BMJ Open

Incidence and Survival of Non-small Cell Lung Cancer in Shanghai: A Population-Based Cohort Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-009419
Article Type:	Research
Date Submitted by the Author:	16-Jul-2015
Complete List of Authors:	Fan, Heng; Fudan University, Department of Biostatistics Shao, Zhen-Yi; Shanghai Municipal Commission of Health and Family Planning, Information Centre Xiao, Yuan-Yuan; Kunming medical university, School of Public Health Xie, Zhi-Hui; Shanghai Municipal Commission of Health and Family Planning, Information Centre Chen, Wen; Shanghai Municipal Commission of Health and Family Planning, Information Centre Xie, Hua; Shanghai Municipal Commission of Health and Family Planning, Information Centre Qin, Guo-You; Fudan University, Department of Biostatistics Zhao, Nai-Qing; Fudan University, Department of Biostatistics
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Oncology, Public health, Smoking and tobacco, Epidemiology
Keywords:	Epidemiology < ONCOLOGY, Respiratory tract tumours < ONCOLOGY, SURGERY, CHEMOTHERAPY

Incidence and Survival of Non-small Cell Lung Cancer in Shanghai: A Population-Based Cohort Study

Heng Fan[1¶], Zhen-Yi Shao[2¶], Yuan-Yuan Xiao[1], Zhi-Hui Xie[2], Wen Chen[2], Hua Xie[2], Guo-You Qin[1], Nai-Qing Zhao[1]*

[1]Department of biostatistics, School of Public Health, Fudan University, Shanghai, China

[2]Information Centre, Shanghai Municipal Commission of Health and Family Planning, Shanghai, China

[¶]These authors contributed equally to this work.

*Corresponding to:

Nai-Qing Zhao, Department of Biostatistics, School of Public Health, Fudan University, No.130 Dong'an Road, Xuhui District, Shanghai 200032, P.R. China;

E-mail: nqzhao@fudan.edu.cn. Tel: +86-21-33563927

Key Words: Non-small cell lung cancer, Incidence, Survival, Surgery, TNM staging

Word count: 3301 words

ABSTRACT

Objectives: Large population-based studies on non-small cell lung cancer (NSCLC) are lacking in mainland China. This study aimed to depict the NSCLC incidence, demographic features and survival as well as prognostic factors affecting survival of patients with NSCLC in urban China, for the first time.

Design: Prospective observational cohort study.

Setting: Baseline information was collected from a mega inpatients database originated from Shanghai metropolitan area.

Participants: All NSCLC cases identified from the database from 2011 to 2013 were recruited (15,020 patients).

Main Results: The crude and age-adjusted incidence rates of NSCLC were 54.20 per 100,000 (55.90 per 100,000 for men, 52.39 per 100,000 for women) and 28.64 per 100, 000, respectively. The median survival time was 22.7 months (95% confidence interval, CI: 21.8-24.2 months) with an overall 1-year survival rate of 71.8% (95%CI: 69.8%-73.8%). The 1-year survival rate by stage was: stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%) and stage IIIb/IV 58.9% (95%CI: 56.1%-61.7%). In the multivariate analysis, patients received surgical resection (Hazard Ratio (HR) =0.607, 95% CI: 0.511-0.722) and chemotherapy (HR=0.838, 95% CI: 0.709-0.991) had improved survival compared with their counterparts. Factors associated with poor survival included older age, male sex, larger tumor size, positive lymph node, distant metastasis and squamous cell carcinoma.

Conclusions: A higher incidence rate and better survival for NSCLC patients were identified compared with previously published foreign studies (and domestic studies, according to estimate), which may provide a fresh perspective for evaluating NSCLC

epidemic and survival.

Strengths and limitations of this study

- This is the first population-based study which depicted the incidence, demographic features and survival of non-small cell lung cancer (NSCLC) in mainland China.
- The current study included a large number of patients with NSCLC from a mega database based on the city's inpatients system, which reflected the entire NSCLC population and allowed adequate statistical analyses in each stage.
- Long-term outcome is complete in Shanghai as the vital status of all citizens is registered in the municipal death registration system.
- One limitation of the study is that some important features of NSCLC patients such as their smoking status, performance status and weight were not available in the database.

INTRODUCTION

Lung cancer remains the most frequently diagnosed cancer worldwide and the leading cause of cancer death in China.[1, 2] Among 85% of the diagnosed lung cancer patients are non–small-cell lung cancer (NSCLC).[3] According to the Surveillance, Epidemiology and End Results (SEER) registry, the incidence rate for NSCLC was 42.6 per 100,000, 49.7 per 100,000 for men and 37.2 per 100,000 for women (adjusted to the United States standard population, 2011).[4] In contrast to the decreasing trend of lung cancer incidence rate in more developed countries, the incidence continues to increase in less developed countries, especially in China.[5] For early stage NSCLC patients, including stage I and II and a subset of stage III disease, the standard and potentially curative treatment is radical resection.[6] The majorities of NSCLC patients are diagnosed at an advanced stage, and are usually not suitable for curative surgical resection. Large population-based studies in western countries indicated that the overall 1-year survival rate for NSCLC was 30%-46%.[7, 8] SEER registry reported the 5-year survival rate for NSCLC was 19%.[4]

Although some population-based studies on the epidemiology and prognosis of NSCLC in western countries have been published, there are few studies investigated the characteristics of NSCLC in China. Existing domestic Chinese studies are mainly annual cancer registry reports, which analyzed lung cancer as a whole (including small cell lung cancer) and reported incidence and mortality only.[1, 9] To our knowledge, there are no published population-based studies in Chinese population focusing on estimating NSCLC incidence and overall survival rates as well as associated demographic and prognostic clinical factors.

Thus, we performed a population-based study using the data from a mega inpatients clinical information database originated from Shanghai metropolitan area to

describe the epidemiological features, survival and explore the prognostic factors of overall survival in patients with NSCLC.

MATERIALS AND METHODS

Ethic statement

This study was approved by the Ethics Committee of the School of Public Health, Fudan University, Shanghai, China. No written consent form was required, since a unique ID was allocated to each patient in advance, in order to replace identifiable personal information. No written consent was given by the patients for their information to be stored in the database and used for research since it was specifically waived by the approving Institutional Review Board.

Data source

Data analyzed in this study were from a mega population-based inpatients database, which covers county-level and above hospitals which are qualified to diagnose cancer in Shanghai. The endpoint mortality data was matched from death registration system. The population demographics in this study were obtained from the Shanghai Statistical Year Book of 2012 and 2013 on the official site of Shanghai Statistics Bureau. Rates are age-standardized (per 100,000 person-years) using the World Standard Population as proposed by Segi and modified by Doll et al.[10, 11]

Case selection and inclusion criteria

From 2011 to 2013, all patients who were diagnosed as NSCLC in Shanghai were recruited to depict the epidemiology features of NSCLC at diagnosis. All these cases were identified by the coding system of the International Classification of Disease for

Oncology, 3rd Revision (ICD-10), from the World Health Organization. In order to confirm the diagnosis and examine the impact of changes on coding, we also retrospectively collected and checked records of diagnosis and pathological reports of these cases. We reviewed the patients' ID, birth date, date of diagnosis and removed duplicated data carefully. Both clinical and pathological tumor, node, metastasis (TNM) data is allowed and was coded to the TNM classification based on the seventh edition of the TNM classification of malignant tumors.[12] We prioritized coded TNM stage where both coded stage and recorded stage were available. Treatment information was also extracted.

Since this mega inpatients database was established online in 2011 and not until 2013 it had covered all qualified hospitals for cancer diagnosis in the city, we selected patients who were diagnosed in 2013 to calculate the yearly incidence rate for NSCLC.

In survival analysis, we excluded patients from all identified NSCLC patients 1) with missing vital status at the deadline of the study(January 31, 2015); 2) with unspecified T, N, or M stage, while cases with stage specified as "x" or "cannot be accessed" is included; 3) with missing baseline information(gender, age or histological type). Survival time was calculated by subtracting the date of diagnosis from the date of death or the deadline of the study, whichever came first.

Statistical methods

The incidence rates of NSCLC in 2013 were calculated by dividing the number of newly-diagnosed NSCLC patients identified from the inpatients database by the number of Shanghai permanent residents provided in Shanghai Statistical Yearbook. We used chi square test to compare percentages of baseline characteristics between

patients with and without surgical resection. Kaplan-Meier survival estimates were used to depict survival rates of NSCLC by cancer stage and treatment (surgical resection vs. non-surgical resection). To investigate the impact on OS, the following clinicopathologic factors were investigated in the univariate Cox proportional hazard model analyses: age, gender, tumor size (T), regional lymph nodes (N), metastasis (M), TNM stage, histology and whether patients received surgery or chemotherapy. Since TNM stage is basically a combination of T, N and M scores, TNM stage was previously excluded from multivariable Cox model while T, N and M scores were still considered as potential prognosis factors. To further explore the influence of surgery and on NSCLC survival, we stratified the analyses by TNM stage. The proportional hazard hypothesis was visually checked with log-log curves. Statistical tests were 2-sided and considered statistically significant for $P<0.05$. Statistical analyses were conducted using SAS version 9.4 (SAS Institute, Inc., Cary, North Carolina).

RESULTS

The incidence rate of non-small cell lung cancer

There were 12,996 newly identified NCSLC cases in Shanghai, 2013, with a crude NSCLC incidence rate of 54.20 per 100,000. Male had higher incidence rate of NSCLC than female, with the incidence rates 55.90 and 52.39 per 100,000, respectively. The age-adjusted incidence rate was 28.64 per 100,000 based on World Standard Population.

Demographic, tumor characteristics and treatment

From January 1, 2011 to December 31, 2013, 15,020 patients with NSCLC were identified. Approximately 53.3% of patients were men. The mean age at diagnosis

was 61.9±10.9 years (range: 15-98 years), and half of the patients were between 55 year-old and 69 year-old upon diagnosis. Tumor stages were available in 34% of patients (5099 patients). Among patients with known stage, 17.9% are with stage I, 5.7% stage II, 12.9% stage IIIa, and 63.5% stage IIIb/IV, indicating advanced disease at diagnosis. We identified 70.6 % of the patients were adenocarcinoma while 27.7 % were squamous cell carcinoma among patients with known histological type. Patients underwent surgical resection had younger age, lower TNM score on tumor size, regional lymph node and metastasis (Table 1).

Among the total study sample, 33.7% underwent a surgical resection (N=5069). This rate ranged from 94.0% among stage I patients to 20.8% among stage IV patients. The total chemotherapy rate is 52.5%, ranged from 55.5% in stage I patients to 82.9% in stage IIIa patients (Table 2).

Table 1. Demographics and tumor characteristics of newly identified NSCLC cases in Shanghai, 2011-2013 (n = 15020)

Characteristics		All subjects (n = 15020)	Surgical resection(n = 5069)	Without surgical resection(n = 9951)	P value
Sex	Male	8002 (53.3%)	2457 (48.5%)	5545 (55.7%)	<.0001
	Female	7018 (46.7%)	2612 (51.5%)	4406 (44.3%)	
Age groups	<55	3396 (22.6%)	1213 (23.9%)	2183 (21.9%)	<.0001
	55-70	7935 (52.8%)	2831 (55.8%)	5104 (51.3%)	
	>=70	3689 (24.6%)	1025 (20.2%)	2664 (26.8%)	
TNM-Tumor*	T1	570 (3.8%)	393 (19.6%)	177 (6.6%)	<.0001
	T2	1630 (10.9%)	1052 (52.4%)	578 (21.4%)	
	T3	646 (4.3%)	254 (12.6%)	392 (14.5%)	
	T4	1771 (11.8%)	286 (14.2%)	1485 (55.1%)	
	Tx	88 (0.6%)	24 (1.2%)	64 (2.4%)	
	Unspecified/unknown	10315 (68.7%)	3060 (-)	7255 (-)	
TNM-Node	N0	1395 (9.3%)	1141 (56.7%)	254 (9.4%)	<.0001
	N1	563 (3.7%)	213 (10.6%)	350 (13.0%)	
	N2	1590 (10.6%)	465 (23.1%)	1125 (41.7%)	
	N3	1020 (6.8%)	151 (7.5%)	869 (32.2%)	
	Nx	144 (1.0%)	44 (2.2%)	100 (3.7%)	
	Unspecified/unknown	10308 (68.6%)	3055 (-)	7253 (-)	
TNM-Metastasis	M0	2390 (15.9%)	1672 (75.4%)	718 (24.4%)	<.0001
	M1	2671 (17.8%)	523 (23.6%)	2148 (73.1%)	
	Mx	95 (0.6%)	22 (1.0%)	73 (2.5%)	
	Unspecified/unknown	9864 (65.7%)	2852 (-)	7012 (-)	
Stage	Ia/Ib	912 (6.1%)	857 (39.5%)	55 (1.9%)	<.0001
	IIa/IIb	292 (1.9%)	253 (11.6%)	39 (1.3%)	
	IIIa	659 (4.4%)	389 (17.9%)	270 (9.2%)	
	IIIb/IV	3236 (21.5%)	673 (31.0%)	2563 (87.6%)	
	Unspecified/unknown	9921 (66.1%)	2897 (-)	7024 (-)	
Histology	Adenocarcinoma	2976 (19.8%)	1408 (70.0%)	1568 (71.1%)	<.0001
	Squamous cell carcinoma	1168 (7.8%)	545 (27.1%)	623 (28.3%)	
	Other(adeno-squamous carcinoma and large cell carcinoma)	73 (0.4%)	59 (2.9%)	14 (0.7%)	
	Unspecified/unknown	10803 (70.3 %)	3057 (-)	7746 (-)	
Chemotherapy	Yes	7134 (47.5%)	2182 (43.0%)	4952 (49.8%)	<.0001
	No	7886 (52.5%)	2887 (57.0%)	4999 (50.2%)	

*TNM: tumor, node and metastasis score.

Table 2. Proportions and rates of surgical resection and chemotherapy, grouping by stage

Stage	N	Surgical resection (n = 5069)		Without surgical resection (n = 9951)		Surgical resection rate	Chemotherapy rate
		Chemotherapy	Without chemotherapy	Chemotherapy	Without chemotherapy		
Ia/Ib	912 (100%)	465 (51.0%)	392 (43.0%)	41 (4.5%)	14 (1.5%)	94.0%	55.5%
IIa/IIb	292 (100%)	184 (63.0%)	69 (23.6%)	27 (9.2%)	12 (4.1%)	86.6%	72.3%
IIIa	659 (100%)	317 (48.1%)	72 (10.9%)	229 (34.7%)	41 (6.2%)	59.0%	82.9%
IIIb/IV	3236 (100%)	508 (15.7%)	165 (5.1%)	1980 (61.2%)	583 (18.0%)	20.8%	76.9%
Unknown	9921 (100%)	1413 (14.2%)	1484 (15.0%)	2722 (27.4%)	4302 (43.4%)	29.2%	41.7%
Total	15020	2887	2182	4999	4952	33.7%	52.5%

Overall survival and prognostic factors

In our study of survival, a total of 2,013 NSCLC patients were included from the date of diagnosis until death or January 31, 2015. 1009 patients (50.1%) experienced death during this period.

The median survival time was shown in Table 3 and plots of the product-limit estimates for survival rate were depicted by stage and surgical resection, respectively (Figure 1, Figure 2). The median survival time for all NSCLCs was 22.7 months (95% confidence interval, CI: 21.8-24.2 months) with 1-year survival rate of 71.8% (95%CI: 69.8%-73.8%). The median survival time was unavailable for stage I and stage II patients, while for stage IIIa and stage IIIb/IV patients the median survival was 24.3 months (95% CI: 21.4-26.2 months) and 16.0 months (95% CI: 14.8-16.7 months). The 1-year survival rate by stage was: stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%) and stage IIIb/ IV 58.9% (95%CI: 56.1%-61.7%) (Table 3, Figure 1).

Table 3. The median overall survival and prognostic factors of newly diagnosed NSCLC cases in Shanghai, 2011-2013 (n = 2013)

Characteristics		Median (95%CI)*	Crude HR (95%CI) **	Adjusted HR (95%CI)***	*P* value***
Surgery	Yes	34.4 (29.5-38.1)	0.276 (0.240 -0.318)	0.607 (0.511-0.722)	0.000
	No	15.4 (14.1-16.5)	1.00	1.00	-
Chemotherapy	Yes	22.2 (21.2-23.2)	1.145 (0.974 -1.347)	0.838 (0.709-0.991)	0.039
	No	26.9 (23.2-32.1)	1.00	1.00	-
Gender	Male	19.2 (17.7-20.4)	1.721 (1.508 -1.964)	1.751 (1.521-2.015)	0.000
	Female	26.2 (25.7-29.1)	1.00	1.00	-
Age (yrs)	<55	24.9 (22.7-26.2)	1.00	1.00	-
	55-70	25.1 (23.7-26.0)	1.048 (0.885 -1.241)	1.111 (0.936-1.318)	0.228
	>=70	16.1 (14.6-18.3)	1.850 (1.539 -2.224)	1.727 (1.426-2.091)	0.000
T	T1	30.3 (28.7-48.2)	1.00	1.00	-
	T2	27.5 (25.9-33.9)	1.437 (1.122 -1.841)	1.214 (0.945-1.561)	0.129
	T3	18.4 (15.5-21.5)	2.847 (2.186 -3.707)	1.461 (1.111-1.920)	0.007
	T4	16.3 (14.5-16.9)	3.545 (2.807 -4.477)	1.385 (1.083-1.772)	0.009
	Tx	10.3 (6.8-16.3)	3.715 (2.400 -5.751)	1.571 (0.872-2.830)	0.133
N	N0	-	1.00	1.00	-
	N1	21.6 (19.0-24.2)	3.890 (3.022 -5.007)	1.949 (1.483-2.563)	0.000
	N2	17.1 (15.8-19.0)	5.221 (4.240 -6.428)	2.845 (2.263-3.576)	0.000
	N3	13.6 (11.8-15.4)	6.927 (5.567 -8.620)	3.527 (2.762-4.504)	0.000
	Nx	12.0 (8.4-18.3)	5.898 (4.073 -8.541)	2.482 (1.489-4.139)	0.000
M	0	30.3 (27.6-35.1)	1.00	1.00	-
	1	15.6 (14.3-16.7)	3.000 (2.627 -3.427)	1.722 (1.456-2.037)	0.000
	x	-	3.223 (1.920 -5.411)	1.458 (0.859-2.476)	0.162
Stage	Ia/Ib	-	1.00	-	-
	IIa/IIb	-	3.578 (2.144 -5.971)	-	-
	IIIa	24.3 (21.4-26.2)	8.094 (5.508 -11.892)	-	-
	IIIb/IV	16.0 (14.8-16.7)	14.594 (10.247 -20.785)	-	-
Histology	Adenocarcinoma	24.4 (23.1-25.7)	1.00	1.00	-
	Squamous cell carcinoma	19.0 (16.8-21.1)	1.370 (1.195 -1.571)	1.172 (1.003-1.369)	0.045
	Other	25.3 (22.7-38.6)	0.768 (0.468 -1.262)	1.058 (0.639-1.752)	0.827

* CI, confidence interval; Median and 95% CI were estimated using Kaplan-Meier curve by each variable individually.

**HR, Hazard ratio; Crude hazard ratio and 95% CI was estimated using univariate Cox regression model.

*** Adjusted hazard ratio, 95% CI and *P* value were estimated using multiple Cox regression model adjusted by surgical resection, chemotherapy, sex, age group, TNM score and histology.

The patients underwent surgical resection had better survival than patients who did not, with median survival 34.4 months (95% CI: 29.5-38.1 months) vs. 15.4 months (95% CI: 14.1-16.5 months) and 1-year survival rate 87.8% (95%CI: 85.7%-89.9%) vs. 57.9% (95%CI: 55.0%-60.8%), respectively (Table 3).

In univariate analysis, patients who were female, younger, with smaller tumor size, negative lymph node, non-metastasis, lower stage, with surgical resection and with adenocarcinoma had longer survival time than their counterparts, while patients received chemotherapy did not show any survival benefit compared with patients who did not receive chemotherapy. However, after adjusted by demographic factors and tumor characteristics in multivariate analysis, patients received chemotherapy showed significantly better survival (HR=0.838, 95% CI: 0.709-0.991) than their counterparts. Patients received surgical resection also had improved survival (HR=0.607, 95% CI: 0.511-0.722) compared with patients who did not receive resection. In this multivariable Cox proportional hazard model, factors associates with poor survival included male sex (HR=1.751, 95% CI: 1.521-2.015), older age at diagnosis (age>=70 years vs. age <55 years: HR=1.727, 95% CI: 1.426-2.091), larger tumor size (T4 vs. T1: HR=1.385, 95% CI: 1.083-1.772), positive lymph node (N3 vs. N0: HR=3.527, 95% CI: 2.762-4.504), distant metastasis (HR=1.722, 95% CI: 1.456-2.037) and squamous cell carcinoma (HR=1.172, 95% CI: 1.003-1.369) (Table 3).

In order to further evaluate the prognostic role of surgery in NSCLC patients, we performed additional multivariable analyses according to each TNM stage. T, N and M scores were excluded since TNM stage was set in each stratum. The survival benefit of surgery was observed among stage IIIa patients (HR=0.513, 95% CI: 0.352-0.748) and stage IIIb/IV patients (HR=0.646, 95% CI: 0.536-0.779) (Table 4, Figure 2).

Table 4. Multivariate Hazard Ratio of overall survival according to surgical resection by stage (n = 2013)

Stage	n = 2013	Surgical resection vs. without surgical resection (ref.)	
		Adjusted HR (95% CI)*	*P* value*
Ia/Ib	451	0.360 (0.104-1.237)	0.105
IIa/IIb	110	0.723 (0.205-2.542)	0.613
IIIa	288	0.513 (0.352-0.748)	0.001
IIIb/IV	1164	0.646 (0.536-0.779)	0.000

*HR, Hazard ratio; Adjusted hazard ratio, 95% CI and *P* value were estimated using multiple Cox regression model adjusted by sex, age group and histology type.

DISCUSSION

Incidence

To our knowledge, this is the first population-based with the largest sample size study to describe the epidemiological characteristics of NSCLC specifically in mainland China. Our study showed the crude incidence rate of NSCLC in 2013 was 54.20 per 100,000 (55.90 per 100,000 for men and 52.39 per 100,000 for women) with an age-adjusted incidence of 28.64 per 100,000 (adjusted by World Standard Population). Compared with studies of SEER registry, the crude incidence in this study was higher than that of all races in SEER registry (42.6 per 100,000 overall, 49.7 per 100,000 for men and 37.2 per 100,000 for women, adjusted to the United States standard population, 2011)[4] and that of Chinese ethnic group in the registry (52.0 per 100,000 for men and 29.9 per 100,000 for women, 2004-2008).[13] Compared with the NSCLC incidence of the American population, one possible explanation to the higher crude NSCLC incidence in our study could be the deteriorating ageing of Chinese population, as an older age has been identified as independent risk factor to NSCLC.[6] Ageing problem is especially serious in Shanghai, where 27% of its population was beyond 60 years old in 2013,[14] while this proportion was merely 16.5% in 2000 US population.[15]

Most existing population-based studies analyzed lung cancer as a whole, including both NSCLC and small cell lung cancer. Compare with the adjusted incidence rate of lung cancer of GLOBOCAN 2012 (30.0 per 100,000 for men and 11.1 per 100,000 for women, incidence adjusted by World Standard Population, 2012)[1] and crude incidence rate of lung cancer of National Central Cancer Registry (48.3 per 100,000, 2011)[2] and estimate the NSCLC incidence as 85% of lung cancer incidence, the NSCLC incidence rate of our study was higher than the global average and Chinese national average. According to an official report by Shanghai Municipal center of Disease Control and Prevention (Shanghai Cancer Report 2012, Shanghai Municipal Center Disease Control and Prevention), the adjusted lung cancer incidence was 29.8 per 100,000 in 2010. Estimating as 85% of lung cancer incidence, the NSCLC incidence was approximately 25.3 per 100,000. With the same standard population adjusted by, we saw a 4.2% increase in NSCLC incidence per year from 2010 to 2013 in Shanghai. These results coincide with the prediction in rising lung cancer incidence in China as described in many researches.[1, 16, 17] Several factors may contribute to this higher and increasing NSCLC incidence in our study except for ageing problem. First, the smoking prevalence has been dramatically increased in the past 2 decades in China. Although cigarette smoking rate has peaked and decreased in the United States and several other areas these years, the prevalence of smoking in China remains high and has become one of the highest in the world, where 53% of men aged 15 years and above are current smokers, according to the 2010 report of China Global Adults Smoking Survey (GATS).[16, 18, 19] Considering that smoking is the main risk factor of NSCLC,[6] this high smoking prevalence lasts in the past three decades in China has no doubt played an important role of its lung cancer epidemic. Another factor that may lead to this higher incidence is that as one of the most

developed cities in China, Shanghai enjoys improved quality of oncology services which is similar to that in Western countries, thus a higher diagnose rate than other Chinese cities can be expected.[20]

It is noteworthy that a higher ratio of NSCLC incidence among women compared to men was observed in this study (0.93), while this ratio was 0.75 and 0.58 in the former mentioned SEER study of whole population and Chinese ethnicity, respectively.[4, 13] A relatively higher lung cancer incidence rate for women in China was also identified by Boffetta et al. compared with other ethnic groups among less developed countries.[21] Besides, 2012 GLOBLECAN reports also pointed out lung cancer rates among women in China were higher than that in some European countries despite a lower prevalence of smoking[1]. Except for possible age constituent ratio difference between countries, this higher crude NSCLC incidence of women in our study may possibly reflect the impact of household air pollution from cooking fumes and unventilated coal-fueled heating stoves.[22-24] Besides, considering the high overall smoking prevalence in China, second-hand-smoking may also be a critical hazard of NSCLC for non-smokers, typically women. A nationwide cross-sectional survey including 15,540 Chinese adults showed that in 2000–2001 more than 49.2% of adult female non-smokers reported exposure to tobacco smoke exposure,[25] while this rate was only 35% in a 2004 worldwide study including 192 countries.[26] This suggests an additional risk Chinese women are experienced with.

Survival

A better survival (overall and stage-specified) was observed in this study compared with previously published population-based studies of non-Asian ethnicity, though different population-based lung cancer databases showed discordant survival

outcomes. According to databases from Australia, Canada, Denmark, Norway, Sweden and the UK, their 1-year overall survival rates of NSCLC in 2004-2007 ranged from 30% to 46%, with stage-specified 1-year survival rates of 71.1%-86.2% for stage I, 58.6%-79.0% for stage II, 34.4%-37.1% for stage III and 15.5%-25.9% for stage IV[7]. Lower survival rates were also observed in the study of SEER program in 1998-2003 (stage IV 1-year survival rate 15.9% (95% CI: 15.6%-16.3%) and the study of Rasco et al[27, 28]. However, Asian ethnicity shows improved survival. Lin et al. reported 2-year survival rates for 30,069 Taiwanese patients of 80.0%-96.2%, 64.4%-80.2% and 57.5%-67.4% for stage I, stage II and stage IIIa patients, respectively[29]. In a study of 4622 Korean patients, the median overall survival for stage I, stage II, stage III and stage IV were 100, 41, 14 and 7 months.[30] No population-based study in mainland China published in English can be found.

The better survival outcome observed in this study may root in several reasons. Firstly, Asian ethnicity has been recognized to be an independent favorable prognostic factor for OS outcomes among NSCLC patients.[31, 32] Asian patients with NSCLC may differ in their response to cytotoxic chemotherapy, when compared with white patients. For example, Gandara DR et al. reported a 3-month increase of median overall survival of Japanese patients than white patients undergoing the same paclitaxel plus carboplatin regimen.[33] At the same time, epidermal growth factor receptor (EGFR) mutations confer survival benefit independent of treatment in NSCLC[34, 35], while the East Asian population has the highest frequency of EGFR mutation in the world.[36, 37] Secondly, in recent years, several advances in diagnostic techniques and therapeutics have influenced outcomes of NSCLC, especially adenocarcinoma. The formerly mentioned EGFR mutation is only presented in lung adenocarcinoma and related to a favorable response to EGFR tyrosine kinase inhibitors (TKI). Furthermore, the new

salvage therapies, pemetrexed is active in non-squamous NSCLC only.[38, 39] In this study, a better survival has been observed in lung adenocarcinoma compared with squamous cell carcinoma. As 72.3% of the NSCLC patients with known histological types were lung adenocarcinoma in this study, while this rate is less than 45% in former mentioned western studies,[7, 27] a longer survival time may therefore be expected.

Similar to previous studies, this study found that female gender, younger age, smaller tumor size, negative lymph node and non-metastasis were related to better survival. Within the evaluation of the impact of surgery in each TNM stage, we found that stage IIIa and stage IIIb/IV patients who underwent surgical resection had improved survival than their counterparts. This suggests active treatment may prolong survival, even for stage IIIa and more advanced patient, though more details of patients' therapy modality are still needed to decide the impact of surgery. Currently there are controversies centered on the role of surgery for stage IIIa patients. According to the Chinese lung cancer guidelines, surgical resection is the current standard treatment for patients with stage I to stage IIIA disease; some stage IV patients with single metastasis are also suitable for surgery.[40] Thanks to the improvements in diagnostic imaging and endoscopic techniques, "current evidence supports an expansion in surgery as part of multimodality management of patients with N2 disease, and greater uptake in patients who are willing to accept higher risks", according to Goldstraw et al.[41]. In multivariate analysis, receiving chemotherapy also showed positive prognostic effect, suggesting that confounding may exists in univariate analysis.

Our study has several strengths. First, this is the first study, to our knowledge, to evaluate the incidence, survival and prognostic factors of NSCLC based on a large

population in mainland China. Existing domestic studies, mainly annual cancer registry reports, analyzed lung cancer as a whole and reported incidence and mortality only. Second, our data were obtained from a citizen healthcare information network which is the most advanced and comprehensive nationwide, thus could be quite representative of the composition of NSCLC population in urban China. Furthermore, this study was based on data through 2013, whereas the most recently NSCLC population-based study among those of other Asian countries or districts using data of 2010.[23, 29, 42] Last, our study reported a higher incidence and better survival of NSCLC patients compared with previous studies, which may provide a fresh and meaningful perspective for evaluating NSCLC diagnose and treatment, considering the ethnic difference, smoking prevalence and treatment improvement.

However, this study also suffers from several limitations. First, as basically a retrospective study, some important features of NSCLC patients such as their performance status, weight and details of treatment were not available in the database. Specifically, patients' smoking status is unavailable as well. As a known prognostic factor,[43] its absence may lead to residual confounding. Also, certain percentage of missing in some variables was observed, especially in TNM classification. However, we ruled out the possibility of selection bias in survival analysis by restricted the sample to patients with known potential prognostic factors.

CONCLUSIONS

In this study, we identified a higher incidence and better survival for NSCLC patients compared with previously published foreign and domestic studies. High smoking prevalence and the consequently high environment tobacco exposure may play an important role in the higher NSCLC incidence both overall and in women. In

addition to female gender and younger age, receiving surgical resection is a positive prognosis factor for NSCLC patients with stage IIIa and above.

Acknowledgements

The authors thank the staffs of the database and death registration system for their efforts in data collection and preprocessing.

Contributors

Conceived and designed the experiments: NZ GQ ZS. Performed the experiments: HF ZS YX ZX WC HX. Analyzed the data: HF. Wrote the paper: HF YX ZS. Receive the manuscript: NZ GQ.

Funding

This research was supported by National Natural Science Foundation of China (Grant 81273187), National Science and Technology Major Project of the People's Republic of China (Major Grant 2012ZX09303-013-014) and Key Lab of Health Technology Assessment, Ministry of Health of the People's Republic of China.

Competing interests

None declared.

Ethics approval

The present study complies with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Public Health, Fudan University, Shanghai, China.

Data sharing statement

Additional data are held by the corresponding author.

REFERENCES

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin 2015; 65: 87-108.
2. Chen W, Zheng R, Zeng H, et al. Annual report on status of cancer in China, 2011. Chin J Cancer Res 2015; 27: 2-12.
3. Oser MG, Niederst MJ, Sequist LV, et al. Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. Lancet Oncol 2015; 16: e165 – e172.
4. SEER Cancer Statistics Review, 1975-2010. Available at: http://seer.cancer.gov. Bethesda, MD: National Cancer Institute;2012. Updated April 2013.
5. Youlden DR, Cramb SM, Baade PD. The International Epidemiology of Lung Cancer: geographical distribution and secular trends. J Thorac Oncol 2008; 3: 819-31.
6. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Non-Small Cell Lung Cancer. V3, 2014. Available at: www.nccn.org. Accessed March 1, 2015
7. Walters S, Maringe C, Coleman MP, et al. Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004-2007. Thorax 2013; 68: 551-64.
8. Coleman MP, Forman D, Bryant H, et al. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. Lancet 2011; 377: 127-38.
9. Chen W, Zheng R, Zhang S, et al. Annual report on status of cancer in China, 2010. Chin J Cancer Res 2014; 26: 48-58.
10. Segi M. Cancer Mortality for Selected Sites in 24 Countries (1950-57). Japan: Department of Public Health, Tohoku University of Medicine 1960.
11. Doll R, Cook P. Summarizing indices for comparison of cancer incidence data. Int J Cancer 1967;

2: 269-79.

12. Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. J Thorac Oncol 2007; 2: 706-14.
13. Gomez SL, Noone AM, Lichtensztajn DY, et al. Cancer incidence trends among Asian American populations in the United States, 1990-2008. J Natl Cancer Inst 2013; 105: 1096-110.
14. Gerontological Society of Shanghai. Nov 2010. Available at: http://www.shanghaigss.org.cn/news_view.asp?newsid=9689. Accessed March 1, 2015
15. Standard Populations - 19 Age Groups, Surveillance, Epidemiology, and End Results Program [National Cancer Institute Web site]. Available at: http://seer.cancer.gov/stdpopulations/stdpop.19ages.html. Accessed March 1, 2015.
16. Yang L, Parkin DM, Li L, et al. Time trends in cancer mortality in China: 1987-1999. Int J Cancer 2003; 106: 771-83.
17. Molina JR, Yang P, Cassivi SD, et al. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc 2008; 83: 584-94.
18. Zhang H, Cai B, The impact of tobacco on lung health in China. Respirology 2003; 8: 17-21.
19. Zhang J, Ou JX, Bai CX, Tobacco smoking in China: prevalence, disease burden, challenges and future strategies. Respirology 2011; 16: 1165-72.
20. Yang LL, Zhang XC, Yang XN, et al., Lung cancer treatment disparities in china: a question in need of an answer. Oncologist 2014; 19(10): 1084-90.
21. Boffetta, P., Parkin, D. M., Cancer in developing countries. Ca : A Cancer Journal for Clinicians, 1994;44(2), 81. Available at: http://search.proquest.com/docview/211963725?accountid=10025
22. Cancer IAFR (2012) Personal Habits and Indoor Combustions. IARC Monnograohs on the Evaluation of Carcinogenic Risks to Humans 2012;100(Pt E):1-538.
23. Wang BY, Huang JY, Cheng CY, et al. Lung cancer and prognosis in taiwan: a population-based cancer registry. J Thorac Oncol 2013; 8: 1128-35.
24. Feng G, Jiang Y, Zhao L, et al. Degree of exposure to secondhand smoking and related knowledge, attitude among adults in urban China. Zhonghua Liu Xing Bing Xue Za Zhi [in Chinese]2014; 35: 998-1001.
25. Gu D, Wu X, Reynolds K, et al. Cigarette smoking and exposure to environmental tobacco smoke

in China: the international collaborative study of cardiovascular disease in Asia. Am J Public Health 2004; 94: 1972-6.

26. Oberg M, Jaakkola MS, Woodward A, et al. Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries. Lancet 2011; 377: 139-46.
27. Cetin K, Ettinger DS, Hei YJ, et al. Survival by histologic subtype in stage IV nonsmall cell lung cancer based on data from the Surveillance, Epidemiology and End Results Program. Clin Epidemiol 2011; 3: 139-48.
28. Rasco DW, Yan J, Xie Y, et al. Looking beyond surveillance, epidemiology, and end results: patterns of chemotherapy administration for advanced non-small cell lung cancer in a contemporary, diverse population. J Thorac Oncol 2010; 5: 1529-35
29. Lin ZZ, Shau WY, Shao YY, et al. Survival following surgery with or without adjuvant chemotherapy for stage I-IIIA non-small cell lung cancer: an east asian population-based study. Oncologist 2012; 17: 1294-302.
30. Ahn MJ, Lee J, Park YH, et al. Korean ethnicity as compared with white ethnicity is an independent favorable prognostic factor for overall survival in non-small cell lung cancer before and after the oral epidermal growth factor receptor tyrosine kinase inhibitor era. J Thorac Oncol 2010; 5: 1185-96.
31. Ou SH, Ziogas A, Zell JA, Asian ethnicity is a favorable prognostic factor for overall survival in non-small cell lung cancer (NSCLC) and is independent of smoking status. J Thorac Oncol 2009; 4: 1083-93.
32. Tannenbaum SL, Koru-Sengul T, Zhao W, et al. Survival disparities in non-small cell lung cancer by race, ethnicity, and socioeconomic status. Cancer J 2014; 20: 237-45.
33. Gandara DR, Kawaguchi T, Crowley J, et al. Japanese-US common-arm analysis of paclitaxel plus carboplatin in advanced non-small-cell lung cancer: a model for assessing population-related pharmacogenomics. J Clin Oncol 2009; 27: 3540-6.
34. Bell DW, Lynch TJ, Haserlat SM, et al. Epidermal growth factor receptor mutations and gene amplification in non-small-cell lung cancer: molecular analysis of the IDEAL/INTACT gefitinib trials. J Clin Oncol 2005; 23: 8081-92.
35. Eberhard DA, Johnson BE, Amler LC, et al. Mutations in the epidermal growth factor receptor

and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. J Clin Oncol 2005; 23: 5900-9.

36. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004; 350: 2129-39.
37. Calvo E, Baselga J, Ethnic differences in response to epidermal growth factor receptor tyrosine kinase inhibitors. J Clin Oncol 2006; 24: 2158-63.
38. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol 2008; 26: 3543-51.
39. Scagliotti G, Hanna N, Fossella F, et al. The differential efficacy of pemetrexed according to NSCLC histology: a review of two Phase III studies. Oncologist 2009; 14: 253-63.
40. XY Z, YK S, JM Y, Standards for the diagnosis and treatment of primary lung cancer (2015 version) in China. Chinese Journal of Oncology [in Chinese]2015; 37.
41. Goldstraw P, Ball D, Jett JR, et al. Non-small-cell lung cancer. Lancet 2011; 378: 1727-4

Figure legends

Figure 1. The overall survival (OS) of NSCLC patients in Shanghai identified in 2011-2013 (n=2013). 1-year OS rate: whole population 71.8% (95% CI: 69.8%-73.8%), stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%), stage IIIb/IV 58.9% (95%CI: 56.1%-61.7%). The survival difference was significant ($p < 0.0001$).

Figure 2. The overall survival (OS) of NSCLC cases in Shanghai identified in 2011-2013 according to surgery by stage (n=2013). 1-year OS rate of stage I patients: with surgery 96.3% (95%CI: 94.5%-98.1%), without surgery 100.0%; stage II: with surgery 90.0% (95%CI: 84.1%-95.9%), without surgery 80.0% (95%CI:

55.2%-100.0%); stage IIIa: with surgery 84.3% (95%CI: 79.1%-89.5%), without surgery 68.9% (95%CI: 60.0%-77.8%); stage IIIb/IV: with surgery 73.1% (95% CI: 67.2%-79.0%), without surgery 55.7% (95% CI: 52.5%-58.9%). The survival benefit of surgery was observed among stage IIIa patients (adjusted HR=0.513, 95% CI: (0.352-0.748) and stage IIIb/IV patients (adjusted HR=0.646, 95% CI: 0.536-0.779).

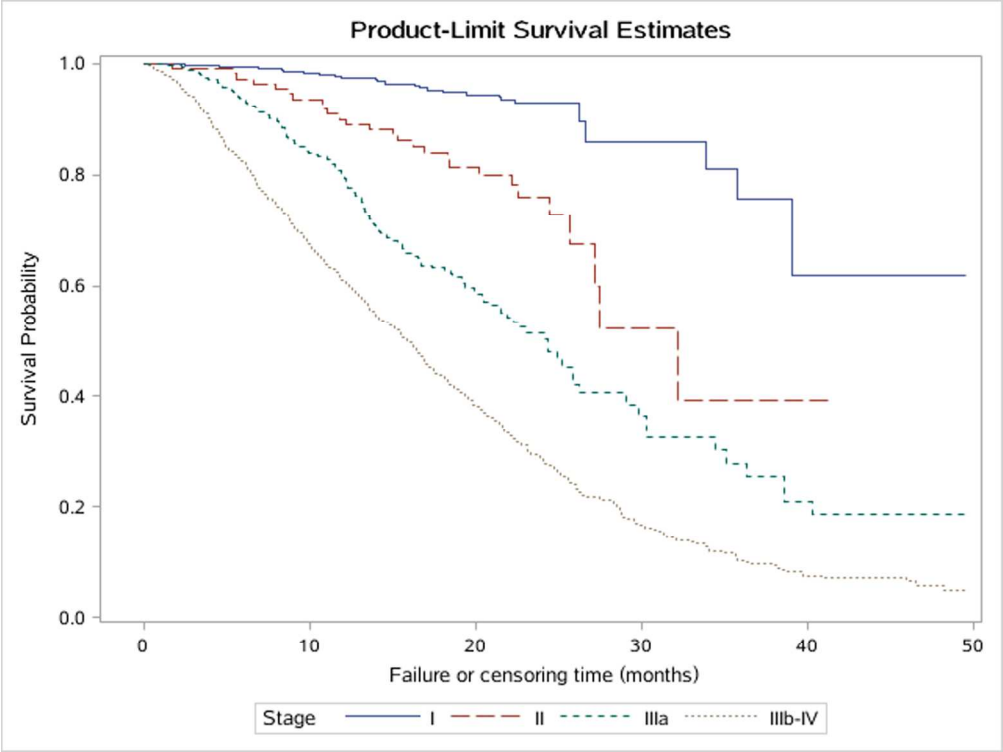

Figure 1. The overall survival (OS) of NSCLC patients in Shanghai identified in 2011-2013 (n=2013). 1-year OS rate: whole population 71.8% (95% CI: 69.8%-73.8%), stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%), stage IIIb/IV 58.9% (95%CI: 56.1%-61.7%). The survival difference was significant ($p < 0.0001$).
127x95mm (600 x 600 DPI)

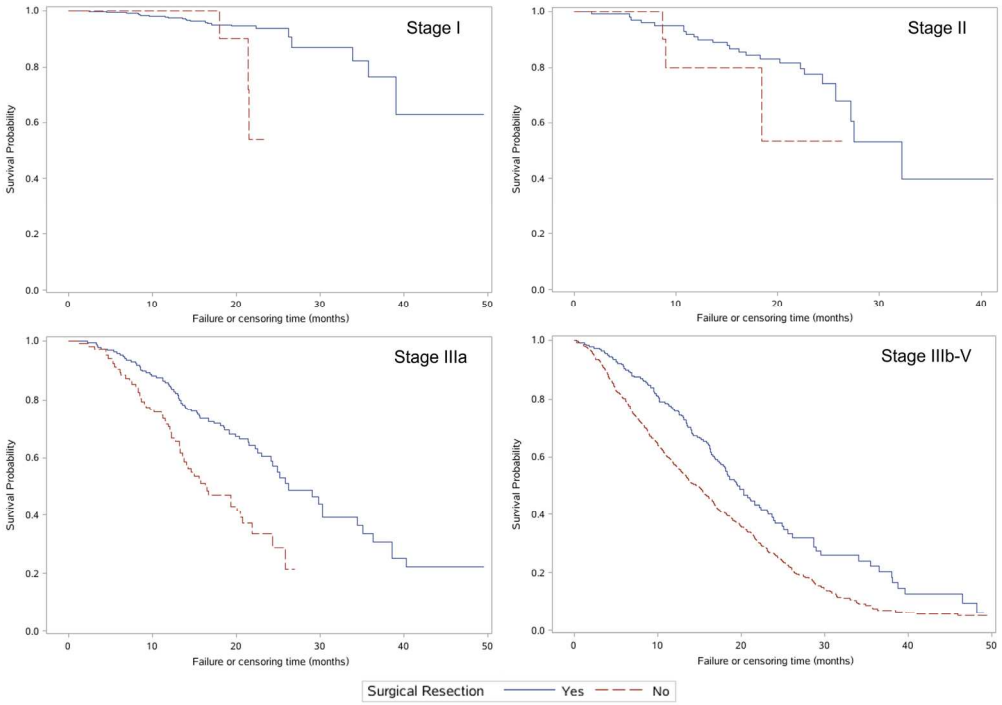

Figure 2. The overall survival (OS) of NSCLC cases in Shanghai identified in 2011-2013 according to surgery by stage (n=2013). 1-year OS rate of stage I patients: with surgery 96.3% (95%CI: 94.5%-98.1%), without surgery 100.0%; stage II: with surgery 90.0% (95%CI: 84.1%-95.9%), without surgery 80.0% (95%CI: 55.2%-100.0%); stage IIIa: with surgery 84.3% (95%CI: 79.1%-89.5%), without surgery 68.9% (95%CI: 60.0%-77.8%); stage IIIb/IV: with surgery 73.1% (95% CI: 67.2%-79.0%), without surgery 55.7% (95% CI: 52.5%-58.9%). The survival benefit of surgery was observed among stage IIIa patients (adjusted HR=0.513, 95% CI: (0.352-0.748) and stage IIIb/IV patients (adjusted HR=0.646, 95% CI: 0.536-0.779).
177x127mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Incidence and Survival of Non-small Cell Lung Cancer in Shanghai: A Population-Based Cohort Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-009419.R1
Article Type:	Research
Date Submitted by the Author:	02-Sep-2015
Complete List of Authors:	Fan, Heng; Fudan University, Department of Biostatistics Shao, Zhen-Yi; Shanghai Municipal Commission of Health and Family Planning, Information Centre Xiao, Yuan-Yuan; Kunming medical university, School of Public Health Xie, Zhi-Hui; Shanghai Municipal Commission of Health and Family Planning, Information Centre Chen, Wen; Shanghai Municipal Commission of Health and Family Planning, Information Centre Xie, Hua; Shanghai Municipal Commission of Health and Family Planning, Information Centre Qin, Guo-You; Fudan University, Department of Biostatistics Zhao, Nai-Qing; Fudan University, Department of Biostatistics
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Oncology, Public health, Smoking and tobacco, Epidemiology
Keywords:	Epidemiology < ONCOLOGY, Respiratory tract tumours < ONCOLOGY, SURGERY, CHEMOTHERAPY

Incidence and Survival of Non-small Cell Lung Cancer in Shanghai: A Population-Based Cohort Study

Heng Fan[1¶], Zhen-Yi Shao[2¶], Yuan-Yuan Xiao[1], Zhi-Hui Xie[2], Wen Chen[2], Hua Xie[2], Guo-You Qin[1], Nai-Qing Zhao[1]*

[1]Department of Biostatistics, School of Public Health, Fudan University, Shanghai, China

[2]Information Centre, Shanghai Municipal Commission of Health and Family Planning, Shanghai, China

[¶]These authors contributed equally to this work.

*Corresponding to:

Nai-Qing Zhao, Department of Biostatistics, School of Public Health, Fudan University, No.130 Dong'an Road, Xuhui District, Shanghai 200032, P.R. China;

E-mail: nqzhao@fudan.edu.cn. Tel: +86-21-33563927

Key Words: Non-small cell lung cancer, Incidence, Survival, Surgery, TNM staging

Word count: 3622 words

ABSTRACT

Objectives: Large population-based studies on the incidence and outcome of non-small cell lung cancer (NSCLC) are lacking in mainland China. This study aimed to investigate the NSCLC incidence, demographic features and survival as well as factors affecting survival of NSCLC patients in Shanghai.

Design: Prospective observational cohort study.

Setting: Baseline information was collected from Shanghai Health Information Network, which is based on the Health Information Systems from all the comprehensive hospitals and specialist hospitals qualified for cancer diagnosis in Shanghai metropolitan area.

Participants: All NSCLC cases identified from the database between 2011 and 2013 were recruited (15,020 patients).

Main Results: The crude and age-adjusted incidences of NSCLC were 54.20 per 100,000 people (55.90 per 100,000 for men, 52.39 per 100,000 for women) and 39.05 per 100, 000 people (41.43 per 100,000 for men and 37.13 per 100,000 for women), respectively. The median survival time was 22.7 months (95% confidence interval [CI]: 21.8-24.2 months) with an overall 1-year survival rate of 71.8% (95%CI: 69.8%-73.8%). The 1-year survival rate was 96.5% (95% CI: 94.0%-98.6%) in stage I NSCLC patients, 89.1% (95%CI: 83.3%-94.9%) in stage II NSCLC patients, 78.8% (95%CI: 74.1%-83.5%) in stage IIIa NSCLC patients and 58.9% (95%CI: 56.1%-61.7%) in stage IIIb/IV NSCLC patients. Multivariate analysis showed surgical resection (Hazard Ratio [HR] =0.607, 95% CI: 0.511-0.722) and chemotherapy (HR=0.838, 95% CI: 0.709-0.991) significantly improved survival. Factors associated with a poor survival included older age, male sex, larger tumor size, lymph node metastasis, distant metastasis and squamous cell carcinoma.

Conclusion: A higher incidence and better survival for NSCLC patients were identified when compared with previously published, which may provide evidence on the incidence and survival of NSCLC in China.

Strengths and limitations of this study

- This is a population-based study which provides the incidence, demographic features and survival of non-small cell lung cancer (NSCLC) in mainland China.
- The present study also provides treatments and reported outcomes (overall and by clinical stage) of NSCLC in China, which is different from previous study.
- One of limitations is that some important features of NSCLC patients such as the smoking status, performance status and body weight were not available in the database.

INTRODUCTION

Lung cancer remains the most frequently diagnosed cancer worldwide and the leading cause of cancer related death in China.[1,2] Non–small-cell lung cancer (NSCLC) accounts for about 85% of lung cancer.[3] According to the Surveillance, Epidemiology and End Results (SEER) registry, the incidence of NSCLC is 42.6 per 100,000 people (49.7 per 100,000 people for men and 37.2 per 100,000 people for women; adjusted to the United States standard population, 2011).[4] In contrast to the decreasing trend of lung cancer incidence in developed countries, its incidence continues to increase in developing countries, especially in China.[5] For patients with early stage NSCLC, including stage I and II and a subset of stage III disease, the standard and potentially curative treatment is radical resection.[6] In a majority of patients, NSCLC is usually diagnosed at an advanced stage, and curative surgical resection is often impossible. Large population-based studies in western countries have indicated that the overall 1-year survival rate of NSCLC is 30%-46%.[7,8] SEER registry reports the 5-year survival rate of NSCLC is 19%.[4]

Although some population-based studies on the epidemiology and prognosis of NSCLC in western countries have been published, few studies have been conducted to investigate the characteristics of NSCLC in China. Available Chinese studies on NSCLC are mainly based on the national or local cancer registry of China, which analyzes lung cancer as a whole (including small cell lung cancer) and only reports the incidence and mortality, such as National Central Cancer Registry (NCCR) and Sihui Cancer Registry.[2,9] To our knowledge, population-based studies in Chinese population have never been conducted to estimate the NSCLC incidence and overall survival (OS) as well as the demographic features and prognostic factors of NSCLC.

In this population-based study, information was collected from Shanghai Health

Information Network. The epidemiological features, survival and prognostic factors of OS were investigated in patients with NSCLC.

MATERIALS AND METHODS

Ethics statement

This study was approved by the Ethics Committee of the School of Public Health, Fudan University, Shanghai, China. Written informed consent was not required, since a unique ID was allocated to each patient to replace identifiable personal information by the source database administrator before analysis. Written informed consent was not obtained from patients since it was specifically waived by the Institutional Review Board.

Data source

Data analyzed in this study were obtained from the Shanghai Health Information Network, which is organized and funded by the Shanghai Municipal Commission of Health and Family Planning (former Shanghai Municipal Bureau of Health).[10] This network automatically and dynamically integrates the data of Health Information Systems (HIS) from all the public healthcare facilities of Shanghai, which aims to facilitate a comprehensive utilization of health records by patients, health care professionals and health management organizations. Therefore, comprehensive healthcare data including demographic, diagnostic and treatment information for each patient are available from this network database. This network was initiated in 2011 and it has covered all the comprehensive hospitals and specialist hospitals qualified for cancer diagnosis in Shanghai metropolitan area in 2013. Only the network database administrator is authorized to extract information from the database as a

third party.

NSCLC cases were identified using the primary site coding system of the International Classification of Disease for Oncology 3rd Revision (ICD-10) from the World Health Organization and the pathological findings in the medical records. The diagnosis of NSCLC was confirmed by tissue diagnosis.

Age, sex, histological subtype, treatments, and tumor, node and metastasis (TNM) score were also collected from the database. Both clinical and pathological TNM information was accepted and coded to the TNM classification based on the TNM classification of malignant tumors (seventh edition).[11] We prioritized coded TNM stage where both coded stage and recorded stage were available.

The deadline of the follow-up was set on January 31, 2015. The endpoint mortality data were matched from municipal death registration system. The population demographics in this study were obtained from the Shanghai Statistical Year Book 2012 and 2013 of Shanghai Statistics Bureau.[12] Incidence was age-standardized (per 100,000 person-years) using the World Standard Population as proposed by Segi and modified by Doll et al.[13,14]

Case selection and inclusion criteria

From January 1, 2011 to December 31, 2013, 15,020 patients with NSCLC were identified in Shanghai. All of these patients were recruited to depict the epidemiological features of NSCLC at diagnosis.

Since this network was established in 2011, but covered all the qualified hospitals for cancer diagnosis in Shanghai in 2013, most of patients identified in the database were diagnosed in 2013 (n=12,996). Patients who were diagnosed in 2013 were selected to calculate the yearly incidence of NSCLC.

In survival analysis, patients were excluded if following conditions were present: (1) Patients had missing vital status before January 31, 2015; (2) patients had unspecified T, N, or M stage, while patients with stage specified as "X" or "cannot be accessed" were included; 3) patients had incomplete baseline information (gender, age or histological subtype). Survival time was calculated by subtracting the date of diagnosis from the date of death or the deadline of the study. Finally, 2,013 patients were included in survival analysis (Figure 1).

Statistical analysis

The incidence of NSCLC in 2013 was calculated by dividing the number of newly-diagnosed NSCLC patients identified from the inpatients database by the number of Shanghai permanent residents in the Shanghai Statistical Yearbook. Chi square test was employed to compare the baseline characteristics between patients with and without surgical resection. Kaplan-Meier method was used to evaluate the survival rate of NSCLC by cancer stage and treatment (surgical resection vs. no surgical resection). To investigate the factors affecting the OS, following clinicopathologic factors were included in the univariate Cox proportional hazard model analysis: age, gender, tumor size (T), regional lymph node status (N), metastasis (M), TNM stage, histology, surgery and chemotherapy. Since TNM stage is basically a combination of T, N and M scores, TNM stage was excluded from the multivariable Cox model while T, N and M scores were still included as potential prognostic factors. To further explore the influence of surgery on the survival, patients were stratified by TNM stage. The proportional hazard hypothesis was visually checked with log-log curves. A value of two-sided $P<0.05$ was considered statistically significant. Statistical analysis was conducted using SAS version 9.4 (SAS Institute,

Inc., Cary, North Carolina).

RESULTS

Incidence of NSCLC

There were 12,996 newly identified NCSLC cases in Shanghai in 2013, with a crude NSCLC incidence of 54.20 per 100,000 people. The crude incidence of NSCLC in males was higher than that in females (55.90 per 100,000 people and 52.39 per 100,000 people, respectively). The age-adjusted incidence was 39.05 per 100,000 people overall, 41.43 per 100,000 people in males and 37.13 per 100,000 people in females based on the World Standard Population (Table 1).

Table 1. Incidence of NSCLC by sex and age

Age Group	Gender	No	Crude Rate ($1/10^5$)	ASR* ($1/10^5$)
<55	Both	2958	15.93	15.94
	Male	1307	13.46	13.35
	Female	1651	18.64	18.85
>=55, <70	Both	6904	187.91	204.23
	Male	3762	200.77	219.86
	Female	3142	174.54	187.93
>=70	Both	3134	180.60	199.59
	Male	1834	239.97	250.99
	Female	1300	133.87	156.96
Overall	Both	12996	54.20	39.05
	Male	6903	55.90	41.43
	Female	6093	52.39	37.13

*ASR: age-standardized rates by world standard population

Demographic and tumor characteristics and treatments

From January 1, 2011 to December 31, 2013, 15,020 patients with NSCLC were identified. Approximately 53.3% of patients were men. The mean age at diagnosis

was 61.9±10.9 years (range: 15-98 years), and half of the patients were between 55 year-old and 69 year-old upon diagnosis. Tumor stages were available in 34% of patients (n=5099). Among patients with known tumor stage, 17.9% had stage I NSCLC, 5.7% stage II NSCLC, 12.9% stage IIIa NSCLC, and 63.5% stage IIIb/IV NSCLC. In addition, adenocarcinoma was found in 70.6 % of patients, while 27.7 % were diagnosed with squamous cell carcinoma among patients with known histological type. Patients undergoing surgical resection were younger, had lower TNM score on tumor size, lymph node metastasis and distant metastasis (Table 2).

Among 15,020 patients, 33.7% underwent a surgical resection (n=5069), and this proportion ranged from 94.0% in stage I NSCLC patients to 20.8% in stage IIIb/IV NSCLC patients. The proportion of patients receiving chemotherapy was 52.5%, ranged from 55.5% in stage I NSCLC patients to 82.9% in stage IIIa NSCLC patients (Table 3).

Table 2. Demographics, tumor characteristics and treatments of newly identified NSCLC cases in Shanghai between 2011 and 2013 (n = 15020)

Characteristics		All subjects (n = 15020)	Surgical resection(n = 5069)	No surgical resection(n = 9951)	P value
Sex	Male	8002 (53.3%)	2457 (48.5%)	5545 (55.7%)	<.0001
	Female	7018 (46.7%)	2612 (51.5%)	4406 (44.3%)	
Age groups	<55	3396 (22.6%)	1213 (23.9%)	2183 (21.9%)	<.0001
	55-70	7935 (52.8%)	2831 (55.8%)	5104 (51.3%)	
	>=70	3689 (24.6%)	1025 (20.2%)	2664 (26.8%)	
TNM-Tumor*	T1	570 (3.8%)	393 (19.6%)	177 (6.6%)	<.0001
	T2	1630 (10.9%)	1052 (52.4%)	578 (21.4%)	
	T3	646 (4.3%)	254 (12.6%)	392 (14.5%)	
	T4	1771 (11.8%)	286 (14.2%)	1485 (55.1%)	
	Tx	88 (0.6%)	24 (1.2%)	64 (2.4%)	
	Unspecified/unknown	10315 (68.7%)	3060 (-)	7255 (-)	
TNM-Node	N0	1395 (9.3%)	1141 (56.7%)	254 (9.4%)	<.0001
	N1	563 (3.7%)	213 (10.6%)	350 (13.0%)	
	N2	1590 (10.6%)	465 (23.1%)	1125 (41.7%)	
	N3	1020 (6.8%)	151 (7.5%)	869 (32.2%)	
	Nx	144 (1.0%)	44 (2.2%)	100 (3.7%)	
	Unspecified/unknown	10308 (68.6%)	3055 (-)	7253 (-)	
TNM-Metastasis	M0	2390 (15.9%)	1672 (75.4%)	718 (24.4%)	<.0001
	M1	2671 (17.8%)	523 (23.6%)	2148 (73.1%)	
	Mx	95 (0.6%)	22 (1.0%)	73 (2.5%)	
	Unspecified/unknown	9864 (65.7%)	2852 (-)	7012 (-)	
Stage	Ia/Ib	912 (6.1%)	857 (39.5%)	55 (1.9%)	<.0001
	IIa/IIb	292 (1.9%)	253 (11.6%)	39 (1.3%)	
	IIIa	659 (4.4%)	389 (17.9%)	270 (9.2%)	
	IIIb/IV	3236 (21.5%)	673 (31.0%)	2563 (87.6%)	
	Unspecified/unknown	9921 (66.1%)	2897 (-)	7024 (-)	
Histology	Adenocarcinoma	2976 (19.8%)	1408 (70.0%)	1568 (71.1%)	<.0001
	Squamous cell carcinoma	1168 (7.8%)	545 (27.1%)	623 (28.3%)	
	Other(adeno-squamous carcinoma and large cell carcinoma)	73 (0.4%)	59 (2.9%)	14 (0.7%)	
	Unspecified/unknown	10803 (70.3 %)	3057 (-)	7746 (-)	
Chemotherapy	Yes	7134 (47.5%)	2182 (43.0%)	4952 (49.8%)	<.0001
	No	7886 (52.5%)	2887 (57.0%)	4999 (50.2%)	

*TNM: tumor, node and metastasis score.

Table 3. Proportions and rates of surgical resection and chemotherapy, grouping by stage

Stage	n	Surgical resection (n = 5069)		No surgical resection (n = 9951)		Surgical resection	Chemotherapy
		Chemotherapy	Without chemotherapy	Chemotherapy	No chemotherapy		
Ia/Ib	912 (100%)	465 (51.0%)	392 (43.0%)	41 (4.5%)	14 (1.5%)	94.0%	55.5%
IIa/IIb	292 (100%)	184 (63.0%)	69 (23.6%)	27 (9.2%)	12 (4.1%)	86.6%	72.3%
IIIa	659 (100%)	317 (48.1%)	72 (10.9%)	229 (34.7%)	41 (6.2%)	59.0%	82.9%
IIIb/IV	3236 (100%)	508 (15.7%)	165 (5.1%)	1980 (61.2%)	583 (18.0%)	20.8%	76.9%
Unknown	9921 (100%)	1413 (14.2%)	1484 (15.0%)	2722 (27.4%)	4302 (43.4%)	29.2%	41.7%
Total	15020	2887	2182	4999	4952	33.7%	52.5%

Overall survival and prognostic factors

In our study, 2,013 NSCLC patients had complete information from the date of diagnosis until death or January 31, 2015, of whom 1009 patients (50.1%) died during this period.

The median OS is shown in Table 4 and plots of the survival rate was depicted by stage and surgical resection, independently (Figure 2 and 3). The median duration of follow-up for all NSCLC patients was 21.5 months (95% confidence interval, [CI]: 21.2-21.8 months). The median survival time for all NSCLC patients was 22.7 months (95%CI: 21.8-24.2 months), and the 1-year survival rate was 71.8% (95%CI: 69.8%-73.8%). The median survival time was unavailable for stage I and stage II NSCLC patients. For stage IIIa and stage IIIb/IV NSCLC patients, the median survival time was 24.3 months (95% CI: 21.4-26.2 months) and 16.0 months (95% CI: 14.8-16.7 months). The 1-year survival rate was 96.5% (95% CI: 94.0%-98.6%) in stage I NSCLC patients, 89.1% (95%CI: 83.3%-94.9%) in stage II NSCLC patients, 78.8% (95%CI: 74.1%-83.5%) in stage IIIa NSCLC patients and 58.9% (95%CI: 56.1%-61.7%) in stage IIIb/ IV NSCLC patients (Table 4, Figure 2).

Table 4. Median overall survival and prognostic factors of newly diagnosed NSCLC cases in Shanghai between 2011 and 2013 (n = 2013)

Characteristics		Median (95%CI)*	Crude HR (95%CI) **	Adjusted HR (95%CI)***	*P* value***
Surgery	Yes	34.4 (29.5-38.1)	0.276 (0.240 -0.318)	0.607 (0.511-0.722)	0.000
	No	15.4 (14.1-16.5)	1.00	1.00	-
Chemotherapy	Yes	22.2 (21.2-23.2)	1.145 (0.974 -1.347)	0.838 (0.709-0.991)	0.039
	No	26.9 (23.2-32.1)	1.00	1.00	-
Gender	Male	19.2 (17.7-20.4)	1.721 (1.508 -1.964)	1.751 (1.521-2.015)	0.000
	Female	26.2 (25.7-29.1)	1.00	1.00	-
Age (yrs)	<55	24.9 (22.7-26.2)	1.00	1.00	-
	55-70	25.1 (23.7-26.0)	1.048 (0.885 -1.241)	1.111 (0.936-1.318)	0.228
	>=70	16.1 (14.6-18.3)	1.850 (1.539 -2.224)	1.727 (1.426-2.091)	0.000
T	T1	30.3 (28.7-48.2)	1.00	1.00	-
	T2	27.5 (25.9-33.9)	1.437 (1.122 -1.841)	1.214 (0.945-1.561)	0.129
	T3	18.4 (15.5-21.5)	2.847 (2.186 -3.707)	1.461 (1.111-1.920)	0.007
	T4	16.3 (14.5-16.9)	3.545 (2.807 -4.477)	1.385 (1.083-1.772)	0.009
	Tx	10.3 (6.8-16.3)	3.715 (2.400 -5.751)	1.571 (0.872-2.830)	0.133
N	N0	-	1.00	1.00	-
	N1	21.6 (19.0-24.2)	3.890 (3.022 -5.007)	1.949 (1.483-2.563)	0.000
	N2	17.1 (15.8-19.0)	5.221 (4.240 -6.428)	2.845 (2.263-3.576)	0.000
	N3	13.6 (11.8-15.4)	6.927 (5.567 -8.620)	3.527 (2.762-4.504)	0.000
	Nx	12.0 (8.4-18.3)	5.898 (4.073 -8.541)	2.482 (1.489-4.139)	0.000
M	0	30.3 (27.6-35.1)	1.00	1.00	-
	1	15.6 (14.3-16.7)	3.000 (2.627 -3.427)	1.722 (1.456-2.037)	0.000
	x	-	3.223 (1.920 -5.411)	1.458 (0.859-2.476)	0.162
Stage	Ia/Ib	-	1.00	-	-
	IIa/IIb	-	3.578 (2.144 -5.971)	-	-
	IIIa	24.3 (21.4-26.2)	8.094 (5.508 -11.892)	-	-
	IIIb/IV	16.0 (14.8-16.7)	14.594 (10.247 -20.785)	-	-
Histological subtype	Adenocarcinoma	24.4 (23.1-25.7)	1.00	1.00	-
	Squamous cell carcinoma	19.0 (16.8-21.1)	1.370 (1.195 -1.571)	1.172 (1.003-1.369)	0.045
	Other	25.3 (22.7-38.6)	0.768 (0.468 -1.262)	1.058 (0.639-1.752)	0.827

* CI, confidence interval; Median and 95% CI were estimated using Kaplan-Meier method.

**HR, Hazard ratio; Crude hazard ratio and 95% CI were estimated using univariate Cox regression model.

*** Adjusted hazard ratio, 95% CI and *P* value were estimated using multiple Cox regression model adjusted by surgical resection, chemotherapy, sex, age group, TNM score and histology.

Patients underwent surgical resection had better survival than those without surgical intervention, with median survival time of 34.4 months (95% CI: 29.5-38.1 months) vs. 15.4 months (95% CI: 14.1-16.5 months) and 1-year survival rate of 87.8% (95%CI: 85.7%-89.9%) vs. 57.9% (95%CI: 55.0%-60.8%) (Table 4).

Univariate analysis showed patients who were female or younger, had smaller tumor size, no lymph node metastasis, no distal metastasis, lower stage, received surgical resection or had adenocarcinoma showed a longer survival time than their counterparts, while chemotherapy failed to benefit patients on the survival. However, after adjustment for the demographic factors and tumor characteristics in multivariate analysis, patients receiving chemotherapy showed a significantly longer survival time (HR=0.838, 95% CI: 0.709-0.991). Patients receiving surgical resection also had improved survival (HR=0.607, 95% CI: 0.511-0.722) as compared to those without surgical intervention. In this multivariable Cox proportional hazard model, factors associated with a poor survival included male sex (HR=1.751, 95% CI: 1.521-2.015), older age at diagnosis (age>=70 years vs. age <55 years: HR=1.727, 95% CI: 1.426-2.091), larger tumor size (T4 vs. T1: HR=1.385, 95% CI: 1.083-1.772), lymph node metastasis (N3 vs. N0: HR=3.527, 95% CI: 2.762-4.504), distant metastasis (HR=1.722, 95% CI: 1.456-2.037) and squamous cell carcinoma (HR=1.172, 95% CI: 1.003-1.369) (Table 4).

In order to further evaluate the prognostic role of surgery in NSCLC patients, additional multivariable analysis was performed according to TNM stages. T, N and M scores were excluded since TNM stage was a combination of them. The survival benefit of surgery was observed in stage IIIa NSCLC patients (HR=0.513, 95% CI: 0.352-0.748) and stage IIIb/IV NSCLC patients (HR=0.646, 95% CI: 0.536-0.779) (Table 5, Figure 3).

Table 5. Multivariate Hazard Ratio of overall survival according to surgical resection by stage (n = 2013)

Stage	n = 2013	Surgical resection vs. no surgical resection (ref.)	
		Adjusted HR (95% CI)*	*P* value*
Ia/Ib	451	0.360 (0.104-1.237)	0.105
IIa/IIb	110	0.723 (0.205-2.542)	0.613
IIIa	288	0.513 (0.352-0.748)	0.001
IIIb/IV	1164	0.646 (0.536-0.779)	0.000

*HR, Hazard ratio; Adjusted hazard ratio, 95% CI and *P* value were estimated using multiple Cox regression model adjusted by sex, age group and histology type.

DISCUSSION

Incidence

To our knowledge, this was the first population-based study to describe the epidemiological characteristics of NSCLC in mainland China. Our results showed the crude incidence of NSCLC in 2013 was 54.20 per 100,000 people (55.90 per 100,000 people for men and 52.39 per 100,000 people for women) with an age-adjusted incidence of 39.05 per 100,000 people (41.43 per 100,000 people for men and 37.13 per 100,000 people for women). Compared with the SEER registry, a population-based national cancer registries covering approximately 28% of the United States (US) population and 50% of Asians in the US, the crude incidence in our study was higher than that of all races in the SEER registry (42.6 per 100,000 people overall, 49.7 per 100,000 people for men and 37.2 per 100,000 people for women, adjusted to the United States standard population in 2011) and that of Chinese group (52.0 per 100,000 people for men and 29.9 per 100,000 people for women, 2004-2008).[4,15] One possible explanation to the higher crude NSCLC incidence in our study could be the aging of Chinese population, as an older age has been identified as an independent risk factor for NSCLC.[6] Population aging is especially obvious in Shanghai, where 27% of its population was older than 60 years in 2013,[16] while merely 16.5% of US

population was older than 60 years in 2000.[17]

Most available population-based studies investigate the lung cancer as a whole, including both NSCLC and small cell lung cancer. GLOBOCAN database, which is from population-based cancer registries worldwide and referenced by World Health Organization, reports an incidence of lung cancer of 50.4 per 100,000 people for men and 19.2 for women in East Asia in 2012. Sihui, a city in south China, is reported to has the yearly incidence of lung cancer of 37.98 per 100,000 people overall, 60.26 per 100,000 people for men and 20.29 people for women between 2007 and 2011, based on the local cancer registry[9]. Both of the incidences are adjusted by the Segi's World Standard Population. Estimating the NSCLC incidence as 85% of lung cancer incidence, the NSCLC incidence in our study was slightly higher than these above, which was largely due to the higher incidence in women in our study. According to the National Central Cancer Registry 2010 in China, the incidence of lung cancer in China was 36.39 per 100,000 people.[18] When compared with the estimated NSCLC incidences, there was a 8% increase in NSCLC incidence per year from 2010 to 2013. This was consistent with previous findings that the incidence of lung cancer is increasing in China.[1,19-20] For example, in the Sihui study, a 6% increase was reported in the annual incidence of lung cancer for women and 11% for men from 2005 to 2010.[9] Except for the effect of population aging, several other factors may contribute to this higher and increased NSCLC incidence in our study. First, the smoking prevalence has dramatically increased in the past 2 decades in China. Although cigarette smoking rate has peaked and decreased in the United States and several other areas in recent years, the prevalence of smoking in China remains at a high level and China has become one of the countries with the highest smoking prevalence in the world.[19,21] According to the 2010 report of China Global Adults Smoking Survey

(GATS), 53% of men aged 15 years and above are current smokers.[22] Considering that smoking is the main risk factor of NSCLC, this high smoking prevalence in the past three decades in China is closely related to the increasing prevalence of lung cancer.[6] The relationship between smoking and lung cancer is also confirmed in the study of Gomez et al. Gomez et al found a significant decline in the incidence of squamous cell lung cancer among foreign-born Chinese Americans from 1900 to 2004, accompanied by a temporal decline in current smoking prevalence among them, while the incidence was stable for adenocarcinoma, which is less closely associated with tobacco smoke than squamous cell lung cancer.[23] Another factor related to this higher incidence is the higher diagnosis rate due to the improved oncology services in Shanghai as one of the most developed cities in China.[24]

Of note, a higher ratio of NSCLC incidence was observed in women as compared to men in this study (0.90), while this ratio was 0.75 and 0.40 in the SEER study and GLOBLECAN report, respectively.[1,4] The higher risk for lung cancer in Chinese women after considering smoking status was also found by Boffetta et al and Epplein et al.[25,26] The reasons for the higher incidence of lung cancer among Chinese women are unclear, and might be partly ascribed to the household air pollution due to cooking fumes and unventilated coal-fueled heating stoves.[27-29] Besides, considering the high overall smoking prevalence in China, the second-hand-smoking may also be a critical risk factor of NSCLC in non-smokers, typically women. A nationwide cross-sectional survey conducted in 15,540 Chinese adults showed that, in 2000–2001, more than 49.2% of adult female non-smokers reported exposure to tobacco smoke, while this proportion was only 35% according to an international data from 192 countries in 2004.[30,31] This suggests an additional risk for lung cancer in Chinese women.

Survival

A better survival (overall and stage-specified) was observed in this study as compared to that in previously published population-based studies on non-Asian ethnicity, though different population-based lung cancer databases showed different outcomes. According to the databases from Australia, Canada, Denmark, Norway, Sweden and UK, the 1-year OS of NSCLC in 2004-2007 ranged from 30% to 46%, with stage-specified 1-year survival rate of 71.1%-86.2% for stage I NSCLC, 58.6%-79.0% for stage II NSCLC, 34.4%-37.1% for stage III NSCLC and 15.5%-25.9% for stage IV NSCLC.[7] Lower survival rate was also observed in SEER registry (overall 1-year survival rate of 46.6% in 2011; 1-year survival rate of 15.9% for stage IV NSCLC in 1998-2003) and the study of Rasco et al.[4,32,33] However, Asian population shows improved survival. Lin et al reported the 2-year survival rate was 80.0%-96.2%, 64.4%-80.2% and 57.5%-67.4% for stage I, stage II and stage IIIa NSCLC patients, respectively, among 30,069 Taiwanese patients between 2004 and 2007.[34] In a study on 4622 Korean patients between 1998 and 2005, the median OS for stage I, stage II, stage III and stage IV NSCLC was 100, 41, 14 and 7 months, respectively.[35] No population-based study has been conducted to investigate the characteristics of NSCLC in mainland China.

The better survival outcome observed in this study may be related to several factors. First, Asian ethnicity has been recognized as an independent favorable prognostic factor for OS among NSCLC patients.[36,37] Asian NSCLC patients showed distinct response to cytotoxic chemotherapy when compared with white patients. For example, Gandara et al. reported a 3-month increase in the median OS of Japanese patients than in white patients receiving chemotherapy with the same paclitaxel plus

carboplatin regimen, while this regimen is also one of the routine regimens for chemotherapy for advanced NSCLC in China.[38] At the same time, epidermal growth factor receptor (EGFR) mutation confers survival benefit independent of treatment in NSCLC,[39,40] while the East Asian population has the highest incidence of EGFR mutation.[41,42] Meanwhile, advances in treatment in recent years, such as the introduction of target agents and adjuvant chemotherapy after complete resection, may improve the survival of NSCLC patients.

Similar to previous studies, our results showed that female gender, younger age, smaller tumor size, no lymph node metastasis and no distant metastasis were related to a better survival. The evaluation of impact of surgery on the survival of patients at different TNM stage showed that patients with stage IIIa or stage IIIb/IV NSCLC who underwent surgical resection had improved survival. This suggests surgical intervention may improve the survival, even for advanced NSCLC patients, though the details of therapeutic modality are still needed to be investigated. Currently there are controversies on the role of surgery in stage IIIa NSCLC. According to the Chinese guidelines for lung cancer, surgical resection is the current standard treatment for patients with stage I to stage IIIA NSCLC; some stage IV NSCLC patients with single metastasis are also suitable for surgery.[43] Goldstraw et al proposed that "current evidence supports an expansion in surgery as part of multimodality management of patients with N2 disease, and greater uptake in patients who are willing to accept higher risks", which may be ascribed to the improvements in diagnostic imaging and endoscopic techniques.[44] In multivariate analysis, chemotherapy was also shown as a protective prognostic factor, suggesting that confounding factor exists in the univariate analysis.

Our study had several strengths. First, this was the first study, to our knowledge,

to evaluate the incidence, survival and prognostic factors of NSCLC based on a large population in mainland China. Existing Chinese studies, mainly the national and local annual cancer registry reports, investigate lung cancer as a whole and only report incidence and mortality, because limited information is offered by the cancer registration report cards used by the registry system. By contrast, based on the HIS system within Shanghai Health Information Network, not only NSCLC cases can be specifically identified, clinicopathological information and treatments are also available. At the same time, our study offered a higher but comparable incidence to that of existing cancer registration system, with consistent constitutions of gender and TNM staging in NSCLC cases with other studies, which confirms the reliability of our findings. Furthermore, this study was based on data through 2013, whereas the most recently NSCLC population-based studies from other Asian countries or districts recruited data of 2010.[28,34] Last, our study reported a higher incidence and better survival of NSCLC as compared to previous studies, which may provide a fresh and meaningful perspective for the evaluation of NSCLC diagnosis and treatment, considering the ethnic difference, smoking prevalence and treatment improvement.

However, this study also had several limitations. First, as a retrospective study, some important features of NSCLC patients such as performance status, body weight and details of treatment were not available in the database. Specifically, patients' smoking status was unavailable. As a known prognostic factor, its absence may lead to residual confounding.[45] In addition, since the network database is newly established, though important variables such as diagnosis and demographic information are available, the TNM classification and histological subtype were still unavailable in several patients. Therefore, it was difficult to calculate the incidence stratified by or adjusted for these variables. Missing in more detailed records of diagnosis and

treatment also prohibited us from further analysis. However, selection bias can be considerably diminished in the survival analysis as only patients with known potential prognostic factors were included. At last, the duration of follow-up time was short (median: 21.5 months) because the network database is newly established. Thus, long term follow up is required to determine the survival of NSCLC patients, especially for those with early stage NSCLC.

CONCLUSIONS

The present study shows a higher incidence and a better survival for Chinese NSCLC patients. High smoking prevalence and the consequent high environment tobacco exposure may be related to the higher NSCLC incidence both overall and in women. In addition to female gender and younger age, surgical resection is found as a protective prognosis factor for NSCLC at stage IIIa and above.

Acknowledgements

The authors thank the staffs of the database and death registration system for their efforts in data collection and preprocessing. The authors also acknowledge the English language editing work of Shanghai LancetSCI Biotechnology Co., Ltd

Contributors

NZ GQ and ZS conceived and designed the study; HF ZS YX ZX WC and HX conducted this study; HF analyzed the data. HF YX and ZS drafted this paper; NZ and GQ revised this paper.

Funding

This study was supported by the National Natural Science Foundation of China (No 81273187), National Science and Technology Major Project of the People's Republic of China (2012ZX09303-013-014) and Key Lab of Health Technology Assessment, Ministry of Health of the People's Republic of China.

Competing interest

None declared.

Ethics approval

The present study complied with the Declaration of Helsinki and was approved by the Ethics Committee of the School of Public Health, Fudan University, Shanghai, China.

Data sharing statement

Additional data were held by the corresponding author.

REFERENCES

1. Torre LA, Bray F, Siegel RL, et al. Global Cancer Statistics, 2012. CA Cancer J Clin 2015;65:87-108
2. Chen W, Zheng R, Zeng H, Zhang S, He J. Annual Report On Status of Cancer in China, 2011. Chin J Cancer Res 2015;27:2-12
3. Oser MG, Niederst MJ, Sequist LV, Engelman JA. Transformation From Non-Small-Cell Lung Cancer to Small-Cell Lung Cancer: Molecular Drivers and Cells of Origin. The Lancet Oncology 2015;16:e165-e172
4. Surveillance, Epidemiology, and End Results (SEER) Program. May 2015. Available at: http://seer.cancer.gov/

5. Youlden DR, Cramb SM, Baade PD. The International Epidemiology of Lung Cancer: Geographical Distribution and Secular Trends. J Thorac Oncol 2008;3:819-831
6. National Comprehensive Cancer Network: NCCN Clinical Practice Guidelines in Oncology: Non-Small Cell Lung Cancer. V. 2. Dec 2014. Available at: www.nccn.org/professionals/physician_gls/PDF/nscl.pdf
7. Walters S, Maringe C, Coleman MP, et al. Lung Cancer Survival and Stage at Diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: A Population-Based Study, 2004-2007. Thorax 2013;68:551-564
8. Coleman MP, Forman D, Bryant H, et al. Cancer Survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (The International Cancer Benchmarking Partnership): An Analysis of Population-Based Cancer Registry Data. Lancet 2011;377:127-138
9. Du JL, Lin X, Zhang LF, et al. Secular Trend Analysis of Lung Cancer Incidence in Sihui City, China Between 1987 and 2011. Chin J Cancer 2015;34:33
10. Shanghai Health Information Network. May 2015. Available at: http://www.shhs.org.cn
11. Goldstraw P, Crowley J, Chansky K, et al. The IASLC Lung Cancer Staging Project: Proposals for the Revision of the TNM Stage Groupings in the Forthcoming (Seventh) Edition of the TNM Classification of Malignant Tumours. J Thorac Oncol 2007;2:706-714
12. Shanghai Statistical Year Book (2013). Dec 2014. Available at:http://www.stats-sh.gov.cn/data/toTjnj.xhtml?y=2013e
13. Segi M. Cancer Mortality for Selected Sites in 24 Countries (1950-57). Japan: Department of Public Health, Tohoku University of Medicine; 1960
14. Doll R, Cook P. Summarizing Indices for Comparison of Cancer Incidence Data. Int J Cancer 1967;2:269-279
15 Gomez SL, Noone AM, Lichtensztajn DY, et al. Cancer Incidence Trends Among Asian American Populations in the United States, 1990-2008. J Natl Cancer Inst 2013;105:1096-1110
16. Gerontological Society of Shanghai. Dec 2014. Available at: http://www.shanghaigss.org.cn/
17. Standard Populations - 19 Age Groups, Surveillance, Epidemiology, and End Results Program. Dec 2014. Available at: http://seer.cancer.gov/stdpopulations/stdpop.19ages.html
18. Chen W, Zheng R, Zeng H, Zhang S. Epidemiology of Lung Cancer in China. Thorac Cancer 2015;6:209-215

19. Yang L, Parkin DM, Li L, Chen Y. Time Trends in Cancer Mortality in China: 1987-1999. Int J Cancer 2003;106:771-783

20. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship. Mayo Clin Proc 2008;83:584-594

21. Zhang H, Cai B. The Impact of Tobacco On Lung Health in China. Respirology 2003;8:17-21

22. Zhang J, Ou JX, Bai CX. Tobacco Smoking in China: Prevalence, Disease Burden, Challenges and Future Strategies. Respirology 2011;16:1165-1172

23. Gomez SL, Yang J, Lin SW, et al. Incidence Trends of Lung Cancer by Immigration Status among Chinese Americans. Cancer Epidemiol Biomarkers Prev 2015;24:1157-1164

24. Yang LL, Zhang XC, Yang XN, et al. Lung Cancer Treatment Disparities in China: A Question in Need of an Answer. Oncologist 2014;19:1084-1090

25. Boffetta P, Parkin D. Cancer in Developing Countries. Ca : A Cancer Journal for Clinicians 1994;2:81

26. Epplein M, Schwartz SM, Potter JD, Weiss NS. Smoking-Adjusted Lung Cancer Incidence Among Asian-Americans (United States). Cancer Causes Control 2005;16:1085-1090

27. Cancer IAFR. Personal Habits and Indoor Combustions. In, IARC Monnograohs on the Evaluation of Carcinogenic Risks to Humans. Lyon, France: IARC Press; 2012

28. Wang BY, Huang JY, Cheng CY, et al. Lung Cancer and Prognosis in Taiwan: A Population-Based Cancer Registry. J Thorac Oncol 2013;8:1128-1135

29. Feng G, Jiang Y, Zhao L, et al. [Degree of Exposure to Secondhand Smoking and Related Knowledge, Attitude Among Adults in Urban China][in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2014;35:998-1001

30. Gu D, Wu X, Reynolds K, et al. Cigarette Smoking and Exposure to Environmental Tobacco Smoke in China: The International Collaborative Study of Cardiovascular Disease in Asia. Am J Public Health 2004;94:1972-1976

31. Oberg M, Jaakkola MS, Woodward A, Peruga A, Pruss-Ustun A. Worldwide Burden of Disease From Exposure to Second-Hand Smoke: A Retrospective Analysis of Data From 192 Countries. Lancet 2011;377:139-146

32. Cetin K, Ettinger DS, Hei YJ, O'Malley CD. Survival by Histologic Subtype in Stage IV Nonsmall Cell Lung Cancer Based On Data From the Surveillance, Epidemiology and End Results Program.

Clin Epidemiol 2011;3:139-148

33. Rasco DW, Yan J, Xie Y, Dowell JE, Gerber DE. Looking Beyond Surveillance, Epidemiology, and End Results: Patterns of Chemotherapy Administration for Advanced Non-Small Cell Lung Cancer in a Contemporary, Diverse Population. J Thorac Oncol 2010;5:1529-1535
34. Lin ZZ, Shau WY, Shao YY, et al. Survival Following Surgery with Or without Adjuvant Chemotherapy for Stage I-IIIA Non-Small Cell Lung Cancer: An East Asian Population-Based Study. Oncologist 2012;17:1294-1302
35. Ahn MJ, Lee J, Park YH, et al. Korean Ethnicity as Compared with White Ethnicity is an Independent Favorable Prognostic Factor for Overall Survival in Non-Small Cell Lung Cancer Before and After the Oral Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor Era. J Thorac Oncol 2010;5:1185-1196
36. Ou SH, Ziogas A, Zell JA. Asian Ethnicity is a Favorable Prognostic Factor for Overall Survival in Non-Small Cell Lung Cancer (NSCLC) and is Independent of Smoking Status. J Thorac Oncol 2009;4:1083-1093
37. Tannenbaum SL, Koru-Sengul T, Zhao W, Miao F, Byrne MM. Survival Disparities in Non-Small Cell Lung Cancer by Race, Ethnicity, and Socioeconomic Status. Cancer J 2014;20:237-245
38. Gandara DR, Kawaguchi T, Crowley J, et al. Japanese-US Common-Arm Analysis of Paclitaxel Plus Carboplatin in Advanced Non-Small-Cell Lung Cancer: A Model for Assessing Population-Related Pharmacogenomics. J Clin Oncol 2009;27:3540-3546
39. Bell DW, Lynch TJ, Haserlat SM, et al. Epidermal Growth Factor Receptor Mutations and Gene Amplification in Non-Small-Cell Lung Cancer: Molecular Analysis of the IDEAL/INTACT Gefitinib Trials. J Clin Oncol 2005;23:8081-8092
40. Eberhard DA, Johnson BE, Amler LC, et al. Mutations in the Epidermal Growth Factor Receptor and in KRAS are Predictive and Prognostic Indicators in Patients with Non-Small-Cell Lung Cancer Treated with Chemotherapy Alone and in Combination with Erlotinib. J Clin Oncol 2005;23:5900-5909
41. Lynch TJ, Bell DW, Sordella R, et al. Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-Small-Cell Lung Cancer to Gefitinib. N Engl J Med 2004;350:2129-2139
42. Calvo E, Baselga J. Ethnic Differences in Response to Epidermal Growth Factor Receptor Tyrosine

Kinase Inhibitors. J Clin Oncol 2006;24:2158-2163

43. XY Z, YK S, JM Y. Standards for the Diagnosis and Treatment of Primary Lung Cancer (2015 Version) in China. Chinese Journal of Oncology 2015;37

44. Goldstraw P, Ball D, Jett JR, et al. Non-Small-Cell Lung Cancer. LANCET 2011;378:1727-1740

45. Ebbert JO, Yang P, Vachon CM, et al. Lung Cancer Risk Reduction After Smoking Cessation: Observations From a Prospective Cohort of Women. J Clin Oncol 2003;21:921-926

Figure legends

Figure 1. Flow chart of study population and analysis groups.

Figure 2. The overall survival (OS) of NSCLC patients in Shanghai identified in 2011-2013 (n=2013). 1-year OS rate: whole population 71.8% (95% CI: 69.8%-73.8%), stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%), stage IIIb/IV 58.9% (95%CI: 56.1%-61.7%). The survival difference was significant ($p < 0.0001$).

Figure 3. The OS of NSCLC cases in Shanghai identified in 2011-2013 according to surgery by stage (n=2013). 1-year OS rate of stage I patients: with surgery 96.3% (95%CI: 94.5%-98.1%), without surgery 100.0%; stage II: with surgery 90.0% (95%CI: 84.1%-95.9%), without surgery 80.0% (95%CI: 55.2%-100.0%); stage IIIa: with surgery 84.3% (95%CI: 79.1%-89.5%), without surgery 68.9% (95%CI: 60.0%-77.8%); stage IIIb/IV: with surgery 73.1% (95% CI: 67.2%-79.0%), without surgery 55.7% (95% CI: 52.5%-58.9%). The survival benefit of surgery was observed among stage IIIa patients (adjusted HR=0.513, 95% CI: (0.352-0.748) and stage IIIb/IV patients (adjusted HR=0.646, 95% CI: 0.536-0.779).

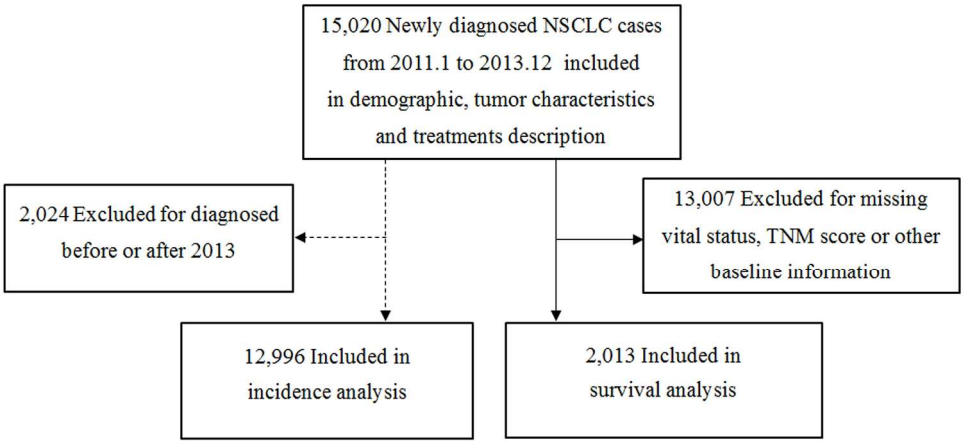

Figure 1. Flow chart of study population and analysis groups.

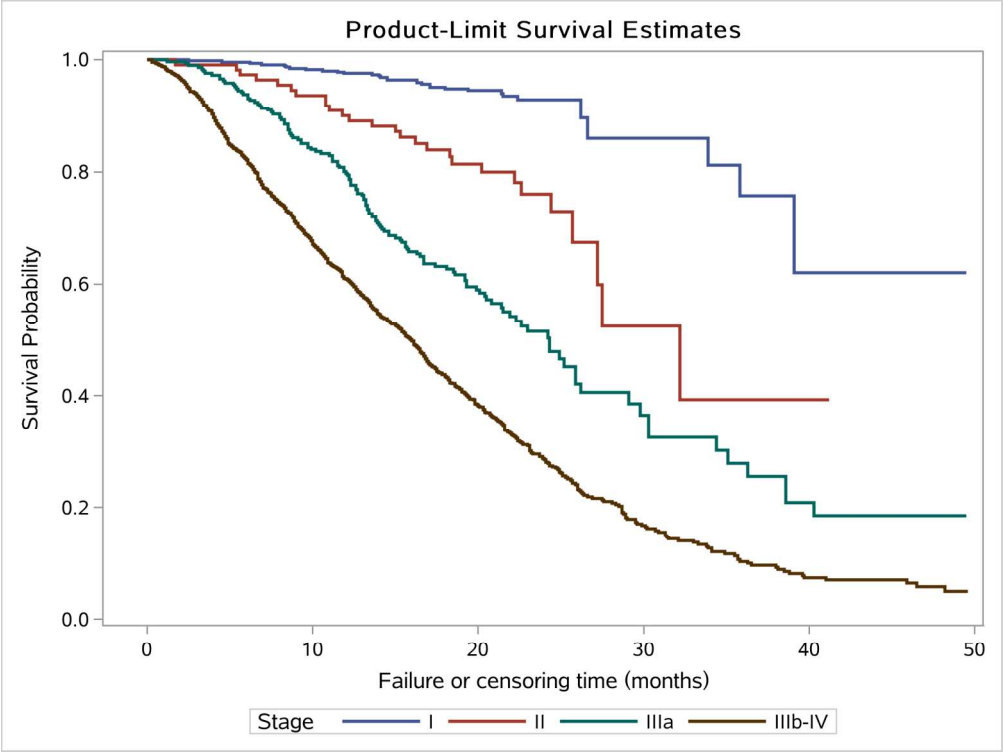

Figure 2. The overall survival (OS) of NSCLC patients in Shanghai identified in 2011-2013 (n=2013). 1-year OS rate: whole population 71.8% (95% CI: 69.8%-73.8%), stage I 96.5% (95% CI: 94.0%-98.6%), stage II 89.1% (95%CI: 83.3%-94.9%), stage IIIa 78.8% (95%CI: 74.1%-83.5%), stage IIIb/IV 58.9% (95%CI: 56.1%-61.7%). The survival difference was significant ($p < 0.0001$).
162x121mm (300 x 300 DPI)

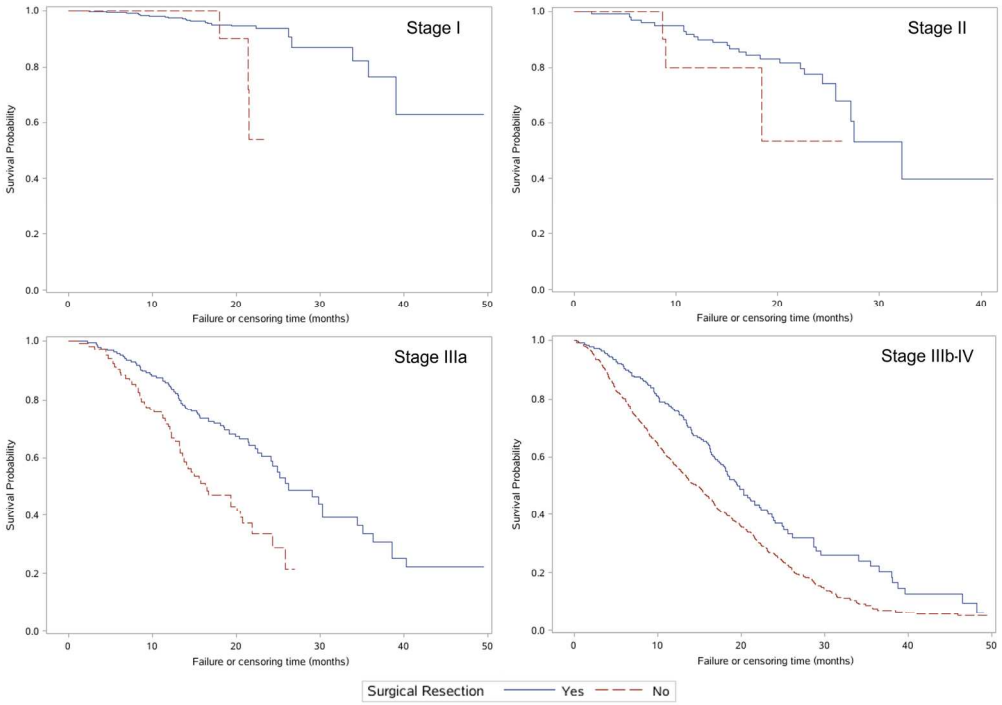

Figure 3. The OS of NSCLC cases in Shanghai identified in 2011-2013 according to surgery by stage (n=2013). 1-year OS rate of stage I patients: with surgery 96.3% (95%CI: 94.5%-98.1%), without surgery 100.0%; stage II: with surgery 90.0% (95%CI: 84.1%-95.9%), without surgery 80.0% (95%CI: 55.2%-100.0%); stage IIIa: with surgery 84.3% (95%CI: 79.1%-89.5%), without surgery 68.9% (95%CI: 60.0%-77.8%); stage IIIb/IV: with surgery 73.1% (95% CI: 67.2%-79.0%), without surgery 55.7% (95% CI: 52.5%-58.9%). The survival benefit of surgery was observed among stage IIIa patients (adjusted HR=0.513, 95% CI: (0.352-0.748) and stage IIIb/IV patients (adjusted HR=0.646, 95% CI: 0.536-0.779).
177x127mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses

Continued on next page

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.