4–1/2 pack/day (5–9 cigarettes/day), 1/2 pack/day (10 cigarettes/day), 1 pack/day (20 cigarettes/day), and 2 packs/day (40 cigarettes/day) were analyzed (Table 4). The results showed that a patient who smoked 2 packs/day had an increased risk of 58% for lung cancer

relapse and metastasis than a non-smoking patient. On the other end of the spectrum, the risk for lung cancer recurrence was increased by 5% in a light smoker (≤ 4 cigarettes/day) than in a non-smoking patient. These results further confirmed that smoking intensity at diagnosis is a useful prognostic factor of lung cancer in clinical decision-making.

Discussion and Conclusion

The dynamic and diverse nature of lung cancer poses tremendous challenges for physicians. The current treatment regimens are based on the clinical staging system (18). Surgical resection is the major intervention option for early stage non-small cell lung cancer. Nevertheless, the high recurrence ratio in clinics greatly reduces lung cancer overall survival rate. Moreover, patients with the same disease stage may have remarkably distinct treatment response and clinical outcomes. It remains an unsolved critical issue to identify patients at high-risk for tumor recurrence to receive more aggressive therapy.

Cigarette smoking is a well known environmental risk factor of lung carcinogenesis. Tobacco smoke contains many mutagenetic and carcinogenic chemicals (22) that might be associated with mutations in tumor suppressor genes such as p53 and oncogenes such as K-ras (23–27). It is still unclear how smoking affects lung tumor progression, recurrence and metastasis, and whether smoking could be used as a predictor of lung cancer treatment outcome. With the completion of human genome project, it has become increasingly important to identify novel disease marker genes for the clinical management of lung cancer (5;15–17;28–31). Gene-environment interaction studies have identified genetic susceptibility locus associated with lung cancer risk (4) and prognosis (9;10) in smokers. In this study, we retrieved clinically annotated DNA microarray data from previous publications (5;15–17) and investigated the impact and interactions between smoking and traditional clinicopathological factors in lung cancer prognosis. In a separate study using the same data, gene-environment interaction analyses have been carried out to identify smoking-related biomarkers that could be used to predict lung cancer recurrence based on our previously identified prognostic gene signatures (31).

To explore whether smoking intensity at diagnosis is an independent prognostic factor of non-small cell lung cancer, multiple statistical models were employed in this study. First, the smallest significant cutoff of smoking intensity was identified for patient stratification by using Cox proportional hazard model. The univariate analysis indicated that patients who smoked more than 61 packs/year had an increased risk of 41% for lung cancer metastases than those who smoked less than 61 packs/year. Kaplan-Meier analysis confirmed that this smoking intensity-stratified patient groups had distinct postoperative overall survival periods (log-rank $p = 0.033$). Second, the identified smoking intensity index (61 PYI) was analyzed together with traditional lung cancer prognostic factors using ANOVA for Cox modeling. The analysis showed that smoking intensity and tumor clinical stage are the only significant predictors of lung cancer overall survival, whereas patient age, gender, and tumor differentiation were not significant in the statistical analysis. Third, the interactions between smoking and other clinicopathological factors in lung cancer progression were investigated using ANOVA for Cox modeling of overall survival. Interactions between smoking and tumor stage as well as patient age and tumor differentiation were found to be significantly associated with lung cancer clinical outcome. Finally, smoking induced risk for lung cancer relapse and metastasis was quantified by using Cox hazard model. Different smoking intensities ranging from 1/5 packs/day (≤ 4 cigarettes/day) to 2 packs/day (40 cigarettes/day) were analyzed. Compared with a non-smoking patient, the risk for postoperative treatment failure (represented by death from lung cancer) increases from 5% in a light smoker (≤ 4 cigarettes/day) to 58% in a heavy smoker (2 packs/day). Collectively, this study proved the hypothesis that smoking is an independent and significant prognostic factor of non-small cell lung cancer. A clinically relevant indicator

of smoking intensity (61 PYI) generated significant prognostic categorization for non-small cell lung cancer patients. This effect is most significant in stage 1 non-small cell lung cancer patients, and is not specific to cell types (lung adenocarcinoma or squamous cell lung cancer). Currently, surgical resection is the major treatment option for stage 1 non-small cell lung cancer. It remains a critical clinical challenge to identify early stage patients at high risk for metastasis and recurrence to receive adjuvant chemotherapy. To our knowledge, this is the first study to demonstrate that smoking intensity at diagnosis is an independent and significant prognostic factor. This factor could potentially, upon further validation in larger cohort study, be used in patient selection for postoperative chemotherapy.

The epidemiology and demographics of lung cancer have altered dramatically in the past 50 years, reflected by worldwide changes in the incidences, gender ratios, smoking associations, and pathologic types of lung cancer (12). The smoking rate and incidence rate of lung cancer in women are rising worldwide, more rapidly than in men (12;32). In recent years, lung adenocarcinoma, a rare tumor type in the early 20th century, has replaced squamous cell lung cancer as the most frequent cell type of non-small cell lung cancer (33). The observations in the US and abroad suggest that increases in lung adenocarcinoma since 1950 are more consistent with changes in smoking behavior and cigarette design than with diagnostic advances or histologic interpretation (13;34;35). In this study, significant association was found between smoking intensity and patient age, gender, histologic subtypes, tumor differentiation and stage in four non-small cell lung cancer patient cohorts. The summarized distribution of smoking intensity on the demographic and clinicopathological factors of lung cancer reflects the recent trend observed in clinics.

Tobacco use is the most important risk factor for lung cancer as well as many other cancers (36). Second hand tobacco smoke is also an established cause of lung cancer (37). The efficacy of antismoking campaign is beginning to result in lower cancer incidences in some part of the world. Nevertheless, the smoking epidemic continues to spread at an accelerated rate in underdeveloped and developing countries (38–40). The evidence on the reversal of health risks after quitting smoking cigarettes was recently assessed (41). Lung cancer risk is lower in former smokers than in current smokers (42), and the lower risk becomes apparent after 5–9 years after quitting smoking. The Lung Health Study reported an increase in forced expiratory volume in smokers with mild-to-moderate chronic obstructive pulmonary disease during the first year of smoking cessation (43;44). Symptoms of chronic bronchitis decrease rapidly within a few months of smoking cessation, and becomes the same as in never smokers within 5 years of sustained smoking abstinence (45). Quitting smoking avoids the further increase in risk for death from cancer, cardiovascular disease, and pulmonary disease caused by continued smoking (41). This study demonstrated that higher smoking intensity significantly increased the risk for death from lung cancer, and smoking intensity at diagnosis is an independent prognostic factor of lung cancer progression. Nevertheless, the effect of smoking cessation was not evaluated in the current study. Our future research will assess the reversal of risk in lung cancer metastases after quitting smoking.

Acknowledgments

This work is supported by the NIH/NCRR P20 RR16440-03 (Dr. Nancy L. Guo). We thank Dr. Geri Dino from West Virginia University for the thoughtful discussion and suggestion for this study. We appreciate the information provided by Dr. Carolyn Dresler at the Department of Health and Human Services for the state of Arkansas.

References

1. Christiani DC. Genetic susceptibility to lung cancer. *J Clin Oncol* 2006 Apr 10;24(11):1651–2. [PubMed: 16549817]

2. Hoffman PC, Mauer AM, Vokes EE. Lung cancer. *Lancet* 2000 Feb 5;355(9202):479–85. [PubMed: 10841143]
3. Spira A, Ettinger DS. Multidisciplinary management of lung cancer. *N Engl J Med* 2004 Jan 22;350(4):379–92. [PubMed: 14736930]
4. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008 Apr 3;452(7187):633–7. [PubMed: 18385738]
5. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002 Aug;8(8):816–24. [PubMed: 12118244]
6. Potti A, Mukherjee S, Petersen R, Dressman HK, Bild A, Koontz J, Kratzke R, Watson MA, Kelley M, Ginsburg GS, et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006 Aug 10;355(6):570–80. [PubMed: 16899777]
7. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, Hsieh LJ, Begg CB. Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003 Mar 19;95(6):470–8. [PubMed: 12644540]
8. Gorlova OY, Amos C, Henschke C, Lei L, Spitz M, Wei Q, Wu X, Kimmel M. Genetic susceptibility for lung cancer: interactions with gender and smoking history and impact on early detection policies. *Hum Hered* 2003;56(1–3):139–45. [PubMed: 14614248]
9. Spinola M, Meyer P, Kammerer S, Falvella FS, Boettger MB, Hoyal CR, Pignatiello C, Fischer R, Roth RB, Pastorino U, et al. Association of the PDCD5 locus with lung cancer risk and prognosis in smokers. *J Clin Oncol* 2006 Apr 10;24(11):1672–8. [PubMed: 16549820]
10. Miura K, Bowman ED, Simon R, Peng AC, Robles AI, Jones RT, Katagiri T, He P, Mizukami H, Charboneau L, et al. Laser capture microdissection and microarray expression analysis of lung adenocarcinoma reveals tobacco smoking- and prognosis-related molecular profiles. *Cancer Res* 2002 Jun 1;62(11):3244–50. [PubMed: 12036940]
11. Yoshino I, Kawano D, Oba T, Yamazaki K, Kometani T, Maehara Y. Smoking status as a prognostic factor in patients with stage I pulmonary adenocarcinoma. *Ann Thorac Surg* 2006 Apr;81(4):1189–93. [PubMed: 16564239]
12. Gazdar AF, Minna JD. Cigarettes, sex, and lung adenocarcinoma. *J Natl Cancer Inst* 1997 Nov 5;89(21):1563–5. [PubMed: 9362148]
13. Thun MJ, Lally CA, Flannery JT, Calle EE, Flanders WD, Heath CW Jr. Cigarette smoking and changes in the histopathology of lung cancer. *J Natl Cancer Inst* 1997 Nov 5;89(21):1580–6. [PubMed: 9362155]
14. Fielding JE. Smoking: health effects and control (1). *N Engl J Med* 1985 Aug 22;313(8):491–8. [PubMed: 3894970]
15. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, Ladd C, Beheshti J, Bueno R, Gillette M, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci USA* 2001 Nov 20;98(24):13790–5. [PubMed: 11707567]
16. Larsen JE, Pavay SJ, Passmore LH, Bowman RV, Hayward NK, Fong KM. Gene Expression Signature Predicts Recurrence in Lung Adenocarcinoma. *Clin Cancer Res* 2007 May 15;13(10):2946–54. [PubMed: 17504995]
17. Raponi M, Zhang Y, Yu J, Chen G, Lee G, Taylor JM, Macdonald J, Thomas D, Moskaluk C, Wang Y, et al. Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res* 2006 Aug 1;66(15):7466–72. [PubMed: 16885343]
18. AJCC Cancer Staging Manual. Springer; 2002.
19. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 1952;47(260):583–621.
20. Cox D. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B, Methodological* 1972;34:187–220.
21. Everitt, B.; Hothorn, T. *A Handbook of Statistical Analyses Using R*. Boca Raton, FL: Chapman & Hall/CRC; 2006.
22. Carbone D. Smoking and cancer. *Am J Med* 1992 Jul 15;93(1A):13S–7S. [PubMed: 1496998]

23. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* 1996 Oct 18;274(5286):430–2. [PubMed: 8832894]
24. Vineis P, Caporaso N. Tobacco and cancer: epidemiology and the laboratory. *Environ Health Perspect* 1995 Feb;103(2):156–60. [PubMed: 7737063]
25. Westra WH, Slebos RJ, Offerhaus GJ, Goodman SN, Evers SG, Kensler TW, Askin FB, Rodenhuis S, Hruban RH. K-ras oncogene activation in lung adenocarcinomas from former smokers. Evidence that K-ras mutations are an early and irreversible event in the development of adenocarcinoma of the lung. *Cancer* 1993 Jul 15;72(2):432–8. [PubMed: 8319174]
26. Westra WH, Offerhaus GJ, Goodman SN, Slebos RJ, Polak M, Baas IO, Rodenhuis S, Hruban RH. Overexpression of the p53 tumor suppressor gene product in primary lung adenocarcinomas is associated with cigarette smoking. *Am J Surg Pathol* 1993 Mar;17(3):213–20. [PubMed: 8434702]
27. Westra WH, Baas IO, Hruban RH, Askin FB, Wilson K, Offerhaus GJ, Slebos RJ. K-ras oncogene activation in atypical alveolar hyperplasias of the human lung. *Cancer Res* 1996 May 1;56(9):2224–8. [PubMed: 8616876]
28. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006 Jan 19;439(7074):353–7. [PubMed: 16273092]
29. Hayes DN, Monti S, Parmigiani G, Gilks CB, Naoki K, Bhattacharjee A, Socinski MA, Perou C, Meyerson M. Gene expression profiling reveals reproducible human lung adenocarcinoma subtypes in multiple independent patient cohorts. *J Clin Oncol* 2006 Nov 1;24(31):5079–90. [PubMed: 17075127]
30. Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, Eschrich S, Jurisica I, Giordano TJ, Misek DE, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008 Aug;14(8):822–7. [PubMed: 18641660]
31. Guo L, Ma Y, Ward R, Castranova V, Shi X, Qian Y. Constructing molecular classifiers for the accurate prognosis of lung adenocarcinoma. *Clin Cancer Res* 2006 Jun 1;12(11):3344–54. [PubMed: 16740756]
32. Center for Disease Control and Prevention. Morbidity and Mortality Weekly Report. Tobacco Use Among Adults---U.S., 2005. 2006. 1145 p
33. Travis WD, Travis LB, Devesa SS. Lung cancer. *Cancer* 1995 Jan 1;75(1 Suppl):191–202. [PubMed: 8000996]
34. Ernster VL. The epidemiology of lung cancer in women. *Ann Epidemiol* 1994 Mar;4(2):102–10. [PubMed: 8205275]
35. Levi F, Franceschi S, La VC, Randimbison L, Te VC. Lung carcinoma trends by histologic type in Vaud and Neuchatel, Switzerland, 1974–1994. *Cancer* 1997 Mar 1;79(5):906–14. [PubMed: 9041152]
36. Cancer Epidemiology and Prevention. Vol. 2. New York, NY: Oxford University Press; 1996.
37. Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. *BMJ* 1997 Oct 18;315(7114):980–8. [PubMed: 9365295]
38. Peto R, Lopez AD, Boreham J, Thun M, Heath C Jr, Doll R. Mortality from smoking worldwide. *Br Med Bull* 1996 Jan;52(1):12–21. [PubMed: 8746293]
39. Rachtan J, Sokolowski A. Risk factors for lung cancer among women in Poland. *Lung Cancer* 1997 Oct;18(2):137–45. [PubMed: 9316005]
40. Wynder EL, Muscat JE. The changing epidemiology of smoking and lung cancer histology. *Environ Health Perspect* 1995 Nov;103(Suppl 8):143–8. [PubMed: 8741774]
41. Dresler CM, Leon ME, Straif K, Baan R, Secretan B. Reversal of risk upon quitting smoking. *Lancet* 2006 Jul 29;368(9533):348–9. [PubMed: 16876647]
42. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004 Jun 26;328(7455):1519. [PubMed: 15213107]
43. Anthonisen NR, Lindgren PG, Tashkin DP, Kanner RE, Scanlon PD, Connett JE. Bronchodilator response in the lung health study over 11 yrs. *Eur Respir J* 2005 Jul;26(1):45–51. [PubMed: 15994388]

44. Scanlon PD, Connett JE, Waller LA, Altose MD, Bailey WC, Buist AS. Smoking cessation and lung function in mild-to-moderate chronic obstructive pulmonary disease. The Lung Health Study. *Am J Respir Crit Care Med* 2000 Feb;161(2 Pt 1):381–90. [PubMed: 10673175]
45. Brown CA, Crombie IK, Smith WC, Tunstall-Pedoe H. The impact of quitting smoking on symptoms of chronic bronchitis: results of the Scottish Heart Health Study. *Thorax* 1991 Feb;46(2):112–6. [PubMed: 2014491]

Figure 1. Distribution of smoking intensity on clinicopathological factors of non-small cell lung cancer, including (A) patient age, (B) gender, (C) tumor stage, (D) differentiation, and (E) histology.

A Non-small cell lung cancer patients

B**Lung adenocarcinomas**

Figure 2. Smoking intensity-defined patient groups had distinct overall survival in Kaplan-Meier analysis. In all figure panels, the upper curves represent patient group who smoked less than

61 packs/year, and the lower curves represent patient group who smoked more than 61 packs/year. A. Non-small cell lung cancer (NSCLC) patients who smoked less than 61 packs/year had longer ($p = 0.033$, log-rank test) postoperative survival periods than those who smoked more than 61 packs/year. B. Lung adenocarcinomas stratified by 61 packs/year did not have significant ($p = 0.12$) postoperative survival periods. C. Squamous cell lung cancers stratified by 61 packs/year did not have significant ($p = 0.24$) postoperative survival periods. D. Stage 1 NSCLC patients had significantly ($p = 0.046$) better prognosis in those who smoked less than 61 packs/year. E. Stage 2 NSCLC patients had better ($p = 0.083$) prognosis in those who smoked less than 61 packs/year. F. Stage 3 NSCLC patients had better ($p = 0.067$) prognosis in those who smoked less than 61 packs/year.

Table 1Association between smoking intensity and clinicopathological factors evaluated by χ -square tests (two-sided).

Clinicopathological Factors	Number of patients with smoking intensity < 61 packs/ year	Number of patients with smoking intensity $\geq$ 61 packs/ year	P-value
Age			0.011
< 60 y	68	30	
$\geq$ 60 y	123	106	
Gender			9.1e-06
Female	109	43	
Male	82	93	
Tumor differentiation			0.0013
Poor	50	56	
Moderate	86	61	
Well	55	19	
Clinical stage			0.0056
1 (IA or IB)	146	88	
2 (IIA or IIB)	16	28	
3 (IIIA or IIIB)	29	20	
Histology			3.7e-16
Lung adenocarcinoma	159	53	
Squamous cell lung cancer	32	83	

P < 0.05 indicates a significant association between smoking intensity indexed with 61 pack-per-year and the corresponding clinicopathological factors.

Table 2

Smoking intensity and traditional clinicopathological factors in lung cancer prognostication. ANOVA analysis for Cox modeling showed that smoking intensity ($p = 0.03$) and tumor stage ($p = 1.2\text{e-}6$) are the only significant prognostic factors of lung cancer.

	Degree of freedom	Deviance	P-value
Smoking	1	4.52	0.03
Age	1	2.99	0.08
Gender	1	0.63	0.43
Stage	2	27.27	1.2e-06
Differentiation	2	5.26	0.07

Table 3

Interactions between smoking intensity and traditional clinicopathological factors in lung cancer progression. ANOVA analysis for Cox modeling showed that the only significant interactions are between smoking intensity and tumor stage ($p = 0.02$) as well as patient age and tumor differentiation ($p = 0.03$).

	Degree of freedom	Deviance	P-value
Smoking & Stage	2	8.02	0.02
Age & Differentiation	2	7.36	0.03
Smoking & Age	1	1.72	0.19
Smoking & Gender	1	0.08	0.77
Smoking & Differentiation	2	3.06	0.22
Age & Gender	1	0.03	0.86
Age & Stage	2	0.55	0.76
Gender & Stage	2	4.07	0.13
Gender & Differentiation	2	4.41	0.11
Stage & Differentiation	4	3.64	0.46

Table 4

Hazard ratios representing the risk for lung cancer relapse and metastasis associated with different smoking intensity at diagnosis.

Smoking Intensity	Hazard Ratio	95% Confidence Interval
2 packs/day (40 cigarettes/day)	1.5793	(1.5783, 1.5802)
1 pack/day (20 cigarettes/day)	1.2567	(1.2559, 1.2575)
½ pack/day (10 cigarettes/day)	1.1210	(1.1204, 1.1217)
¼–½ pack/day (5–9 cigarettes/day)	1.0588–1.1210	(1.0581, 1.0594)–(1.1204, 1.1217)
1/5 pack/day (≤4 cigarettes/day)	1.0468	(1.0461, 1.0474)

Smoking intensity was fitted in a Cox model of lung cancer overall survival as a continuous variable with R . Computed from the Cox model, the hazard ratio between two patients, X and X^* , is $e^{0.000626(X-X^*)}$ (coefficient $\beta = 0.000626$, $SE(\beta) = 0.000307$, log-rank $p = 0.041$). For instance, the hazard ratio between a patient who smoked 365 packs/year (20 cigarettes/day) and a non-smoking patient is $e^{0.000626(365-0)} = 1.25$. Thus smoking 1 pack/day increases the hazard or risk of lung cancer recurrence by 25% compared with a non-smoking lung cancer patient.