Title
Single Nucleotide Polymorphisms (5p15.33, 6p22.1 and 15q25.1), Smoking and Lung Cancer Survival

Author
ZHANG, WEN

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Single Nucleotide Polymorphisms (5p15.33, 6p22.1 and 15q25.1), Smoking and Lung Cancer Survival

A thesis submitted in partial satisfaction

of the requirement for the degree

Master of Science in Epidemiology

By

Wen Zhang

2015

ABSTACT OF THESIS

Single Nucleotide Polymorphisms (5p15.33, 6p22.1 and 15q25.1), Smoking and Lung Cancer Survival

By

Wen Zhang

Master of Science in Epidemiology

University of California, Los Angeles, 2015

Professor Zuofeng Zhang, Chair

Background: Genome-wide studies of single nucleotide polymorphisms (SNPs) reported associations between chromosomal regions (5p15.33, 6p22.1 and 15q25.1) and lung cancer susceptibility. However, there are few published studies of the relationship between SNPs at these chromosomal regions and survival of lung cancer. **Method**: Six hundred and eleven lung cancer patients were followed up with a median duration of 11 years. Nine SNPs were genotyped using TaqMan assay. Cox proportional hazards regression model was employed to investigate their association with lung cancer survival. We also explored potential interactions of selected SNPs' with smoking on survival. **Results:** Overall, a total of 406 deaths occurred during the follow-up period. In survival analysis, only rs4324798 on chromosome 6p22.1 was associated with survival of lung cancer in dominant model (aHR: 1.51, 95%CI=1.11, 2.07). The association was also observed in non-small cell lung cancer (aHR: 1.56, 95%CI=1.09, 2.23) and large cell

lung cancer (aHR: 4.15, 95%CI=1.63, 10.52). Association of AG+AA of rs4324798 and lung cancer survival was found among smokers (aHR: 1.5, 95%CI: 1.1, 2.1), but not among non-smokers (aHR: 3.5, 95%CI: 0.4, 29.4). **Conclusion**: We observed an association of rs4324798 (6p22.1) with survival of lung cancer patients. The association remained among NSCLC and LCC histological subtypes, and smokers, although small numbers in some strata suggest caution in the interpretation of these findings.

The thesis of Wen Zhang is approved.

Linda B. Bourque

Onyebuchi A. Arah

Zuo-Feng Zhang, Committee Chair

University of California, Los Angeles

2015

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

The incidence and mortality rates of lung cancer rose rapidly since 1930's, primarily due to the popularity of cigarette smoking [1]. According to Globocan 2012 [2], lung cancer was the most common cancer and the leading cause of death from cancer worldwide , accounting for 1.8 million new cases and 1.6 million deaths. In the United States, lung cancer is the leading cause of cancer death among both men and women [3, 4] which causes more deaths than the next three leading causes of cancer death combined (colorectal, breast and pancreatic) [4]. An estimated 158,040 Americans are expected to die from lung cancer in 2015, accounting for approximately 27 percent of all cancer deaths [4]. The number of deaths due to lung cancer has increased approximately 3.5 percent between 1999 and 2012 from 152,156 to 157,499 [1]. The number of deaths among men has reached a plateau but the number is still rising among women [1]. Approximately 402,326 Americans living today have ever been diagnosed with lung cancer [5]. During 2015, an estimated 221,200 new cases of lung cancer were expected to be diagnosed, representing about 13 percent of all cancer diagnoses [4]. Lung cancer is mostly a disease of the elderly. In 2011, 82 percent of those living with lung cancer were 60 years of age or older [5]. The overall 5-year survival of lung cancer is 16.8 present in the U.S. in 2014, which is lower than many other leading cancer sites, such as colon (65.4%), breast (90.5%) and prostate (99.6%) [6]. The 5-year survival rate for lung cancer is 54.0 percent for cases detected when the disease is still localized [4]. However, only 15% of lung cancer cases are diagnosed at an early stage [4]. For distant tumors (spread to other organs) the 5-year survival is only 4.0 percent. Over half of lung cancer patients die within one year of being diagnosed [4].

Lung cancers are broadly classified into two types: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). This classification is based on histological type [7, 8]. SCLC

comprises about 10%-15% of lung cancers. This type of lung cancer is the most aggressive and rapidly growing of all the types. SCLC is strongly related to cigarette smoking with only 1% of these tumors occurring in non-smokers. NSCLC is the most common lung cancer type, accounting for about 85%-90% of all cases. NSCLC can be further divided into 3 subtypes: adenocarcinoma (ADC), squamous cell carcinoma (SQC) and large cell carcinoma (LCC) [8]. ADC is the most common type of NSCLC in the U.S. and comprise up to 40% of NSCLC cases. While ADC is associated with smoking like other lung cancers, this type is also seen in non-smokers [9]. It is increasing in frequency and is very common in non-smoking women and in the Asian population, and may have a better long-term survival [10]. SQC accounts for about 25% of NSCLC cases. This type of lung cancer most often stays within the lung, spreads to lymph nodes, and grows quite large, forming a cavity. LCC accounts for 10%-15% of all lung cancers. This type of cancer has a high tendency to spread to the lymph nodes and distant sites [7, 10].

A substantial amount of clinical and basic science research has focused on the prognostic factors for patients with lung cancer. Early investigations focused on clinical characteristics of the tumor and of the patient, such as stage and weight loss, respectively [11]. A number of clinical laboratory test, such as serum lactate dehydrogenase (LDH) level, were subsequently identified as being relevant, followed by investigation of factors in cellular and molecular biological level [12, 13, 14].

On the clinical level, prognostic factors in NSCLC include presence or absence of pulmonary symptoms, tumor size, histology, stage and metastases to multiple lymph nodes, and vascular invasion. For people with inoperable disease, outcomes are worse in those with poor performance status and weight loss of more than 10% [15]. Prognostic factors in small cell lung

cancer include performance status, gender, stage, and involvement of the central nervous system or liver at the time of diagnosis [16].

Most recently, several genome-wide association studies (GWAS) provided strong evidence of susceptibility SNPs: rs16969968, rs8034191 (15q25.1), rs402710 (5p15.33) [17-19], and rs4324798 (6p22)to be associated with the risk of lung cancer. Following studies provided validation in a Japanese population [20] and some evidence for biological plausibility: the region 15q25.1 (rs16969968 and rs8034191), which is located in the nicotinic cholinergic receptor subunit cluster, has been consistently associated with various aspects of smoking behavior including nicotine addiction initiation, long-term dependence and the transition from light to heavy smoker, in European and Japanese populations [21-23]; rs402710 was also correlated with high tumor DNA adduct levels in a group of NSCLC patients [24].

Extending such associations to lung cancer survival can provide insight into the mechanistic pathways and contribute to knowledge on the prognostic factors of lung cancer. A recent validation study by Yang et al. [25] found that of the top five lung cancer candidate SNPs (rs16969968, rs8034191, rs402710, rs2736100, and rs4324798), only rs4324798 (6p22.1) significantly altered overall survival time in 235 SCLC patients. Another study of European population by Xun et al. [26] found that SNPs at chromosomal regions 15q25.1 (rs8034191 and rs16969968) and 6p22.1 (rs4324798) to be relevant to survival in lung cancer patients who were never-smokers, while the rs402710 SNP at 5p15.33 was associated with survival among current smokers. Previous studies of SNPs at 5p15.33, 6p22.1 and 15q25.1 and their association of lung cancer survival are present in Table 1.

In this study, we aimed to estimate the associations between SNPs identified from lung cancer

GWAS and survival among lung cancer patients. We also explored potential interactions of these SNPs and smoking on lung cancer survival.

Table 1. Previous studies of SNPs (5p15.33, 6p22.1 and 15q25.1) and lung cancer survival

Published year and author	Cancer	Population	Sample Size	SNP	Risk SNPs
2010 Yang et al.	Lung cancer	White	1700/2200	rs402710, rs2736100, rs4324798, rs16969968, rs8034191	rs4324798 among SCLC
2011 Xun et al.	lung cancer	European	763	35 SNPs (including rs16969968, rs8034191, rs402710, rs4324798)	rs8034191, rs16969968and rs4324798 among non-smokers lung cancer; rs402710 among current smokers lung cancer.
2014 Azad et al.	NSCLC	Caucasian	564	20 SNPs	rs4975616 among late stage NSCLC

METHODS

Study population

This study was based on a population-based case-control study conducted in the Los Angeles County between 1999 and 2004 [27]. The study population consisted of pathologically confirmed newly diagnosed lung cancer cases identified through the rapid ascertainment system of the USC Cancer Surveillance Program for Los Angeles County. The study population was chosen based on the following criteria,

(a) were residents of Los Angeles County at the time of diagnosis;

(b) were ages 18 to 65 years during the study period; and

(c) were able to speak either English or Spanish or had a translator available on site.

Recurrence cancer cases were excluded from our study. Recruitment rate was 39%. A total of 611 patients were included in our study, including 297 adenocarcinomas (ADC), 95 squamous carcinomas (SQC), 115 large cell carcinomas (LCC), 75 small cell lung cancers (SCLC) and 29 others.

Data collection

Trained interviewers used study specific standardized questionnaires to collect general information including age, gender, ethnicity, education level, smoking and alcohol drinking. Smokers were those who smoked more than 100 cigarettes in their lifetimes. Alcohol drinkers were those who drank at least one alcoholic drink per month for a period of at least six months. Cumulative level of smoking was measured in pack-years. One pack-year of cigarette use is equivalent to smoking one pack of cigarettes per day for one year. Alcohol drinking was measured by the average number of drinks (including wine, beer or liquor) consumed per day. Interviews occurred within six months of diagnoses for 89% of cases. Eighty-nine percent of lung cancer cases provided buccal cell samples for DNA extraction and SNP genotyping. All specimens were transported and stored at -70℃ freezers in the Molecular Epidemiology Laboratory, Fielding School of Public Health, University of California at Los Angeles (UCLA). The death information was collected through the Social Security Death Index (SSDI). The SSDI is generated from the public Death Master File of the US Social Security Administration and provides death records of qualified social security recipients. The SSDI is accessible through several commercial providers; we used the Social Security Death Index Interactive Search. A death record contains information of a decedent's first and last name, social security number (SSN), last benefit, birth date, death date, last residence, and state issued. We first used the nine-digit SSN with information of name and birth date to retrieve the participant's record. If a record

was unavailable, we used the first three or last four digits of the participant's SSN, birth date, and first/last name. We lastly retrieved these records on October 10th, 2013.

Selection of SNPs and genotyping

Top five SNPs that were identified in previous GWAS on lung cancer were selected: rs402710, rs2736100, rs4324798, rs16969968, and rs8034191. In addition, we included four SNPs, without prior hypothesis, based on their locations. All SNPs were determined from the National Center for Biotechnology Information SNP database [28].

SNPs were genotyped using the TaqMan assay (Applied Biosystems [ABI], Foster City, CA). Briefly, DNA aliquots were randomized onto PCR plates. A total volume of 5 μl of the reaction mix containing fluorescently labeled sequence-specific Taqman primers and probes was added into each well. PCR was performed under the following condition: denaturation at 92°C for 10 min followed by 60 cycles at 92°C for 15 sec and extension at 62°C for 80 sec. Genotype detection was performed using an ABI 7900 Sequence Detection System with SDS 2.3 software. Call rates were >97% for TaqMan and reproducibility was 0.997 (using a 5% random sample). The complete list of SNPs selected is shown in Table 2.

Table 2. List of all SNP investigated in the study by chromosome

Chromosome	Brand	RS number	Allele	
			Reference	Non-reference
5	p15.33	rs402710	G	A
	p15.33	rs2736100	C	A
6	p22.1	rs4324798	G	A
	(Unknown)	rs2256543	G	A
15	q25.1	rs16969968	G	A
	q25.1	rs931794	A	G
	q25.1	rs12914385	C	T
	q25.1	rs1317286	A	G
	q25.1	rs8034191	A	G

Statistical analysis

Kaplan-Meier survival analysis and log-rank tests were used for the initial survival analysis. The survival time was calculated as the interval between the date of diagnosis and the date of death, or the date of the last follow-up which was Oct 10th, 2013. Cox hazards models were used to obtain crude and adjusted hazard ratios (HRs) and corresponding 95% confidence intervals (CIs). We adjusted for age, gender, ethnicity, education level, and smoking pack-years; we also adjusted for cell differentiation and morphology that included SQL, ADC, LCC and SCLC, where appropriate.

The associations between SNPs and survival were analyzed with four genetic models: a 2-df genotype-specific, log-additive, dominant, and recessive models. We further explored interaction between SNPs and smoking by adding a product term (SNP*smoking) in the model. SNPs were in either dominant or recessive model as determined from results of main-effect analysis.

The proportional hazard assumption [29] was checked for all models and no noble violations were detected. All statistical analyses were performed using SAS 9.2 (SAS Institute Inc., Cary, NC).

RESULT

Baseline characteristics

Baseline characteristics of 611 lung cancer patients were shown in Table 3. At the end of the last mortality update (Oct 10th, 2013), 406 deaths occurred. The overall median survival time—defined as the time at which 50% of participates were alive—was 2.5 years. The proportion of deaths among patients who were 55 years old or older (72%) was higher than those who were 55 years younger (62%). Comparing to males, females had lower proportion of deaths and longer median survival years. Among all morphologic types, patients of ADC had lowest proportion of

deaths (56%) and longest median survival years (5.8). SCLC patients had the highest proportion of deaths (80%) and shortest median survival years (1.4). The proportion of deaths increases with cell differentiation, 54% in well to moderately differentiated vs. 69% in poorly differentiated. Median survival years decrease with cell differentiation, 7.1 vs. 2.0 years. Mortality was higher among smokers than non-smokers. Median survival years were shorter among smokers than non-smokers. The proportion of deaths increased slightly as pack-years of smoking increased. Mortality among different levels of education was similar.

SNP analysis

Kaplan-Meier curves of survival probabilities of lung cancer by gene variants of each SNP were shown in Figure 1-9.

Both crude and adjusted hazard ratios were presented in Table 4-6. For overall lung cancer, only rs4324798 was associated with mortality of lung cancer in dominant model (cHR: 1.36, 95%CI=1.02, 1.82). After adjusting for age, gender, ethnicity, education levels, smoking pack-years and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma and small cell carcinoma, the hazard ratio increased a little (aHR: 1.51, 95%CI=1.11, 2.07). The association was also observed in NSCLC (aHR: 1.56, 95%CI=1.09, 2.23) and LCC (aHR: 4.15, 95%CI=1.63, 10.52) after adjusting for potential confounders.

Gene-smoking interactions

We explored potential interaction of smoking and gene polymorphisms in fully adjusted models. Results were shown in Table 7. Association of AG+AA of rs4324798 and lung cancer survival was found among lung cancer smokers (aHR: 1.5, 95%CI: 1.1, 2.1), but not among non-smokers

(aHR: 3.5, 95%CI: 0.4, 29.4). P-value for the ratio of hazardous ratio between smoking and rs4324798 (AG+AA vs. GG) on survival was 0.9.

DISCUSSION

We evaluated the top candidate SNPs at 5p15.33, 6p22.1 and 15q25.1 that have been revealed to be associated with lung cancer susceptibility from multiple GWAS and a few validation studies. Our result showed that only rs4324798 (6p22.1) was associated with mortality in lung cancer patients. The association persisted among NSCLC patients, but not among SCLC patients, which could be due to small sample size of SCLC patients. Among NSCLC subtypes, the association was only observed for LCC. We also explored potential interactions between gene polymorphisms and smoking. The association of AG+AA of rs4324798 and lung cancer survival remained significant among smokers, but not among non-smokers. These findings could be due to small sample size of non-smokers. The power to detect aninteraction in this study is limited. This is an unfortunate characteristic of lung cancer case-control studies which are usually comprised overwhelmingly of past and current smokers.

Since studies of association between 5p, 15q and 6p polymorphisms and lung cancer survival are currently limited, any result will require verification by other independent studies of greater size. The strength of our study is its population-based prospective design, which contributes to reduce potential bias. It was less likely that our study was vulnerable to measurement errors from genotyping or disease diagnosis since we applied state of the art genotyping techniques and repeated 5% of samples for quality control; and the cancers were confirmed by pathology diagnoses.

There are some limitations in this study. There might be recall bias as the questionnaires were taken after diagnoses, which might result in smoking status being exaggerated. The low recruitment rates might reduce our generality to patients with more advanced diseases. We also have limited power detecting an association in subtype analysis and non-smokers. There was no tumor-node-metastasis (TNM) cancer staging information in our dataset. Using cell differentiation grades as a proxy might introduce measurement errors. The biological mechanisms and pathway are still not fully understood and need to be further investigated.

Our study will contribute to knowledge on the determinant of lung cancer survival and help to predict prognosis for individually treatment.

CONCLUSIONS

In conclusion, this study has identified rs4324798 (6p22.1) as possible genetic variants that can alter survival in lung cancer patients. Results also suggest that the association of rs4324798 and lung cancer survival remained among NSCLC patients, LCC histological subtype, and smokers, although small numbers in some strata suggest caution in the interpretation of stratified findings.

Table 3. Demographic characteristics in lung cancer.

	All, n	Death, n (%)	Censored, n (%)	Median survival years
Survival	611	406 (66)	205 (34)	2.5
Age of diagnosis, mean (SD)		52.6 ± 5.3	51.5 ± 5.7	
<45	61	38 (62)	23 (38)	3.0
45-54	301	188 (62)	113 (38)	2.9
55+	249	180 (72)	69 (28)	2.2
Sex				
Male	303	215 (71)	88 (29)	2.0
Female	308	191 (62)	117 (38)	3.7
Ethnicity				
Caucasian	359	245 (68)	114 (32)	2.4
Hispanic	70	44 (63)	26 (37)	2.2
African-American	96	60 (62)	36 (38)	3.0
Asian-American	70	46 (66)	24 (34)	2.8
Other	15	10 (67)	5 (33)	4.1
Morphology				
Adenocarcinoma	95	53 (56)	42 (44)	5.8
Squamous cell	297	186 (63)	111 (34)	3.4
Large cell	115	85 (74)	30 (26)	2.1
Small cell	75	60 (80)	15 (20)	1.4
other	29	22 (76)	7 (24)	1.5
Education (years of schooling)		13.2 ± 3.3	13.4 ± 3.5	
0-12	265	181 (68)	84 (32)	2.6
13-16	275	181 (66)	94 (34)	2.4
>16	71	44 (62)	27 (38)	2.8
Tumor cell differentiation				
Well to moderate	168	90 (54)	78 (46)	7.1
Poor to very poor	222	154 (69)	68 (31)	2.0
Undetermined	219	161 (74)	58 (26)	2.0
Smoking (pack-years)		33.11 ± 24.78	27.89 ± 25.35	
Never	110	61 (55)	49 (44)	5.5
Ever	501	345 (69)	156 (31)	2.2
Less than 20	98	63 (64)	35 (36)	2.6
20-40	201	139 (69)	62 (31)	2.5
40 or more	202	143 (71)	59 (29)	1.9
Alcohol drinking status (drinks-day)		1.62 ± 3.08	1.63 ± 5.06	
Never	170	111 (65)	59 (34)	2.8
Ever	440	294 (67)	146 (33)	2.5
<2 drinks/day	302	200 (66)	102 (34)	2.4
>= 2 drinks/day	138	94 (68)	44 (32)	2.8

Table 4. Crude and adjusted hazard ratio of selected SNPs and lung cancer survival

SNP	death/n	cHR (95% CI)	aHR* (95% CI)
rs16969968			
GG	182/278	1.00	1.00
AG	127/198	0.88 (0.70, 1.10)	0.87 (0.68, 1.13)
AA	44/60	1.25 (0.90, 1.74)	1.20 (0.84, 1.73)
Log-additive		1.04 (0.89, 1.21)	1.03 (0.86, 1.23)
Dominant	171/258	0.95 (0.77, 1.17)	0.94 (0.74, 1.19)
Recessive	44/60	1.33 (0.97, 1.82)	1.29 (0.92, 1.81)
rs931794			
AA	148/219	1.00	1.00
AG	152/243	0.81 (0.65, 1.02)	0.84 (0.66, 1.07)
GG	56/77	1.15 (0.84, 1.56)	1.15 (0.84, 1.59)
Log-additive		1.00 (0.86, 1.17)	1.02 (0.86, 1.19)
Dominant	208/320	0.88 (0.72, 1.09)	0.91 (0.73, 1.14)
Recessive	56/77	1.28 (0.96, 1.70)	1.26 (0.94, 1.70)
rs12914385			
CC	147/224	1.00	1.00
CT	141/220	0.90 (0.71, 1.12)	0.92 (0.72, 1.19)
TT	52/78	1.04 (0.76, 1.43)	1.05 (0.75, 1.47)
Log-additive		0.99 (0.85, 1.15)	1.00 (0.85, 1.18)
Dominant	193/298	0.93 (0.75, 1.15)	0.95 (0.76, 1.20)
Recessive	52/78	1.10 (0.82, 1.48)	1.09 (0.80, 1.49)
rs1317286			
AA	155/231	1.00	1.00
AG	151/240	0.82 (0.65, 1.02)	0.85 (0.67, 1.08)
GG	49/66	1.17 (0.85, 1.62)	1.21 (0.86, 1.71)
Log-additive		1.00 (0.85, 1.17)	1.02 (0.86, 1.21)
Dominant	200/306	0.88 (0.72, 1.10)	0.91 (0.73, 1.14)
Recessive	49/66	1.30 (0.96, 1.76)	1.33 (0.97, 1.83)
rs8034191			
AA	169/257	1.00	1.00
AG	141/222	0.83 (0.66, 1.04)	0.82 (0.64, 1.06)
GG	45/60	1.26 (0.91, 1.75)	1.15 (0.81, 1.63)
Log-additive		1.01 (0.87, 1.19)	0.99 (0.83, 1.18)
Dominant	186/282	0.91 (0.74, 1.12)	0.89 (0.70, 1.12)
Recessive	45/60	1.38 (1.01, 1.88)	1.27 (0.91, 1.76)
rs4324798			
GG	303/468	1.00	1.00
AG	50/66	1.39 (1.03, 1.88)	1.48 (1.07, 2.04)
AA	3/5	0.96 (0.31, 3.00)	2.22 (0.70, 7.04)
Log-additive		1.27 (0.98, 1.65)	1.48 (1.12, 1.97)
Dominant	53/71	1.36 (1.02. 1.82)	1.51 (1.11, 2.07)
Recessive	3/5	0.92 (0.30, 2.88)	2.12 (0.67, 6.74)
rs2256543			
GG	132/198	1.00	1.00

AG	163/243	0.99 (0.79, 1.25)	0.96 (0.76, 1.22)
AA	57/94	0.80 (0.59, 1.10)	0.83 (0.60, 1.14)
Log-additive		0.91 (0.79, 1.06)	0.92 (0.79, 1.07)
Dominant	220337	0.93 (0.75, 1.12)	0.92 (0.73, 1.15)
Recessive	57/94	0.80 (0.60, 1.07)	0.84 (0.63, 1.13)
rs402710			
GG	153/225	1.00	1.00
AG	160/243	0.92 (0.74, 1.15)	0.96 (0.76, 1.21)
AA	34/52	0.91 (0.63, 1.32)	0.89 (0.60, 1.30)
Log-additive		0.94 (0.80, 1.11)	0.95 (0.80, 1.12)
Dominant	194295	0.92 (0.74, 1.14)	0.95 (0.76, 1.18)
Recessive	34/52	0.95 (0.66, 1.35)	0.90 (0.62, 1.30)
rs2736100			
CC	106/149	1.00	1.00
AC	177/257	0.98 (0.77, 1.24)	1.00 (0.78, 1.29)
AA	69/127	0.72 (0.53, 0.98)	0.74 (0.54, 1.02)
Log-additive		0.86 (0.75, 1.00)	0.88 (0.75, 1.02)
Dominant	246/384	0.89 (0.71, 1.12)	0.91 (0.72, 1.16)
Recessive	69/127	0.73 (0.56, 0.95)	0.74 (0.56, 0.98)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma and small cell carcinoma, if applicable.

Table 5. Crude and adjusted hazard ratio of selected SNPs and lung cancer survival, stratified by morphological types (NSCLC vs. SCLC).

	NSCLC			SCLC		
SNP	death/n	cHR (95%)	aHR* (95%)	death/n	cHR (95%)	aHR* (95%)
rs16969968						
GG	153/240	1.00	1.00	29/38	1.00	1.00
AG	107/173	0.89 (0.70, 1.14)	0.87 (0.68, 1.13)	20/25	0.75 (0.42, 1.33)	0.61 (0.32, 1.16)
AA	39/55	1.26 (0.89, 1.79)	1.20 (0.84, 1.73)	5/5	1.70 (0.64, 4.48)	1.06 (0.37, 3.01)
Log-additive		1.05 (0.89, 1.24)	1.05 (0.87, 1.27)		1.00 (0.63, 1.57)	
Dominant	143/228	0.97 (0.77, 1.22)	1.00 (0.77, 1.30)	25/30	0.85 (0.49, 1.45)	0.68 (0.37, 1.23)
Recessive	39/55	1.32 (0.95, 1.85)	1.22 (0.85, 1.75)	5/5	1.92 (0.75, 4.94)	1.32 (0.48, 3.63)
rs931794						
AA	123/187	1.00	1.00	25/32	1.00	1.00
AG	130/213	0.86 (0.67, 1.10)	0.84 (0.66, 1.07)	22/30	0.57 (0.32, 1.01)	0.52 (0.28, 0.98)
GG	50/71	1.18 (0.85, 1.63)	1.15 (0.84, 1.59)	6/6	1.32 (0.54, 3.25)	1.21 (0.47, 3.09)
Log-additive		1.03 (0.87, 1.22)	1.04 (0.87, 1.24)		0.84 (0.53, 1.34)	0.82 (0.51, 1.31)
Dominant	180/284	0.93 (0.73, 1.17)	0.97 (0.76, 1.24)	28/36	0.65 (0.38, 1.11)	0.62 (0.35, 1.10)
Recessive	50/71	1.28 (0.94, 1.73)	1.22 (0.89, 1.68)	6/6	1.78 (0.75, 4.22)	1.60 (0.64, 4.02)
rs12914385						
CC	123/193	1.00	1.00	24/31	1.00	1.00
CT	120/193	0.93 (0.72, 1.19)	0.92 (0.72, 1.19)	21/27	0.67 (0.37, 1.20)	0.63 (0.33, 1.18)
TT	46/70	1.09 (0.78, 1.53)	1.05 (0.75, 1.47)	6/8	0.84 (0.34, 2.06)	0.59 (0.22, 1.56)
Log-additive		1.02 (0.86, 1.20)	1.05 (0.88, 1.25)		0.82 (0.53, 1.26)	0.72 (0.46, 1.12)
Dominant	166/263	0.97 (0.77, 1.22)	1.02 (0.79, 1.32)	27/35	0.70 (0.40, 1.21)	0.62 (0.34, 1.12)
Recessive	46/70	1.13 (0.83, 1.55)	1.13 (0.81, 1.58)	6/8	1.04 (0.44, 2.43)	0.74 (0.29, 1.89)
rs1317286						
AA	127/196	1.00	1.00	28/35	1.00	1.00
AG	132/213	0.87 (0.68, 1.12)	0.85 (0.67, 1.08)	19/27	0.56 (0.31, 1.01)	0.41 (0.21, 0.80)
GG	43/60	1.20 (0.85, 1.70)	1.21 (0.86, 1.71)	6/6	1.27 (0.52, 3.09)	1.00 (0.39, 2.57)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 5. cont. Crude and adjusted hazard ratio of selected SNPs and lung cancer survival, stratified by morphological types (NSCLC vs. SCLC).

Log-additive		1.03 (0.89, 1.22)	1.07 (0.89, 1.29)		0.83 (0.53, 1.31)	0.71 (0.43, 1.16)
Dominant	175/273	0.94 (0.75, 1.18)	1.01 (0.78, 1.29)	25/33	0.65 (0.38, 1.12)	0.50 (0.27, 0.92)
Recessive	43/60	1.29 (0.94, 1.79)	1.29 (0.91, 1.82)	6/6	1.65 (0.70, 3.90)	1.48 (0.59, 3.68)
rs8034191						
AA	141/221	1.00	1.00	28/36	1.00	1.00
AG	119/193	0.86 (0.68, 1.10)	0.82 (0.64, 1.06)	22/29	0.63 (0.36, 1.10)	0.44 (0.23, 0.86)
GG	41/56	1.30 (0.92, 1.84)	1.15 (0.81, 1.63)	4/4	1.71 (0.59, 4.95)	0.90 (0.29, 2.78)
Log-additive		1.05 (0.88, 1.24)	1.02 (0.85, 1.22)		0.84 (0.52, 1.37)	0.64 (0.38, 1.10)
Dominant	160/249	0.94 (0.75, 1.18)	0.95 (0.74, 1.22)	26/33	0.70 (0.41, 1.19)	0.49 (0.26, 0.92)
Recessive	41/56	1.39 (1.00, 1.93)	1.20 (0.84, 1.70)	4/4	2.13 (0.75, 6.02)	1.33 (0.44, 4.02)
rs4324798						
GG	260/412	1.00	1.00	43/56	1.00	1.00
AG	39/54	1.33 (0.95, 1.86)	**1.48 (1.07, 2.04)**	11/12	1.52 (0.78, 2.95)	1.63 (0.79, 3.39)
AA	3/5	1.02 (0.33, 3.17)	2.22 (0.70, 7.05)	0	--	--
Log-additive		1.22 (0.92, 1.63)	**1.50 (1.09, 2.06)**		1.52 (0.78, 2.95)	1.45 (0.98, 2.16)
Dominant	42/59	1.30 (0.94, 1.80)	**1.56 (1.09, 2.23)**	11/12	1.52 (0.78, 2.95)	1.63 (0.79, 3.39)
Recessive	3/5	0.98 (0.32, 3.06)	2.04 (0.64, 6.48)	0	--	--
rs2256543						
GG	116/174	1.00	1.00	16/24	1.00	1.00
AG	136/211	0.93 (0.72, 1.19)	0.96 (0.76, 1.22)	27/32	1.50 (0.81, 2.78)	1.90 (0.98, 3.67)
AA	47/82	0.74 (0.53, 1.04)	0.83 (0.60, 1.14)	10/12	1.24 (0.56, 2.73)	1.93 (0.82, 4.53)
Log-additive		0.87 (0.75, 1.03)	0.85 (0.72, 1.00)		1.15 (0.80, 1.65)	1.45 (0.98, 2.16)
Dominant	183/293	0.87 (0.69, 1.10)	0.81 (0.63, 1.04)	37/44	1.42 (0.79, 2.55)	1.90 (1.02, 3.55)
Recessive	47/82	0.77 (0.57, 1.06)	0.78 (0.57, 1.08)	10/12	0.98 (0.49, 1.95)	1.37 (0.64, 2.95)
rs402710						
GG	131/197	1.00	1.00	22/28	1.00	1.00
AG	135/213	0.89 (0.70, 1.14)	0.96 (0.76, 1.21)	25/30	1.15 (0.65, 2.05)	0.88 (0.47, 1.63)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 5. cont. Crude and adjusted hazard ratio of selected SNPs and lung cancer survival, stratified by morphological types (NSCLC vs. SCLC).

AA	31/46	1.02 (0.69, 1.51)	0.88 (0.60, 1.30)	3/6	0.39 (0.12, 1.32)	0.29 (0.08, 1.09)
Log-additive		0.97 (0.81, 1.15)	0.99 (0.82, 1.19)		0.81 (0.54, 1.23)	0.68 (0.42, 1.10)
Dominant	166/259	0.92 (0.73, 1.15)	0.96 (0.75, 1.22)	28/36	0.96 (0.55, 1.67)	0.76 (0.41, 1.40)
Recessive	31/46	1.08 (0.75, 1.57)	1.08 (0.73, 1.60)	3/6	0.37 (0.11, 1.18)	0.32 (0.09, 1.12)
rs2736100						
CC	91/132	1.00	1.00	15/17	1.00	1.00
AC	147/219	0.98 (0.75, 1.27)	1.00 (0.78, 1.29)	30/38	0.80 (0.43, 1.49)	0.64 (0.29, 1.43)
AA	61/115	0.75 (0.54, 1.03)	0.74 (0.54, 1.02)	8/12	0.55 (0.23, 1.30)	0.38 (0.15, 1.01)
Log-additive		0.87 (0.75, 1.02)	0.90 (0.77, 1.07)		0.75 (0.50, 1.13)	0.62 (0.38, 1.00)
Dominant	208/334	0.90 (0.70, 1.15)	0.95 (0.73, 1.24)	38/50	0.73 (0.40, 1.33)	0.53 (0.25, 1.13)
Recessive	61/115	0.76 (0.57, 1.01)	0.78 (0.58, 1.05)	8/12	0.64 (0.30, 1.36)	0.51 (0.22, 1.16)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 6. Crude and adjusted hazard ratio of selected SNPs and NSCLC survival, stratified by NSCLC, stratified by NSCLC subtypes.

	ADC			SQC			LCC		
SNP	death/ n	cHR (95%)	aHR* (95%)	death /n	cHR (95%)	aHR* (95%)	death/n	cHR (95%)	aHR* (95%)
rs16969968									
GG	84/139	1.00	1.00	21/41	1.00	1.00	36/46	1.00	1.00
AG	54/91	0.91 (0.65, 1.28)	1.08 (0.74, 1.57)	21/36	1.14 (0.62, 2.08)	1.14 (0.50, 2.63)	28/39	0.83 (0.50, 1.35)	0.85 (0.49, 1.49)
AA	22/31	1.38 (0.86, 2.20)	1.28 (0.76, 2.16)	4/6	1.33 (0.46, 3.89)	1.67 (0.52, 5.42)	10/14	1.01 (0.50, 2.04)	1.09 (0.51, 2.31)
Log-additive		1.09 (0.87, 1.37)	1.12 (0.87, 1.44)		1.15 (0.73, 1.81)	1.25 (0.72, 2.18)		0.95 (0.68, 1.33)	0.99 (0.69, 1.44)
Dominant	76/122	1.01 (0.74, 1.38)	1.12 (0.79, 1.60)	25/42	1.16 (0.65, 2.08)	1.25 (0.57, 2.74)	38/53	0.87 (0.55, 1.37)	0.91 (0.54, 1.52)
Recessive	22/31	1.43 (0.91, 2.24)	1.23 (0.76, 2.01)	4/6	1.25 (0.45, 3.50)	1.58 (0.51, 4.84)	10/14	1.10 (0.57, 2.15)	1.18 (0.58, 2.40)
rs931794									
AA	68/107	1.00	1.00	15/32	1.00	1.00	32/39	1.00	1.00
AG	67/117	0.84 (0.60, 1.18)	0.95 (0.66, 1.36)	25/42	1.25 (0.66, 2.37)	0.97 (0.47, 2.00)	30/42	0.80 (0.49, 1.33)	0.94 (0.54, 1.65)
GG	28/40	1.20 (0.77, 1.86)	1.21 (0.76, 1.93)	8/11	1.86 (0.79, 4.41)	1.78 (0.69, 4.57)	11/16	0.92 (0.46, 1.82)	1.17 (0.56, 2.45)
Log-additive		1.04 (0.83, 1.30)	1.07 (0.85, 1.35)		1.34 (0.88, 2.06)	1.25 (0.77, 2.03)		0.92 (0.65, 1.28)	1.05 (0.74, 1.51)
Dominant	95/157	0.92 (0.68, 1.26)	1.01 (0.72, 1.41)	33/52	1.36 (0.74, 2.50)	1.13 (0.58, 2.21)	41/58	0.83 (0.52, 1.32)	1.00 (0.60, 1.68)
Recessive	28/40	1.31 (0.87, 1.97)	1.25 (0.81, 1.90)	8/11	1.63 (0.77, 3.50)	1.80 (0.76, 4.26)	11/16	1.03 (0.54, 1.95)	1.21 (0.61, 2.39)
rs12914385									
CC	67/114	1.00	1.00	15/31	1.00	1.00	31/37	1.00	1.00
CT	64/106	0.99 (0.70, 1.40)	1.12 (0.78, 1.60)	22/37	1.32 (0.68, 2.54)	1.36 (0.64, 2.91)	29/42	0.70 (0.42, 1.17)	0.76 (0.44, 1.33)
TT	24/36	1.27 (0.79, 2.02)	1.24 (0.76, 2.03)	8/13	1.30 (0.55, 3.07)	1.60 (0.63, 4.03)	10/15	0.84 (0.41, 1.72)	1.09 (0.52, 2.28)
Log-additive		1.09 (0.87, 1.37)	1.11 (0.88, 1.41)		1.17 (0.78, 1.75)	1.27 (0.81, 2.00)		0.85 (0.60, 1.22)	0.97 (0.66, 1.40)
Dominant	88/142	1.06 (0.77, 1.45)	1.15 (0.82, 1.60)	30/50	1.31 (0.71, 2.44)	1.43 (0.70, 2.91)	39/57	0.74 (0.46 1.18)	0.83 (0.50, 1.40)
Recessive	24/36	1.27 (0.82, 1.97)	1.17 (0.74, 1.85)	8/13	1.12 (0.52, 2.40)	1.35 (0.60, 3.04)	10/15	1.02 (0.52, 1.99)	1.25 (0.62, 2.49)
rs1317286									
AA	69/113	1.00	1.00	16/32	1.00	1.00	31/39	1.00	1.00
AG	67/114	0.90 (0.64, 1.26)	0.98 (0.69, 1.39)	25/42	1.23 (0.66, 2.01)	0.96 (0.47, 1.95)	35/48	0.81 (0.50, 1.31)	0.98 (0.57, 1.68)
GG	24/32	1.43 (0.90, 2.28)	1.34 (0.81, 2.21)	7/11	1.31 (0.54, 3.19)	1.53 (0.56, 4.19)	9/13	0.94 (0.45, 1.97)	1.25 (0.56, 2.75)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 6. cont. Crude and adjusted hazard ratio of selected SNPs and NSCLC survival, stratified by NSCLC, stratified by NSCLC subtypes

Log-additive		1.11 (0.88, 1.40)	1.11 (0.87, 1.41)		1.16 (0.77, 1.77)	1.15 (0.71, 1.86)		0.91 (0.64, 1.30)	1.08 (0.74, 1.57)
Dominant	91/146	1.00 (0.73, 1.36)	1.04 (0.75, 1.45)	32/53	1.25 (0.69, 2.28)	1.08 (0.56, 2.07)	43/61	0.83 (0.52, 1.32)	1.03 (0.62, 1.71)
Recessive	24/32	1.51 (0.98, 2.34)	1.35 (0.85, 2.15)	7/11	1.16 (0.52, 2.59)	1.55 (0.59, 4.06)	9/13	1.06 (0.53, 2.12)	1.26 (0.60, 2.63)
rs8034191									
AA	80/131	1.00	1.00	16/35	1.00	1.00	36/44	1.00	1.00
AG	60/103	0.86 (0.61, 1.20)	0.88 (0.62, 1.26)	24/39	1.30 (0.69, 2.45)	1.32 (0.62, 2.82)	28/40	0.75 (0.45, 1.23)	0.92 (0.51, 1.64)
GG	21/29	1.34 (0.83, 1.26)	1.22 (0.72, 2.07)	7/10	1.77 (0.73, 4.32)	2.29 (0.86, 6.12)	10/14	0.92 (0.46, 1.86)	0.97 (0.47, 2.01)
Log-additive		1.05 (0.83, 1.33)	1.03 (0.80, 1.33)		1.32 (0.86, 2.03)	1.48 (0.90, 2.43)		0.89 (0.63, 1.25)	0.97 (0.68, 1.38)
Dominant	81/132	0.94 (0.69, 1.29)	0.94 (0.67, 1.31)	31/49	1.38 (0.76, 2.53)	1.49 (0.73, 3.07)	38/54	0.79 (0.50, 1.24)	0.94 (0.56, 1.57)
Recessive	21/29	1.43 (0.91, 2.27)	1.31 (0.80, 2.14)	7/10	1.53 (0.68, 3.41)	1.95 (0.82, 4.66)	10/14	1.06 (0.55, 2.07)	1.00 (0.50, 2.00)
rs4324798									
	134/22								
GG	4	1.00	1.00	44/78	1.00	1.00	67/91	1.00	1.00
AG	25/36	1.21 (0.79, 1.86)	1.25 (0.80, 1.96)	4/7	1.43 (0.51, 3.98)	2.85 (0.84, 0.70)	6/6	**2.54 (1.10, 5.90)**	**4.66 (1.66, 13.1)**
AA	2/3	1.37 (0.34, 5.52)	2.01 (0.48, 8.43)	0	--	--	1/1	2.30 (0.31, 16.8)	2.68 (0.30, 23.7)
Log-additive		1.20 (0.83, 1.74)	1.29 (0.87, 1.91)		1.43 (0.51, 3.98)	2.85 (0.84, 9.70)		**1.95 (1.06, 3.61)**	**2.54 (1.26, 5.11)**
Dominant	27/39	1.22 (0.81, 1.85)	1.29 (0.83, 1.99)	4/7	1.43 (0.51, 4.00)	2.85 (0.84, 9.70)	7/7	**2.51 (1.14, 5.50)**	**4.15 (1.63, 10.5)**
Recessive	2/3	1.33 (0.33, 5.36)	1.95 (0.47, 8.17)	0	--	--	1/1	2.16 (0.30, 15.7)	2.82 (0.32, 24.8)
rs2256543									
GG	60/98	1.00	1.00	18/34	1.00	1.00	30/32	1.00	1.00
AG	71/112	1.00 (0.71, 1.42)	1.02 (0.72, 1.46)	25/42	1.20 (0.65, 2.20)	1.00 (0.50, 2.00)	30/45	0.48 (0.29, 0.81)	0.52 (0.30, 0.90)
AA	28/49	0.84 (0.53, 1.31)	0.78 (0.50, 1.24)	5/9	1.16 (0.43, 3.12)	0.90 (0.31, 2.60)	14/22	0.40 (0.21, 0.76)	0.47 (0.24, 0.91)
Log-additive		0.93 (0.75, 1.15)	0.90 (0.73, 1.12)		1.12 (0.72, 1.72)	0.96 (0.59, 1.58)		0.60 (0.43, 0.84)	0.66 (0.47, 0.92)
Dominant	99/161	0.95 (0.69, 1.31)	0.94 (0.67, 1.31)	30/51	1.19 (0.67, 2.14)	0.98 (0.50, 1.93)	44/67	0.45 (0.28, 0.73)	0.50 (0.31, 0.83)
Recessive	28/49	0.84 (0.56, 1.26)	0.77 (0.51, 1.17)	5/9	1.05 (0.41, 2.64)	0.90 (0.35, 2.33)	14/22	0.62 (0.34, 1.10)	0.68 (0.37, 1.25)
rs402710									
GG	71/104	1.00	1.00	21/39	1.00	1.00	29/40	1.00	1.00
AG	68/122	0.72 (0.51, 1.00)	0.77 (0.54, 1.08)	21/36	1.10 (0.60, 2.01)	1.04 (0.55, 2.00)	38/46	1.21 (0.75, 1.97)	1.26 (0.75, 2.11)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 6. cont. Crude and adjusted hazard ratio of selected SNPs and NSCLC survival, stratified by NSCLC, stratified by NSCLC subtypes

AA	18/26	1.11 (0.66, 1.87)	0.96 (0.56, 1.66)	6/8	1.59 (0.64, 3.95)	1.80 (0.68, 4.79)	6/10	0.64 (0.26, 1.53)	0.67 (0.24, 1.91)
Log-additive		0.91 (0.71, 1.17)	0.89 (0.69, 1.15)		1.21 (0.79, 1.85)	1.24 (0.77, 1.98)		0.93 (0.66, 1.30)	1.00 (0.67, 1.49)
Dominant	86/148	0.78 (0.57, 1.06)	0.80 (0.58, 1.11)	27/44	1.18 (0.67, 2.09)	1.15 (0.62, 2.12)	44/56	1.08 (0.67, 1.72)	1.18 (0.71, 1.97)
Recessive	18/26	1.33 (0.81, 2.17)	1.10 (0.66, 1.85)	6/8	1.52 (0.65, 3.58)	1.76 (0.70, 4.40)	6/10	0.57 (0.25, 1.32)	0.58 (0.21, 1.54)
rs2736100									
CC	52/79	1.00	1.00	14/18	1.00	1.00	20/28	1.00	1.00
AC	81/124	0.98 (0.69, 1.40)	1.01 (0.71, 1.44)	16/34	0.44 (0.21, 0.90)	0.59 (0.25, 1.39)	40/48	1.49 (0.87, 2.55)	**1.98 (1.07, 3.66)**
AA	28/59	0.65 (0.41, 1.02)	0.64 (0.40, 1.02)	16/30	0.66 (0.32, 1.35)	1.00 (0.40, 2.50)	14/22	1.00 (0.50, 1.97)	1.15 (0.56, 2.33)
Log-additive		0.83 (0.67, 1.02)	0.83 (0.67, 1.03)		0.83 (0.55, 1.25)	1.07 (0.66, 1.72)		1.03 (0.76, 1.40)	1.11 (0.81, 1.53)
Dominant	109/183	0.87 (0.62, 1.21)	0.88 (0.63, 1.23)	32/64	0.53 (0.28, 0.99)	0.71 (0.32, 1.57)	54/70	1.32 (0.79, 2.21)	1.60 (0.92, 2.77)
Recessive	28/59	0.65 (0.43, 0.98)	0.63 (0.42, 0.96)	16/30	1.11 (0.61, 2.04)	1.47 (0.75, 2.91)	14/22	0.78 (0.44, 1.40)	0.81 (0.44, 1.50)

*Adjusted by age, gender, ethnicity, education level, smoking as packyears and tumor cell differentiation and morphology including adenocarcinoma, squamous carcinoma and large cell carcinoma.

Table 7. Interactions between selected SNPs and smoking status in lung cancer

	Nonsmokers		Smokers		SNP-Smoking product terms
SNP*	Death/n	aHR (95% CI)	Death/n	aHR (95% CI)	P-value
rs16969968					
GG+AG	51/88	1.00	258/388	1.00	
AA	1/4	0.4 (0.05, 2.9)	43/56	1.3 (0.9, 1.8)	0.4
rs931794					
AA+AG	58/83	1.00	152/379	1.00	
GG	4/9	0.8 (0.3, 2.4)	52/68	1.2 (0.9, 1.7)	0.5
rs12914385					
CC+CT	48/78	1.00	240/366	1.00	
TT	3/12	0.4 (0.1, 1.2)	49/66	1.2 (0.9, 1.7)	0.1
rs1317286					
AA+AG	51/86	1.00	255/385	1.00	
GG	2/5	0.6 (0.1, 2.7)	47/61	1.3 (1.0, 1.8)	0.7
rs8034191					
AA+AG	52/88	1.00	158/391	1.00	
GG	2/5	0.5 (0.1, 2.3)	43/55	1.3 (0.9, 1.8)	0.5
rs4324798					
GG	52/91	1.00	254/443	1.00	
AG+AA	4/6	3.5 (0.4, 29.4)	49/65	1.5 (1.1, 2.1)	0.9
rs2256543					
GG+AG	44/73	1.00	251/368	1.00	
AA	9/18	0.7 (0.3, 1.5)	48/76	0.9 (0.6, 1.2)	0.6
rs402710					
GG+AG	43/74	1.00	270/394	1.00	
AA	5/9	0.8 (0.3, 2.2)	29/43	0.9 (0.6, 1.4)	0.8
rs2736100					
CC+AC	42/71	1.00	241/335	1.00	
AA	12/21	1.1 (0.6, 2.3)	57/106	0.7 (0.5, 0.9)	0.1

*All the selected SNPs were analyzed in recessive model, except for rs4324798 (dominant model)

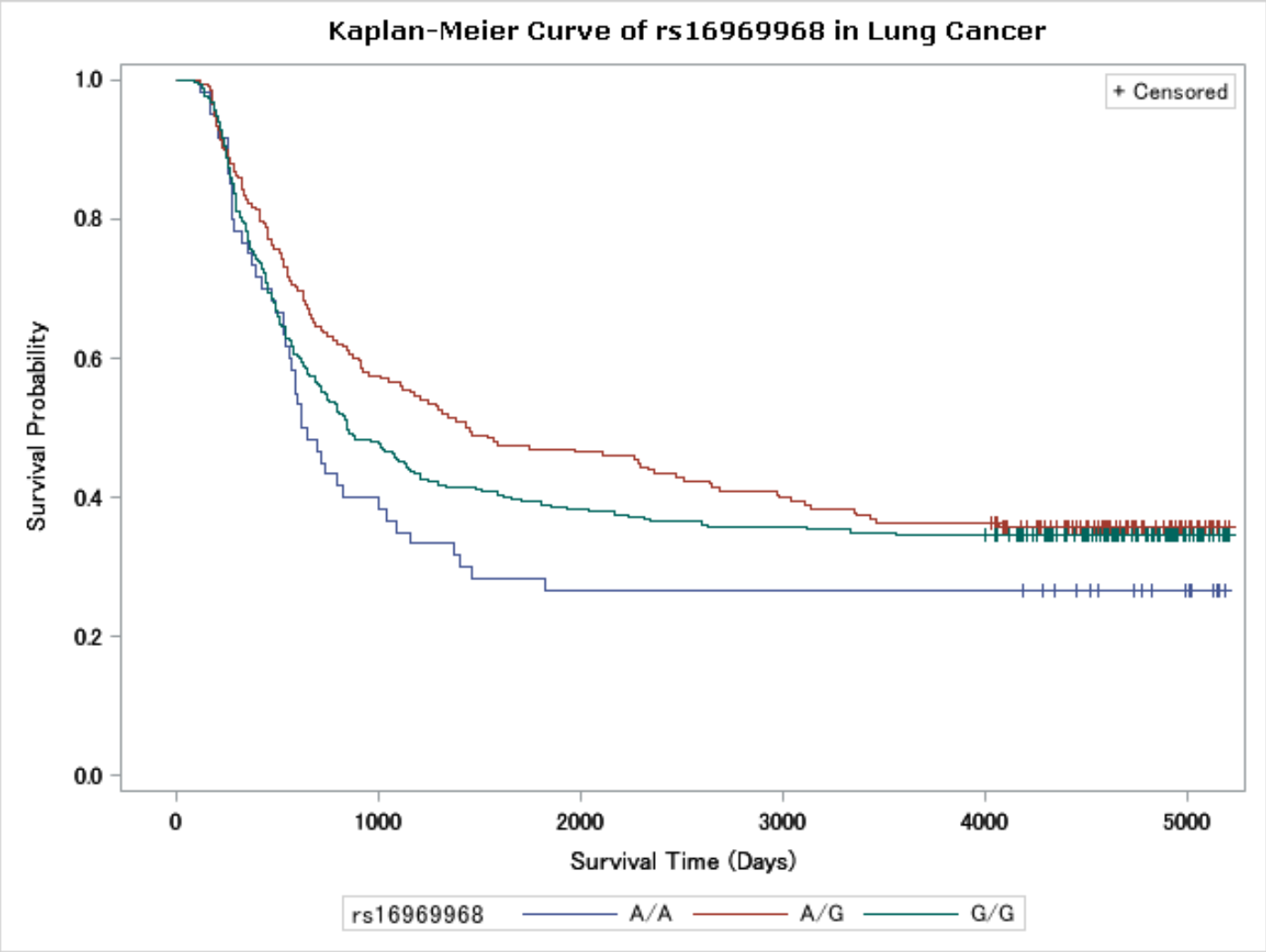

Figure 1. Kaplan-Meier curve of rs16969968 in lung cancer

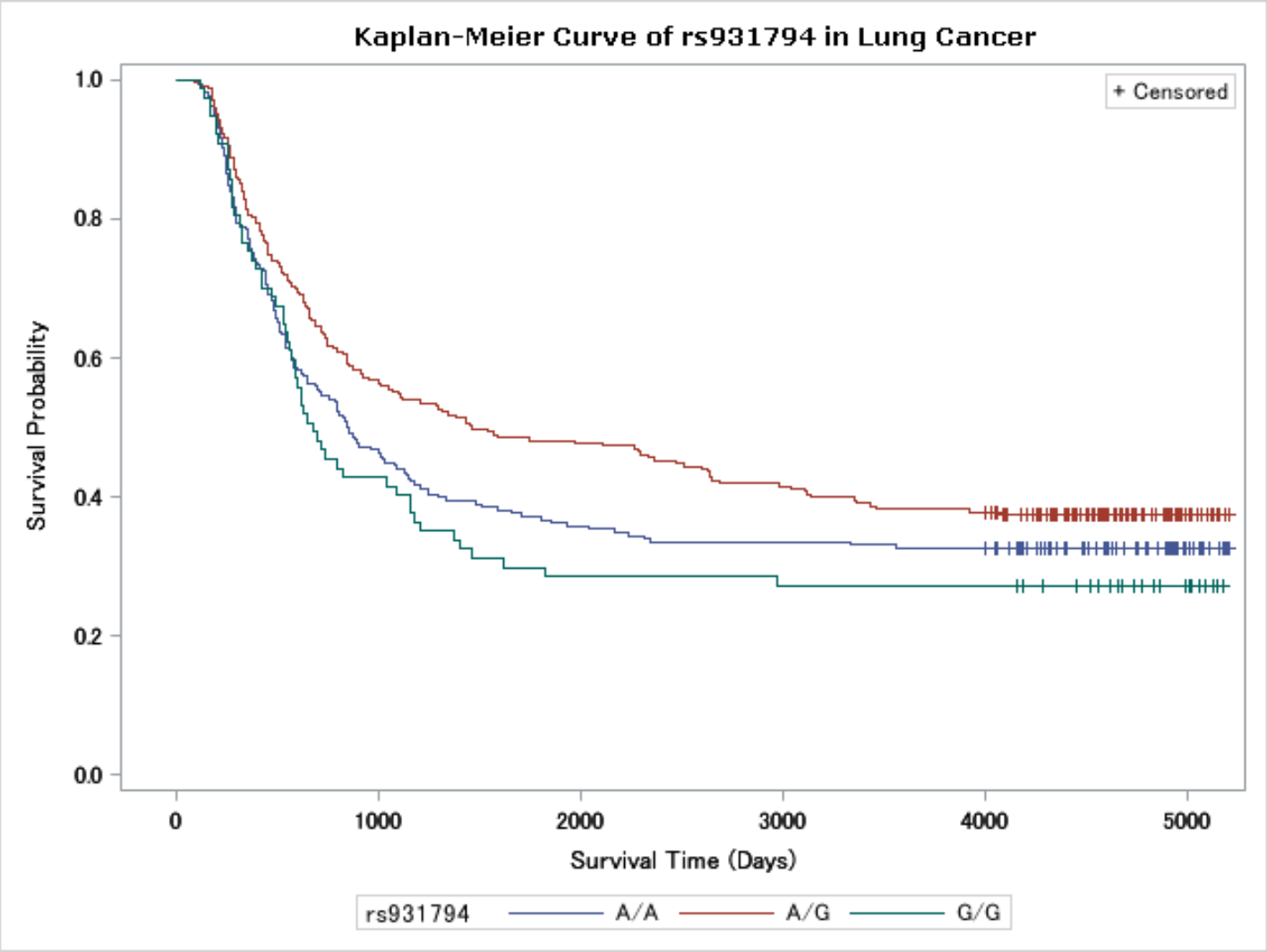

Figure 2. Kaplan-Meier curve of rs931794 in lung cancer

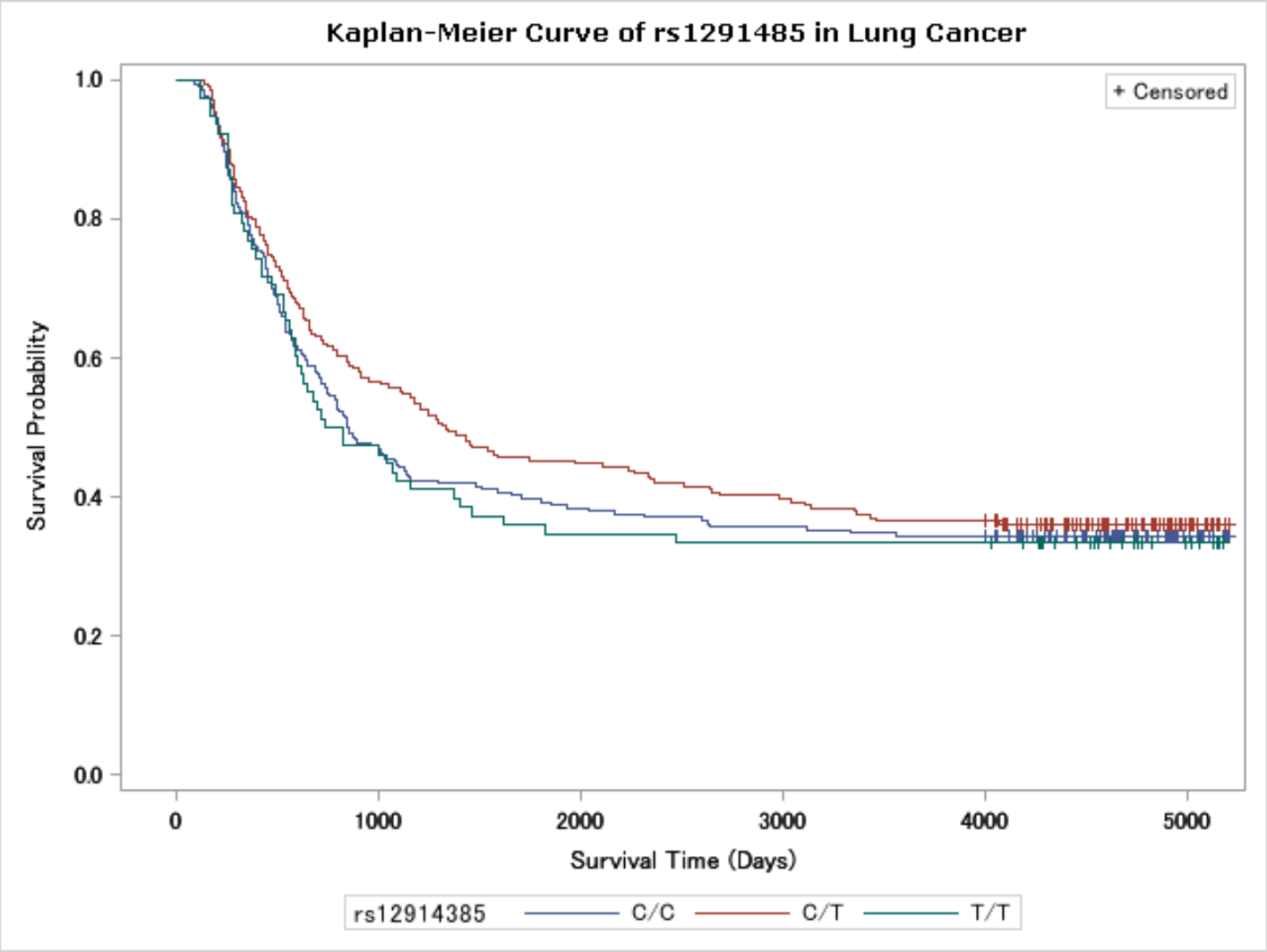

Figure 3. Kaplan-Meier curve of rs12914385 in lung cancer

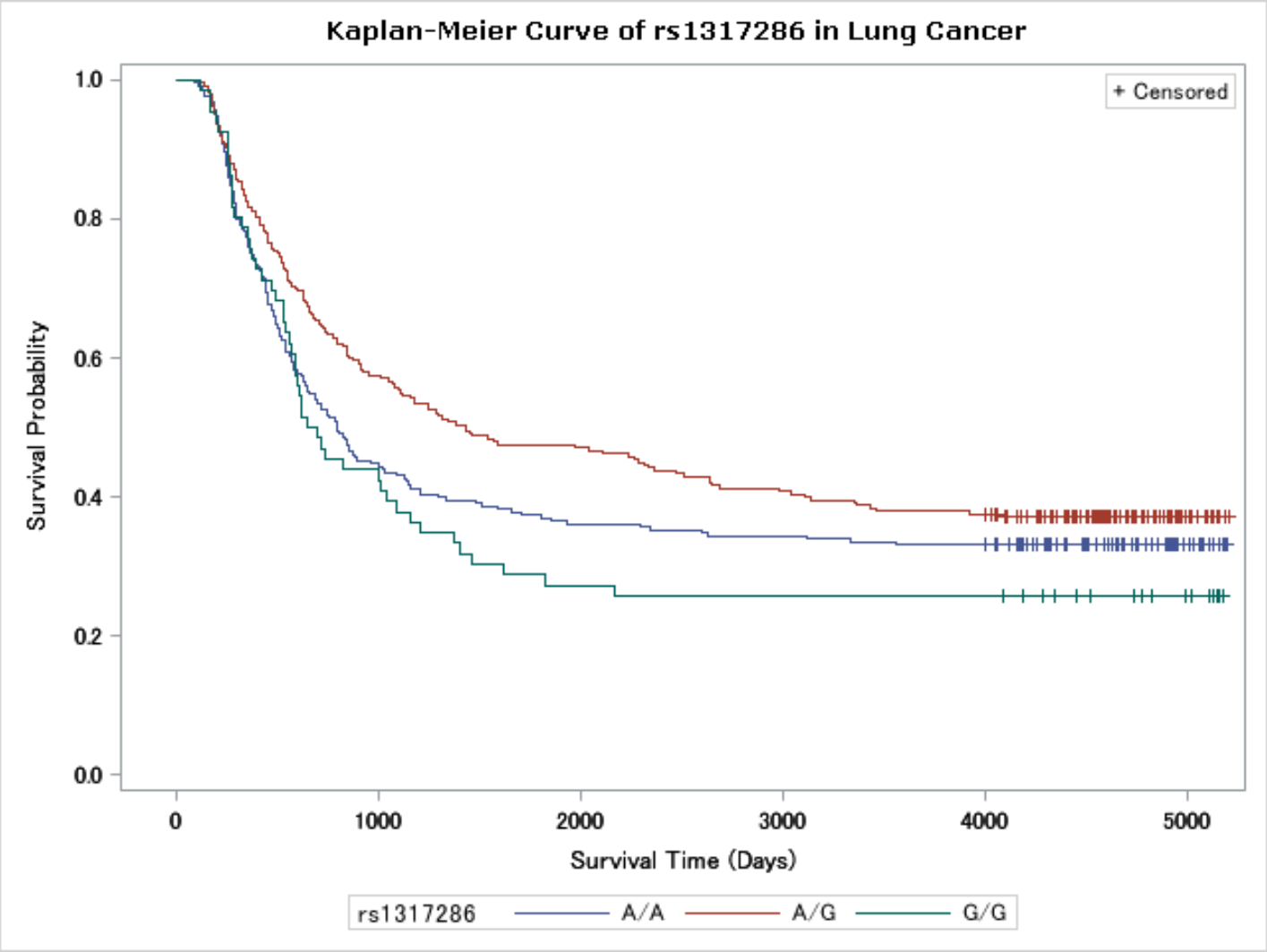

Figure 4. Kaplan-Meier curve of rs1317286 in lung cancer

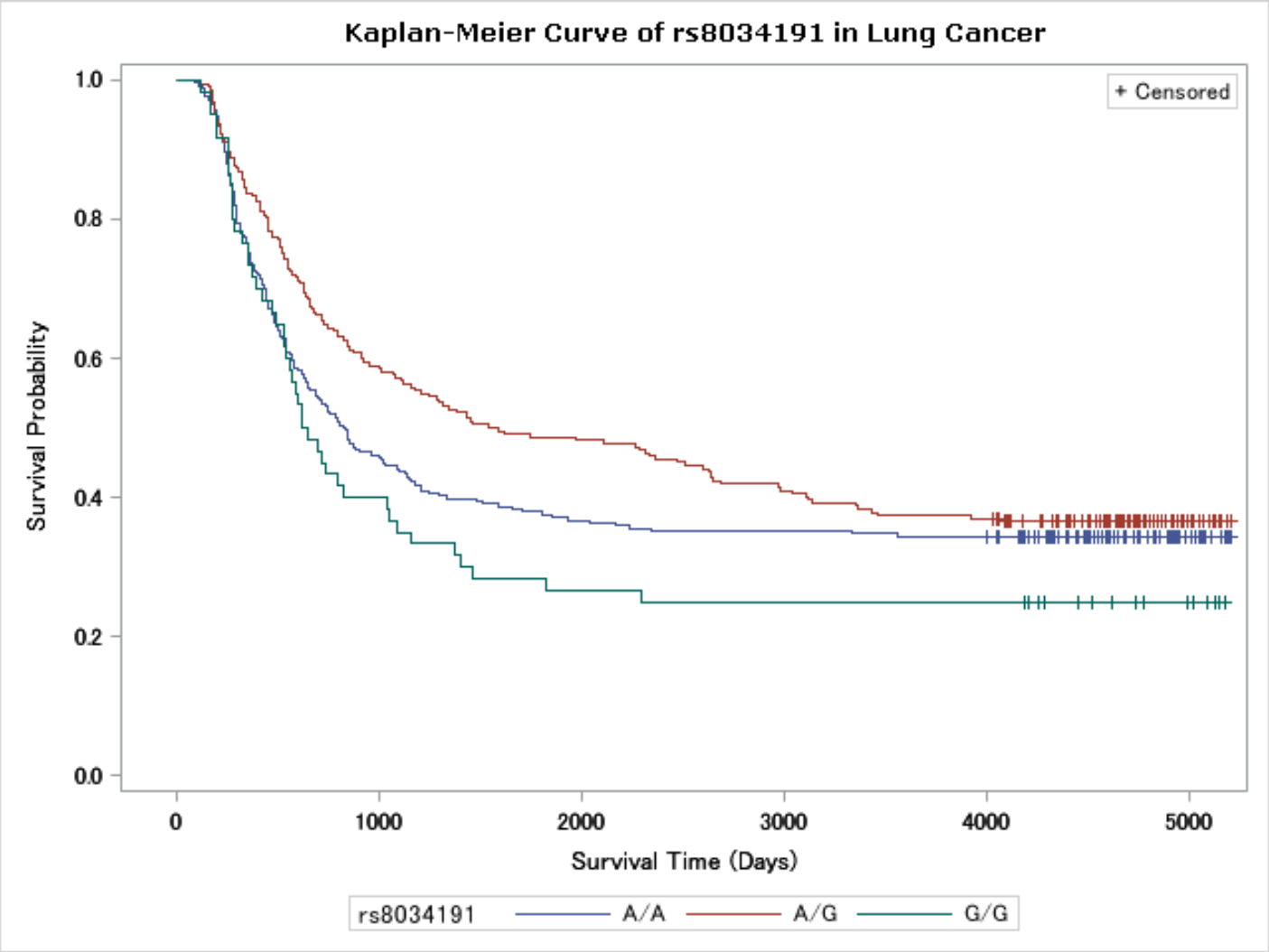

Figure 5. Kaplan-Meier curve of rs8034191 in lung cancer

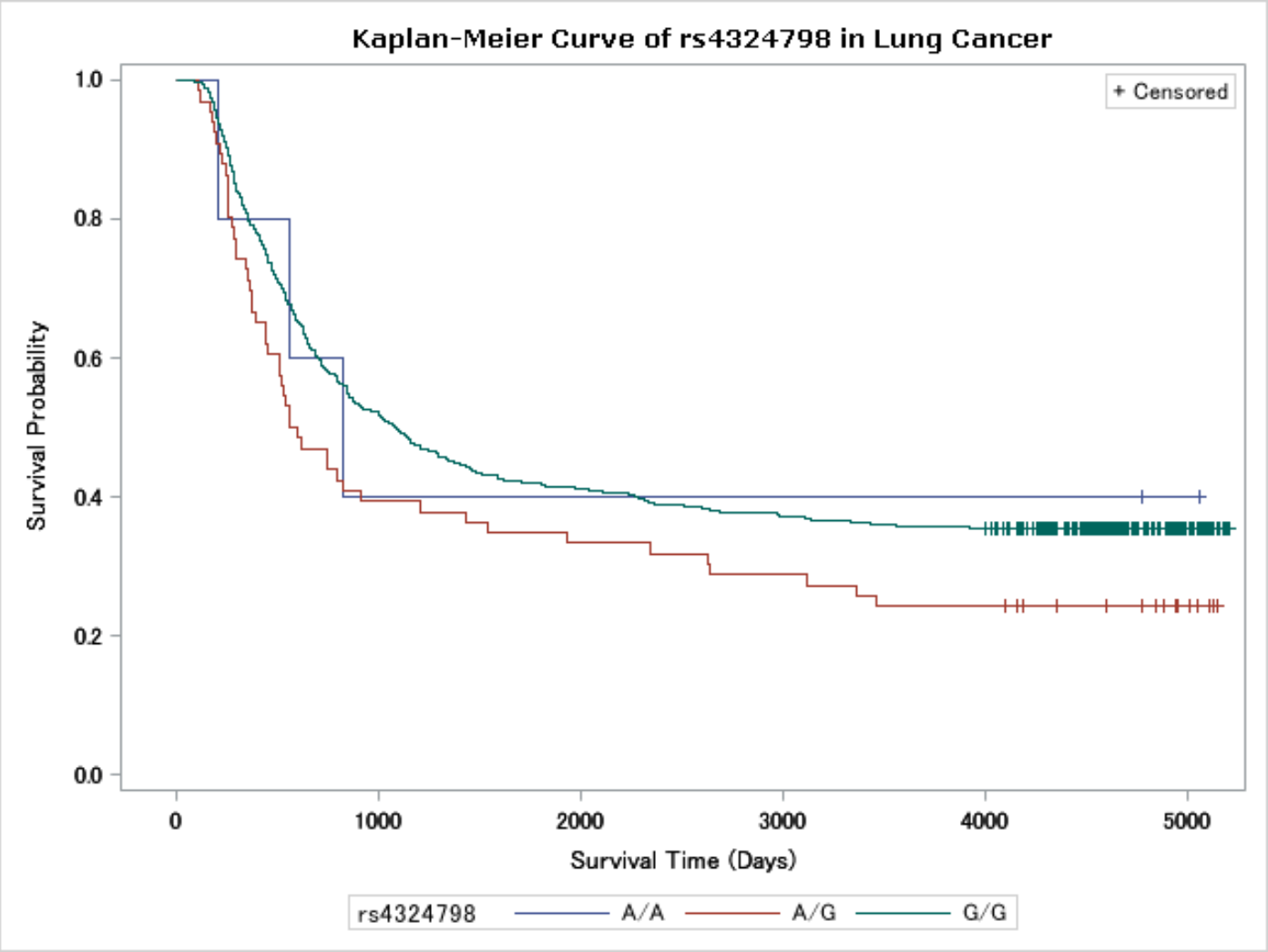

Figure 6. Kaplan-Meier curve of rs4324798 in lung cancer

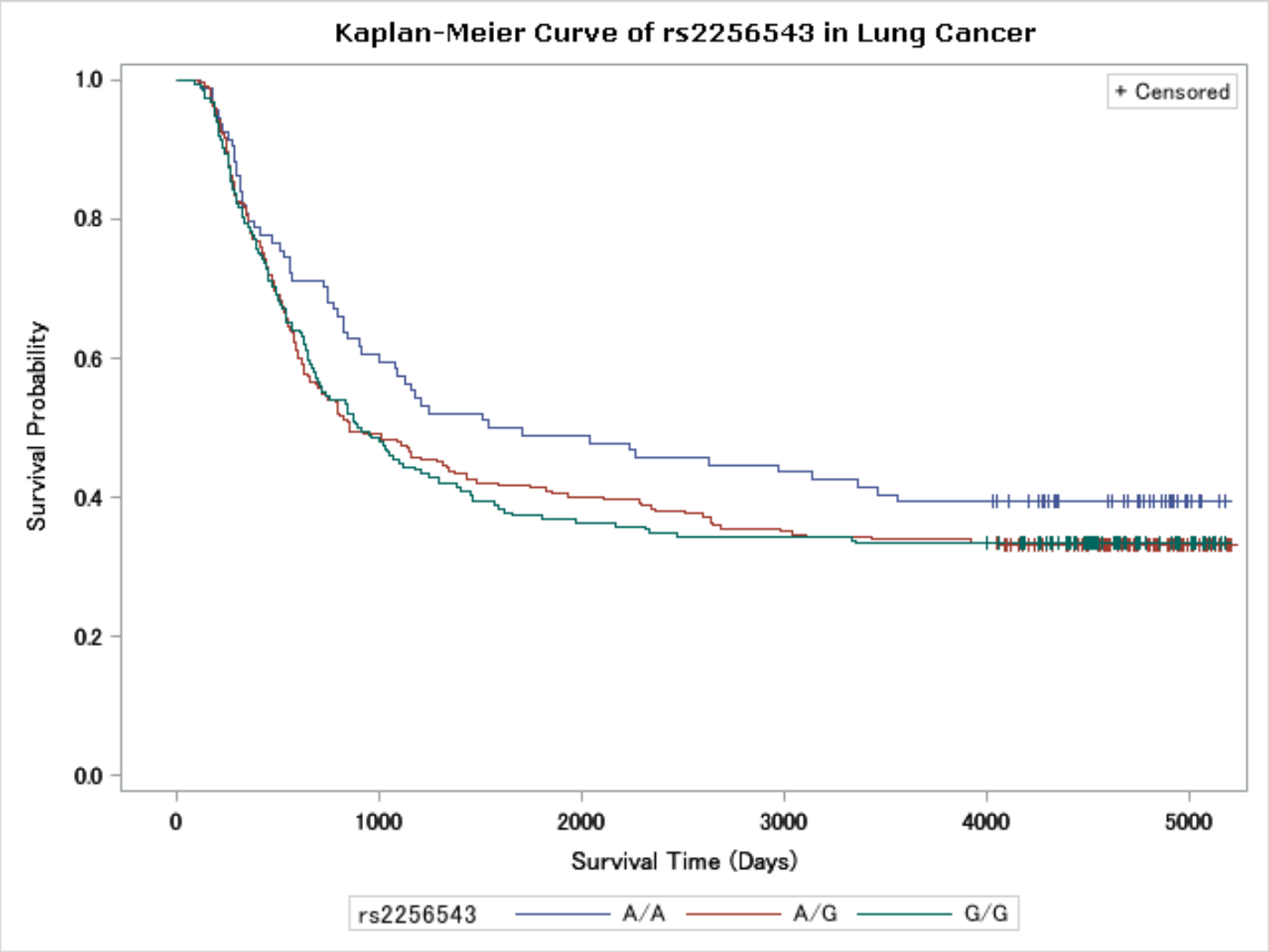

Figure 7. Kaplan-Meier curve of rs2256543 in lung cancer

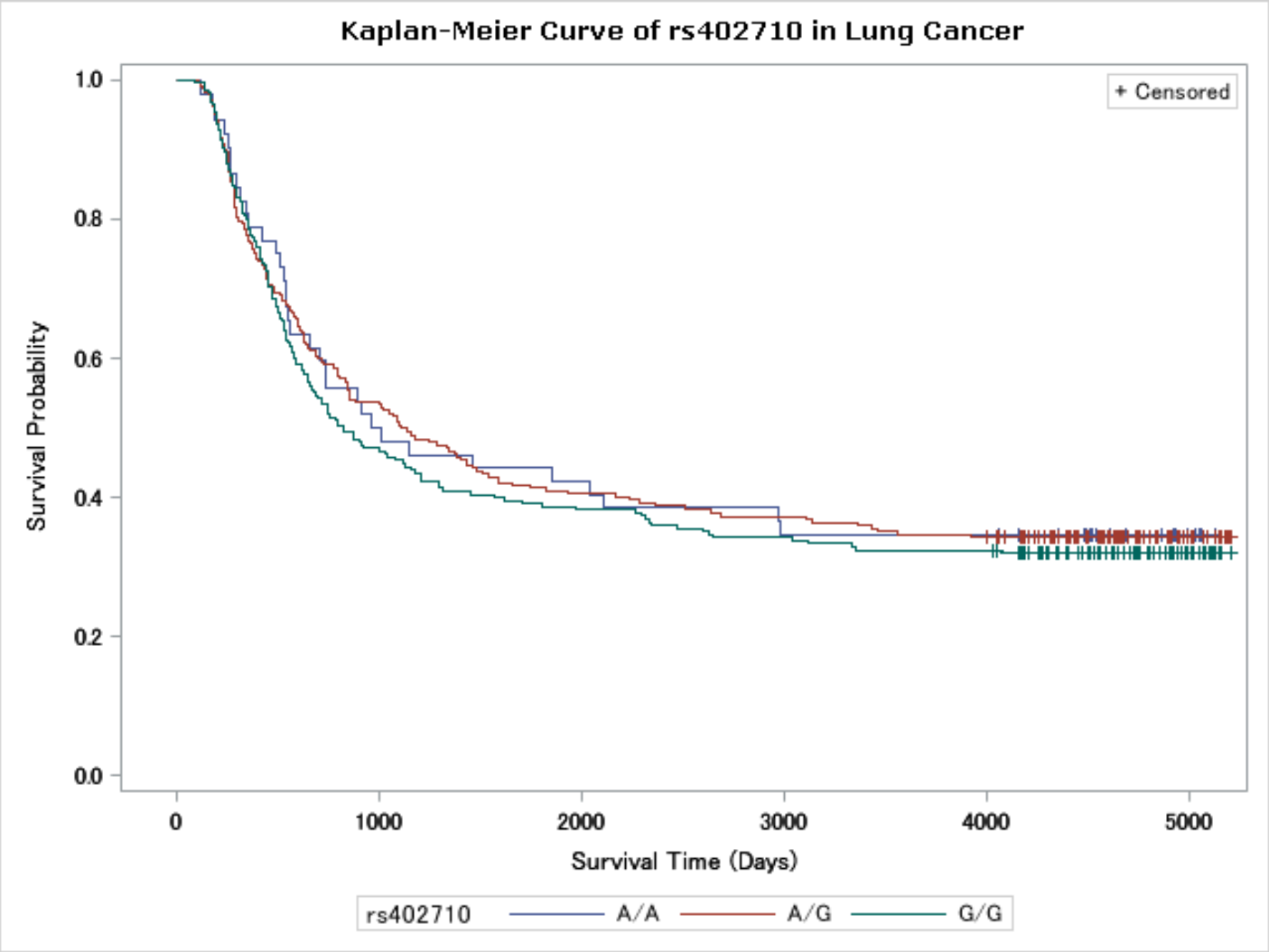

Figure 8. Kaplan-Meier curve of rs402710 in lung cancer

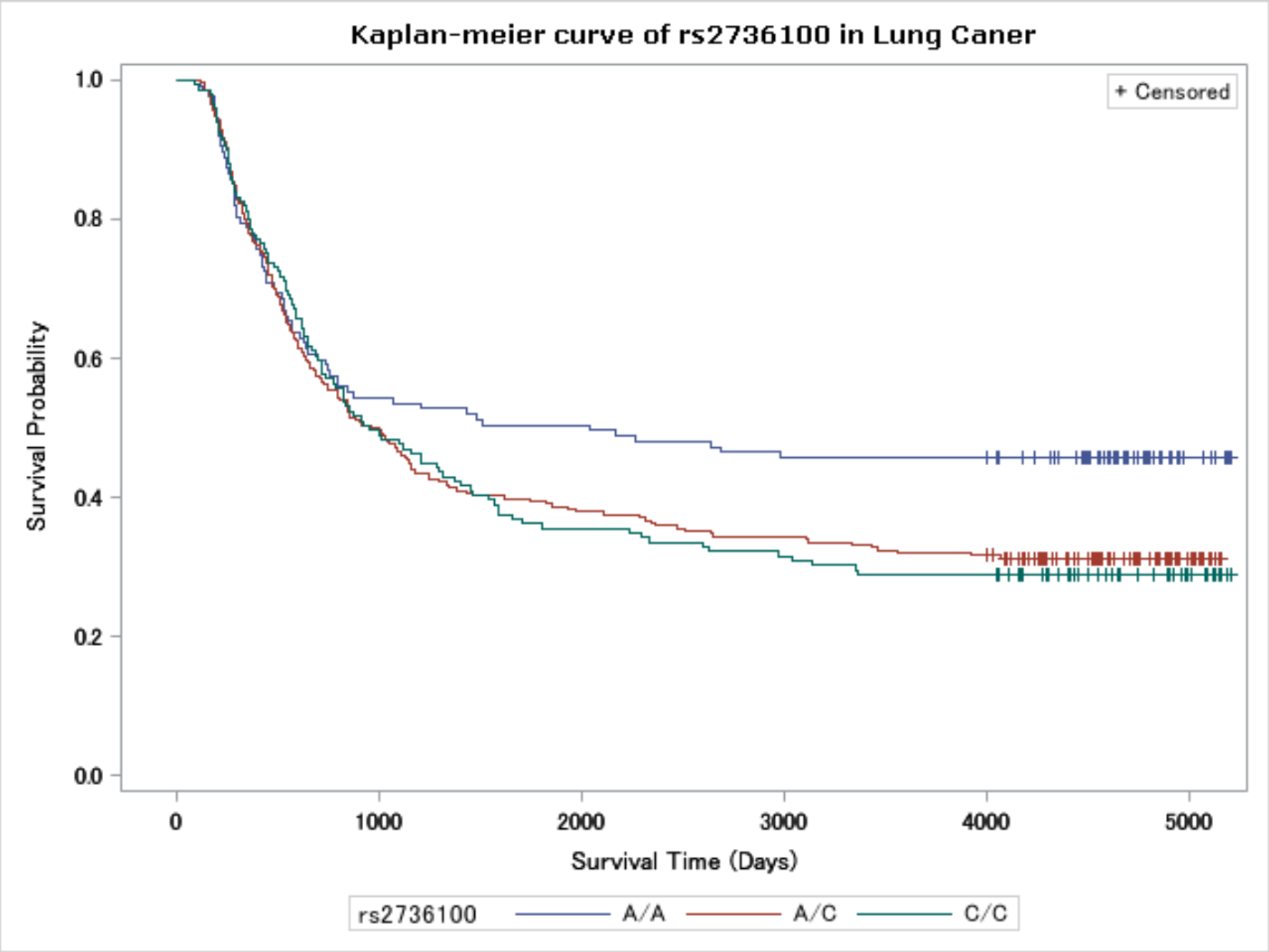

Figure 9. Kaplan-Meier curve of rs2736100 in lung cancer

REFERENCES

[1] Alberg, Anthony J., and Jonathan M. Samet. Epidemiology of lung cancer, Chest Journal 123.1_suppl (2003): 21S-49S.

[2] Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F.

GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet].

Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr, accessed on 02/01/2015.

[3] Centers for Disease Control and Prevention. National Center for Health Statistics. CDC WONDER On-line Database, compiled from Compressed Mortality File 1999-2012 Series 20 No. 2R, 2014

[4] American Cancer Society. Cancer Facts and Figures, 2015.

[5] U.S. National Institutes of Health. National Cancer Institute. SEER Cancer Statistics Review, 1975-2011.

[6] SEER Cancer Statistics Factsheets: Lung and Bronchus Cancer. National Cancer Institute. Bethesda, MD: http://seer.cancer.gov/statfacts/html/lungb.html.

[7] Travis, W.D., Pathology and genetics of tumours of the lung, pleura, thymus and heart. Vol. 7. 2004: Iarc.

[8] Yesner, R., Classification of lung-cancer histology. The New England journal of medicine, 1985. 312(10): p. 652-653.

[9] Subramanian, J., & Govindan, R. (2007). Lung cancer in never smokers: a review. Journal of Clinical Oncology, 25(5), 561-570.

[10] Raz, D. J., He, B., Rosell, R., & Jablons, D. M. (2006). Bronchioloalveolar carcinoma: a review. Clinical lung cancer, 7(5), 313-322.

[11] Buccheri G, Ferrigno D. Prognostic factors in lung cancer: tables and comments. Eur Respir J 1994; 7:1350–1364.

[12] Feld R, Borges M, Giner V, et al. Prognostic factors in non-small cell lung cancer. Lung Cancer 1994; 11(suppl): S19–S23

[13] Smit EF, Groen HJ, Splinter TA, et al. New prognostic factors in resectable non-small cell lung cancer. Thorax 1996; 51:638–646

[14] Mountain CF. New prognostic factors in lung cancer: biologic prophets of cancer cell aggression. Chest 1995; 108:246–254

[15] National Cancer Institute. PDQ for Health Professionals: Non-Small Cell Lung Cancer Treatment. 2008.

[16] National Cancer Institute. Small Cell Lung Cancer Treatment. PDQ for Health Professionals. 2012.

[17] Huang, Y. T., Heist, R. S., Chirieac, L. R., Lin, X., Skaug, V., Zienolddiny, S., ... & Christiani, D. C. (2009). Genome-wide analysis of survival in early-stage non–small-cell lung cancer. *Journal of Clinical Oncology*, *27*(16), 2660-2667.

[18] Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifi es a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008; 40: 616–22.

[19] Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008; 452: 638–42.

[20] Shiraishi, K., Kohno, T., Kunitoh, H., Watanabe, S. I., Goto, K., Nishiwaki, Y., ... & Yokota, J. (2009). Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. *Carcinogenesis*, *30*(1), 65-70.

[21] Laura Jean Bierut, M. D., Stitzel, J. A., Wang, J. C., Hinrichs, A. L., Grucza, R. A., Xuei, X., ... & Goate, A. M. (2008). Variants in nicotinic receptors and risk for nicotine dependence. *The American journal of psychiatry*, *165*(9), 1163-1171.

[22] Grucza, R. A., Wang, J. C., Stitzel, J. A., Hinrichs, A. L., Saccone, S. F., Saccone, N. L., ... & Bierut, L. J. (2008). A risk allele for nicotine dependence in CHRNA5 is a protective allele for cocaine dependence. *Biological psychiatry*,*64*(11), 922-929.

[23] Saccone, N. L., Saccone, S. F., Hinrichs, A. L., Stitzel, J. A., Duan, W., Pergadia, M. L., ... & Bierut, L. J. (2009). Multiple distinct risk loci for nicotine dependence identified by dense coverage of the complete family of nicotinic receptor subunit (CHRN) genes. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *150*(4), 453-466.

[24] Zienolddiny, S., Skaug, V., Landvik, N. E., Ryberg, D., Phillips, D. H., Houlston, R., & Haugen, A. (2009). The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. *Carcinogenesis*,*30*(8), 1368-1371.

[25] Yang, P., Li, Y., Jiang, R., Cunningham, J. M., Zhang, F., & de Andrade, M. (2010). A rigorous and comprehensive validation: common genetic variations and lung cancer. *Cancer Epidemiology Biomarkers & Prevention*, *19*(1), 240-244.

[26] Xun, W. W., Brennan, P., Tjonneland, A., Vogel, U., Overvad, K., Kaaks, R., ... & Vineis, P. (2011). Single-nucleotide polymorphisms (5p15. 33, 15q25. 1, 6p22. 1, 6q27 and 7p15. 3) and lung cancer survival in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Mutagenesis*, ger030.

[27] Hashibe, M., Morgenstern, H., Cui, Y., Tashkin, D. P., Zhang, Z. F., Cozen, W., ... & Greenland, S. (2006). Marijuana use and the risk of lung and upper aerodigestive tract cancers: results of a population-based case-control study.*Cancer Epidemiology Biomarkers & Prevention*, *15*(10), 1829-1834.

[28] Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, *29*(1), 308-311.

[29] Wei, L. J., Ying, Z., & Lin, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika*, *77*(4), 845-851.