

# Prognostic nomograms for predicting overall survival of stage I lung cancer patients: a large population-based study

Liwei Wu (✉ [docwuliwei@foxmail.com](mailto:docwuliwei@foxmail.com))

Shanghai Public Health Clinical Center <https://orcid.org/0000-0002-3150-2260>

Laiyi Wan

Shanghai Public Health Clinical Center

Yanzheng Song

Shanghai Public Health Clinical Center

Lin Wang

Shanghai Public Health Clinical Center

---

## Primary research

**Keywords:** non-small cell lung cancer, lymph node dissection, surgery, nomogram

**Posted Date:** June 16th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-35108/v1>

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

[Read Full License](#)

---

## **Abstract**

## **Background**

Postoperative prognosis of stage I non-small cell lung cancer (NSCLC) undergoing lymph node dissection is heterogeneous. Therefore, we sought to construct a novel survival prediction model for stage I NSCLC undergoing lymph node dissection.

## **Method:**

Based on the data from the Surveillance, Epidemiology, and End Results (SEER) program, we successfully determined and incorporated independent prognostic markers to construct the nomogram. The constructed nomogram was further subjected to external validation with an independent cohort of patients. The performance of the survival prediction model was assessed by concordance index, calibration plots, and risk subgroup classification.

## **Results**

A total of 23394 patients from SEER registries (training cohort), as well as 7016 patients (verification cohort) was included. Age, gender, race, histology, differentiation grade, stage, and lymph node dissection mode were identified as significant prognostic parameters. A novel nomogram was developed and externally validated. The model displayed good discrimination with a C-index of 0.666 (95% CI, 0.652 to 0.681) and good calibration. Calibration plots demonstrated an optimal consistency between the nomogram predicted and actual observed probability of survival.

## **Conclusion**

This novel nomogram incorporating the age, gender, race, histology, differentiation grade, stage, and lymph node dissection mode could be conveniently used to facilitate the individual prognosis prediction in stage I NSCLC patients.

## **Introduction**

Non-small cell lung cancer (NSCLC) is still a major health problem in the world[1]. Although smoking control measures have been introduced in recent years, the incidence of NSCLC has not decreased significantly[2, 3]. The number of deaths from NSCLC is also on the rise every year[4]. Although stage I NSCLC patients can be cured by surgery, there are still differences in prognosis between patients and patients. At present, there is a lack of an accurate tool for personalized evaluation of the prognosis of patients with stage I NSCLC, which is an urgent clinical issue to be solved.

The National Comprehensive Cancer Network (NCCN) has released the latest NSCLC treatment guidelines. Surgical treatment and lymph node dissection or sentinel lymph node biopsy are recommended in the latest guidelines for patients with stage I NSCLC. The prognosis of patients with NSCLC is related to many factors. Some studies[5, 6] have shown that patient age may be an independent factor related to prognosis. Other studies[7–11] have shown that race differences that may be caused by genetic differences may also be independent factors affecting prognosis. Pathological type, pathological grade and clinical stage are all independent factors affecting the prognosis of patients[12]. The mode of operation is also an independent factor affecting the prognosis of patients[13].

Considering so many related factors, it is very necessary to establish an accurate prediction model including all these factors. The purpose of this study was to develop a valid but simple prediction tool for stage I NSCLC to assess the prognosis using only characteristics easily available when starting follow-up.

## **Patients And Method**

### **Patients**

For this study, stage I NSCLC patients who were performed surgery were eligible. Patients who underwent surgery included lymph node dissection, no lymph node dissection and sentinel lymph node biopsy were included. Patients with unknown or missing information were excluded. A total of 300410 patients with early lung cancer underwent surgery from the Surveillance, Epidemiology, and End Results (SEER) database between 2003 and 2015. The analysis included age, gender, race, histology, differentiation grade, stage, lymph node dissection mode, and months of survival.

### **Statistical analysis**

All data including demographic, disease, and treatment characteristics were expressed as count (%). Statistical analysis was performed using the R software (Version 3.6.2; <https://www.R-project.org>).

The study included two cohorts, the training cohort and the verification cohort. The training cohort was used for establishing Cox proportional hazards model, the nomogram and the survival curves[14]. The verification cohort demonstrated accuracy. Multivariate Cox proportional hazard regression was used to determine independent predictors of mortality[15].  $P < .05$  was considered statistically significant. The predicting model was established with independent predictors which were determined by multivariate Cox proportional hazard regression[15]. Calibration curves were plotted to assess the calibration of the nomogram. To quantify the discrimination performance of the nomogram, Harrell's C-index was measured[16]. The verification cohort was subjected to calculate a relatively corrected C-index. Survival curves of the use of different therapies in stage IA and IB, younger and older age, and squamous cell carcinoma and adenocarcinoma were evaluated by the Kaplan-Meier method.

## **Results**

## **Patients' characteristics**

In the training cohort, of 23394 patients with stage I NSCLC who received surgery during the study period. 20421(87.3%) patients underwent lymph node dissection, 2860(12.2%) patients did not undergo lymph node dissection, and 113(0.5%) patients underwent sentinel lymph node biopsy. The characteristics of patients in the training and verification cohorts are listed in Table 1.

## **Development of an individualized prognosis model**

The results of the multivariate Cox proportional hazard regression among age, sex, race, histology, differentiation grade, stage, and lymph node dissection mode are given in Table 2. The model that incorporated the above independent predictors was developed and presented as the prognosis nomogram (Figure 1).

## **Apparent performance of the prognosis nomogram**

The C-index for OS prediction was 0.666 (95% CI, 0.652 to 0.681). The calibration plot for the probability of survival at 3 or 5 year after surgery showed an optimal agreement between the prediction by nomogram and actual observation (Figure 2A and 2B).

In verification cohort, the C-index of the nomogram for predicting OS was 0.658 (95% CI, 0.650 to 0.665), and a calibration curve showed good agreement between prediction and observation in the probability of 3- or 5-year survival (Figure 2C and 2D).

## **Survival curves**

The OS curves are shown in Figure 3 based on different cohorts. As presented in the total sample, there is a significant difference between lymph node dissection (LND) and non-lymph node dissection (NLND), indicating that the survival rate of LND is better than that of NLND. However, there was no significant difference between LND and sentinel lymph node biopsy (SLNB). At the same time, there was significant difference between SLNB and NLND. Further analysis was performed to identify the effects of other factors in Figure 3. Patients were divided into six subgroups based on age, histology and stage. In stage IB, older age ( $\geq 65$ ) and SCC subgroups, there was no significant difference between SLNB and NLND curves. The survival rate of SLNB was better than that of NLND in stage IA, younger age ( $< 65$ ) and adenocarcinoma subgroups. There was no significant difference between the curves of SLNB and LND in all subgroups, but there was significant difference between the curves of LND and NLND in all subgroups.

## **Discussion**

Nowadays, nomograms are widely used as prognostic devices in oncology and medicine[17-20]. Nomograms depended on user-friendly digital interfaces, increased accuracy, and more easily understood prognoses to aid better clinical decision making[20]. The purpose of our study is to establish nomogram to predict the prognosis of stage I NSCLC and to guide clinical decision making.

We developed and validated a novel prediction tool for prognosis among stage I NSCLC patients merely using seven easily available variables. Incorporating demographic, disease, and therapy features' risk factors into an easy-to-use nomogram facilitates the stage I NSCLC individualized prediction of prognosis. This study provided a relatively accurate prediction tool of prognosis for stage I NSCLC patients. Validation in the cohort demonstrated good discrimination and calibration power; especially our C-index in the validation identified that this nomogram can be widely and accurately used for its large sample size.

In this study, age was an independent predictor. It can be seen that the risk in the older age group ( $\geq 65$ ) is higher than that in the younger ( $< 65$ ) age group. The weight score of the older age group ( $\geq 65$ ) was 48 in nomogram, while the weight score of the younger ( $< 65$ ) age group was only 0. It demonstrated that for stage I NSCLC, patients in the younger age group can get a better prognosis.

Race was also an independent predictor. However, race had no statistical significance ( $P=0.864$ ) in Cox proportional hazard regression. This factor was included only to make the prediction model evaluation more comprehensive. Another possible reason was that the white group accounts for an excessive proportion (84.4%) of the entire sample, which may have an impact on the statistics. According to another study[21] based on the SEER database, it was pointed out that there were differences in prognosis among different races for advanced lung cancer with bone metastasis. However, there was still no convincing evidence that whether there was a difference in prognosis among different races in early lung cancer.

Gender was another independent predictor. The prognosis of male group was worse than that of female group. In nomogram, the weight score of the male group was 42, while that of the female group was only 0. This may be because men are more likely to smoke than women, and smoking was an independent predictor of lung cancer prognosis.

Pathological grade was also an independent predictor. The pathological grade was divided into four grades from highly differentiated to undifferentiated. The weighted score of undifferentiation in nomogram was 100, indicating that the prognosis of undifferentiated patients was the worst. The scores of low differentiation, middle differentiation and high differentiation were 68, 32 and 0 respectively. It demonstrated that the prognosis of high differentiation was the best. This was also consistent with NCCN lung cancer treatment guidelines. At the same time, pathological classification was another independent predictor. The weighted score of adenocarcinomas was 0 in nomogram and 32 in squamous cell carcinoma. It is suggested that the prognosis of adenocarcinoma group was better than that of squamous cell carcinoma group.

As an independent predictor of clinical stage, the prognosis of IB group was worse than that of IA group. In nomogram, the weight score of IB group was 32 and that of IA group was 0. The tumor size of patients in IB group is larger than that in IA group. According to the results of a study, patients in IB group may have tumor micro metastasis, so the prognosis was worse than that in IA group.

The mode of lymph node dissection was an important independent predictor. Current treatment guidelines for stage I NSCLC recommend lymph node dissection or sentinel lymph node biopsy. In nomogram, the weight score of lymph node dissection group was 0 and that of non-lymph node dissection group was 52. This demonstrated that the prognosis of the lymph node dissection group was better than that of the non-lymph node dissection group. At the same time, in the KM survival curves, in different subgroups, the prognosis of the lymph node dissection group was better than that of the non-lymph node dissection group, and the results were statistically significant. However, the current cancer treatment is becoming more and more standardized and personalized, so lymph node dissection after sentinel lymph node biopsy may become a more useful way. In our study, the results of the sentinel lymph node biopsy group were not statistically significant. This suggests that sentinel lymph node biopsy and lymph node dissection may have similar effects on prognosis. Some studies[22] have shown that patients with positive results of sentinel lymph node biopsy followed by lymph node dissection may have a better prognosis. Direct lymph node dissection without sentinel lymph node biopsy may cause more complications and affect the long-term prognosis of patients[23, 24]. Based on these results, sentinel lymph node biopsy after surgical resection of the tumor may be a better way.

## Conclusion

This study developed a novel nomogram with a relatively good accuracy to help clinicians access the prognosis in stage I NSCLC patients when starting follow-up. For better prognosis, sentinel lymph node biopsy was recommended for patients with stage I NSCLC.

## Declarations

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

### Competing interests

None of the authors has any conflict of interests.

### Funding

Not applicable

## Author contributions

Study design: Yanzheng Song.

Data collection: Lin Wang, Liwei Wu, Laiyi Wan.

Data analysis: Liwei Wu, Lin Wang.

Writing: Lin Wang and Liwei Wu.

## Acknowledgments

I would like to thank Dr. Zhu for her generous and sincere contribution and for providing very practical suggestions in the process of writing the article.

## References

- 1 Raez LE, Cardona AF, Santos ES, et al. The burden of lung cancer in Latin-America and challenges in the access to genomic profiling, immunotherapy and targeted treatments. Lung cancer (Amsterdam, Netherlands) 2018;119:7-13.
- 2 Maisonneuve P, Raminelli C, Bertolotti R, et al. Low-dose computed tomography screening for lung cancer in people with workplace exposure to asbestos. Lung cancer (Amsterdam, Netherlands) 2019;131:23-30.
- 3 Roach MC, Rehman S, DeWees TA, et al. It's never too late: Smoking cessation after stereotactic body radiation therapy for non-small cell lung carcinoma improves overall survival. Practical radiation oncology 2016;6(1):12-8.
- 4 Ho JC, Leung CC. Management of co-existent tuberculosis and lung cancer. Lung cancer (Amsterdam, Netherlands) 2018;122:83-7.
- 5 Löfling L, Karimi A, Sandin F, et al. Clinical characteristics and survival in non-small cell lung cancer patients by smoking history: a population-based cohort study. Acta oncologica (Stockholm, Sweden) 2019;58(11):1618-27.
- 6 Oskarsdottir GN, Halldorsson H, Sigurdsson MI, et al. Lobectomy for non-small cell lung carcinoma: a nationwide study of short- and long-term survival. Acta oncologica (Stockholm, Sweden) 2017;56(7):936-42.
- 7 Sukrithan V, Barbaro A, Chergui A, et al. Differential Efficacy of Anti-VEGF Antibodies Based on Sex and Race in a Diverse Cohort of Advanced Nonsquamous Non-Small Cell Lung Cancer. American journal of clinical oncology 2020;43(1):64-8.

- 8 Erickson P, Gardner LD, Loffredo CA, et al. Racial and Ethnic Differences in the Relationship between Aspirin Use and Non-Small Cell Lung Cancer Risk and Survival. *Cancer epidemiology, biomarkers & prevention* : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2018;27(12):1518-26.
- 9 Sineshaw HM, Wu XC, Flanders WD, et al. Variations in Receipt of Curative-Intent Surgery for Early-Stage Non-Small Cell Lung Cancer (NSCLC) by State. *Journal of thoracic oncology* : official publication of the International Association for the Study of Lung Cancer 2016;11(6):880-9.
- 10 Ferguson MK, Demchuk C, Wroblewski K, et al. Does Race Influence Risk Assessment and Recommendations for Lung Resection? A Randomized Trial. *The Annals of thoracic surgery* 2018;106(4):1013-7.
- 11 Gibson AJW, D'Silva A, Elegbede AA, et al. Impact of Asian ethnicity on outcome in metastatic EGFR-mutant non-small cell lung cancer. *Asia-Pacific journal of clinical oncology* 2019;15(6):343-52.
- 12 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. *Journal of clinical oncology* : official journal of the American Society of Clinical Oncology 2015;33(8):861-9.
- 13 Zhang Y, Sun Y, Wang R, et al. Meta-analysis of lobectomy, segmentectomy, and wedge resection for stage I non-small cell lung cancer. *Journal of surgical oncology* 2015;111(3):334-40.
- 14 Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. *The Lancet Oncology* 2015;16(4):e173-80.
- 15 Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. *Journal of clinical oncology* : official journal of the American Society of Clinical Oncology 2008;26(8):1364-70.
- 16 Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics in medicine* 2004;23(13):2109-23.
- 17 Callegaro D, Miceli R, Mariani L, et al. Soft tissue sarcoma nomograms and their incorporation into practice. *Cancer* 2017;123(15):2802-20.
- 18 Garcia P, Ruiz W, Loza Munárriz C. Warfarin initiation nomograms for venous thromboembolism. *The Cochrane database of systematic reviews* 2016(1):CD007699.
- 19 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. *Critical reviews in oncology/hematology* 2018;126:13-8.

20 Caulfield S, Menezes G, Marignol L, et al. Nomograms are key decision-making tools in prostate cancer radiation therapy. *Urologic oncology* 2018;36(6):283-92.

21 Xu G, Cui P, Zhang C, et al. Racial disparities in bone metastasis patterns and targeted screening and treatment strategies in newly diagnosed lung cancer patients. *Ethnicity & health* 2020;1-14.

22 Melfi FM, Chella A, Menconi GF, et al. Intraoperative radioguided sentinel lymph node biopsy in non-small cell lung cancer. *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery* 2003;23(2):214-20.

23 Karamustafaoglu YA, Yoruk Y, Yanik F, et al. Sentinel lymph node mapping in patients with operable non-small cell lung cancer. *Journal of thoracic disease* 2013;5(3):317-20.

24 Digesu CS, Hachey KJ, Gilmore DM, et al. Long-term outcomes after near-infrared sentinel lymph node mapping in non-small cell lung cancer. *The Journal of thoracic and cardiovascular surgery* 2018;155(3):1280-91.

## Tables

Table 1 Patients characteristics

Characteristic	Training cohort (n)	Verification cohort (n)
Age		
<65	7222 (30.9)	2152 (30.7)
=>65	16172 (69.1)	4864 (69.3)
Race		
Black	1874 (8.0)	566 (8.1)
White	19736 (84.4)	5942 (84.7)
Other	1784 (7.6)	508 (7.2)
Gender		
Female	12787 (54.7)	3878 (55.3)
Male	10607 (45.3)	3138 (44.7)
Grade		
Grade I well differentiated	5612 (24.0)	1643 (23.4)
Grade II moderately differentiated	11106 (47.5)	3330 (47.5)
Grade III poorly differentiated	6504 (27.8)	1987 (28.3)
Grade IV undifferentiated	172 (0.7)	56 (0.8)
Histology		
SCC	6631 (28.3)	2009 (28.6)
Adenocarcinoma	16763 (71.7)	5007 (71.4)
Stage		
IA	14371 (61.4)	4264 (60.8)
IB	9023 (38.6)	2752 (39.2)
Surgery		
Removed	20421 (87.3)	6112 (87.1)
None	2860 (12.2)	870 (12.4)
Sentry	113 (0.5)	34 (0.5)

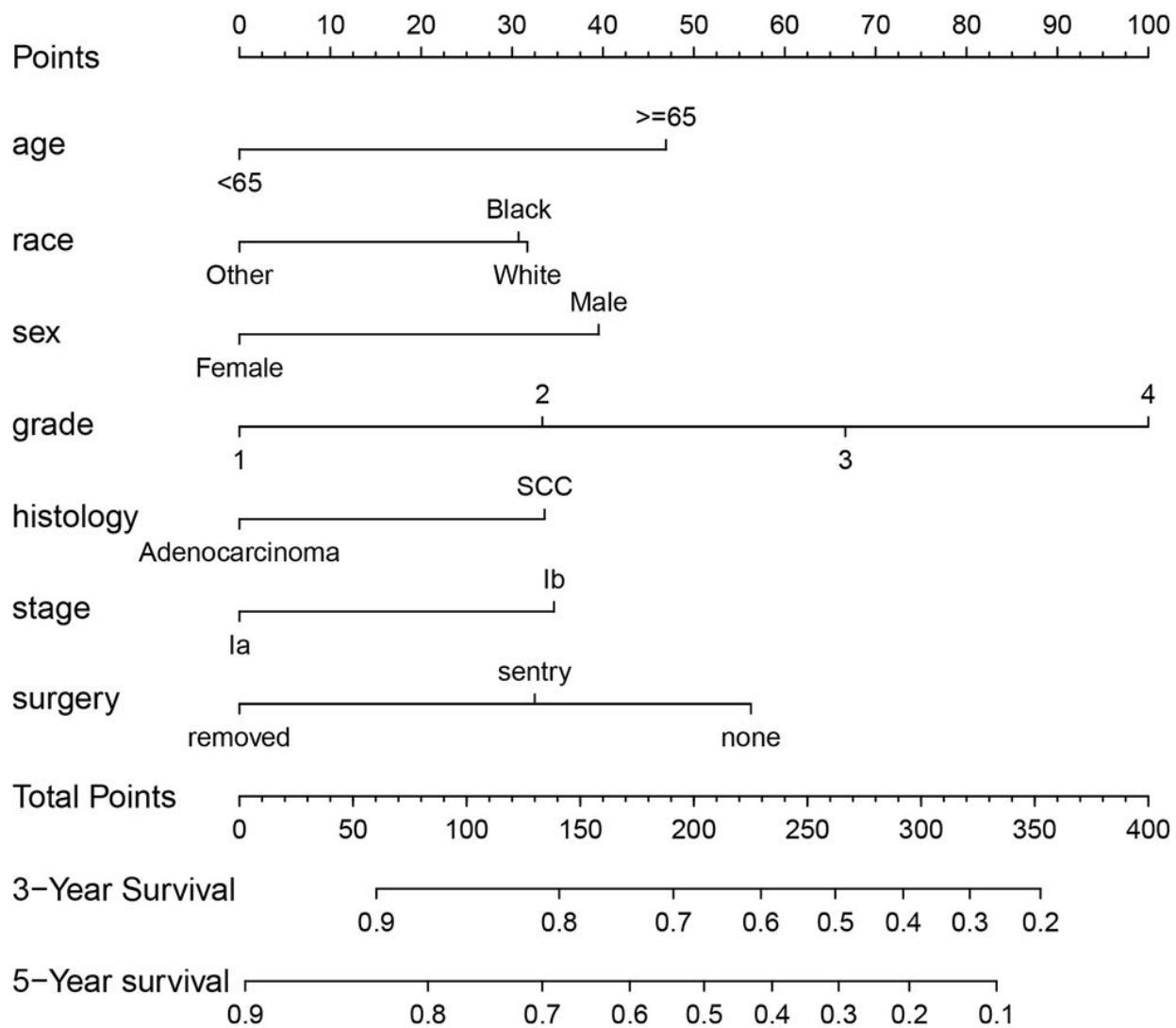
Notes: Removed means lymph node dissection; None means no lymph node dissection; Sentry means sentinel lymph node biopsy.

Table 2 Identification of independent prognostic factors for the development cohort

Variable	Overall Survival		
	P	HR	95% CI
Age (<65 as ref.)			
>=65	<0.001	1.549	1.448-1.657
Race (black as ref.)			
White	0.864	1.009	0.909-1.119
Other	<0.001	0.751	0.642-0.877
Gender (female as ref.)			
Male	<0.001	1.446	1.366-1.530
Grade (I as ref.)			
II	<0.001	1.696	1.553-1.853
III	<0.001	2.084	1.896-2.290
IV	<0.001	2.212	1.639-2.986
Histology (SCC as ref.)			
Adenocarcinoma	<0.001	0.737	0.694-0.783
Stage (IA as ref.)			
IB	<0.001	1.374	1.298-1.459
Surgery (removed as ref.)			
None	<0.001	1.701	1.578-1.833
Sentry	0.105	1.346	0.940-1.929

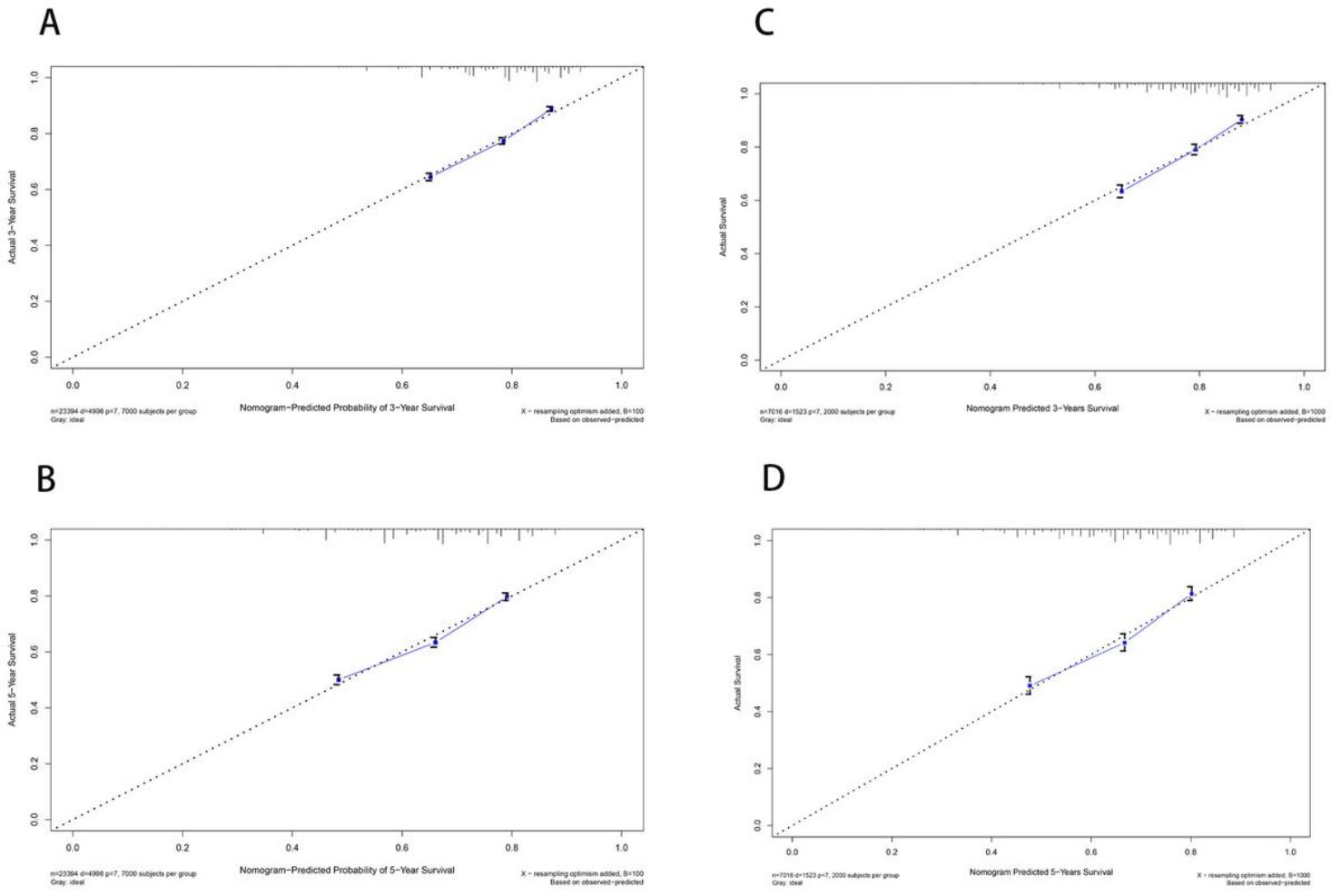
Notes: Removed means lymph node dissection; None means no lymph node dissection; Sentry means sentinel lymph node biopsy.

## Figures



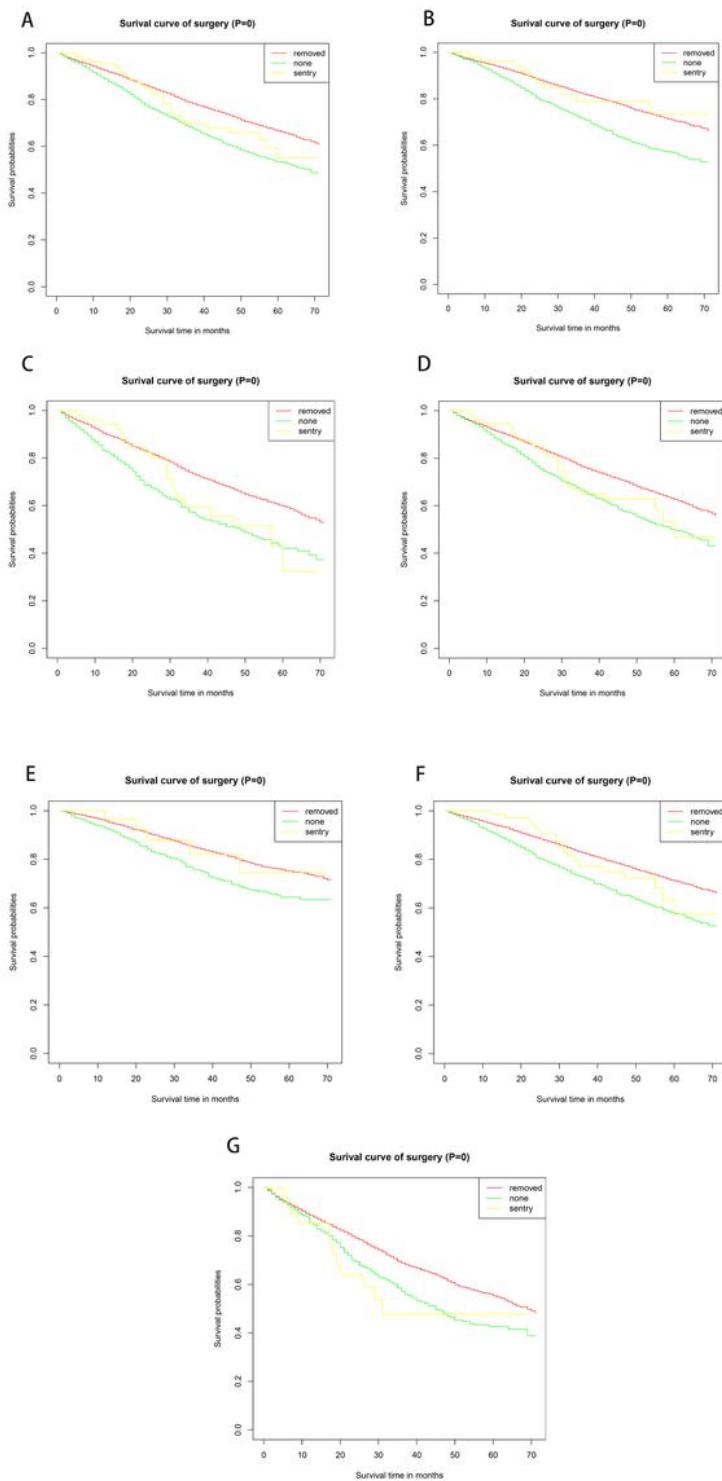
**Figure 1**

Constructed nomogram for predicting survival in patients with stage I non-small cell lung cancer undergoing lymph node dissection. SC, squamous cell carcinoma.



**Figure 2**

Calibration plots demonstrating the consistency between the nomogram predicted and actual observed 3- and 5-year survival in (A, B) development cohort and (C, D) validation cohort.



**Figure 3**

Overall survival of removed, none, sentry for the total sample(A), stage IA(B), stage IB(C),  $\geq 65$ (D),  $< 65$ (E), adenocarcinoma(F) and squamous cell carcinoma(G).