BMJ Open

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]

To cite: Fan H, Shao Z-Y, Xiao Y-Y, *et al*. Incidence and survival of non-small cell lung cancer in Shanghai: a population-based cohort study. *BMJ Open* 2015;**5**: e009419. doi:10.1136/bmjopen-2015-009419

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2015-009419).

HF and ZY-S contributed equally.

Received 16 July 2015
Revised 2 September 2015
Accepted 26 October 2015

[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

Correspondence to
Professor Nai-Qing Zhao; nqzhao@fudan.edu.cn

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 patients with NSCLC 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 the 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% CI 21.8 to 24.2 months) with an overall 1-year survival rate of 71.8% (95% CI 69.8% to 73.8%). The 1-year survival rate was 96.5% (95% CI 94.0% to 98.6%) in patients with stage I NSCLC, 89.1% (95% CI 83.3% to 94.9%) in patients with stage II NSCLC, 78.8% (95% CI 74.1% to 83.5%) in patients with stage IIIa NSCLC and 58.9% (95% CI 56.1% to 61.7%) in patients with stage IIIb/IV NSCLC. Multivariate analysis showed surgical resection (HR=0.607, 95% CI 0.511 to 0.722) and chemotherapy (HR=0.838, 95% CI 0.709 to 0.991) significantly improved survival. Factors associated with poor survival included older age, male sex, larger tumour size, lymph node metastasis, distant metastasis and squamous cell carcinoma.

Conclusions: A higher incidence and better survival rates for patients with NSCLC were identified when compared with previously published studies, 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 that 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, not provided in previous studies.
- One limitation is that some important features of patients with NSCLC such as 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 for men and 37.2 per 100 000 for women; adjusted to the US 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] The SEER registry reports the 5-year survival rate of NSCLC as being 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, such as National Central Cancer Registry (NCCR) and Sihui Cancer Registry, which analyse lung cancer as a whole (including small cell lung cancer), and only report the incidence and mortality.[2 9] To the best of our knowledge, population-based studies in the 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, and survival and prognostic factors of OS were investigated in patients with NSCLC.

MATERIALS AND METHODS

Ethics statement

Written informed consent was not obtained from patients as it was not required, since a unique ID was allocated to each patient to replace identifiable personal information by the source database administrator before analysis, and also since it was specifically waived by the Institutional Review Board.

Data source

Data analysed in this study were obtained from the Shanghai Health Information Network, which is organised 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, aiming to facilitate a comprehensive utilisation of health records by patients, healthcare professionals and health management organisations. 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 the Shanghai metropolitan area in 2013. Only the network database administrator, as a third party, is authorised to extract information from the database.

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

Age, sex, histological subtype, treatments, and tumour, node and metastasis (TNM) score were also collected from the database. Clinical as well as pathological TNM information was accepted and coded to the TNM classification based on the TNM classification of malignant tumours (seventh edition).[11] We prioritised the coded TNM stage where coded stage and recorded stage were both available.

The deadline of the follow-up was set on 31 January 2015. The end point mortality data were matched using the 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-standardised (per 100 000 person-years) using the World Standard Population as proposed by Segi,[13] and modified by Doll and Cook.[14]

Case selection and inclusion criteria

From 1 January 2011 to 31 December 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 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 the following conditions were present: (1) patients had missing vital status before 31 January 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, 2013 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 patients with NSCLC identified from the inpatients database by the number of Shanghai permanent residents in the Shanghai Statistical Yearbook. The χ^2 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, the following clinicopathological factors were included in the univariate Cox proportional hazard model analysis: age, gender, tumour 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 V.9.4 (SAS Institute, Inc, Cary, North Carolina, USA).

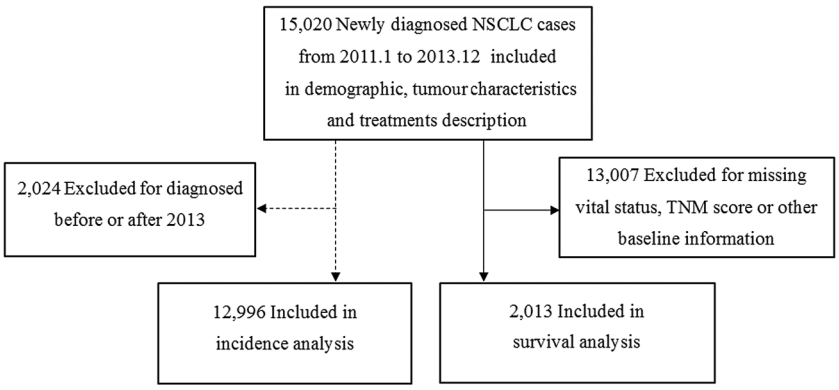

Figure 1 Flow chart of study population and analysis groups (NSCLC, non-small cell lung cancer; TNM, tumour, node and metastasis score).

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 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).

Demographic and tumour characteristics and treatments

From 1 January 2011 to 31 December 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 aged between 55 and 69 years on diagnosis. Tumour stages were available in 34% of patients (n=5099). Among patients with known tumour 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 tumour 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% in patients with stage I NSCLC to 20.8% in patients with stage IIIb/IV NSCLC. The proportion of patients receiving chemotherapy was 52.5%, and ranged from 55.5% in patients with stage I NSCLC to 82.9% in patients with stage IIIa NSCLC (table 3).

Table 1 Incidence of NSCLC by sex and age

Age group, years	Gender	N	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	12 996	54.20	39.05
	Male	6903	55.90	41.43
	Female	6093	52.39	37.13

*ASR, age-standardised rates by world standard population; NSCLC, non-small cell lung cancer.

OS and prognostic factors

In our study, 2013 patients with NSCLC had complete information from the date of diagnosis until death or 31 January 2015, of whom 1009 patients (50.1%) died during this period.

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

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

Univariate analysis showed that patients who were female or younger and had smaller tumour size, no

Table 2 Demographics, tumour characteristics and treatments of newly identified NSCLC cases in Shanghai between 2011 and 2013 (n=15 020)

Characteristics	All subjects (n=15 020)	Surgical resection (n=5069)	No surgical resection (n=9951)	p Value
Sex				
Male	8002 (53.3%)	2457 (48.5%)	5545 (55.7%)	<0.0001
Female	7018 (46.7%)	2612 (51.5%)	4406 (44.3%)	
Age groups, years				
<55	3396 (22.6%)	1213 (23.9%)	2183 (21.9%)	<0.0001
55–70	7935 (52.8%)	2831 (55.8%)	5104 (51.3%)	
≥70	3689 (24.6%)	1025 (20.2%)	2664 (26.8%)	
TNM tumour				
T1	570 (3.8%)	393 (19.6%)	177 (6.6%)	<0.0001
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	10 315 (68.7%)	3060 (–)	7255 (–)	
TNM node				
N0	1395 (9.3%)	1141 (56.7%)	254 (9.4%)	<0.0001
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	10 308 (68.6%)	3055 (–)	7253 (–)	
TNM metastasis				
M0	2390 (15.9%)	1672 (75.4%)	718 (24.4%)	<0.0001
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%)	<0.0001
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%)	<0.0001
Squamous cell carcinoma	1168 (7.8%)	545 (27.1%)	623 (28.3%)	
Other (adenosquamous carcinoma and large cell carcinoma)	73 (0.4%)	59 (2.9%)	14 (0.7%)	
Unspecified/unknown	10 803 (70.3%)	3057 (–)	7746 (–)	
Chemotherapy				
Yes	7134 (47.5%)	2182 (43.0%)	4952 (49.8%)	<0.0001
No	7886 (52.5%)	2887 (57.0%)	4999 (50.2%)	

NSCLC, non-small cell lung cancer; TNM, tumour, node and metastasis score.

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

In order to further evaluate the prognostic role of surgery in patients with NSCLC, additional multivariable analysis was performed according to TNM stages. T, N

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	15 020	2887	2182	4999	4952	33.7	52.5

and M scores were excluded since TNM stage was a combination of them. The survival benefit of surgery was observed in patients with stage IIIa NSCLC (HR=0.513, 95% CI 0.352 to 0.748) and stage IIIb/IV NSCLC (HR=0.646, 95% CI 0.536 to 0.779) (table 5 and figure 3).

DISCUSSION

Incidence

To the best of 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 for men and 52.39 per 100 000 for women) with an age-adjusted incidence of 39.05 per 100 000 people (41.43 per 100 000 for men and 37.13 per 100 000 for women). Compared with the SEER registry, a population-based national cancer registry covering approximately 28% of the population in the USA and 50% of Asians in the USA, 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 for men and 37.2 per 100 000 for women, adjusted to the US standard population in 2011) and that of the Chinese group (52.0 per 100 000 for men and 29.9 per 100 000 for women, 2004–2008).[4 15] One possible explanation for the higher crude NSCLC incidence in our study could be the ageing of the Chinese population, as an older age has been identified as an independent risk factor for NSCLC.[6] Population ageing is especially obvious in Shanghai, where 27% of the population was older than 60 years of age in 2013,[16] while a mere 16.5% of the US population was older than 60 years of age in 2000.[17]

Most available population-based studies investigate 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 WHO, reports an incidence of lung cancer of 50.4 per 100 000 for men and 19.2 per 100 000 for women in East Asia in 2012. Sihui, a city in south China, is reported to have had a yearly incidence of lung cancer of 37.98 per 100 000 people overall, 60.26 per 100 000 for men and 20.29 per 100 000 for women, between 2007 and 2011, based on the local cancer registry.[9] Both examples are adjusted by 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 those listed 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 an 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 an 11% increase for men, from 2005 to 2010.[9] Except for the effect of population ageing, several other factors may contribute to this higher and increased NSCLC incidence in our study. First, the smoking prevalence in China has dramatically increased in the past two decades. Although the cigarette smoking rate has peaked and decreased in the USA and several other areas in recent years, the prevalence of smoking in China remains at a high level, and China has become one of those 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 a study by Gomez *et al*. Gomez *et al*[23] 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 within the same group, while the incidence was stable for adenocarcinoma, which is less closely associated with tobacco smoke than squamous cell lung cancer. Another factor related to this higher incidence is the higher diagnosis rate due to improved oncology services in Shanghai, as it is 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

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

*Median and 95% CI were estimated using the Kaplan-Meier method.
† Crude HR and 95% CI were estimated using the univariate Cox regression model.
‡Adjusted HR, 95% CI and p value were estimated using the multiple Cox regression model adjusted by surgical resection, chemotherapy, sex, age group, TNM score and histology.
NSCLC, non-small cell lung cancer; TNM, tumour, node and metastasis score.

(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 and Parkin,[25] and Epplein *et al*. [26] The reasons for the higher incidence of lung cancer among Chinese women are unclear, but might be partly ascribed to household air pollution due to cooking fumes and unventilated coal-fuelled heating stoves.[27–29] Besides, considering the high overall smoking prevalence in China, secondhand smoke may also be a critical risk factor for 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 international data from 192 countries in 2004.[30 31] This suggests an additional risk for lung cancer in Chinese women.

Survival

A better survival rate (overall and stage-specified) was observed in this study as compared to that in previously published population-based studies on groups of non-Asian ethnicity, though different population-based lung cancer databases showed different outcomes.

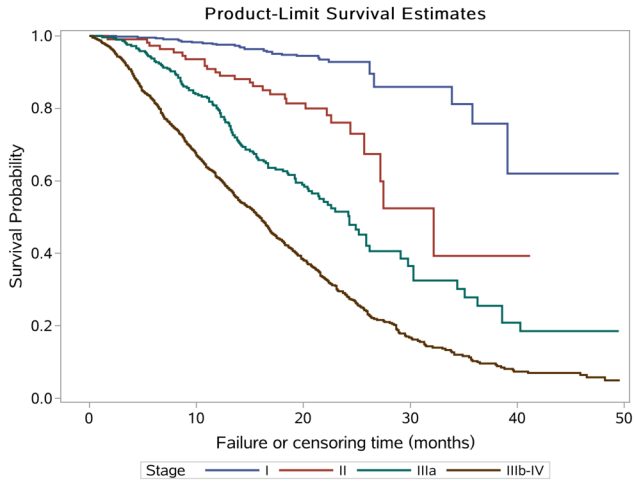

Figure 2 The overall survival (OS) of patients with non-small cell lung cancer (NSCLC) in Shanghai identified in 2011–2013 (n=2013). A 1-year OS rate: whole population 71.8% (95% CI 69.8% to 73.8%), stage I 96.5% (95% CI 94.0% to 98.6%), stage II 89.1% (95% CI 83.3% to94.9%), stage IIIa 78.8% (95% CI 74.1% to 83.5%) and stage IIIb/IV 58.9% (95% CI 56.1% to 61.7%). The survival difference was significant (p<0.0001).

According to the databases from Australia, Canada, Denmark, Norway, Sweden and the 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] A lower survival rate was also observed in the 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, the Asian population shows improved survival. Lin *et al*[34] reported the 2-year survival rate was 80.0–96.2%, 64.4–80.2% and 57.5–67.4% for patients with stages I, II and IIIa NSCLC, respectively, among 30 069 Taiwanese patients, between 2004 and 2007. In a study on 4622 Korean patients between 1998 and 2005, the median OS for stages I, II, III and 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 recognised as an independent favourable prognostic factor for OS among patients with NSCLC.[36 37] Asian patients with NSCLC showed distinct response to cytotoxic chemotherapy when compared with white patients. For example, Gandara *et al*[38] reported a 3-month increase in the median OS of Japanese patients over white patients receiving chemotherapy with the same paclitaxel plus carboplatin regimen; this regimen is also one of the routine regimens for chemotherapy for advanced NSCLC in China. 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 patients with NSCLC.

Similar to previous studies, our results showed that female gender, younger age, smaller tumour size, no lymph node metastasis and no distant metastasis, were related to a better survival rate. The evaluation of impact of surgery on the survival of patients at different TNM stages showed that patients with stage IIIa or IIIb/IV NSCLC who underwent surgical resection had improved survival. This suggests surgical intervention may improve the survival, even for patients with advanced NSCLC, though the details of therapeutic modality still need 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 patients with stage IV NSCLC with single metastasis are also suitable for surgery.[43] Goldstraw *et al*[44] 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. In multivariate analysis, chemotherapy was also shown as a protective prognostic factor, suggesting that a confounding factor does exist in the univariate analysis.

Our study had several strengths. First, this was the first study, to the best of 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 the Shanghai Health Information Network, not only can NSCLC cases be specifically identified, but clinicopathological information and treatments are also available. At the same time, our study offered a higher but comparable incidence to that of the existing cancer registration systems, 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 of and better survival rates for NSCLC as compared to previous studies, which may provide a fresh and meaningful perspective for the evaluation of NSCLC diagnosis and treatment,

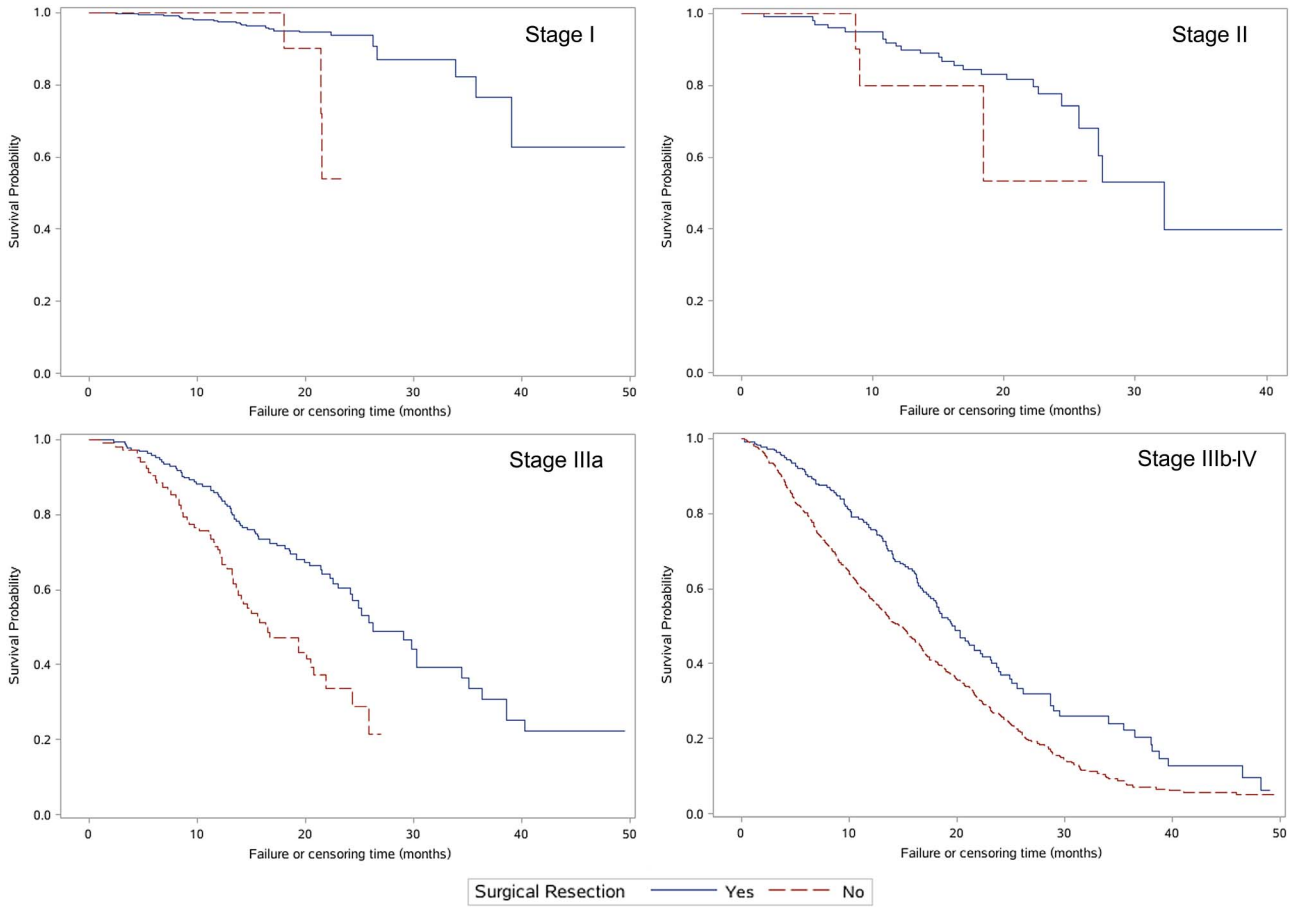

Figure 3 The overall survival (OS) of non-small cell lung cancer (NSCLC) cases in Shanghai identified in 2011–2013, according to surgery by stage (n=2013). A 1-year OS rate of stage I patients: with surgery 96.3% (95% CI 94.5% to 98.1%), without surgery 100.0%; stage II: with surgery 90.0% (95% CI 84.1%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

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 patients with NSCLC, 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. Non-availability of detailed records of diagnosis and treatment also handicapped us in conducting 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 has been newly established. Thus, long-term follow-up is required to determine the survival of patients with NSCLC, especially of those with early-stage NSCLC.

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

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

*Adjusted HR, 95% CI and p value were estimated using the multiple Cox regression model adjusted by sex, age group and histology type.

CONCLUSION

The present study shows a higher incidence and a better survival rate for Chinese patients with NSCLC. 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 prognostic factor for NSCLC at stage IIIa and above.

Acknowledgements The authors thank the staff 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 Lancet SCI Biotechnology Co, Ltd.

Contributors N-QZ, G-YQ and Z-YS conceived and designed the study; HF, Z-YS, Y-YX, Z-HX, WC and HX conducted this study; HF analysed the data. HF, Y-YX and Z-YS drafted the paper; N-QZ and G-YQ revised the paper.

Funding This study was supported by the National Natural Science Foundation of China (number 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 interests 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.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

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:e165–72.
4. Surveillance, Epidemiology, and End Results (SEER) Program. May 2015. 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;3:819–31.
6. National Comprehensive Cancer Network. NCCN Clinical Practice Guidelines in Oncology: Non-Small Cell Lung Cancer. V. 2. Dec 2014. http://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;68:551–64.
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. 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. 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;2:706–14.
12. Shanghai Statistical Year Book (2013). Dec 2014. 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;2:269–79.
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;105:1096–110.
16. Gerontological Society of Shanghai. Dec 2014. http://www.shanghaigss.org.cn/
17. Standard Populations—19 Age Groups, Surveillance, Epidemiology, and End Results Program. Dec 2014. http://seer.cancer.gov/stdpopulations/stdpop.19ages.html
18. Chen W, Zheng R, Zeng H, *et al*. Epidemiology of lung cancer in China. *Thorac Cancer* 2015;6:209–15.
19. Yang L, Parkin DM, Li L, *et al*. Time trends in cancer mortality in China: 1987–1999. *Int J Cancer* 2003;106:771–83.
20. 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.
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;16:1165–72.
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;24:1157–64.
24. Yang LL, Zhang XC, Yang XN, *et al*. Lung cancer treatment disparities in China: a question in need of an answer. *Oncologist* 2014;19:1084–90.
25. Boffetta P, Parkin D. Cancer in developing countries. *CA Cancer J Clin* 1994;2:81–90.
26. Epplein M, Schwartz SM, Potter JD, *et al*. Smoking-adjusted lung cancer incidence among Asian-Americans (United States). *Cancer Causes Control* 2005;16:1085–90.
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;8:1128–35.
29. 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* 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;94:1972–6.
31. 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.
32. 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.
33. 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.
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;17:1294–302.
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;5:1185–96.
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;4:1083–93.
37. 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.
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;27:3540–6.
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;23:8081–92.
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;23:5900–9.
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;350:2129–39.
42. Calvo E, Baselga J. Ethnic differences in response to epidermal growth factor receptor tyrosine kinase inhibitors. *J Clin Oncol* 2006;24:2158–63.
43. XY Z, YK S, JM Y. Standards for the diagnosis and treatment of primary lung cancer (2015 Version) in China. *Chin J Oncol* 2015;37:67–78.
44. Goldstraw P, Ball D, Jett JR, *et al*. Non-small-cell lung cancer. *Lancet* 2011;378:1727–40.
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;21:921–6.