

# Nomogram Predicting the Prognosis of Patients with Surgically Resected Stage IA Non-Small Cell Lung Cancer

**Quan-Xing Liu**

Xinqiao Hospital

**Zi-Qi Huang**

Xinqiao Hospital

**Dong Zhou**

Xinqiao Hospital

**Hong Zheng** (✉ [ziecoe@163.com](mailto:ziecoe@163.com))

Xinqiao Hospital

**Ji-Gang Dai**

Xinqiao Hospital

---

## Research

**Keywords:** Nomogram, non-small cell lung cancer, SEER, prognosis.

**Posted Date:** September 27th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-927146/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The AJCC 8th stage system was limited in accuracy for predicting prognosis of stage IA non-small cell lung cancer (NSCLC) patients. This study aimed to establish and validate two nomograms that predict overall survival (OS) and lung cancer specific survival (LCSS) in surgically resected stage IA NSCLC patients.

**Methods:** Postoperative patients with stage IA NSCLC in SEER database between 2004 and 2015 were examined. Survival and clinical information according to the inclusion and exclusion criteria was collected. All patients were randomly divided into the training cohort and validation cohort with a ratio of 7:3. Independent prognosis factors were evaluated using univariate and multivariate Cox regression analyses, and predictive nomogram was established based on these factors. Nomogram performance was measured using the C-index, calibration plots, and decision curve analysis (DCA). Patients were grouped by quartiles of nomogram scores and survival curves were plotted by Kaplan-Meier analysis.

**Results:** In total, 33533 patients were included in the study. The nomogram of OS and LCSS contained 12 and 10 prognostic factors respectively. The C-index of nomogram showed a relative good performance which was significantly superior than AJCC 8th stage both in training set and validating set ( $P < 0.001$ ). The calibration curve results showed that the actual survival rate was consistent with the predicted survival rate. Nomogram scores related risk stratification revealed statistically significant difference which have better discrimination than AJCC 8th stage.

**Conclusions:** The two established nomograms can accurately predict OS and LCSS in surgical resected patients with stage IA NSCLC.

## Background

In 2018, more than 1.7 million cancer deaths worldwide were caused by lung cancer that remains the most common cancer (11.6% of all cancers) and the main cause of cancer-related deaths[1]. Non-small cell lung cancer (NSCLC) is the major pathological type, consisting of 85% of all types of lung cancer and was classified by WHO into 3 main types: adenocarcinoma, squamous cell carcinoma, and large cell[2]. The AJCC staging system is used worldwide and considered has a great impact on prognosis of NSCLC.

Even for patients with stage IA NSCLC in AJCC 8th stage, the 5-year OS is reported 77–92%[3]. The postoperative adjuvant chemotherapy and targeted therapy were not recommended according to the National Comprehensive Cancer Network (NCCN) guidelines. Find a more effective method to stratify IA NSCLC patients with obvious prognostic difference may be helpful to identify patients who have the potential to receive benefit from postoperative adjuvant therapy. Studies have proposed that other independent prognostic factors, aside from AJCC staging system could significantly contribute to individualized prediction of overall survival (OS) and lung cancer-specific survival (LCSS) in early-stage NSCLC[4–6].

Nomogram is a graphical illustration of a mathematical model, in which different factors are pooled together to predict a definite endpoint, and has been utilized as a convenient and reliable tool to predict the outcome of cancer patients[7]. Several studies revealed that the constructed nomograms were shown to be more accuracy in prediction of prognosis than AJCC staging system[8–11]. However, nomograms for predicting survival outcomes in surgically resected stage IA NSCLC have not been published.

In this study, we aimed to develop clinical nomograms for predicting OS and LCSS of the current stage IA surgically resected NSCLC patients using the surveillance, epidemiology and end results (SEER) database.

## Methods

### Ethics statement

All the data used in our study came from the publicly available SEER database with permission granted to access these research data (SEER\*Stat version 8.3.8 username: 16809-Nov2019). The Nov 2018 Sub (1973–2016 varying) datasets were selected for analysis. Participant informed consent was not required for this study as SEER data is publicly available and de-identified.

### Data Source And Study Population

The SEER data set diagnosed between 2004 and 2015 in our study was restricted to NSCLC in stage I and tumor size  $\leq 3$ cm. The extracted clinical information included the following: patient ID, age at diagnosis, sex, marital status at diagnosis, laterality, primary site, histologic type, grade, surgery at primary site, scope of regional lymph node surgery, the number of regional lymph nodes examined, the 6/7th edition TNM stages, T stage, N stage, M stage, Visceral Pleural Invasion, CS extension, Separate Tumor Nodules-Ipsilateral Lung, survival months, vital status recode, first malignant primary indicator, and sequence number.

### Data Processing

Patients who met all of the eligibility criteria were enrolled in the study: (1) age > 18 years; (2) pathologically confirmed NSCLC (histologic types were classified as squamous carcinoma, adenocarcinoma, large cell carcinoma, and NSCLC); (3) diagnosis of stage IA NSCLC according to the the 8th TNM classification of American Joint Committee on Cancer (AJCC) staging manua[12]. The AJCC 8th stage was calculated according to tumor size, extension, visceral pleural invasion and 6/7th TNM stages. (4) surgical performed by lobectomy, segmentectomy or wedge resection. The exclusion criteria include: (1) pathologically confirmed small cell lung cancer or all subtypes of sarcoma, (2) age < 18, (3) tumor located in the main bronchus and have positive visceral pleural invasion (VPI), which could not be classified in stage IA according to the 8th edition AJCC manual.

The study variables obtained from the datasets include: age, sex, race, marital status, tumor size, histologic type, grade, first malignancy, malignancy sequence, surgery, primary site, laterality, lymph node scope region, lymph node examined number. Tumor size and lymph node examined number were numerical variable, and other variables were treated as factor variable.

## Statistical Analysis

The whole eligible patient cohort was randomly divided into a training set, and a validating set at a ratio of 7:3. The training set was used to develop the model and the validation set was used to evaluate the prediction accuracy of the model. In the training set, survival prognosis was investigated with univariate Cox regression model followed by multivariate Cox regression model to recognize the independent prognostic factors. In addition, the VIF and correlation coefficient were calculated between all independent prognostic factors to identify multicollinearity. When two independent variables were identified multicollinearity, only one could be included in the final factors. Finally, independent prognosis risk factors included in constructing Cox regression model and nomogram for Stage IA NSCLC patients were identified.

During validation of the nomograms, the total points for each patient were calculated according to the established nomograms and then Cox regression was performed using the total points as predictor in the validation cohort. The performance of the nomogram was evaluated by concordance index (C-index)[13], calibration curve and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. ROC curves were used for the sensitivity and specificity of the nomogram. Finally, decision curve analysis (DCA) performed with the Decision Curve package[14] was used to evaluate the clinical benefits and utility of the nomogram compared with AJCC 8th TNM stage alone in predicting of 1-, 3-, and 5-year survival rates. To clarify the descriptive power of the nomogram, all patients in training and validation set were divided into four subgroups according to quartiles of nomogram-derived total scores. Then survival analyses were carried out using Kaplan – Meier method and differences in OS and LCSS were examined using the Log-Rank test.

In addition, we compared the prognostic prediction power between our established nomogram and AJCC 8th stage (IA1, IA2 and IA3) by the Cox regression analysis and C-index. We also compared the benefit of OS and LCSS between chemotherapy and non-chemotherapy for stage IA patients in all subgroups according to nomogram scores. To minimize the potential selection bias, propensity score matching (PSM) was performed to balance the confounding factors between the two groups using the nearest matching method with a 1:1 ratio. Patients were matched by the variables as following: age, sex, marital status, year diagnosis, race, Primary site, Histologic type, Grade, Surgery, Laterality, LN Scope Reg, LN exam Num, tumor, size and Stage eight.

All statistical computations were done using the statistical programming language R (R version 3.6.3, <https://www.r-project.org>). Statistical differences of distribution in demographic and clinical characteristics between the training and validation cohorts were evaluated using the CBCgrps package.

The rms, foreign, survival packages were used to construct the nomogram. The survminer package was used to cumulative survival time by Kaplan-Meier method and compare the differences in survival curves by log-rank test. Spearman correlation matrix (R package PerformanceAnalytics) and variance inflation factors (vif function in R package car) were computed to evaluate possible collinearity among explanatory variables. Multicollinearity was explored using the VIF and correlation coefficient, considering values of VIF > 5 or correlation coefficient of > 0.7 indications of collinearity between variables[15, 16]. The MatchIt package was used to complete the PSM between two group of chemotherapy and non-chemotherapy. P < 0.05 was considered as statistically significant.

## Results

### Patient characteristics

A total of 33533 patients in stage IA NSCLC (AJCC 8th edition) were finally included based on the inclusion and exclusion criteria, which were randomly divided into a training set of 23589 patients and a validation set of 9944 patients (Fig. 1). Table 1 summarizes baseline demographics and clinical characteristics for the study population.

Table 1

The demographics, surgery and pathological characteristics of the included patients in the training and validation sets.

<b>Variables</b>	<b>Total (n = 33533)</b>	<b>train (n = 23589)</b>	<b>test (n = 9944)</b>	<b>P value</b>
<b>Age, Median (IQR)</b>	69 (62, 75)	69 (63, 75)	69 (62, 75)	0.513
<b>Age range, n (%)</b>				0.219
18–40	82 (0.24)	60 (0.25)	22 (0.22)	
41–60	6476 (19.31)	4513 (19.13)	1963 (19.74)	
61–80	24051 (71.72)	16918 (71.72)	7133 (71.73)	
>80	2924 (8.72)	2098 (8.89)	826 (8.31)	
<b>Year diagnosis, n (%)</b>				0.735
2004–2007	9840 (29.34)	6900 (29.25)	2940 (29.57)	
2008–2011	11727 (34.97)	8279 (35.1)	3448 (34.67)	
2012–2015	11966 (35.68)	8410 (35.65)	3556 (35.76)	
<b>Sex, n (%)</b>				0.412
Female	18842 (56.19)	13220 (56.04)	5622 (56.54)	
Male	14691 (43.81)	10369 (43.96)	4322 (43.46)	
<b>Race, n (%)</b>				0.139
White	28830 (85.98)	20313 (86.11)	8517 (85.65)	
Black	2595 (7.74)	1782 (7.55)	813 (8.18)	
Others	2108 (6.29)	1494 (6.33)	614 (6.17)	
<b>Marital status, n (%)</b>				0.456
Unmarried	14563 (43.43)	10213 (43.3)	4350 (43.74)	
Married	18970 (56.57)	13376 (56.7)	5594 (56.26)	
<b>Tumor size, Median (IQR)</b>	18 (13, 23)	18 (13, 23)	18 (13, 23)	0.474
<b>AJCC 8th Stage, n (%)</b>				0.698
IA1	4433 (13.22)	3127 (13.26)	1306 (13.13)	
IA2	17669 (52.69)	12394 (52.54)	5275 (53.05)	
IA3	11431 (34.09)	8068 (34.2)	3363 (33.82)	
<b>Histologic type, n (%)</b>				0.722

<b>Variables</b>	<b>Total (n = 33533)</b>	<b>train (n = 23589)</b>	<b>test (n = 9944)</b>	<b>P value</b>
AC	8356 (24.92)	5852 (24.81)	2504 (25.18)	
SCC	23677 (70.61)	16670 (70.67)	7007 (70.46)	
LCC	703 (2.1)	506 (2.15)	197 (1.98)	
NSCLC	797 (2.38)	561 (2.38)	236 (2.37)	
<b>Grade, n (%)</b>				0.821
I	7998 (23.85)	5620 (23.82)	2378 (23.91)	
II	16113 (48.05)	11360 (48.16)	4753 (47.8)	
III or IV	9422 (28.1)	6609 (28.02)	2813 (28.29)	
<b>First Malignant, n (%)</b>				0.178
No	11563 (34.48)	8080 (34.25)	3483 (35.03)	
Yes	21970 (65.52)	15509 (65.75)	6461 (64.97)	
<b>Malignant Sequence, n (%)</b>				0.281
First	21419 (63.87)	15119 (64.09)	6300 (63.35)	
Second	9229 (27.52)	6473 (27.44)	2756 (27.72)	
Third or more	2885 (8.6)	1997 (8.47)	888 (8.93)	
<b>Surgery, n (%)</b>				0.606
Wedge	8082 (24.1)	5690 (24.12)	2392 (24.05)	
Segmentectomy	1914 (5.71)	1327 (5.63)	587 (5.9)	
Lobectomy	23537 (70.19)	16572 (70.25)	6965 (70.04)	
<b>Primary Site, n (%)</b>				0.307
Upper lobe	20842 (62.15)	14617 (61.97)	6225 (62.6)	
Middle lobe	1766 (5.27)	1275 (5.41)	491 (4.94)	
Lower lobe	10820 (32.27)	7621 (32.31)	3199 (32.17)	
Overlapping lesion	105 (0.31)	76 (0.32)	29 (0.29)	
<b>Laterality, n (%)</b>				0.491
Left	13941 (41.57)	9778 (41.45)	4163 (41.86)	
Right	19592 (58.43)	13811 (58.55)	5781 (58.14)	

Variables	Total (n = 33533)	train (n = 23589)	test (n = 9944)	P value
<b>LN Scope Region, n (%)</b>				0.733
None	5217 (15.56)	3672 (15.57)	1545 (15.54)	
1 to 3 regional	6788 (20.24)	4803 (20.36)	1985 (19.96)	
4 or more regional	21013 (62.66)	14745 (62.51)	6268 (63.03)	
Others	515 (1.54)	369 (1.56)	146 (1.47)	
<b>LN examined Number, Median (IQR)</b>	5 (2, 10)	5 (2, 10)	5 (2, 10)	0.758
IQR: Inter-quartile range; LN: Lymph node.				

## Cox Regression Analysis

Multivariate Cox regression analysis was performed by using the significant prognostic factors identified in univariate analysis. In the multivariate Cox regression analyses 13 factors were significantly associated with OS as shown in Fig. 2B. While 10 variables were identified as independent prognostic factors for LCSS in multivariate Cox regression analysis (Fig. 3B). Among the 13 factors associated with OS, correlation coefficient between first malignant and malignant sequence was 0.9 (eFigure 1 A). Thus, the variant first malignant was removed for the less significant statistic compared with malignant sequence. There was no significant correlation the 10 selected independent variables of LCSS for the training set (eFigure 1B). Finally, 12 and 10 variables were included in the construction of the nomogram of OS and LCSS respectively. The results of univariate and multivariate Cox regression analysis of OS are shown in Fig. 2, and those for LCSS are shown in Fig. 3.

## Construction And Validation Of The Nomogram

The nomogram of OS was constructed of the above 12 variables (Fig. 4A). The nomogram of LCSS was constructed of 10 variables (Fig. 5A). In the training set, bootstrap validation revealed that the C-index for OS prediction was 0.656 (95%CI, 0.650–0.662) and the C-index for for LCSS prediction was 0.642 (95%CI, 0.633–0.650). The nomogram-derived total score of OS and LCSS for each patients in training and validating set was calculated by adding the scores of each selected factor. Cox regression analysis was performed using the total scores as a unique factor in the validation set, the C-index was 0.652 (95% CI, 0.643–0.662) for OS and 0.651 (95% CI, 0.639–0.664) for LCSS. In the training set, the 1-, 3-, and 5-year OS AUC of the nomogram were 0.659, 0.67, and 0.673, respectively (eFigure 4A) and the 1-, 3-, and 5-year LCSS AUC of the nomogram were 0.635, 0.659, and 0.648, respectively (eFigure 4B). In the validation set, the 1-, 3- and 5-year OS AUC of the nomogram were 0.651, 0.665 and 0.667, respectively (eFigure 4C) and the 1-, 3- and 5-year LCSS AUC of the nomogram were 0.671, 0.665 and 0.652, respectively (eFigure 4D).

The C-index of nomogram showed a relative good performance which was significantly superior than AJCC 8th stage both in training set and validating set ( $P < 0.001$ ) (**eTable 1**). Also, the final established multivariate Cox regression model was identified superior than AJCC 8th stage Cox regression model in terms of the discriminability for OS (Fig. 4D **and E**) and LCSS (Fig. 5D **and E**) in both training set and validation set. The predicted calibration curves for the 1-, 3- and 5-year OS (Fig. 4B **and C**) and LCSS (Fig. 5B **and C**) showed favorable calibration of the nomogram in the training set and validation set. In validation set, the results of 1-, 3-, and 5- year DCA for both OS (Fig. 4F-H) and LCSS (Fig. 5F-H) demonstrated that the predictive power of our nomogram had obvious clinical net benefit which was superior than that of the AJCC 8th stage.

The total scores in the training sets, was divided into quartile for OS (Min-212.65, 212.65–238.35, 238.35–264.69, 264.69-Max) and LCSS (Min-257.76, 257.76-293.58, 293.58-332.52, 332.52-Max). The Log-rank test revealed significant differences between the four groups, in both OS ( $p < 0.001$ ) and LCSS ( $p < 0.001$ ) which shown higher distinguish power than AJCC 8th stage groups (Figs. 6A-H).

## Prognosis Of Adjuvant Chemotherapy

We extracted the patients in our data set who suffered primary lung cancer only once and the information of whether received chemotherapy. There were 16864 patients included with 488 patients received adjuvant chemotherapy and 16376 did not. According to the nomogram scores those patients were divided into four risk stratification groups as the previous described in OS (**eTable 2**) and LCSS (**eTable 3**). No significant difference in the OS was observed between the adjuvant chemotherapy and non-adjuvant chemotherapy for patients with high nomogram scores (264.69-Max) and (238.35–264.69) groups (**eFigure 2 C-D**). However, for patients in other nomogram scores groups, non-adjuvant chemotherapy revealed a superior OS than adjuvant chemotherapy (**eFigure 2 A-B**). Additionally, non-adjuvant chemotherapy showed significantly superior LCSS than adjuvant chemotherapy in each nomogram scores sub-groups (**eFigure 2 E-H**). Those results was also consistent through propensity matching analysis (**eFigure 3 A-H**). These results suggest that for surgical resected patients in stage IA received adjuvant chemotherapy have poor prognosis than those who received non-chemotherapy.

## Discussion

In the study, we evaluate the OS and LCSS for patients with surgically resected stage IA NSCLC from the SEER database diagnosed between 2004 and 2015. Two nomograms were built to predict the probability of OS and LCSS, as an easily used clinical tool. The good prediction power and clinical utility of the two nomograms were also supported by C-index, calibration curves and DCA curves.

In the final Cox regression model, independent prognostic factors for OS was 12 and for LCSS was 10. Compared with LCSS nomogram the OS nomogram additional included the factor of malignant sequence and age range. This means the patient with prior cancer diagnosis and high age were at greater risk of overall death, but no increase in the risk of lung cancer specific death. As previous reported, the risks of

increasing treatment intolerance due to prior cancer treatment are more likely to worsen the survival of patients with early stage lung cancer[17]. Due to the declining health surgery and disease resistance in elderly patients is poorer than that in young patients; thus resulted poor prognosis[18]. However, in our study the age was detected no significant influence in LCSS for patients with stage IA NSCLC. This reason may be the over excellent prognosis in NSCLC patients with stage IA, with a median LCSS over 10 years.

Interestingly, our study found that for stage IA NSCLC patients married status had superior OS and LCSS than unmarried status. The reason may be married patients have a better psychosocial environment to fight against lung cancer. As reported that marriage could significantly reduce the mortality rate of women and social and psychological factors could influence the prognosis of patients with cancer[19]. Notably, primary site with overlapping lesion was found associated with both poor OS and LCSS in our study. Similarly, Dariusz Dziedzic also reported that lung cancer with adjacent lobe invasion is a separate group of tumor that lies between stages T2 and T3 for its poor survival rate[20]. Recently, several studies demonstrated that lobectomy compared with sublobar and segmentectomy compared with wedge resection both showed significant superior OS and LCSS for stage IA NSCLC, which was in line with our finding[21]. In our study we found both LN Scope Region and LN examined Number should be considered as important prognostic factors even for stage IA NSCLC without positive LN. This finding was consistent with the published study, in which the authors recommended at least 16 LN should be resected for NSCLC patients[22]. But in their study the LN Scope Region was not considered in analysis.

We additionally analyzed the OS and LCSS by dividing the patient population into four sub-groups according to quartiles of nomogram scores. The high scores sub-groups could be considered as high risk patients in stage IA, which may receive benefit from adjuvant therapy. However, to our surprise, the trends of adjuvant chemotherapy for stage IA patients were noticed that the better the prognosis sub-groups have, the more adverse effect of OS and LCSS affected. One recent study found surgically resected EGFR-Mutated NSCLC patients in IB stage could receive significant benefit from adjuvant targeted therapy based on Osimertinib[23]. Whether patients in identified high risk stratification of stage IA stage could also receive the similar benefit is highly expected.

To our knowledge, this is the first retrospective study of large cohort to construct the nomogram of NSCLC patients in AJCC 8th edition IA stage. However, it should be noted that there were several limitations in this study. First, this study was retrospective, various biases are introduced including selection bias, loss to follow-up, and missing data. Second, other known risk factors for prognosis, such as information on chemotherapy, lymphovascular invasion, spread through air space and the latest classification of adenocarcinoma were hard to obtain from the SEER database. Finally, there was a lack of external validation independent of the SEER database, and the generalization prognostic model to the global population is still unclear.

## Conclusion

We developed practical nomograms with good prediction accuracy and discrimination that can help clinicians make personalized survival predictions for NSCLC patients with AJCC 8th stage IA NSCLC. This nomogram can help to identify high risk stage IA patients after surgical resection.

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **Availability of data and materials**

The datasets generated and analyzed during the current study are available in the [SEER database] repository, [<https://seer.cancer.gov/>].

### **Competing interests**

The authors declare that they have no competing interests.

### **Funding**

The study was supported by funding from the Clinical research Foundation of Army Medical University (2018XLC2014) for Dai, and the medical and scientific research project of science and health joint department of Chongqing for Liu (2020MSXM031)

### **Authors' contributions**

Quan-Xing Liu and Zhi-Qi Huang are co-first authors they contributed equally to this study.

(I) Conception and design: Quan-Xing Liu, Zhi-Qi Huang, Hong Zheng and Ji-Gang Dai

(II) Administrative support: Quan-Xing Liu and Ji-Gang Dai

(III) Provision of study materials or patients: Xiao-Qing Liu and Zhi-Qi Huang

(IV) Collection and assembly of data: Quan-Xing Liu, and Dong Zhou

(V) Data analysis and interpretation: Quan-Xing Liu and Hong Zheng

(VI) Manuscript writing: All authors

(VII) Final approval of manuscript: All authors

## Acknowledgements

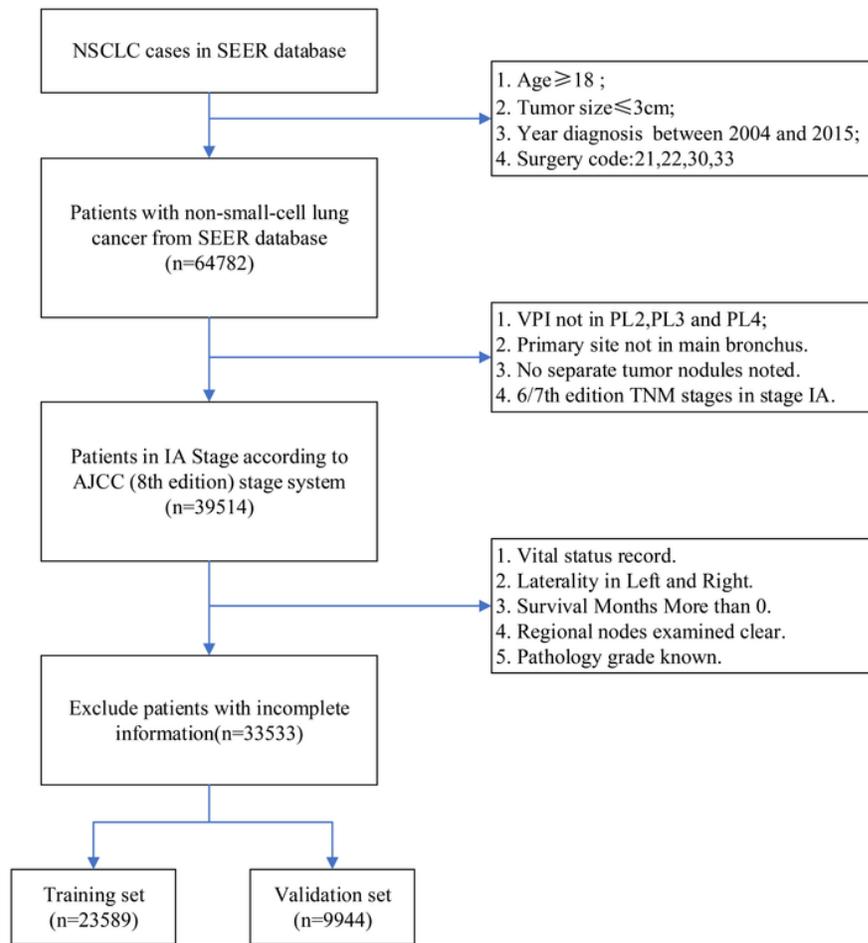
The authors thank the study participants for their contribution to the research, as well as current and past researchers and staff.

## References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424.
2. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol*. 2015;10(9):1243–60.
3. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: Proposals for revision of the TNM stage groupings in the forthcoming (Eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11(1):39–51.
4. Zhou H, Zhang Y, Qiu Z, et al. Nomogram to predict Cause-Specific mortality in patients with surgically resected stage i Non-Small-Cell lung cancer: A competing risk analysis. *Clin Lung Cancer*. 2018;19(2):e195–203.
5. Zeng Y, Mayne N, Yang CJ, et al. A nomogram for predicting Cancer-Specific survival of TNM 8th edition stage i non-small-cell lung cancer. *Ann Surg Oncol*. 2019;26(7):2053–62.
6. Zhang Y, Sun Y, Xiang J, et al. A clinicopathologic prediction model for postoperative recurrence in stage Ia non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2014;148(4):1193–99.
7. Nieder C, Mehta MP, Geinitz H, et al. Prognostic and predictive factors in patients with brain metastases from solid tumors: A review of published nomograms. *Crit Rev Oncol Hematol*. 2018;126:13–8.
8. Yang J, Chen S, Li Y, et al. Incidence rate and risk factors for suicide death in patients with skin malignant melanoma: A Surveillance, Epidemiology, and End Results analysis. *Melanoma Res*. 2020;30(4):402–09.
9. He Y, Ong Y, Li X, et al. Performance of prediction models on survival outcomes of colorectal cancer with surgical resection: A systematic review and meta-analysis. *Surg Oncol*. 2019;29:196–202.
10. Wang Y, Li J, Xia Y, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol*. 2013;31(9):1188–95.
11. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol*. 2015;33(8):861–69.
12. Nicholson AG, Chansky K, Crowley J, et al. The international association for the study of lung cancer lung cancer staging project: Proposals for the revision of the clinical and pathologic staging of small cell lung cancer in the forthcoming eighth edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11(3):300–11.

13. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92–105.
14. Kerr KF, Brown MD, Zhu K, et al. Assessing the clinical impact of risk prediction models with decision curves: Guidance for correct interpretation and appropriate use. *J Clin Oncol*. 2016;34(21):2534–40.
15. Waldmann P. On the use of the pearson correlation coefficient for model evaluation in Genome-Wide prediction. *Front Genet*. 2019;10:899.
16. Marcoulides KM, Raykov T. Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educ Psychol Meas*. 2019;79(5):874–82.
17. Janssen-Heijnen M, van Erning FN, De Ruyscher DK, et al. Variation in causes of death in patients with non-small cell lung cancer according to stage and time since diagnosis. *Ann Oncol*. 2015;26(5):902–07.
18. Wang BY, Huang JY, Cheng CY, et al. Lung cancer and prognosis in taiwan: A population-based cancer registry. *J Thorac Oncol*. 2013;8(9):1128–35.
19. Galvin A, Delva F, Helmer C, et al. Sociodemographic, socioeconomic, and clinical determinants of survival in patients with cancer: A systematic review of the literature focused on the elderly. *J Geriatr Oncol*. 2018;9(1):6–14.
20. Dziedzic D, Rudzinski P, Langfort R, et al. Results of surgical treatment and impact on T staging of non-small-cell lung cancer adjacent lobe invasion. *Eur J Cardiothorac Surg*. 2016;50(3):423–27.
21. Zheng YZ, Zhai WY, Zhao J, et al. Oncologic outcomes of lobectomy vs. Segmentectomy in non-small cell lung cancer with clinical T1N0M0 stage: A literature review and meta-analysis. *J Thorac Dis*. 2020;12(6):3178–87.
22. Liang W, He J, Shen Y, et al. Impact of examined lymph node count on precise staging and Long-Term survival of resected Non-Small-Cell lung cancer: A population study of the US SEER database and a chinese Multi-Institutional registry. *J Clin Oncol*. 2017;35(11):1162–70.
23. Wu YL, Tsuboi M, He J, et al. Osimertinib in resected EGFR-Mutated Non-Small-Cell lung cancer. *N Engl J Med*. 2020;383(18):1711–23.

## Figures



**Figure 1**

Patient flow chart.

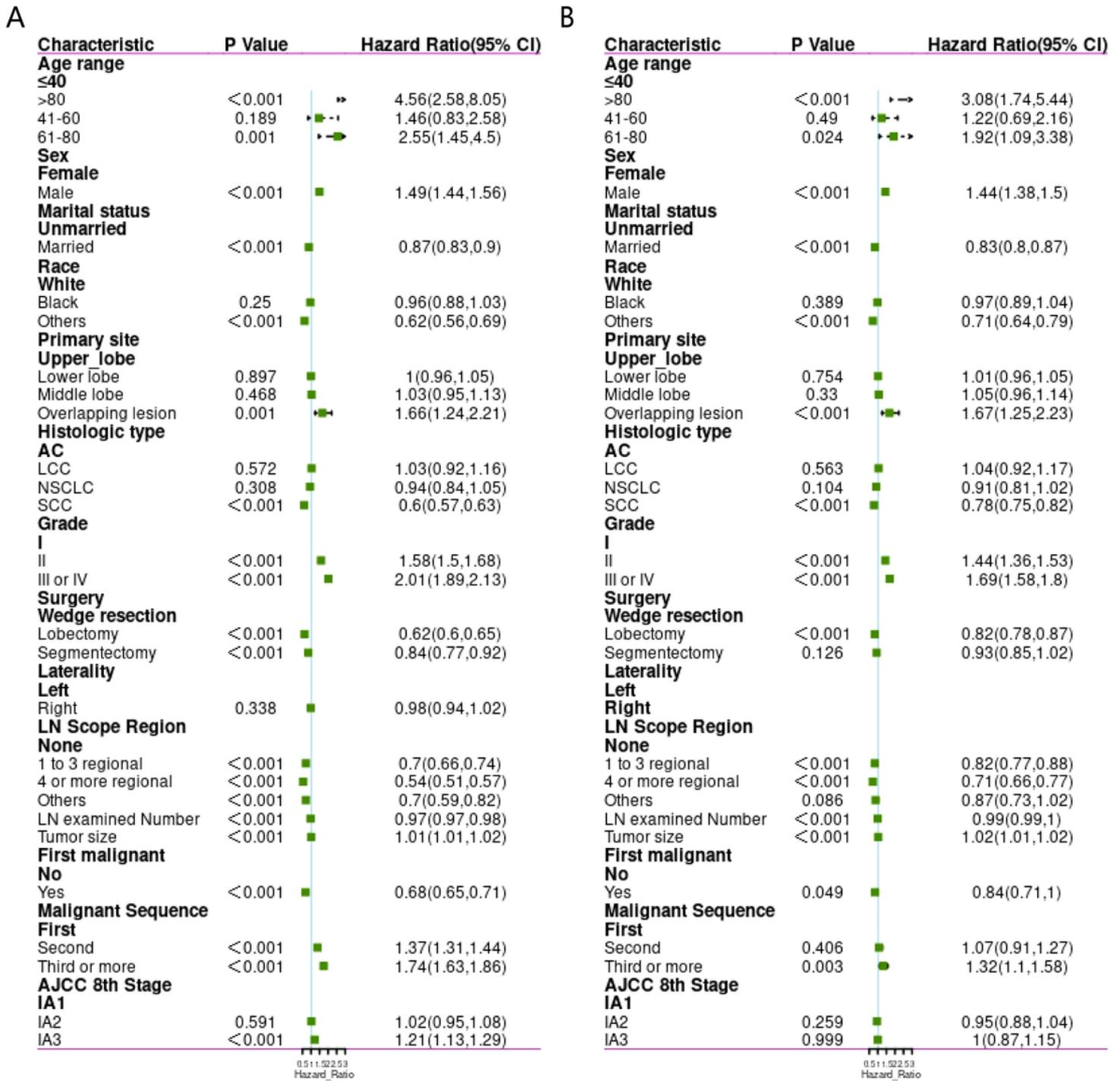
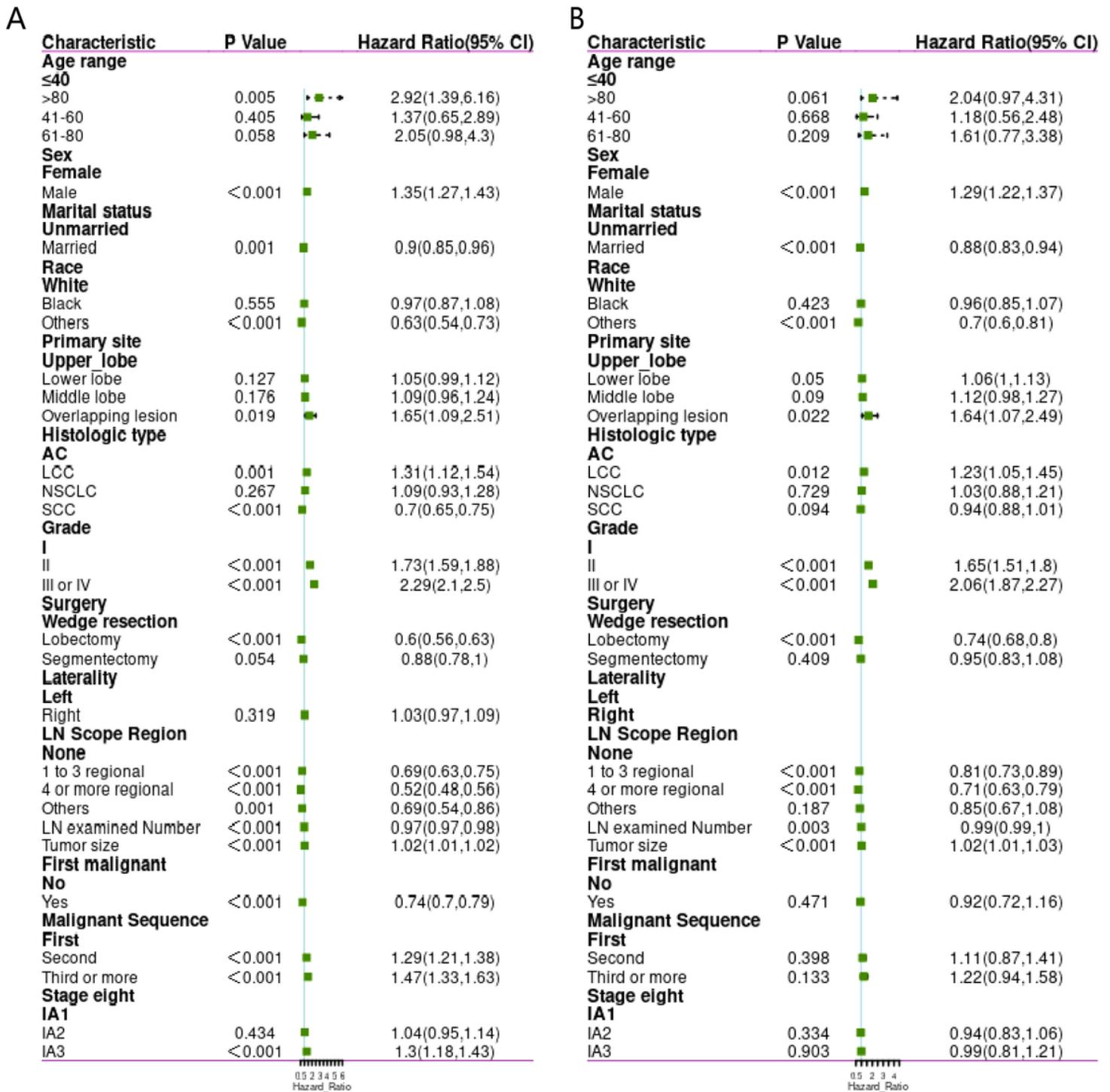


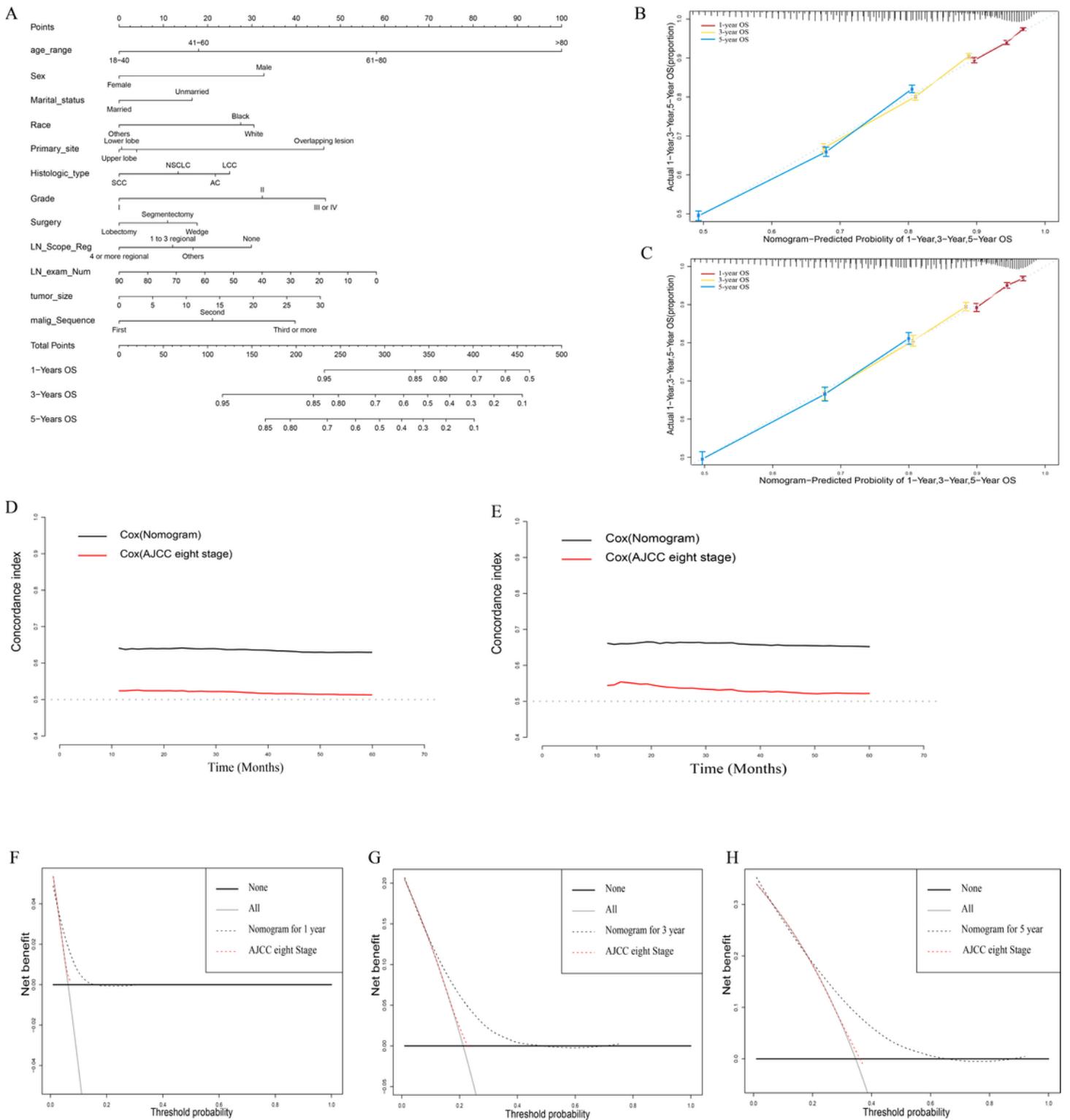
Figure 2

The forest map of Cox regression analysis. Univariate (A) and Multivariate (B) Cox regression analyses estimating the risk factors for OS.



**Figure 3**

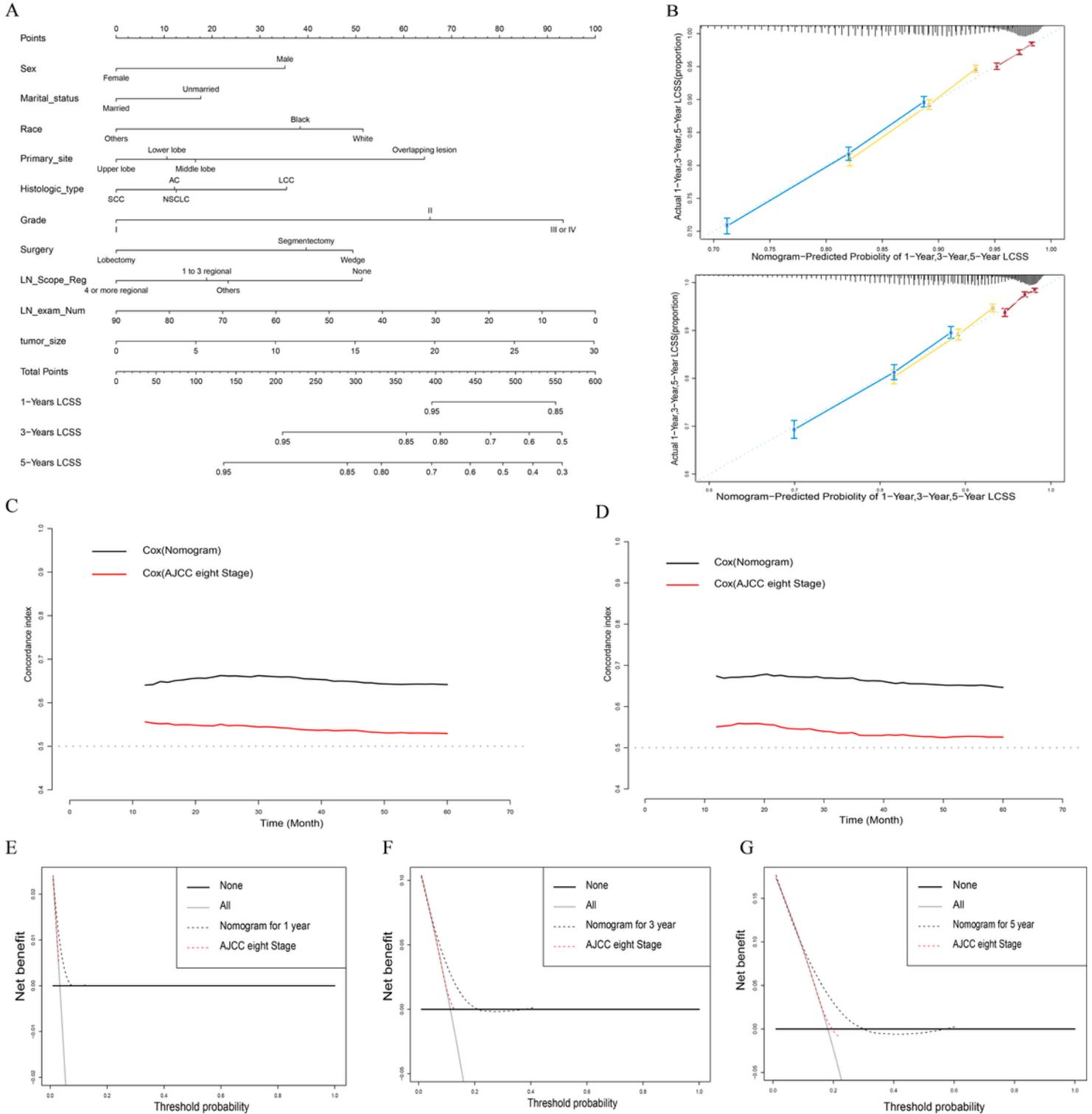
The forest map of Cox regression analysis. Univariate (A) and Multivariate (B) Cox regression analyses estimating the risk factors for LCSS.



**Figure 4**

(A) Nomogram used to predict the 1, 3 and 5 year OS rates of patients with stage IA NSCLC. Calibration curve of the nomogram for predicting the 1, 3 and 5 year OS rates of patients with stage IA NSCLC from the training set (B) and validation set (C). Model Cox of OS (12 variables) is better than Cox (1 variable, AJCC 8th stage) in terms of the discriminability of the regression model in training set (D) and validation

set (E). Decision curve analysis of the AJCC 8th stage and nomogram the 1 (F), 3 (G) and 5 year (H) OS rates of patients with stage IA NSCLC from the validation set. OS, overall survival



**Figure 5**

(A) Nomogram used to predict the 1, 3 and 5 year LCSS rates of patients with stage IA NSCLC. Calibration curve of the nomogram for predicting the 1, 3 and 5 year LCSS rates of patients with stage IA NSCLC from the training set (B) and validation set (C). Model Cox of LCSS (10 variables) is better than Cox (1

variable, AJCC 8th stage) in terms of the discriminability of the regression model in training set (D) and validation set (E). Decision curve analysis of the AJCC 8th stage and nomogram the 1 (F), 3 (G) and 5 year (H) LCSS rates of patients with stage IA NSCLC from the validation set. LCSS, lung cancer-specific survival.

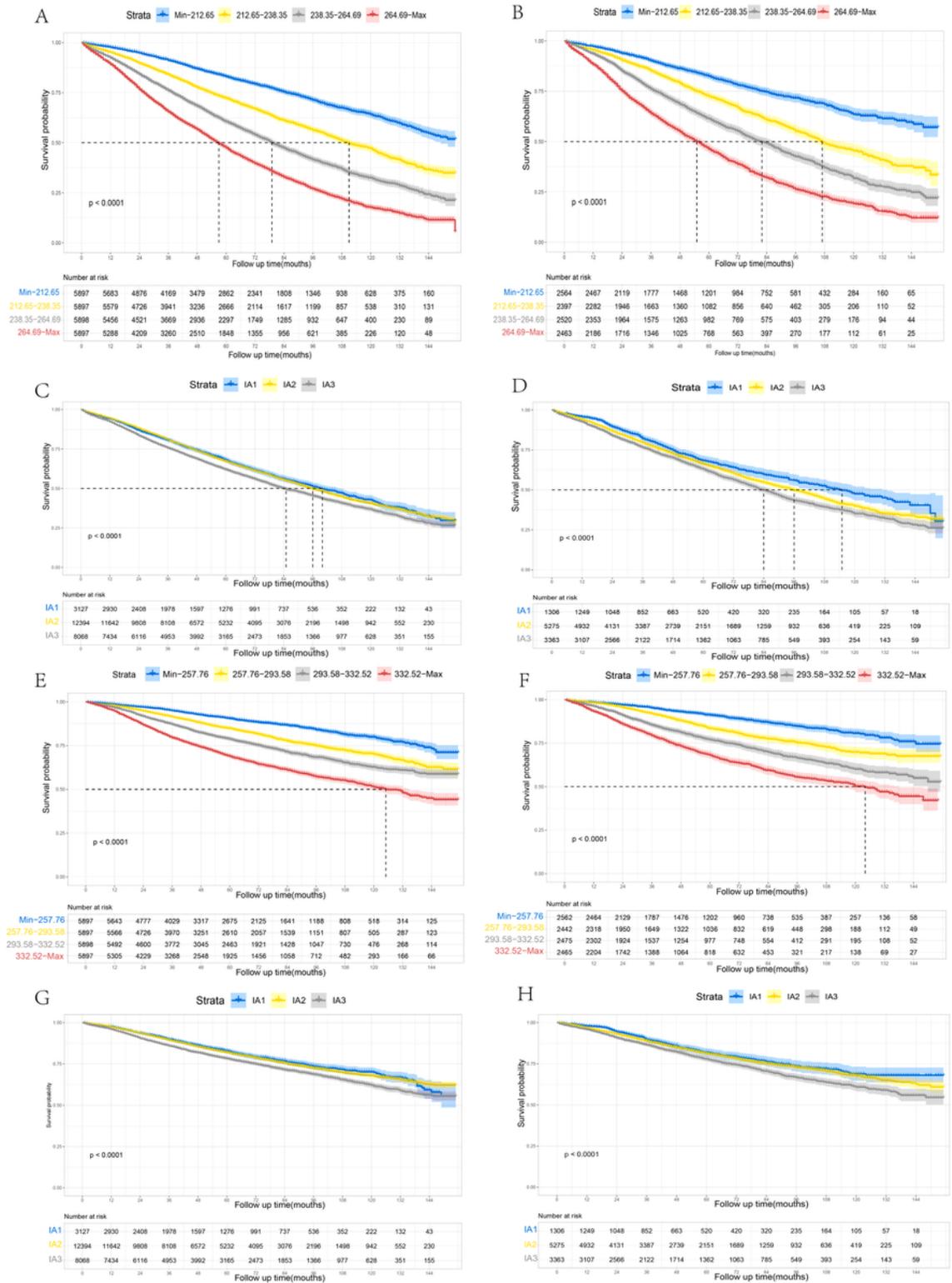


Figure 6

Kaplan-Meier curves of risk group stratification within the training set (A) and validation set (B) for OS. Kaplan-Meier curves of AJCC 8th stage group stratification within the training set (C) and validation set (D) for OS. Kaplan-Meier curves of risk group stratification within the training set (E) and validation set (F) for LCSS. Kaplan-Meier curves of AJCC 8th stage group stratification within the training set (G) and validation set (H) for OS.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigurelegends.docx](#)
- [renamedef28d.tif](#)
- [eFigure2Survivalanalysisofchemotherapy.tif](#)
- [eFigure3chro.group.PSM.tif](#)
- [eFigure4AUC.tif](#)
- [renamed4a817.docx](#)
- [eTable2.PSM.os.pdf](#)
- [eTable3.PSM.css.pdf](#)