


Establishment and Validation of a Nomogram Based on Negative Lymph Nodes to Predict Survival in Postoperative Patients with non-Small Cell Lung Cancer

Technology in Cancer Research & Treatment
Volume 21: 1-11
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15330338221074506
journals.sagepub.com/home/tct


Xinyi Huang, PhD^{1,2,3,*} , Pingping Hu, PhD^{3,*}, Fei Yan, MD⁴, and Jiandong Zhang, PhD³

Abstract

Background: The importance of the negative lymph node (NLN) count has recently attracted attention. This study aimed to determine the prognostic value of NLN count in patients with non-small cell lung cancer (NSCLC) after radical surgery by constructing NLN-based prognostic models. **Methods:** This study included 33 756 patients pooled from the case listing session of the US Surveillance, Epidemiology, and End Results (SEER) database from 2004 to 2015 and 545 patients collected from The First Affiliated Hospital of Shandong First Medical University between 2012 and 2016. X-tile software was used to calculate the optimal cutoff value for the NLN count. The associated clinical factors were determined using univariate and multivariate Cox analyses. Nomograms were developed using the SEER database and validated using hospital data. **Results:** The training cohort was divided into high and low NLN count subgroups based on the cancer-specific survival (CSS) and overall survival (OS), respectively. Multivariate analysis showed that NLN count was an independent prognostic factor, and the high NLN count subgroup had better CSS and OS than those of the low NLN count subgroup (HR = 0.632, 95% CI 0.551-0.724, $P < .001$ for CSS and HR = 0.641, 95% CI 0.571-0.720, $P < .001$ for OS). Nomograms were established, exhibiting good discrimination ability with a C-index of 0.789 (95% CI 0.778 – 0.798) for CSS and 0.704 (95% CI, 0.694 – 0.714) for OS. The calibration plots of the validation cohorts showed optimal agreement with the training cohort, with a C-index of 0.681 (95% CI 0.646 – 0.716) for CSS and 0.645 (95% CI 0.614 – 0.676) for OS. **Conclusions:** NLN count is a strong prognostic factor for OS and CSS in NSCLC patients and the prognostic model provides a useful risk stratification for NSCLC patients when applied to clinical practice.

Keywords

nomogram, non-small cell lung cancer, negative lymph node count, cancer specific survival, overall survival

Received: September 26, 2021; Revised: December 19, 2021; Accepted: December 24, 2021.

¹ The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China

² Shandong Cancer Hospital and Institute, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China

³ The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Shandong Key Laboratory of Rheumatic Disease and Translational Medicine, Shandong Lung Cancer Institute, Jinan, China

⁴ Dezhou Seventh People's Hospital, Dezhou, China

*These authors contributed equally to this work.

Corresponding Author:

Jiandong Zhang, Department of Oncology, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Shandong Key Laboratory of Rheumatic Disease and Translational Medicine, Shandong Lung Cancer Institute, Jinan, China.

Email: zhangjd2233@126.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Introduction

Lung cancer is one of the most common malignant tumors in the world and the leading cause of cancer-related death in the United States.¹ The incidence and mortality of lung cancer in China have increased significantly in the past 30 years, and non-small cell lung cancer (NSCLC) is the most common pathological type of lung cancer.² Despite numerous advances in early diagnosis and adjuvant therapies, surgical resection is currently the only potentially curative therapeutic means for NSCLC and the 5-year overall survival rate remains only approximately 17%, even after surgical management.^{3,4} In the conventional sense, many factors such as positive lymph node (PLN) count and removed lymph node (LN) count are commonly used to predict survival in cancer outcomes.⁵⁻⁷ However, these predictors fail to clearly distinguish between patients with different negative lymph node (NLN) counts and thus do not predict which patients are more likely to benefit from lymphadenectomy. Therefore, researchers recently recognized the significance of a novel prognostic indicator, NLN count, which showed significant and promising prognostic value for cancer patients. Several studies have highlighted the clinical significance of NLN count in lung cancer and other types of cancer, such as gastric carcinoma, colon cancer, rectal, esophageal, breast, and cholangiocarcinoma.⁸⁻¹³ However, the limitations of these studies included the use of single-center cohorts, small study populations, and lack of validation.

In our study, a multivariate predictive model based on the Surveillance, Epidemiology, and End Results (SEER) database was established, which aimed to provide visualization information and offer extra precision in predicting survival outcomes. In addition, data from The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial

Qianfoshan Hospital was also obtained to identify and validate. We hope that this predictive model could benefit lung patients after surgery and provide more insights into the molecular mechanisms of this disease.

Material and Methods

Patient Selection and Eligibility Criteria

In the current cohort study, 67,001 eligible patients with single primary NSCLC diagnosed between 2004 and 2015 were obtained from the SEER database released on August 17, 2021 (SEER*Stat Database: Incidence-SEER 18 Registries Research Plus Data, Nov. 2019 Sub [2000-2017]). The case listing module function of the SEER*Stat software (Surveillance Research Program, National Cancer Institute, www.seer.cancer.gov/seerstat; Version 8.3.9.2) was used to access the data. Clinicopathological characteristics, including age at diagnosis, sex, the sixth American Joint Committee on Cancer (AJCC) cancer tumor node metastasis (TNM) stage, grade, histology (ICD-O-3, 8046/3), type of surgery, radiation, chemotherapy, cause of death, survival time (months), regional lymph node positivity, and regional lymph node examination, were recorded. Patients without pathological confirmation were excluded from the study. The other specific exclusion processes are shown in Figure 1. Ultimately, 33,756 patients were enrolled, while patients who underwent lymphadenectomy were selected in our training cohort. The regional nodes examined were defined as the total number of regional lymph nodes that were removed and examined by a pathologist. Positive regional nodes refer to the exact number of regional lymph nodes examined by the pathologist that were found to contain metastases. The number of negative regional nodes, also called the NLN

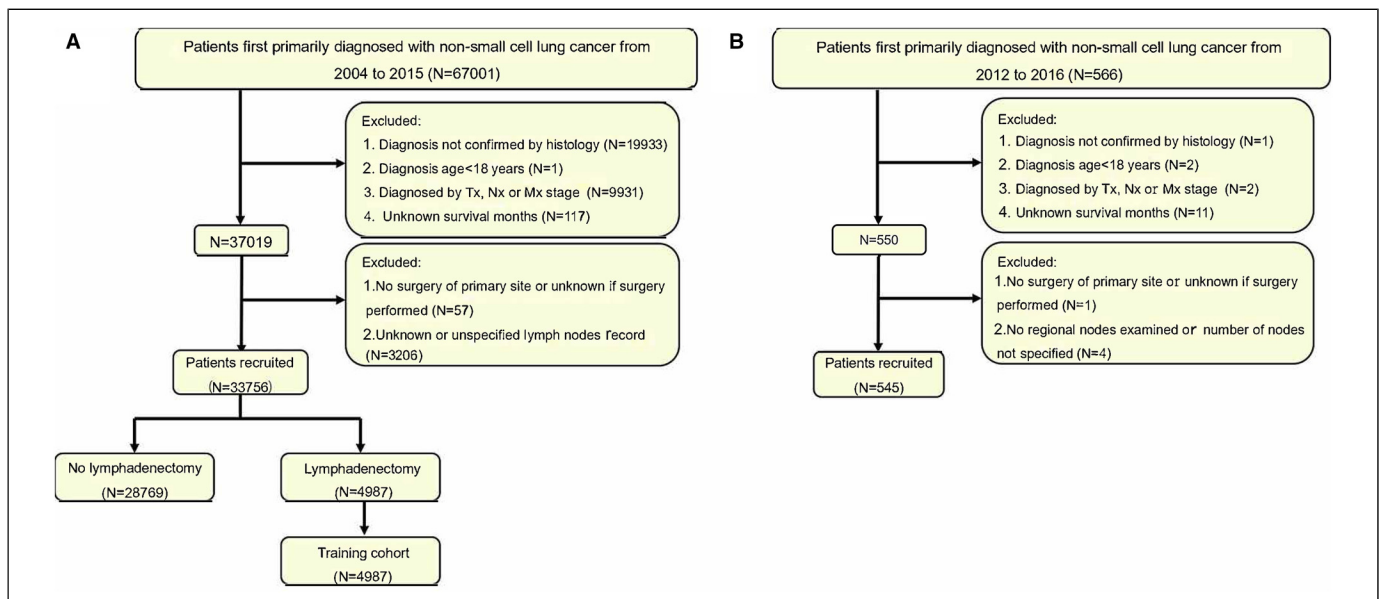


Figure 1. Flow diagrams. (A) Flow diagram of the training cohort. (B) Flow diagram of the validation cohort.

count, was defined as the total number of regional nodes minus the number of positive regional nodes. For cancer-specific survival (CSS), death due to NSCLC was the primary endpoint, and for overall survival (OS), death due to any cause was an event.

Establishment of Nomogram

All variables in the training cohort were analyzed using univariate analysis with a Cox regression model to explore the association between the baseline variables and survival outcomes. Variables with a *P*-value <.10 in the univariate analysis were included in the multivariate analysis. A stepwise forward Cox regression analysis was then used to identify independent predictive factors in the multivariate analysis. Nomograms that predicted 3-year and 5-year survival were constructed based on independent risk factors in the final multivariate Cox proportional hazards model results. A

nomogram is an intuitive and effective method to display the results of risk models that has significant clinical applications in predicting prognosis. The first line of the nomogram panel represents the point values for each variable. The total points were calculated by adding the points identified for each variable. A vertical line can be drawn on the bottom scale at the total point value to indicate the 3- and 5-year probabilities of CSS or OS.

Model Validation

To validate the findings from the SEER database and clarify relevant issues, a dataset containing 545 patients with NSCLC from The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital between 2012 and 2016 was assigned to the validation cohort. The inclusion and exclusion criteria were consistent with those in the SEER database (Figure 1B). The

Table 1. Demographic and Clinicopathological Characteristics of Patients with NSCLC

| Characteristics | Training cohort | | | <i>P</i> -value | Validation cohort | | | <i>P</i> -value |
|-----------------|-----------------|--|--|-----------------|-------------------|---------------------------------------|--|-----------------|
| | All N = 4987 | Low NLN count NO. (%) 4481 (89.9) | High NLN count NO. (%) 506 (10.1) | | All N = 545 | Low NLN count NO. (%) 143(26.3) | High NLN count NO. (%) 402(73.7) | |
| Age | | | | .861 | | | | .150 |
| ≤65 | 2295 (46.0) | 2064 (46.1) | 231 (45.7) | | 373 (68.4) | 91 (63.6) | 282 (70.1) | |
| >65 | 2692 (54.0) | 2417 (53.9) | 275 (43.3) | | 172 (31.6) | 52 (36.4) | 120 (29.9) | |
| Sex | | | | .039 | | | | .009 |
| Female | 2186 (43.8) | 1986 (44.3) | 200 (39.5) | | 152 (27.9) | 52 (36.4) | 100 (24.9) | |
| Male | 2801 (56.2) | 2495 (55.7) | 306 (60.5) | | 393 (72.1) | 91 (63.6) | 302 (75.1) | |
| TNM stage | | | | <.001 | | | | .193 |
| I | 1624 (32.6) | 1392 (31.1) | 232 (45.8) | | 207 (38.0) | 51 (35.7) | 156 (38.8) | |
| II | 519 (10.4) | 427 (9.5) | 92 (18.2) | | 154 (28.3) | 36 (25.2) | 118 (29.4) | |
| III | 1750 (35.1) | 1616 (36.1) | 134 (26.5) | | 155 (28.4) | 44 (30.8) | 111 (27.6) | |
| IV | 1094 (21.9) | 1046 (23.3) | 48 (9.5) | | 29 (5.3) | 12 (8.4) | 17 (4.2) | |
| Tumor grade | | | | <.001 | | | | .169 |
| I | 67 (1.3) | 59 (1.3) | 8 (1.6) | | 41 (7.5) | 16 (11.2) | 25 (6.2) | |
| II | 356 (7.1) | 315 (7.0) | 41 (8.1) | | 399 (73.2) | 99 (69.2) | 300 (74.6) | |
| III | 2497 (50.1) | 2176 (48.6) | 321 (63.4) | | 89 (16.3) | 22 (15.4) | 67 (16.7) | |
| IV | 193 (3.9) | 170 (3.8) | 23 (4.5) | | 16 (2.9) | 6 (4.2) | 10 (2.5) | |
| Unknown | 1874 (37.6) | 1761 (39.3) | 113 (22.3) | | | | | |
| Surgery type | | | | <.001 | | | | .495 |
| Sublobectomy | 373 (7.5) | 358 (8.0) | 15 (3.0) | | 50 (9.2) | 10 (7.0) | 40 (10.0) | |
| Lobectomy | 2331 (46.7) | 1950 (43.5) | 381 (75.3) | | 417 (76.5) | 108 (75.5) | 309 (76.9) | |
| Pneumonectomy | 223 (4.5) | 150 (3.3) | 73 (14.4) | | 75 (13.8) | 24 (16.8) | 51 (12.7) | |
| Others | 2060 (41.3) | 2023 (45.1) | 37 (7.3) | | 3 (0.6) | 1 (0.7) | 2 (0.5) | |
| Radiotherapy | | | | <.001 | | | | .003 |
| Yes | 1988 (39.9) | 2631 (58.7) | 138 (27.3) | | 211 (38.7) | 70 (49.0) | 141 (35.1) | |
| No/Unknown | 2999 (60.1) | 1850 (41.3) | 368 (72.7) | | 334 (61.3) | 73 (51.0) | 261 (64.9) | |
| Chemotherapy | | | | <.001 | | | | .003 |
| Yes | 2531 (50.8) | 2293 (51.2) | 238 (47.0) | | 210 (38.5) | 70 (49.0) | 140 (34.8) | |
| No/Unknown | 2456 (49.2) | 2188 (48.8) | 268 (53.0) | | 335 (61.5) | 73 (51.0) | 262 (65.2) | |

Notes: In the training set: Low NLN count subgroup: 0 to 14 NLN count. High NLN count subgroup: ≥15 NLN count; In the validation set: Low NLN count subgroup: 0 to 6 NLN count. High NLN count subgroup: ≥7 NLN count.

Abbreviations: NLN, negative lymph node; TNM, tumor, node, metastases; Tumor grade: I: highly differentiated; II: moderately differentiated; III: poorly differentiated; IV: undifferentiated; Unknown: unknown tumor differentiation.

Concordance index (C-index) was used to evaluate the discrimination performance of the nomogram. Calibration plots (1000 bootstrap resamples) were used to visualize the performances of the nomograms. The 45° line represents the best prediction.^{14,15}

Statistical Methods

The optimal cutoff value of the NLN count was determined using X-tile software (Version 3.6.1, Yale University School of Medicine, USA). Continuous and categorical variables were compared using the χ^2 test and Mann–Whitney U-test, respectively. The Kaplan–Meier method was used to assess the effect of NLN count on CSS and OS. Univariate and multivariate Cox proportional hazard models with a stepwise forward method were used to analyze the risk factors related to CSS and OS in patients with NSCLC; hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated, with an HR of <1.0 indicating survival. Receiver operating characteristic (ROC) and decision curve analysis (DCA) curves were used to analyze sensitivity and specificity between the nomogram and TNM stage.

All statistical analyses were performed using R ver. 3.5.0 (the R Foundation for Statistical Computing, Vienna, Austria), and SPSS ver. 20.0 (SPSS Inc., Chicago, IL, USA) software. All statistical tests were two-sided, with a *P*-value <.05 considered statistically significant.

Ethics Statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The request for the SEER Research Plus Data was approved to obtain therapy information for research purposes only (reference number: 18323-Nov2020). Institutional review board approval was not required because SEER is an anonymized database that is open to the public. Informed patient consent from the validation cohort was approved by The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital review board (No. S033). The study was conducted in accordance with the principles of the Declaration of Helsinki (as revised in 2013).

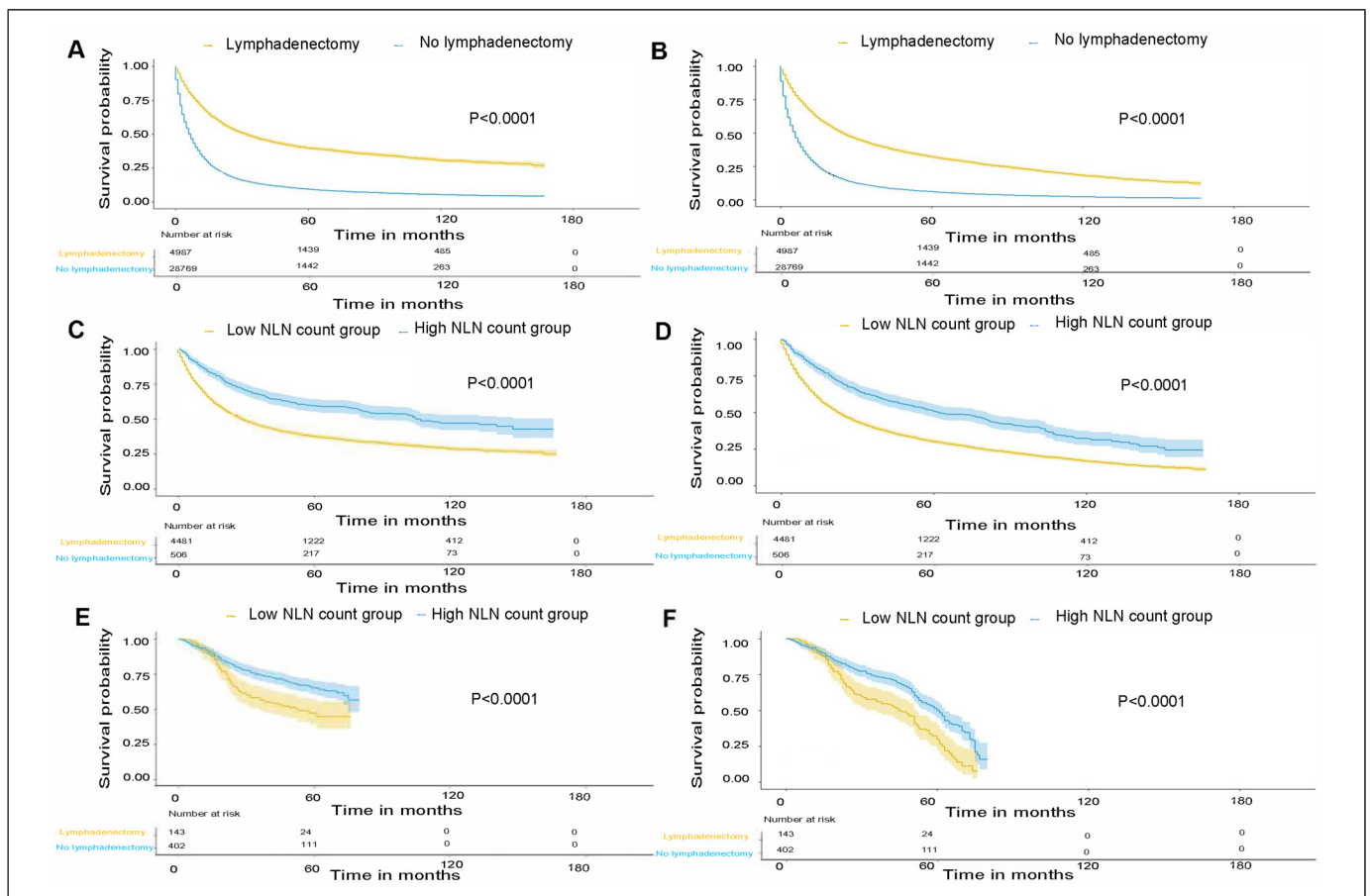


Figure 2. Survival analysis. (A) Prognostic value of lymphadenectomy in CSS in the training set. (B) Prognostic value of lymphadenectomy in OS in the training set. (C) CSS for high (≥ 15) and low (0-14) NLN count subgroups in the training set. (D) OS for high (≥ 15) and low (0-14) NLN count subgroup in the training set. (E) CSS for high (≥ 7) and low (0-6) NLN count subgroup in the validation set. (F) OS for high (≥ 7) and low (0-6) NLN count subgroup in the validation set.

Table 2. Cox proportional hazard model analysis for CSS and OS in the training cohort

| Covariates | CSS | | | | OS | | | |
|---------------|---------------------|---------|-----------------------|---------|------------------------|---------|-----------------------|---------|
| | Univariate analysis | | Multivariate analysis | | Univariate analysis | | Multivariate analysis | |
| | HR (95% CI) | P-value | HR (95% CI) | P-value | HR (95% CI) | P-value | HR (95%CI) | P-value |
| Age | | | | | | | | |
| ≤65 | 1 | | 1 | | 1 | | 1 | |
| >65 | 1.117 (1.040-1.199) | .002 | 1.232 (1.144-1.320) | <.001 | 1.288 (1.209-1.372) | <.001 | 1.370 (1.285-1.461) | <.001 |
| Sex | | | | | | | | |
| Female | 1 | | 1 | | 1 | | 1 | |
| Male | 1.204 (1.121-1.293) | <.001 | 1.229 (1.144-1.320) | <.001 | 1.193 (1.120-1.271) | <.001 | 1.220 (1.145-1.299) | <.001 |
| TNM stage | | | | | | | | |
| I | 1 | | 1 | | 1 | | 1 | |
| II | 1.643 (1.427-1.892) | <.001 | 1.979 (1.709-2.292) | <.001 | 1.355 (1.206-1.523) | <.001 | 1.691 (1.497-1.909) | <.001 |
| III | 2.726 (2.474-3.003) | <.001 | 3.230 (2.886-3.615) | <.001 | 1.999 (1.846-2.164) | <.001 | 2.488 (2.264-2.734) | <.001 |
| IV | 6.136 (5.531-6.808) | <.001 | 7.181 (6.401-8.057) | <.001 | 4.265 (3.904-4.658) | <.001 | 5.211 (4.722-5.751) | <.001 |
| Tumor grade | | | | | | | | |
| I | 1 | <.001 | | | 1 | <.001 | | |
| II | 1.051 (0.698-1.583) | .811 | | | 0.981 (0.711-1.354) | .908 | | |
| III | 1.575 (1.076-2.305) | .019 | | | 1.322 (0.981-1.783) | .067 | | |
| IV | 1.503 (0.984-2.295) | .059 | | | 1.177 (0.838 to 1.654) | .347 | | |
| Unknown | 2.995 (2.046-4.383) | <.001 | | | 2.169 (1.608-2.927) | <.001 | | |
| Surgery type | | | | | | | | |
| Sublobectomy | 1 | <.001 | | | 1 | <.001 | | |
| Lobectomy | 0.686 (0.591-0.796) | <.001 | | | 0.737 (0.650-0.835) | <.001 | | |
| Pneumonectomy | 1.031 (0.824-1.289) | .789 | | | 0.927 (0.764-1.125) | .441 | | |
| Others | 2.778 (2.404-3.210) | <.001 | | | 2.324 (2.053-2.631) | <.001 | | |
| NLN count | | | | | | | | |
| Low* | 1 | | 1 | | 1 | | 1 | |
| High† | 0.542 (0.473-0.620) | <.001 | 0.632 (0.551-0.724) | <.001 | 0.576 (0.514-0.647) | <.001 | 0.641 (0.571-0.720) | <.001 |
| Radiotherapy | | | | | | | | |
| Yes | 1 | | 1 | | 1 | | 1 | |
| No/Unknown | 0.596 (0.556-0.640) | <.001 | 0.849 (0.781-0.923) | <.001 | 0.691 (0.649-0.737) | <.001 | 0.867 (0.803-0.936) | <.001 |
| Chemotherapy | | | | | | | | |
| Yes | 1 | | 1 | | 1 | | 1 | |
| No/Unknown | 0.767 (0.714-0.824) | <.001 | 1.505 (1.380-1.641) | <.001 | 0.907 (0.852-0.965) | .002 | 1.561 (1.444-1.687) | <.001 |

Notes: *Low NLN count for CSS; †High NLNs for CSS; ≥15 NLNs.

Abbreviations: CSS, cancer specific survival; HR, hazard ratio; CI, confidence interval; TNM, tumor, node, metastases; NLN, negative lymph node.

Results

Patient Characteristics

From 2004 to 2015, a total of 33,756 NSCLC patients were collected from the SEER database, of whom 4987 (14.8%) were assigned to the training set (Figure 1A). Of all patients in the training cohort, 2186 (43.8%) were female and 2801 (56.2%) were male. The median age at diagnosis of the included patients was 67 years (range, 20-95) years. The cutoff value for the NLN count was 15 (range, 0-60). We determined the best cutoff value using X-tile software and thus divided the entire cohort into two subgroups (low NLN count [0-14] and high NLN count [≥ 15]). There were 4481 patients (89.9%) with a low NLN count and 506 patients (10.1%) with a high NLN count. The characteristics of the training cohort are shown in Table 1.

In the validation cohort study, a total of 545 eligible patients with lymphadenectomy information from The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital between 2012 and 2016 were also identified, including 152 (27.9%) women and 393 (72.1%) men. The median age at diagnosis was 60 years (range, 32-84 years). NLN count was divided into two subgroups using X-tile software: low NLN count, 0 to 6; high NLN count, ≥ 7 (range, 0-41). Patient demographics and pathological characteristics are shown in Table 1.

Effects of NLN Count on Survival

First, we determined the impact of lymphadenectomy on survival, and found that patients who underwent lymphadenectomy had improved survival in both CSS and OS (Figure 2A and B, all $P < .001$). Next, we analyzed the prognostic effect of NLN count in the training and validation cohorts. Kaplan-Meier analyses showed that NLN count was positively associated with CSS and OS, and a high NLN count group was associated with better survival in both the training and validation cohorts. (Figure 2C-F, all $P < .001$). In addition, a stronger survival advantage of the NLN count was found in patients with worse M and lymph node (N) status. In other words, the survival benefit trend was constant regardless of M (Supplementary Figure 1, all $P < .05$) and N status (Supplementary Figure 2, all $P < .05$).

The results of univariate and multivariate Cox proportional hazard model analyses are shown in Table 2. Univariate analyses revealed that NLN count and other clinicopathological characteristics including age, sex, TNM stage, radiotherapy and chemotherapy were associated with CSS and OS (all $P < .001$). Multivariate analyses were further performed using the above factors and confirmed that NLN count, age, sex, TNM stage, radiotherapy and chemotherapy were independent prognostic factors for CSS and OS, and patients with high NLN

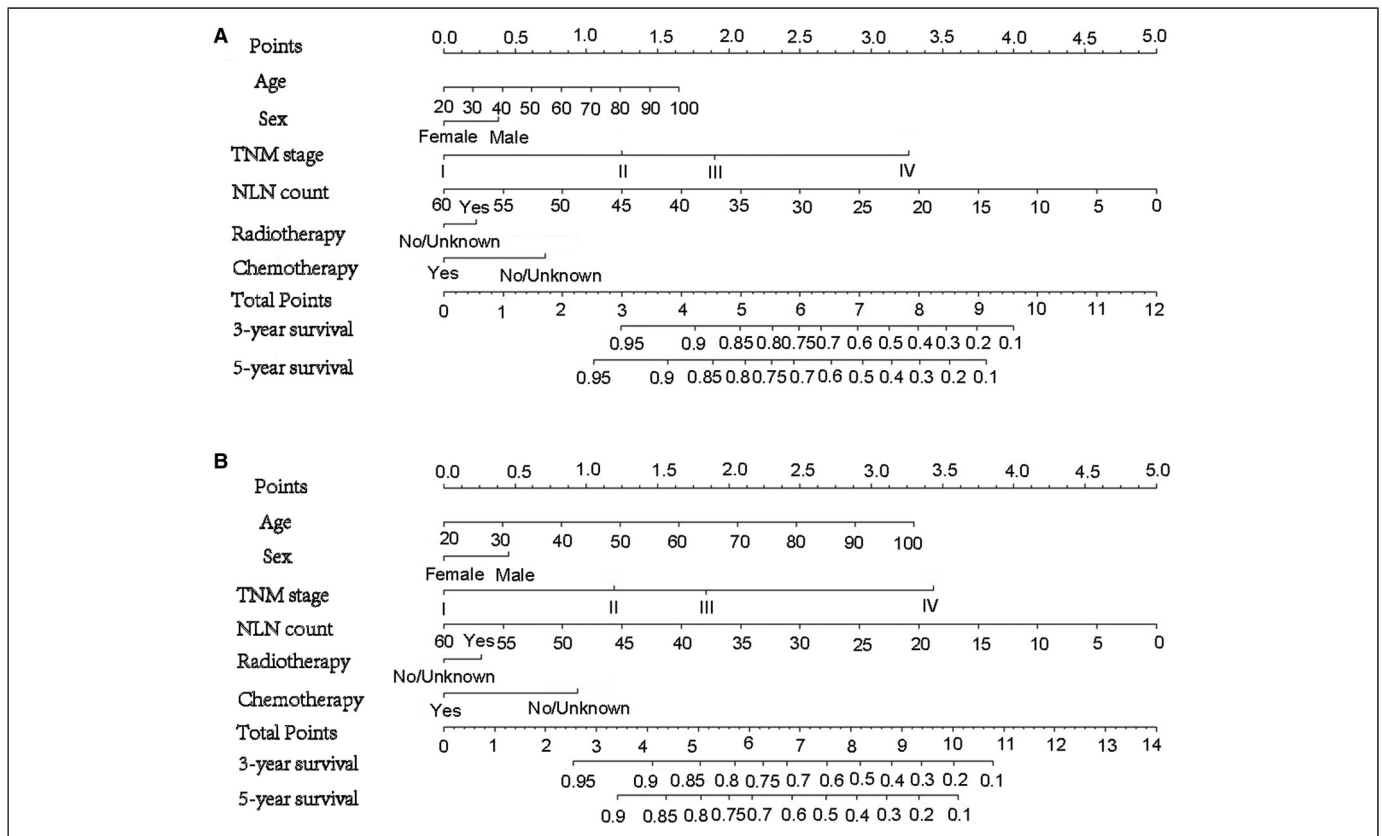


Figure 3. Establishment of the nomogram. (A) Nomogram predicting the 3- and 5-year CSS. (B) Nomogram predicting the 3- and 5-year OS. In the panels, the first row is the points assigned for each variable. The points for each variable are summed to obtain a total. A vertical line is drawn from this point to obtain the 3- and 5-year probabilities of survival.

count had better CSS than those with low NLN count (HR = 0.632, 95% CI: 0.551-0.724, $P < .001$). A consistent trend was also found for OS (HR = 0.641, 95% CI: 0.571-0.720, $P < .001$ [high vs low]). In other words, a higher NLN count was found to have a better effect on survival in NSCLC, independent of sex, age, TNM stage, radiotherapy, and chemotherapy (Table 2).

Construction and Validation of Nomogram

First, we established a nomogram using the results of the final multivariate Cox analysis. (Figure 3A for CSS and Figure 3B

for OS) at 3 and 5 years, respectively. Variables included in the nomogram were the same as those in the multivariate Cox risk model. Next, internal validation was performed, with a C-index estimated by analyzing the SEER database. The c-statistic of the predictive models was 0.789 (95% CI: 0.778 – 0.798) for CSS and 0.704 (95% CI: 0.694 – 0.714) for OS, which indicated a good ability to predict outcome. We then constructed a calibration plot to determine whether the predicted and observed probabilities of survival were in accordance. As expected, the calibration plots for the probability of 3 years and 5 years showed good consistency between the prediction by nomogram and actual observation in survival (Figure 4A

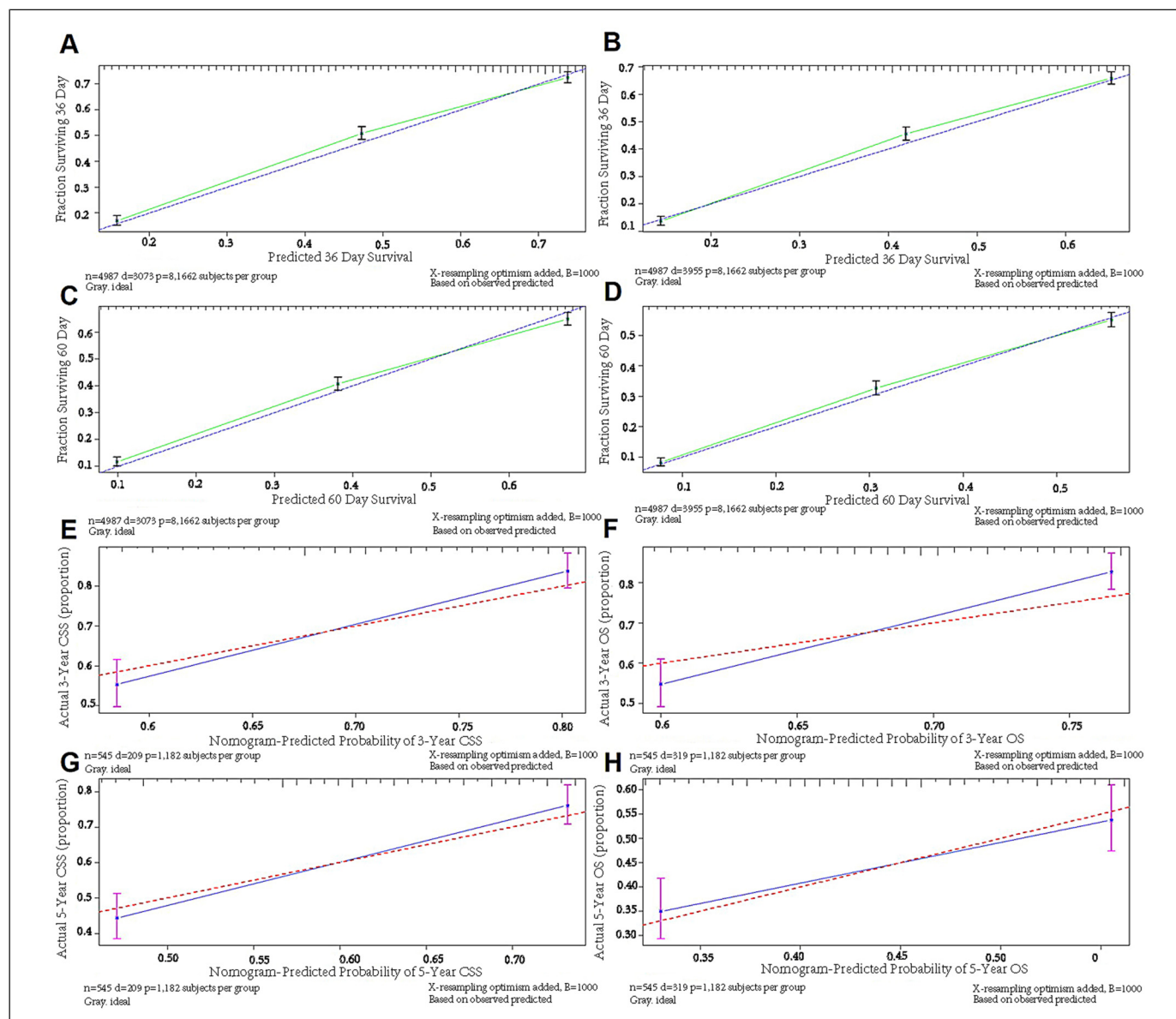


Figure 4. Calibration curves for 3- and 5-year survival. (A) Calibration curve for 3-year CSS based on the training cohort. (B) Calibration curve for 3-year OS based on the training cohort. (C) Calibration curve for 5-year CSS based on the training cohort. (D) Calibration curve for 5-year OS based on the training cohort. (E) Calibration curve for 3-year CSS based on the validation cohort. (F) Calibration curve for 3-year OS based on the validation cohort. (G) Calibration curve for 5-year CSS based on the validation cohort. (H) Calibration curve for 5-year OS based on the validation cohort.

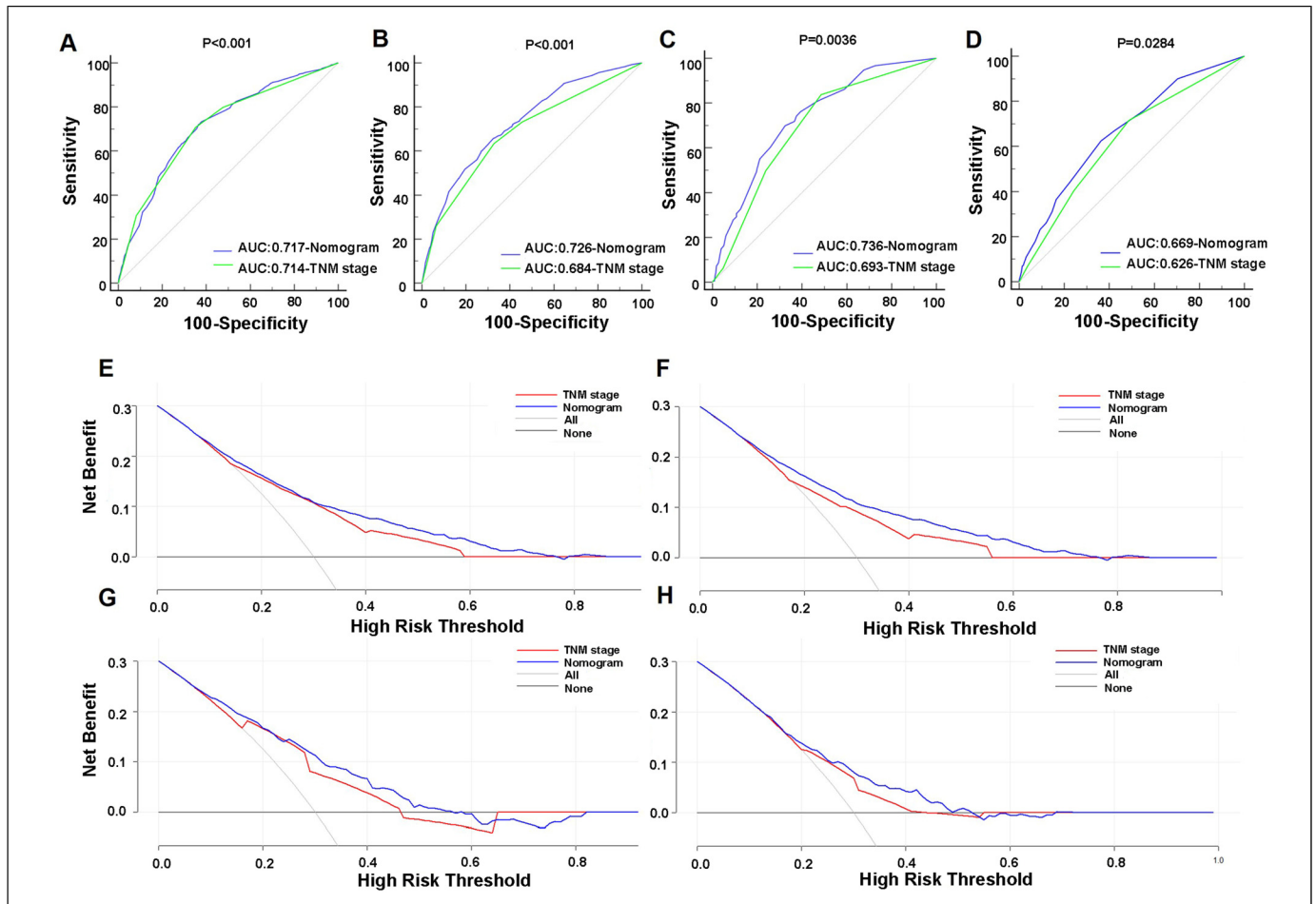


Figure 5. ROC and DCA curves for nomogram and TNM stage. (A) The CSS ROC curve for the nomogram and TNM stage in the training cohort. (B) The OS ROC curve for the nomogram and TNM stage in the training cohort. (C) The CSS ROC curve for the nomogram and TNM stage in the validation cohort. (D) The OS ROC curve for the nomogram and TNM stage in the validation cohort. (E) The CSS ROC curve for the nomogram and TNM stage in the training cohort. (F) The OS ROC curve for the nomogram and TNM stage in the training cohort. (G) The CSS ROC curve for the nomogram and TNM stage in the validation cohort. (H) The OS ROC curve for the nomogram and TNM stage in the validation cohort.

and C for CSS, Figure 4B and D for OS). In addition, the length of the line for each variable in the nomogram indicates the influence and values of the final outcomes. The longer the line, the greater the influence of the variable on the result. In our nomogram, we can see that the length of the line for NLN count was longest for both CSS and OS. Therefore, our nomogram predicts the prognosis of patients with NSCLC intuitively, which is different from traditional clinical predictors, such as TNM stage and other variables.

Finally, external validation was performed using an independent validation cohort of 545 patients from The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital. The C-index measuring predictive accuracy in the test cohort was 0.681 (95% CI: 0.646–0.716) for CSS and 0.645 (95% CI: 0.614–0.676) for OS, which demonstrated a good level of discriminative ability. As expected, the calibration curve suggests that the nomogram was well calibrated and the 3- and 5-year calibration curve showed an agreement between the actual observation and

nomogram prediction (Figure 3E and G for CSS, Figure 3F and H for OS). In addition, we performed ROC curve analysis and DCA curve analysis on nomogram and TNM stage, and found that the performance of the nomogram for CSS and OS was better than that of the TNM stage (Figure 5, $P < .05$). Therefore, the results proved that the nomogram for predicting patient prognosis had excellent advantages both in the short and long term.

Discussion

Although the LN and PLN counts have been considered to be of great value in determining the prognosis of patients with NSCLC,^{16,17} the NLN count is now widely believed to be a better prognostic factor that predicts survival status and provides guidelines for postoperative therapy.^{8,12,18} In our study, we found that NLN count is a significant predictor of survival outcome in a large cohort of NSCLC patients. In addition, we established a nomogram including NLN count based on the

SEER database for the first time. A nomogram is a predictive model that incorporates multiple factors to predict patient survival outcomes. Many studies have shown that the nomogram is more accurate than conventional TNM stage in predicting patient prognosis,^{19,20} which is consistent with our results obtained using ROC and DCA curves. The nomogram was further validated by our own hospital data, which showed good agreement with the prediction and determined the prognostic value and significance of NLN count. These findings contribute to the clinical translation of our results.

As noted above, NLN count positively correlated with survival, and the high NLN count group was found to be associated with better survival with respect to both CSS and OS. Findings from several studies may explain this prognostic phenomenon. First, increasing the number of NLNs examined has been shown to be associated with a significant improvement in survival benefit, thus ensuring accurate staging and adequate treatment, and delivering the optimum adjuvant therapy strategy.^{21,22} Second, extensive lymphadenectomy provides a protective effect for patients, and reduces the risk of understaging.²³ Therefore, better surgery with more extensive lymphadenectomy may potentially mitigate the detrimental effects of NSCLC on survival.

The N status has long been regarded as a vital criterion in cancer staging and the strongest prognostic indicator for survival after surgery. The definition standard of N status in the AJCC TNM classification system depends on localized spread to the regional lymph nodes and lymph node status.^{24,25} In our study, our survival curve showed that NLN count was associated with a stronger survival benefit in patients with worse staging and N status. For a long time, it has been widely accepted that stage IV lung cancer patients are those with distant metastasis regardless of N status. Therefore, the N status of stage IV patients was ignored for the M status (M0 and M1). Palliative surgery is constantly performed in patients with metastatic (M1) disease to improve quality of life, alleviate symptoms, and prevent complications with limited resection or lymph node retrieval. In our study, NLN count was associated with survival benefits in the N3 and M1 stages in our subgroup analysis, which was consistent with some findings in stage IV cancer patients.^{26,27} Therefore, N stage is a significant factor affecting survival outcome, especially in stage IV patients, which indicates that the prognostic role of N status in M1 disease should not be underestimated. Moreover, disease management for NSCLC has recently evolved into multidisciplinary therapy and now requires more accurate selection of patients with advanced disease to maximize the effect of lymphadenectomy and minimize morbidity.

In the current study, we established an NLN-based nomogram to predict survival outcomes at 36 and 60 months. Several clinical parameters have been reported as prognostic markers.^{28–31} In accordance with these findings, our study revealed a strong correlation between survival and these independent prognostic factors in the multivariate Cox proportional hazards model. Based on these variables, the nomogram precisely predicted the survival probability in the internal and external cohorts. Survival prediction is a crucial issue for

NSCLC patients undergoing resection; this predictive model could assist physicians by providing individualized risk estimates of survival outcomes and potentially helping to avoid unnecessary treatments in patients.

The present study had several limitations. First, the SEER database only contained general treatment information, and there was no more detailed clinical treatment information, such as specific chemotherapy and radiotherapy regimens, targeted therapy, and immunotherapy, which may have caused bias in our study. Second, since this is a retrospective analysis, more prospective large randomized clinical trials and multicenter clinical samples are needed to further validate the reliability of our study findings. Finally, the lack of molecular analysis may overlook important prognostic clues for researchers.

Conclusion

In summary, the NLN count was determined to be a promising prognostic factor in patients with NSCLC. This is the first study to establish a novel prognosis prediction model based on NLN count in NSCLC based on a large cohort and validated the prognostic role of NLN count using our own hospital data, which could assist in personalizing the risk classification of patients. Further studies about NLN count are needed to determine optimal lymph node resection and the underlying molecular mechanisms of NSCLC.

Acknowledgements

We gratefully acknowledge the Surveillance, Epidemiology, and End Results (SEER) program tumor registries in the creation of the SEER-Medicare database.

Author's Note

Conception and design: All authors. Administrative and fund support: Jiandong Zhang. Provision of study materials or patients: All authors. Collection and assembly of data: All authors. Data analysis and interpretation: Xinyi Huang and Pingping Hu. Manuscript writing: Xinyi Huang and Pingping Hu. All authors contributed to the article and approved the submitted version.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics Statement

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The request for the SEER Research Plus Data was approved to obtain therapy information for research purposes only (reference number: 18323-Nov2020). Institutional review board approval was not required because SEER is an anonymized database that is open to the public. Informed patient consent from the validation cohort was approved by The First Affiliated Hospital of Shandong First Medical University review board (No. S033). The study was conducted in accordance with the principles of the Declaration of Helsinki (as revised in 2013).


Data Accessibility Statement

Publicly available datasets were analyzed in this study. These data can be found here: The SEER datasets for this study can be found in the <https://seer.cancer.gov/seerstat/software/>.

Funding

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: This study was funded by the National Natural Science Foundation of China [grant no. 81672974].

ORCID iD

Xinyi Huang  <https://orcid.org/0000-0002-2827-1860>

Supplemental material

Supplemental material for this article is available online.

References

- 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.
- Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc*. 2008;83(5):584-594.
- Yang M, Cao HX, Guo X, et al. The number of resected lymph nodes (nLNs) combined with tumor size as a prognostic factor in patients with pathologic N0 and Nx non-small cell lung cancer. *PLoS One*. 2013;8(9):e73220.
- Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin*. 2016;66(4):271-289.
- Chen Y, Zhang L, Tian J, et al. Combining the negative lymph nodes count with the ratio of positive and removed lymph nodes can better predict the postoperative survival in cervical cancer patients. *Cancer Cell Int*. 2013;13(1):6.
- Wong JH, Severino R, Honnebiel MB, et al. Number of nodes examined and staging accuracy in colorectal carcinoma. *J Clin Oncol*. 1999;17(9):2896-2900.
- Yoshida Y, Watanabe S. Primary lung cancer surgery-clinical trial results. *Gan To Kagaku Ryoho*. 2018;45(10):1405-1409.
- Shi RL, Chen Q, Ding JB, et al. Increased number of negative lymph nodes is associated with improved survival outcome in node positive gastric cancer following radical gastrectomy. *Oncotarget*. 2016;7(23):35084-35091.
- Johnson PM, Porter GA, Ricciardi R, Baxter NN. Increasing negative lymph node count is independently associated with improved long-term survival in stage IIIB and IIIC colon cancer. *J Clin Oncol*. 2006;24(22):3570-3575.
- Li QG, Zhuo CH, Cai GX, Li DW, Liang L, Cai SJ. Increased number of negative lymph nodes is associated with improved cancer specific survival in pathological IIIB and IIIC rectal cancer treated with preoperative radiotherapy. *Oncotarget*. 2014;5(23):12459-12471.
- Hsu PK, Huang CS, Wang BY, Wu YC, Chou TY, Hsu WH. The prognostic value of the number of negative lymph nodes in esophageal cancer patients after transthoracic resection. *Ann Thorac Surg*. 2013;96(3):995-1001.
- Wang S, Zhang B, Li C, et al. Prognostic value of number of negative lymph node in patients with stage II and IIIa non-small cell lung cancer. *Oncotarget*. 2017;8(45):79387-79396.
- Wu SG, Wang Y, Zhou J, et al. Number of negative lymph nodes should be considered for incorporation into staging for breast cancer. *Am J Cancer Res*. 2015;5(2):844-853.
- Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med*. 1996; 15(4):361-387.
- Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med*. 2004;23(13):2109-2123.
- Deng WY, Xu T, Wang YF, et al. Log odds of positive lymph nodes may predict survival benefit in patients with node-positive non-small cell lung cancer. *Lung Cancer*. 2018;122:60-66.
- Samayoa AX, Pezzi TA, Pezzi CM, et al. Rationale for a minimum number of lymph nodes removed with non-small cell lung cancer resection: correlating the number of nodes removed with survival in 98,970 patients. *Ann Surg Oncol*. 2016; 23(Suppl 5):1005-1011.
- Zhu Z, Chen H, Yu W, et al. Number of negative lymph nodes is associated with survival in thoracic esophageal squamous cell carcinoma patients undergoing three-field lymphadenectomy. *Ann Surg Oncol*. 2014;21(9):2857-2863.
- Muneoka Y, Akazawa K, Ishikawa T, et al. Nomogram for 5-year relapse-free survival of a patient with advanced gastric cancer after surgery. *Int J Surg*. 2016;35:153-159.
- Montes AF, Lopez CL, Martinez GA, et al. Prognostic nomogram and patterns of use of FOLFIRI-affibercept in advanced colorectal cancer: a real-world data analysis. *Oncologist*. 2019;24(8): e687-e695.
- Dai J, Liu M, Yang Y, et al. Optimal lymph node examination and adjuvant chemotherapy for stage I lung cancer. *J Thorac Oncol*. 2019;14(7):1277-1285.
- Ludwig MS, Goodman M, Miller DL, Johnstone PA. Postoperative survival and the number of lymph nodes sampled during resection of node-negative non-small cell lung cancer. *Chest*. 2005;128(3):1545-1550.
- Stiles BM, Mao JL, Harrison S, et al. Extent of lymphadenectomy is associated with oncological efficacy of sublobar resection for lung cancer ≤ 2 cm. *J Thorac Cardiovasc Surg*. 2019;157(6):2454-2465.
- Shersher DD, Liptay MJ. Status of sentinel lymph node mapping in non-small cell lung cancer. *Cancer J*. 2015;21(1):17-20.
- Taylor MD, LaPar DJ, Thomas CJ, et al. Lymph node ratio predicts recurrence and survival after R0 resection for non-small cell lung cancer. *Ann Thorac Surg*. 2013;96(4):1163-1170.
- Zhuo CG, Ying MG, Lin RR, et al. Negative lymph node count is a significant prognostic factor in patient with stage IV gastric cancer after palliative gastrectomy. *Oncotarget*. 2017;8(41):71197-71205.
- Ishihara S, Hayama T, Yamada H, et al. Prognostic impact of primary tumor resection and lymph node dissection in stage IV colorectal cancer with unresectable metastasis: a propensity score analysis in a multicenter retrospective study. *Ann Surg Oncol*. 2014;21(9):2949-2955.

28. Liu L, Shi M, Wang Z, et al. A molecular and staging model predicts survival in patients with resected non-small cell lung cancer. *BMC Cancer*. 2018;18(1):966.
29. Xie D, Allen MS, Marks R, et al. Nomogram prediction of overall survival for patients with non-small-cell lung cancer incorporating pretreatment peripheral blood markers. *Eur J Cardio-Thorac*. 2018;53(6):1214-1222.
30. 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-e203.
31. Sun F, Ma K, Yang X, et al. A nomogram to predict prognosis after surgery in early stage non-small cell lung cancer in elderly patients. *Int J Surg*. 2017;42:11-16.