Original Article

The construction and validation of the model for predicting the incidence and prognosis of brain metastasis in lung cancer patients

Chunjian Zuo[1], Guanchu Liu[2], Ye Bai[3], Jie Tian[4], Huanwen Chen[1]^

[1]Department of Cardiothoracic Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China; [2]Department of Cardiothoracic Surgery, The Affiliated Hospital of North Sichuan Medical College, Nanchong, China; [3]Department of Epidemiology and Biostatistics, School of Public Health and Management, Chongqing Medical University, Chongqing, China; [4]Department of Thoracic Surgery, The Third Affiliated Hospital of Chongqing Medical University (Gener Hospital), Chongqing, China

Contributions: (I) Conception and design: H Chen, C Zuo, J Tian; (II) Administrative support: H Chen; (III) Provision of study materials or patients: C Zuo, Y Bai; (IV) Collection and assembly of data: G Liu, C Zuo; (V) Data analysis and interpretation: C Zuo, J Tian; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Huanwen Chen, MD, PhD. Department of Cardiothoracic Surgery, The First Affiliated Hospital of Chongqing Medical University, No. 1 Youyi Road, Yuzhong District, Chongqing 400010, China. Email: coolstarchw9527@163.com.

Background: Brain metastasis (BM) causes high morbidity and mortality rates in lung cancer (LC) patients. The present study aims to develop models for predicting the development and prognosis of BM using a large LC cohort.

Methods: A total of 266,522 LC cases diagnosed between 2010 and 2016 were selected from the Surveillance, Epidemiology, and End Results (SEER) Program cohort. Risk factors for developing BM and prognosis were calculated by univariable and multivariable logistic and Cox regression analysis, respectively, and nomograms were constructed based on risk factors. Nomogram performance was evaluated with receiver operating characteristics (ROC) curve, or C-index and calibration curve.

Results: The prevalence of BM was 13.33%. Associated factors for developing BM include: advanced age; Asian or Pacific Islander race; uninsured status; primary tumor site; higher T stage; higher N stage; poorly differentiated grade; the presence of lung, liver, and bone metastases; and adenocarcinoma histology. Median overall survival (OS) was 4 months; associated prognosis factors were similar to risk factors plus female gender, unmarried status, and surgery. The calibration curve showed good agreement between predicted and actual probability, and the AUC/C-index was 73.1% (95% CI: 72.6–73.6%) and 0.88 (95% CI: 0.87–0.89) for risk and prognosis predictive models, respectively.

Conclusions: BM was highly developed in LC patients, and homogeneous and heterogeneous factors were found between risk and prognosis for BM. The nomogram showed good performance in predicting BM development and prognosis.

Keywords: Lung cancer (LC); brain metastases (BM); incidence; prognosis; prediction model

Submitted Aug 16, 2020. Accepted for publication Nov 20, 2020.
doi: 10.21037/tcr-20-2745
View this article at: http://dx.doi.org/10.21037/tcr-20-2745

^ ORCID: 0000-0001-6539-7668.

Introduction

As indicated in the latest GLOBOCAN report, about 2.1 million new diagnosed lung cancer (LC) cases and 1.8 million deaths were predicted worldwide in 2018, representing the leading cause of cancer morbidity and mortality (1,2). Brain metastasis (BM) was the most prevalent intracranial tumor in adults, commonly developing in LC patients (3-5). Generally, BM shows poor prognosis; approximately 90% of patients died within two years after initial diagnosis and median survival ranges from 7 to 10 months, with significant worsening of quality of life (6,7).

It is critical to identify high-risk patients with BM and conduct targeted screening. Yet, the National Comprehensive Cancer Network (NCCN) suggests that brain MRI should only be recommended for patients diagnosed with stage II to IV and high-risk stage 1B non-small-cell lung cancer; no unanimous screening guidelines for identifying BM exists, until now (8,9). Previous studies found several risk and prognostic factors for BM, shedding light on methods to identify high-risk patients, predict survival, and conduct targeted therapy (10-13). However, due to the relatively small sample size, the results were not consistent (7,14,15).

Surveillance, Epidemiology, and End Results (SEER) is an open-access database comprising approximately 30% of the total US population. SEER collected complete social demographics and clinical data for assessment of characteristics of cancer epidemiology worldwide. The present study aims to investigate BM prevalence in LC patients, determine risk and prognostic factors, and construct nomograms for predicting BM development and prognosis in LC patients using the SEER database. We present the following article in accordance with the TRIPOD reporting checklist (available at http://dx.doi.org/10.21037/tcr-20-2745).

Methods

Ethics statement

The data about cancer in the SEER database is continually reported in every state of the United States and retrieved with no need for informed patient consent. The present study complied with the Declaration of Helsinki (as revised in 2013).

Data source

Data in this population-based study were abstracted from the SEER. SEER*Stat Software version 8.3.4 (https://seer.cancer.gov/seerstat/; Information Management Service, Inc., Calverton, MD, USA) was used to generate the case listing.

Cohort selection

A total of 266,522 eligible LC patients were identified between January 1, 2010, and December 31, 2016. The inclusion criteria were as follows: (I) aged 18 years or older; (II) diagnosed as the first and only malignant cancer; (III) only one primary site; (IV) diagnosis not obtained from a death certificate or an autopsy; (V) with definite BM information at primal diagnosis of LC. Patients diagnosed before 2010 were excluded because the SEER did not record BM information until 2010. The flow-chart for the population selection was shown in Figure S1.

Statistical analysis

Absolute numbers and incidence proportions were calculated for patients with LC and BM identified at diagnosis with the classification strategies as: age (<41, 41–50, 51–60, 61–70, 71–80, 81–90 and >90 years), sex (female and male), race [white, black, AI (American Indian/Alaska Native), and API (Asian or Pacific Islander)], marital status (married and unmarried), insurance status (insured and uninsured), primary tumor site (Main bronchus, Upper lobe, Middle lobe, Lower lobe, Overlapping lesion of lung, Lung, NOS), primary tumor T stage (T1, T2, T3, and T4), regional lymph node stage (N0, N1, N2 and N3), tumor metastases stage (M0 and M1), tumor differentiated grade (grade I, grade II, grade III, and grade IV), the presence or absence of lung metastases, liver metastases, or bone metastases, histology [adenocarcinoma (AC), squamous lung cancer (SLC), adenosquamous carcinoma (ASC), large cell lung cancer (LCLC), small cell lung cancer (SCLC) and others]. All the data of "Unknown" in every categorical variable were included. The prevalence of BM was analyzed in the total population and patients with M1 stage, respectively. Subgroup analysis was also conducted concerning the age and pathological types groups. The univariable and multivariable logistic regression model was conducted to determine the associated factors for developing BM. The Kaplan-Meier method was used to analyze the overall survival (defined as the time from diagnosis of LC with BM to all causes of death) of the LC patients with BM. Meanwhile, the surgical treatment of primary site (not or yes) was further added with significant risk factors to identify

the prognostic factors for LC patients with BM by using multivariable Cox proportional hazards regression. The final model selection was performed by a backward selection process using the Akaike information criterion.

Predictive nomograms were formulated based on the results of multivariable logistic analysis and multivariable Cox proportional hazards regression, respectively using the rms package in R version 3.6.0 (R Foundation for Statistical Computing, Vienna, Austria; www.r-project.org). The receiver operating characteristics (ROC) and concordance index (C-index) were utilized to evaluate the discrimination of nomograms severally.

The calibration of the nomograms was assessed by calibration plots, which was generated by adopting a regression smoothing method and bootstrapping with 100 resamples to estimate the calibration of nomograms, where the consistency between observed and predicted probabilities of occurrence of BM and 1, 3 and 5-year survvival was presented graphically.

The randomly splitting method was adopted as the method of internal validation to evaluate the stability of the nomograms. The ROC curve was constructed to evaluate the performance of the construction and validation, and the difference in the AUC was tested by DeLong's test. To evaluate the external performance of survival predicted nomogram, an external validation set that consisted of LC patients with de novo BM in 2016 was selected from the same database, and the calibration curve for evaluating the performance in predicting 1-year was generated based on the conducted program. The construction and validation of the nomogram were conducted using the rms package in R version 3.4.1 (R Foundation for Statistical Computing, Vienna, Austria; www.r-project.org). Other statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 23.0 software package for Windows (SPSS, Inc.). Statistically significant levels were set two-tailed at $P<0.05$.

Results

Incidence of BM

A total of 250,141 eligible patients were included in the present

Table 1 Baseline of the demographic and related clinical characteristics for patients diagnosed with initial lung carcinoma

Subject characteristics	Total population (2010 to 2016)		Construction cohort (2010 to 2015)		Validation cohort (in 2016)	
	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)
Age, years						
<41	363 (17.0)	1,775 (83.0)	303 (16.5)	1,530 (83.5)	60 (19.7)	245 (80.3)
41–50	2,285 (21.2)	8,494 (78.8)	2,011 (21.1)	7,541 (78.9)	274 (22.3)	953 (77.7)
51–60	8,841 (19.1)	37,429 (80.9)	7,586 (19.1)	32,202 (80.9)	1,255 (19.4)	5,227 (80.6)
61–70	11,567 (14.5)	68,374 (85.5)	9,848 (14.4)	58,371 (85.6)	1,719 (14.7)	10,003 (85.3)
71–80	7,601 (10.4)	65,576 (89.6)	6,479 (10.4)	55,723 (89.6)	1,122 (10.2)	9,853 (89.8)
81–90	2,491 (7.3)	31,794 (92.7)	2,145 (7.3)	27,366 (92.7)	346 (7.2)	4,428 (92.8)
>90	184 (5.2)	3,367 (94.8)	154 (5.2)	2,821 (94.8)	30 (5.2)	546 (94.8)
Sex						
Male	17,411 (13.4)	112,307 (86.6)	14,907 (13.4)	96,443 (86.6)	2,504 (13.6)	15,864 (86.4)
Female	15,921 (13.2)	104,502 (86.8)	13,619 (13.3)	89,111 (86.7)	2,302 (13.0)	15,391 (87.0)
Race						
White	26,015 (12.9)	175,568 (87.1)	22,400 (13.0)	150,446 (87.0)	3,615 (12.6)	25,122 (87.4)
Black	4,290 (14.5)	25,223 (85.5)	3,619 (14.3)	21,647 (85.7)	671 (15.8)	3,576 (84.2)
AI	193 (13.8)	1,201 (86.2)	159 (13.5)	1,018 (86.5)	34 (15.7)	183 (84.3)

Table 1 (*continued*)

Table 1 (*continued*)

Subject characteristics	Total population (2010 to 2016)		Construction cohort (2010 to 2015)		Validation cohort (in 2016)	
	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)
API	2,764 (16.3)	14235 (83.7)	2,296 (16.0)	12,030 (84.0)	468 (17.5)	2,205 (82.5)
Unknown	70 (10.7)	582 (89.3)	52 (11.2)	413 (88.8)	18 (9.6)	169 (90.4)
Marital status						
Others	15,401 (13.2)	101,707 (86.8)	13,123 (13.1)	86,870 (86.9)	2,278 (13.3)	14,837 (86.7)
Married	16,557 (13.6)	104,930 (86.4)	14,238 (13.7)	89,838 (86.3)	2,319 (13.3)	15,092 (86.7)
Unknown	1,374 (11.9)	101,72 (88.1)	1,165 (11.6)	8,846 (88.4)	209 (13.6)	1,326 (86.4)
Insurance status						
Insured	25,275 (12.6)	174,934 (87.4)	21,607 (12.6)	149,701 (87.4)	3,668 (12.7)	25,233 (87.3)
Any Medic aid	5,845 (15.5)	31,836 (84.5)	4,945 (15.5)	27,004 (84.5)	900 (15.7)	4,832 (84.3)
Uninsured	1,542 (20.8)	5,855 (79.2)	1,419 (21.2)	5,290 (78.8)	123 (17.9)	565 (82.1)
Unknown	670 (13.8)	4,184 (86.2)	555 (13.5)	3,559 (86.5)	115 (15.5)	625 (84.5)
Site						
MB	1,726 (14.5)	10,146 (85.5)	1,482 (14.4)	8,809 (85.6)	244 (15.4)	1,337 (84.6)
UL	16,803 (13.1)	111,712 (86.9)	14,390 (13.1)	95,484 (86.9)	2,413 (12.9)	16,228 (87.1)
ML	1,300 (12.0)	9,571 (88.0)	1,117 (12.0)	8,160 (88.0)	183 (11.5)	1,411 (88.5)
LL	7,856 (12.2)	56,663 (87.8)	6,638 (12.1)	48,011 (87.9)	1,218 (12.3)	8,652 (87.7)
OL	331 (12.2)	2,391 (87.8)	278 (12.0)	2,035 (88.0)	53 (13.0)	356 (87.0)
NOS	5,316 (16.8)	26,326 (83.2)	4,621 (16.7)	23,055 (83.3)	695 (17.5)	3,271 (82.5)
T stage						
T1	3,812 (7.4)	47,865 (92.6)	3,185 (7.4)	39,793 (92.6)	627 (7.2)	8,072 (92.8)
T2	8,097 (11.9)	59,812 (88.1)	6,849 (11.8)	51,357 (88.2)	1,248 (12.9)	8,455 (87.1)
T3	7,148 (14.8)	41,129 (85.2)	6,243 (14.9)	35,732 (85.1)	905 (14.4)	5,397 (85.6)
T4	9,849 (17.4)	46,686 (82.6)	8,392 (17.3)	40,047 (82.7)	1,457 (18.0)	6,639 (82.0)
Unknown	4,426 (17.2)	21,317 (82.8)	3,857 (17.2)	18,625 (82.8)	569 (17.4)	2,692 (82.6)
N stage						
N0	6,870 (7.2)	88,264 (92.8)	5,885 (7.3)	75,135 (92.7)	985 (7.0)	13,129 (93.0)
N1	2,724 (13.4)	17,656 (86.6)	2,320 (13.2)	15,228 (86.8)	404 (14.3)	2,428 (85.7)
N2	14,854 (16.8)	73,614 (83.2)	12,853 (16.8)	63,673 (83.2)	2,001 (16.8)	9,941 (83.2)
N3	6,612 (19.6)	27,195 (80.4)	5,538 (19.4)	23,033 (80.6)	1,074 (20.5)	4,162 (79.5)
Unknown	2,272 (18.4)	10,080 (81.6)	1,930 (18.5)	8,485 (81.5)	342 (17.7)	1,595 (82.3)
M stage						
M0	2 (0.0)	120,083 (100.0)	0 (0.0)	102,501 (100.0)	2 (0.0)	17,582 (100.0)
M1	33,269 (25.7)	95,939 (74.3)	28,495 (25.6)	82,651 (74.4)	4,774 (26.4)	13,288 (73.6)
Unknown	61 (7.2)	787 (92.8)	31 (7.2)	402 (92.8)	30 (7.2)	385 (92.8)

Table 1 (*continued*)

Table 1 (*continued*)

Subject characteristics	Total population (2010 to 2016)		Construction cohort (2010 to 2015)		Validation cohort (in 2016)	
	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)	With BM, n (%)	Without BM, n (%)
Grade						
Grade I	402 (3.0)	13,064 (97.0)	347 (3.0)	11,084 (97.0)	55 (2.7)	1,980 (97.3)
Grade II	2,447 (6.3)	36,361 (93.7)	2,124 (6.4)	31,043 (93.6)	323 (5.7)	5,318 (94.3)
Grade III	7,564 (13.1)	50,167 (86.9)	6,537 (13.1)	43,491 (86.9)	1,027 (13.3)	6,676 (86.7)
Grade IV	1,044 (14.8)	6,007 (85.2)	942 (15.0)	5,329 (85.0)	102 (13.1)	678 (86.9)
Unknown	21,875 (16.4)	111,210 (83.6)	18,576 (16.4)	94,607 (83.6)	3,299 (16.6)	16,603 (83.4)
Bone Met						
No	21,576 (10.7)	179,603 (89.3)	18,546 (10.8)	153,784 (89.2)	3,030 (10.5)	25,819 (89.5)
Yes	11,016 (23.4)	36,090 (76.6)	9,328 (23.2)	30,794 (76.8)	1,688 (24.2)	5,296 (75.8)
Unknown	740 (39.9)	1,116 (60.1)	652 (40.0)	976 (60.0)	88 (38.6)	140 (61.4)
Liver Met						
No	25,634 (11.8)	191,999 (88.2)	21,931 (11.8)	164,238 (88.2)	3,703 (11.8)	27,761 (88.2)
Yes	6,740 (22.2)	23,657 (77.8)	5,756 (22.1)	20,304 (77.9)	984 (22.7)	3,353 (77.3)
Unknown	958 (45.4)	1,153 (54.6)	839 (45.3)	1,012 (54.7)	119 (45.8)	141 (54.2)
Lung Met						
No	24,048 (11.5)	185,211 (88.5)	20,557 (11.5)	158,472 (88.5)	3,491 (11.5)	26,739 (88.5)
Yes	8,029 (21.7)	29,026 (78.3)	6,860 (21.7)	24,729 (78.3)	1,169 (21.4)	4,297 (78.6)
Unknown	1,255 (32.8)	2,572 (67.2)	1,109 (32.0)	2,353 (68.0)	146 (40.0)	219 (60.0)
Histology						
AC	17,409 (14.8)	100,406 (85.2)	14,710 (14.8)	84,956 (85.2)	2,699 (14.9)	15,450 (85.1)
SLC	3,090 (5.8)	50,113 (94.2)	2,655 (5.8)	42,979 (94.2)	435 (5.7)	7,134 (94.3)
ASLC	349 (11.2)	2,767 (88.8)	306 (11.4)	2,369 (88.6)	43 (9.8)	398 (90.2)
LCLC	660 (17.8)	3,045 (82.2)	573 (17.6)	2,680 (82.4)	87 (19.2)	365 (80.8)
SCLC	5,503 (17.1)	26,675 (82.9)	4,727 (17.0)	23,040 (83.0)	776 (17.6)	3,635 (82.4)
Others	6,321 (15.8)	33,803 (84.2)	5,555 (15.8)	29,530 (84.2)	766 (15.2)	4,273 (84.8)
Surg (pri)						
No	32,382 (16.5)	164,148 (83.5)	27,712 (16.4)	140,781 (83.6)	4,670 (16.7)	23,367 (83.3)
Yes	877 (1.7)	52,056 (98.3)	755 (1.7)	44,261 (98.3)	122 (1.5)	7,795 (98.5)
Unknown	73 (10.8)	605 (89.2)	59 (10.3)	512 (89.7)	14 (13.1)	93 (86.9)

LC, lung carcinoma; BM, brain metastases; Met, metastases; AI, American Indian/Alaska Native; API, Asian or Pacific Islander; MB, main bronchus; UL, upper lobe; ML, middle lobe; LL, lower lobe; OL, overlapping lesion of lung; AC, adenocarcinoma; SLC, squamous lung carcinoma; ASLC, adenosquamous lung carcinoma; LCLC, large cell lung carcinoma; SCLC, small cell lung carcinoma; Surg(pri), surgical treatments of primary site.

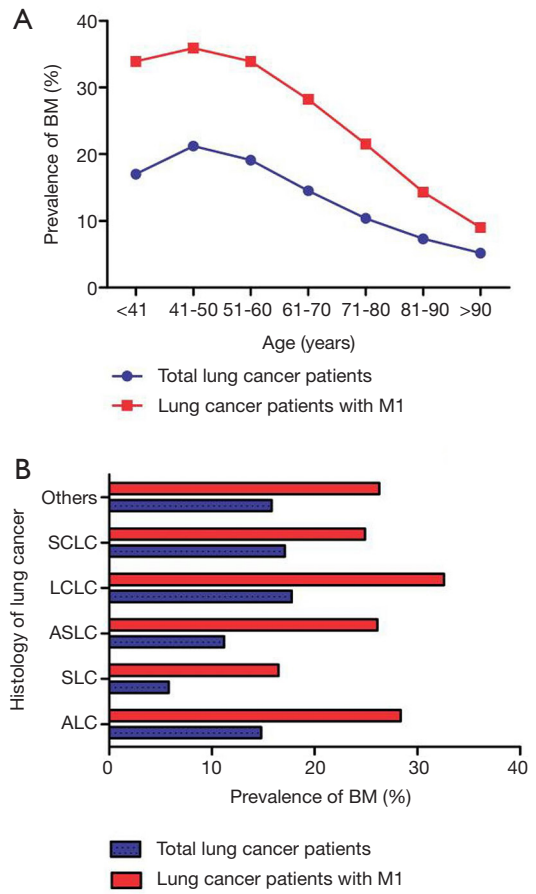

Figure 1 The prevalence of BM in LC patients with metastasis or not. (A) Stratified by age; (B) stratified by histology. BM, brain metastasis; LC, lung cancer.

study; 33,332 (13.33% of the entire cohort) subjects were diagnosed with BM and 28,526 LC patients with BM received follow-up for over 1 year (*Table 1*). BM prevalence increased from 18 years, peaked at 50 years, and gradually decreased thereafter (*Figure 1A*). The variation for BM prevalence with age in metastatic LC patients showed a similar trend. When stratified by different LC histology, the highest BM prevalence was seen in LCLC patients and the lowest prevalence was observed in SCL patients. similar trend was seen in LC patients with distant metastatic disease (*Figure 1B*).

Risk factors for developing BM in LC patients

As is shown in *Table 2*, multivariable logistic regression showed Black and API race (*vs.* White), Medicaid status and uninsured status (*vs.* insured status), higher T stage, N stage, poorly differentiated grade stage, bone metastases, liver metastases, and lung metastases were positively associated with BM, while advanced age and non-AC histology were negatively associated with BM development.

Construction of nomogram to predict BM development and performance assessment

Based on the aforementioned BM risk factors, a nomogram was constructed to accurately predict the probability of BM development (*Figure 2A*). The AUC of the ROC curve 73.1% (95% CI: 72.6–73.6%), indicating good discrimination (*Figure 2B*). The stability of this prediction

Table 2 Univariate and multivariable logistic regression for analyzing the demographic and related clinical characteristics for developing Brain Metastases in patients diagnosed with initial lung carcinoma (diagnosed 2010–2016)

Subject characteristics	Univariate		Multivariate	
	OR (95%CI)	P value	OR (95%CI)	P value
Age, years				
<41	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
41–50	1.32 (1.16–1.49)	<0.001	1.23 (1.08–1.40)	0.002
51–60	1.16 (1.03–1.30)	0.014	1.14 (1.01–1.28)	0.040
61–70	0.83 (0.74–0.93)	0.001	0.89 (0.78–1.00)	0.047
71–80	0.57 (0.51–0.64)	<0.001	0.63 (0.56–0.71)	<0.001
81–90	0.38 (0.34–0.43)	<0.001	0.40 (0.35–0.46)	<0.001
>90	0.27 (0.22–0.32)	<0.001	0.24 (0.20–0.29)	<0.001

Table 2 (*continued*)

Table 2 (*continued*)

Subject characteristics	Univariate		Multivariate	
	OR (95%CI)	P value	OR (95%CI)	P value
Gender				
Male	1.00 (Ref)	1.00 (Ref)	NS	NS
Female	0.98 (0.96–1.01)	0.139	NS	NS
Race				
White	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Black	1.15 (1.11–1.19)	<0.001	1.00 (0.96–1.04)	0.997
AI	1.09 (0.93–1.26)	0.297	1.03 (0.88–1.21)	0.718
API	1.31 (1.26–1.37)	<0.001	1.19 (1.14–1.25)	<0.001
Unknown	0.81 (0.63–1.04)	0.100	0.86 (0.67–1.12)	0.264
Marital status				
Others	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Married	1.04 (1.02–1.07)	<0.001	1.02 (1.00–1.05)	0.080
Unknown	0.89 (0.84–0.95)	0.001	0.88 (0.82–0.93)	<0.001
Insurance status				
Insured	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Any medic aid	1.27 (1.23–1.31)	<0.001	1.04 (1.00–1.07)	0.027
Uninsured	1.82 (1.72–1.93)	<0.001	1.18 (1.11–1.25)	<0.001
Unknown	1.19 (1.02–1.20)	0.015	0.98 (0.90–1.07)	0.696
Site				
MB	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
UL	0.88 (0.84–0.93)	<0.001	1.15 (1.09–1.22)	<0.001
ML	0.80 (0.74–0.86)	<0.001	1.04 (0.96–1.13)	0.299
LL	0.82 (0.77–0.86)	<0.001	1.15 (1.08–1.22)	<0.001
OL	0.81 (0.72–0.92)	0.001	0.89 (0.78–1.01)	0.078
NOS	1.19 (1.12–1.26)	<0.001	1.04 (0.98–1.11)	0.201
T stage				
T1	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
T2	1.70 (1.63–1.77)	<0.001	1.43 (1.37–1.49)	<0.001
T3	2.18 (2.09–2.28)	<0.001	1.48 (1.41–1.54)	<0.001
T4	2.65 (2.55–2.76)	<0.001	1.53 (1.47–1.60)	<0.001
Unknown	2.61 (2.49–2.73)	<0.001	1.54 (1.46–1.63)	<0.001
N stage				
N0	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
N1	1.98 (1.89–2.08)	<0.001	1.54 (1.47–1.62)	<0.001

Table 2 (*continued*)

Table 2 (*continued*)

Subject characteristics	Univariate		Multivariate	
	OR (95%CI)	P value	OR (95%CI)	P value
N2	2.59 (2.52–2.67)	<0.001	1.68 (1.62–1.73)	<0.001
N3	3.12 (3.01–3.24)	<0.001	1.69 (1.62–1.76)	<0.001
Unknown	2.90 (2.75–3.05)	<0.001	1.66 (1.56–1.76)	<0.001
Grade				
Grade I	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Grade II	2.19 (1.96–2.44)	<0.001	2.19 (1.96–2.44)	<0.001
Grade III	4.90 (4.42–5.43)	<0.001	3.75 (3.38–4.17)	<0.001
Grade IV	5.65 (5.01–6.36)	<0.001	3.58 (3.16–4.05)	<0.001
Unknown	6.39 (5.78–7.07)	<0.001	4.23 (3.82–4.69)	<0.001
Bone Met				
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	2.54 (2.48–2.61)	<0.001	1.58 (1.54–1.62)	<0.001
Unknown	5.52 (5.03–6.06)	<0.001	2.22 (1.99–2.48)	<0.001
Liver Met				
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	2.13 (2.07–2.20)	<0.001	1.26 (1.22–1.30)	<0.001
Unknown	6.22 (5.71–6.79)	<0.001	2.72 (2.45–3.01)	<0.001
Lung Met				
No	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
Yes	2.13 (2.07–2.19)	<0.001	1.45 (1.40–1.49)	<0.001
Unknown	3.76 (3.51–4.03)	<0.001	1.83 (1.69–1.98)	<0.001
Histology				
AC	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
SLC	0.36 (0.34–0.37)	<0.001	0.39 (0.37–0.40)	<0.001
ASLC	0.73 (0.65–0.81)	<0.001	0.71 (0.64–0.80)	<0.001
LCLC	1.25 (1.15–1.36)	<0.001	1.04 (0.95–1.14)	0.409
SCLC	1.19 (1.15–1.23)	<0.001	0.81 (0.78–0.84)	<0.001
Others	1.08 (1.05–1.11)	<0.001	0.96 (0.93–0.99)	0.015

LC, lung carcinoma; BM, brain metastases; Met, metastases; AI, American Indian/Alaska Native; API, Asian or Pacific Islander; MB, main bronchus; UL, upper lobe; ML, middle lobe; LL, lower lobe; OL, overlapping lesion of lung; AC, adenocarcinoma; SLC, squamous lung carcinoma; ASLC, adenosquamous lung carcinoma; LCLC, large cell lung carcinoma; SCLC, small cell lung carcinoma; Ref, reference; NS, no significance.

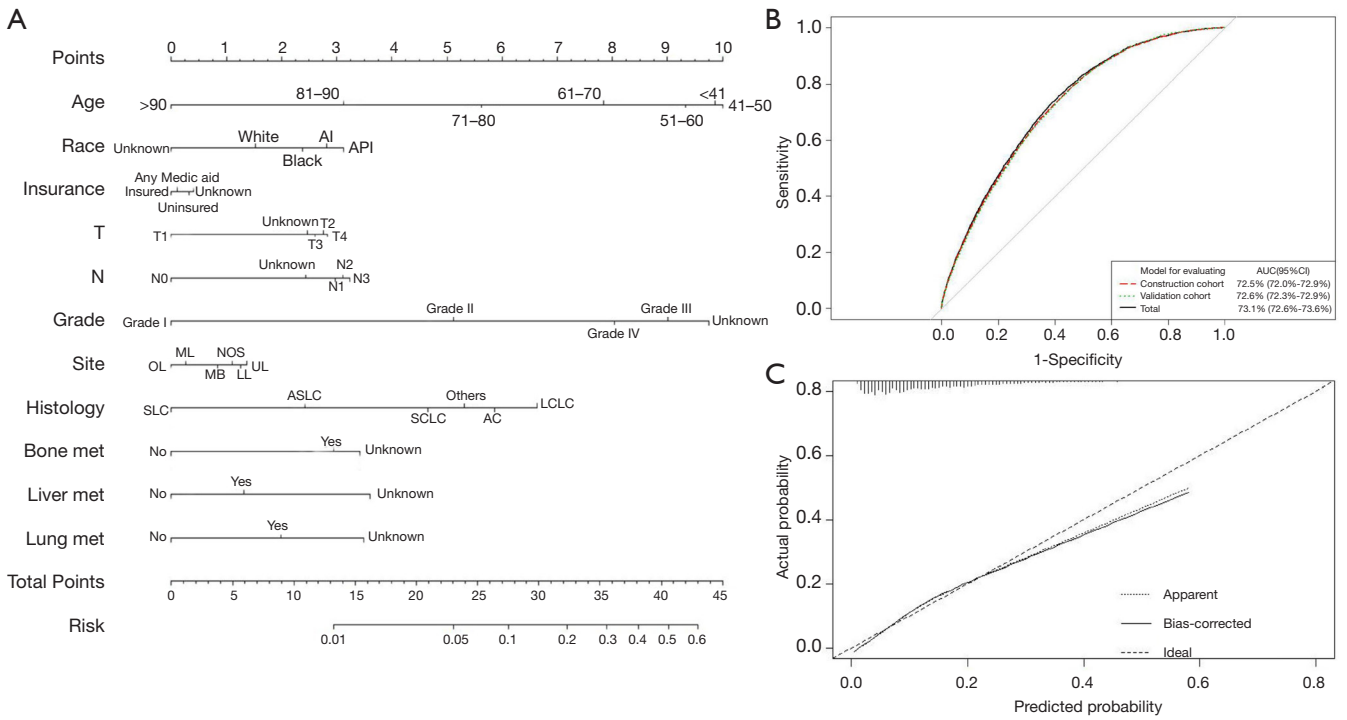

Figure 2 The nomogram for predicting BM in LC (A), ROC curve and randomly splitting method for evaluating the performance and stability of the nomogram. (B) and the calibration curve for estimating the predictive accuracy of the nomogram (C). BM, brain metastasis; LC, lung cancer.

model was validated by random splitting method; the AUC values of the construction and validation models were 72.5% (95% CI: 72.0–72.9%) and 72.6% (95% CI: 72.3–72.9%), respectively, with no significant difference (D =–0.43; P=0.67) (*Figure 2B*). The calibration curve revealed good agreement between predicted and observed probabilities for BM occurrence, indicating good calibration ability of the nomogram (*Figure 2C*).

Survival analysis

For the 28,526 patients who developed BM at initial LC diagnosis, the 1-year, 3-year, and 5-year overall survival (OS) was 21%, 5%, and 3%, respectively; the median OS was 4 months (95% CI: 3.91–4.09 months).

Multivariate Cox regression analysis showed older age, insured status, higher T stage, N stage, poorly differentiated tumor grade, tumor location and the presence of lung, liver, or bone metastases were positively associated with overall mortality, while female gender, API race, marital status, UL site, and surgery on primary site were negatively associated with overall mortality (*Table 3*).

Construction of nomogram to predict prognosis and performance assessment

The survival prediction nomogram was constructed based on the prognostic factors for predicting the 1-year, 3-year, and 5-year OS of LC patients with BM (*Figure 3A*). The overall C-index was 0.88 (95% CI: 0.87–0.89), indicating the nomogram has satisfactory performance in distinguishing those who have terminal events from those who do not. The calibration curve has good performance for predicting 1-year (*Figure 3B*), 3-year (*Figure 3C*), and 5-year (*Figure 3D*) OS. Additionally, LC patients diagnosed in 2016 were selected to validate nomogram performance; the calibration curve suggested good agreement between predicted and actual probability curves (*Figure 3E*).

Discussion

The present study comprehensively describes BM prevalence in patients initially diagnosed with LC in the SEER cohort; then, risk and prognostic factors associated with BM were detected. We constructed two nomograms

Table 3 Multivariable cox regression for analyzing the prognosis factors among primary lung carcinoma with brain metastases (diagnosed 2010–2015)

Subject characteristics	Patients' No. of LC with BM		Median survival (IQR), months	Univariate		Multivariate	
	Overall	Dead, n (%)		HR (95%CI)	P value	HR (95%CI)	P value
Age, years							
<41	303	231 (76.2)	13.00 (10.29–15.72)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
41–50	2,011	1,768 (87.9)	8.00 (7.48–8.52)	1.36 (1.18–1.55)	<0.001	1.36 (1.18–1.56)	<0.001
51–60	7,586	6,885 (90.8)	6.00 (5.77–6.24)	1.66 (1.46–1.89)	<0.001	1.67 (1.47–1.91)	<0.001
61–70	9,848	9,158 (93.0)	4.00 (3.85–4.15)	1.92 (1.69–2.19)	<0.001	1.98 (1.74–2.26)	<0.001
71–80	6,479	6,191 (95.6)	3.00 (2.90–3.11)	2.44 (2.14–2.78)	<0.001	2.59 (2.27–2.96)	<0.001
81–90	2,145	2,094 (97.6)	2.00 (1.90–2.11)	3.25 (2.83–3.72)	<0.001	3.37 (2.94–3.87)	<0.001
>90	154	153 (99.4)	1.00 (0.70–1.30)	4.88 (3.98–5.99)	<0.001	4.52 (3.68–5.55)	<0.001
Sex							
Male	14,907	14,024 (94.1)	3.00 (2.89–3.11)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Female	13,619	12,456 (91.5)	4.00 (3.83–4.17)	0.84 (0.82–0.86)	<0.001	0.84 (0.82–0.86)	<0.001
Race							
White	22,400	20,931 (93.4)	4.00 (3.90–4.10)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Black	3,619	3,406 (94.1)	4.00 (3.77–4.23)	1.00 (0.97–1.04)	0.816	1.00 (0.97–1.04)	0.889
AI	159	154 (96.9)	3.00 (2.05–3.95)	1.17 (1.00–1.37)	0.055	1.16 (0.99–1.36)	0.067
API	2,296	1,951 (85.0)	7.00 (6.30–7.70)	0.68 (0.65–0.71)	<0.001	0.69 (0.65–0.72)	<0.001
Unknown	52	38 (73.1)	7.00 (3.94–10.06)	0.76 (0.55–1.04)	0.090	0.73 (0.53–1.00)	0.053
Marital status							
Others	13,123	12,372 (94.3)	3.00 (2.89–3.11)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Married	14,238	13,014 (91.4)	5.00 (4.84–5.16)	0.82 (0.80–0.84)	<0.001	0.86 (0.83–0.88)	<0.001
Unknown	1,165	1,094 (93.9)	4.00 (3.59–4.41)	0.96 (0.90–1.02)	0.196	0.95 (0.89–1.01)	0.091
Insurance status							
Insured	21,607	20,018 (92.7)	4.00 (3.89–4.11)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Any medic aid	4,945	4,619 (93.4)	4.00 (3.80–4.20)	1.07 (1.04–1.11)	<0.001	1.17 (1.13–1.21)	<0.001
Uninsured	1,419	1,322 (93.2)	4.00 (3.66–4.34)	1.09 (1.04–1.16)	0.001	1.27 (1.20–1.34)	<0.001
Unknown	555	521 (93.9)	3.00 (2.61–3.39)	1.21 (1.11–1.32)	<0.001	1.13 (1.03–1.23)	0.010
Site							
MB	1,482	1,415 (95.5)	3.00 (2.64–3.36)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
UL	14,390	13,307 (92.5)	4.00 (3.85–4.15)	0.83 (0.79–0.88)	<0.001	0.90 (0.86–0.96)	<0.001
ML	1,117	1,025 (91.8)	4.00 (3.54–4.46)	0.81 (0.75–0.88)	<0.001	0.89 (0.82–0.97)	0.006
LL	6,638	6,140 (92.5)	4.00 (3.82–4.19)	0.86 (0.82–0.92)	<0.001	0.94 (0.88–0.99)	0.032
OL	278	263 (94.6)	3.00 (1.90–4.11)	0.94 (0.82–1.07)	0.337	1.01 (0.88–1.15)	0.927
NOS	4,621	4,330 (93.7)	3.00 (2.83–3.17)	0.97 (0.92–1.03)	0.348	0.97 (0.91–1.03)	0.339

Table 3 (*continued*)

Table 3 (*continued*)

Subject characteristics	Patients' No. of LC with BM		Median survival (IQR), months	Univariate		Multivariate	
	Overall	Dead, n (%)		HR (95%CI)	P value	HR (95%CI)	P value
T stage							
T1	3,185	2,818 (88.5)	6.00 (5.59–6.41)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
T2	6,849	6,298 (92.0)	4.00 (3.78–4.22)	1.21 (1.15–1.26)	<0.001	1.17 (1.12–1.22)	<0.001
T3	6,243	5,822 (93.3)	4.00 (3.84–4.16)	1.34 (1.28–1.40)	<0.001	1.25 (1.19–1.31)	<0.001
T4	8,392	7,913 (94.3)	3.00 (2.85–3.15)	1.40 (1.35–1.47)	<0.001	1.29 (1.23–1.34)	<0.001
Unknown	3,857	3,629 (94.1)	3.00 (2.81–3.19)	1.49 (1.42–1.56)	<0.001	1.28 (1.22–1.35)	<0.001
N stage							
N0	5,885	5,304 (90.1)	4.00 (3.75–4.25)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
N1	2,320	2,143 (92.4)	4.00 (3.64–4.36)	1.06 (1.00–1.11)	0.033	1.02 (0.97–1.07)	0.569
N2	12,853	12,042 (93.7)	4.00 (3.87–4.13)	1.16 (1.12–1.20)	<0.001	1.09 (1.06–1.13)	<0.001
N3	5,538	5,162 (93.2)	4.00 (3.80–4.20)	1.15 (1.10–1.19)	<0.001	1.09 (1.05–1.14)	<0.001
Unknown	1,930	1,829 (94.8)	3.00 (2.78–3.22)	1.36 (1.29–1.44)	<0.001	1.09 (1.03–1.16)	0.003
Grade							
Grade I	347	306 (88.2)	7.00 (5.18–8.83)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Grade II	2,124	1,876 (88.3)	6.00 (5.49–6.51)	1.04 (0.92–1.17)	0.552	1.01 (0.89–1.14)	0.904
Grade III	6,537	6,115 (93.5)	4.00 (3.85–4.15)	1.42 (1.26–1.59)	<0.001	1.22 (1.09–1.37)	0.001
Grade IV	942	907 (96.3)	4.00 (3.52–4.49)	1.48 (1.30–1.68)	<0.001	1.21 (1.06–1.38)	0.004
Unknown	18,576	17,276 (93.0)	4.00 (3.88–4.12)	1.42 (1.27–1.59)	<0.001	1.15 (1.03–1.29)	0.017
Bone Met							
No	18,546	17,085 (92.1)	4.00 (3.87–4.13)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Yes	9,328	8,772 (94.0)	3.00 (2.87–3.13)	1.16 (1.13–1.19)	<0.001	1.13 (1.10–1.17)	<0.001
Unknown	652	623 (95.6)	3.00 (2.55–3.45)	1.28 (1.18–1.38)	<0.001	1.03 (0.94–1.14)	0.504
Liver Met							
No	21,931	20,117 (91.7)	4.00 (3.88–4.12)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Yes	5,756	5,556 (96.5)	2.00 (1.88–2.13)	1.45 (1.41–1.49)	<0.001	1.35 (1.31–1.39)	<0.001
Unknown	839	807 (96.2)	3.00 (2.67–3.33)	1.30 (1.22–1.40)	<0.001	1.13 (1.04–1.24)	0.007
Lung Met							
No	20,557	18,943 (92.2)	4.00 (3.88–4.12)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Yes	6,860	6,478 (94.4)	3.00 (2.84–3.16)	1.17 (1.14–1.20)	<0.001	1.07 (1.04–1.11)	<0.001
Unknown	1,109	1,059 (95.5)	3.00 (2.72–3.28)	1.24 (1.16–1.32)	<0.001	1.03 (0.95–1.11)	0.481
Histology							
AC	14,710	13,167 (89.5)	5.00 (4.85–5.15)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
SLC	2,655	2,563 (96.5)	3.00 (2.84–3.16)	1.54 (1.47–1.60)	<0.001	1.40 (1.34–1.46)	<0.001

Table 3 (*continued*)

Table 3 (*continued*)

Subject characteristics	Patients' No. of LC with BM		Median survival (IQR), months	Univariate		Multivariate	
	Overall	Dead, n (%)		HR (95%CI)	P value	HR (95%CI)	P value
ASLC	306	286 (93.5)	4.00 (3.21–4.79)	1.18 (1.05–1.33)	0.005	1.17 (1.04–1.32)	0.008
LCLC	573	539 (94.1)	4.00 (3.43–4.57)	1.26 (1.16–1.37)	<0.001	1.23 (1.13–1.34)	<0.001
SCLC	4,727	4,559 (96.5)	5.00 (4.72–5.29)	1.28 (1.23–1.32)	<0.001	1.11 (1.07–1.15)	<0.001
Others	5,555	5,366 (96.6)	2.00 (1.90–2.10)	1.76 (1.71–1.82)	<0.001	1.59 (1.54–1.64)	<0.001
Surg(pri)							
No	27,712	25,867 (93.3)	4.00 (3.91–4.091)	1.00 (Ref)	1.00	1.00 (Ref)	1.00
Yes	755	560 (74.2)	13.00 (11.41–14.59)	0.89 (0.68–1.17)	0.415	0.55 (0.50–0.60)	<0.001
Unknown	59	53 (89.8)	2.00 (1.20–2.80)	0.39 (0.30–0.52)	<0.001	0.86 (0.65–1.12)	0.261

LC, lung carcinoma; BM, brain metastases; Met, metastases; AI, American Indian/Alaska Native; API, Asian or Pacific Islander; MB, main bronchus; UL, upper lobe; ML, middle lobe; LL, lower lobe; OL, overlapping lesion of lung; AC, adenocarcinoma; SLC, squamous lung carcinoma; ASLC, adenosquamous lung carcinoma; LCLC, large cell lung carcinoma; SCLC, small cell lung carcinoma; Surg(pri), surgical treatments of primary site.

to predict the development and prognosis of LC patients, respectively. Our results showed that BM prevalence in LC patients was 13.33%, consistent with previous studies that ranged from 10.4% to 19.9% (3,16). In accordance with the consequence of risk factors analysis, we specified that BM diagnosis is more likely among LC patients with the following features: 41 to 60 years of age, API race, any Medicaid status and uninsured status, tumor location in upper lobe and lower lobe, higher T stage, more lymphatic metastasis (N), higher grade stage, presence of metastases outside of the brain, and AC histological subtype. Part of the above risk factors has also been determined in previous studies (17-20). It is worth noting that a few previous works found a higher incidence of BM in women with NSCLC than in men (3,21,22). We observed no significant differences in BM incidence of LC without specific histological subtype according to sex. More studies are warranted to determine the potential cause.

A cohort of 28,526 LC patients with BM who received follow-up for over one year was utilized for survival analysis. Results suggest that the median OS is 4 months, correlating with previous findings (4,22). It has been widely discussed in many pieces of research that older age, poor tumor grade, extracranial metastases, non-AC histological subtype, and absence of definitive therapy were negative relevant predictors in the prognosis of LC patients with BM (23-26). In the present study, we found that patients with the following features also had less favorable outcomes: male, uninsured status, higher T stage, lymph node involvement, and main bronchus tumor location. No previous studies reported that API race patients had better survival than other races; however, this was determined in our study. A previous study reported that tumor location in the main bronchus is an independent risk factor for distant metastasis of lung adenocarcinoma. Interestingly, in this study, it was found that patients with LC in the main bronchus are less likely to develop BM than those without main bronchus location, yet have relatively poor prognosis (27). Further research is needed to reveal the potential cause.

National Comprehensive Cancer Network (NCCN) guidelines do not recommend BM screening for patients who are asymptomatic and in clinical stage I (9). While brain imaging has been increasingly applied to newly-diagnosed LC patients, there are still many missed diagnoses of BM in early-stage LC or asymptomatic patients (28). To achieve timely BM screening for the suitable population, an easy operating prediction nomogram was constructed from our data set. A quantified metastatic risk could be generated utilizing the constructed nomogram based on established risk factors, including: age, race, insurance status, primary tumor site, primary tumor, N stage, tumor differentiation grade, presence, or absence of metastases in the liver or bone, and histology. Nomogram performance was evaluated from two aspects; it is qualified for predicting the risk of BM development. Moreover, internal validation results indicated excellent stability of this model's predictive accuracy.

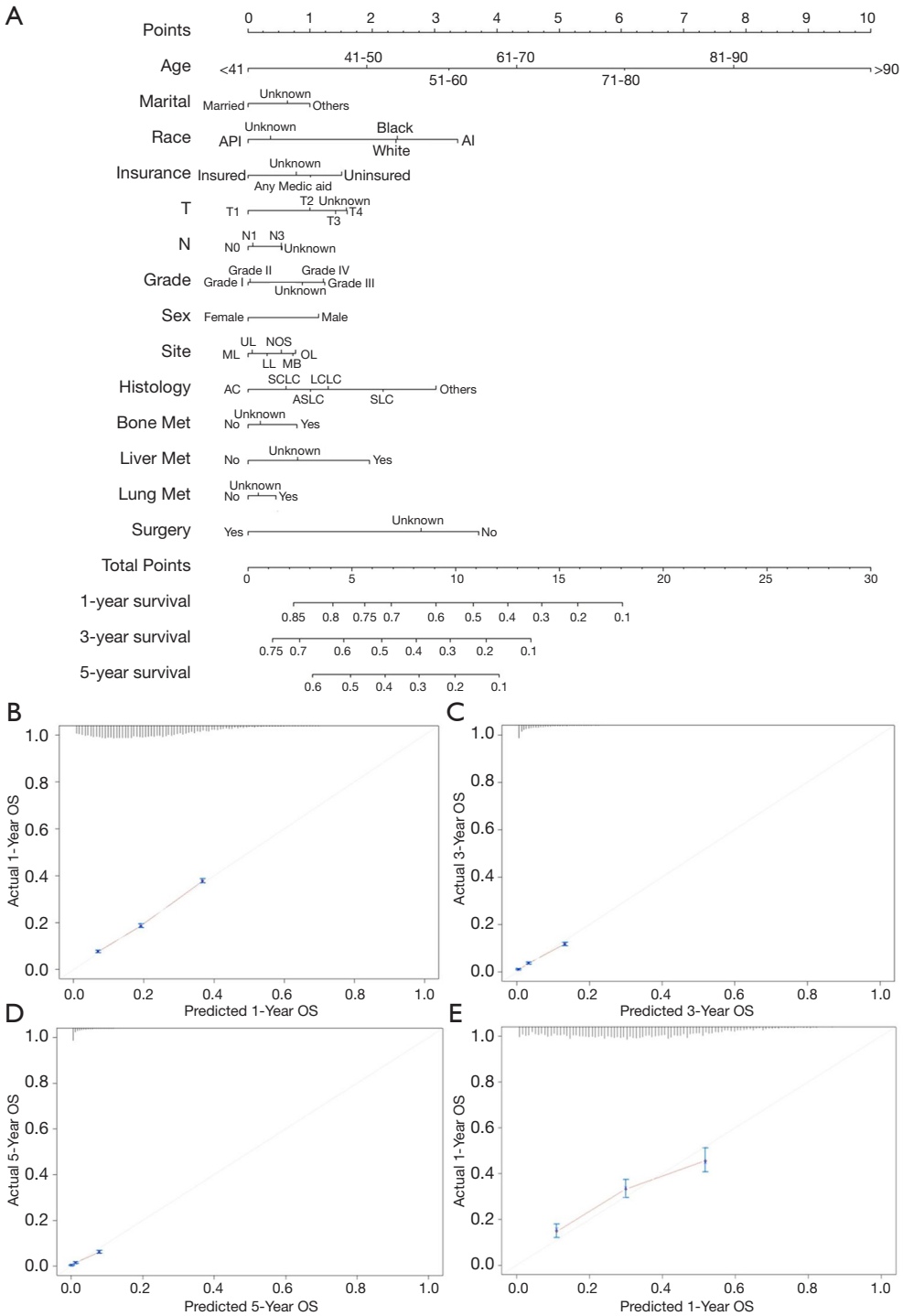

Figure 3 The nomogram for predicting prognosis of LC patients with BM (A) and the calibration curve for evaluating the performance in predicting 1-year (B), 3year (C), 5-year (D) and external performance in 1-year (E) survival probability. BM, brain metastasis; LC, lung cancer.

Many parameters influence the treatment choice for LC patients with BM (25). It is vital to determine more effective treatments to achieve longer survival and better quality of life (23). The nomogram for predicting OS was built based on factors determined by multivariate Cox regression analysis; the C-index was 0.88, with the addition of the calibration curve proving the model's prediction capacity. In some previous studies, different scoring systems to evaluate the prognosis of LC patients with BM were established: Recursive Partitioning Analysis (RPA), Gaspar *et al.* (24); Diagnosis-Specific Graded Prognostic Assessment (DS-GPA), Sperduto *et al.* (29); Graded Prognostic Assessment for Lung Cancer Using Molecular Markers (Lung-molGPA), Sperduto *et al.* (19). From these, we have learned that other factors—such as number of brain metastases, gene alterations, and treatment for brain metastases—contribute to patient survival.

Our study had several limitations arising from the SEER database itself. First, the actual BM rate in LC patients may be underestimated due to non-recording in asymptomatic cases and missed information related to BM development during disease course. Second, patient data regarding comorbidities, smoking history, and molecular biological characteristics are not available; these were supposed to be the underlying risk and prognosis factors. Third, due to data access limitations, the specific treatment program of study subjects could not be obtained. Finally, the follow-up period for the external validation group was short. These limitations may affect the model's predictive accuracy. Still, the calibration and discrimination ability of the prediction model reached a high level, suggesting the prediction model may be applied to a more extensive and heterogeneous population; considering the deficiencies of the external validation set, the research results must be further verified.

Conclusions

In LC patients, BM was highly developed with a prevalence of 13.33%, varying with age. The prognosis of LC patients with BM was relatively poor with a median OS of four months. Our nomograms showed good performance and stability in predicting BM development and prognosis, possibly aiding clinicians in identifying high-risk patients and tailoring treatment regimens for LC patients with BM.

Acknowledgments

We thank Xin Wang for his contribution on collecting data and editing the language of the manuscript. This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors. The authors acknowledge the efforts of the National Cancer Institute; the Office of Research, Development and Information, CMS; Information Management Services (IMS), Inc.; and the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER-Medicare database.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD checklist. Available at http://dx.doi.org/10.21037/tcr-20-2745

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-20-2745). The authors have no conflicts of interest to declare.

Ethical 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 study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The data about cancer in the SEER database is continually reported in every state of the United States and retrieved with no need for informed patient consent.

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;68:394-424.

2. Guo X, Zhang C, Ma W, et al. Patterns of bone metastases in newly diagnosed colorectal cancer: a real-world analysis in the SEER database. Int J Colorectal Dis 2019;34:533-43.
3. Barnholtz-Sloan JS, Sloan AE, Davis FG, et al. Incidence proportions of brain metastases in patients diagnosed (1973 to 2001) in the Metropolitan Detroit Cancer Surveillance System. J Clin Oncol 2004;22:2865-72.
4. Cagney DN, Martin AM, Catalano PJ, et al. Incidence and prognosis of patients with brain metastases at diagnosis of systemic malignancy: a population-based study. Neuro Oncol 2017;19:1511-21.
5. Nayak L, Lee EQ, Wen PY. Epidemiology of brain metastases. Curr Oncol Rep 2012;14:48-54.
6. Nam MW, Kim JM, Cheong JH, et al. Paradoxical relationship between body mass index and bone mineral density in patients with non-small cell lung cancer with brain metastasis. PLoS One 2019;14:e0218825.
7. Pedrosa RMSM, Mustafa DAM, Aerts JGJV, et al. Potential Molecular Signatures Predictive of Lung Cancer Brain Metastasis. Front Oncol 2018;8:159.
8. Ascha MS, Ostrom QT, Wright J, et al. Lifetime Occurrence of Brain Metastases Arising from Lung, Breast, and Skin Cancers in the Elderly: A SEER-Medicare Study. Cancer Epidemiol Biomarkers Prev 2019;28:917-25.
9. Rueda-Clausen CF, Ogunleye AA, Sharma AM. Health Benefits of Long-Term Weight-Loss Maintenance. Annu Rev Nutr 2015;35:475-516.
10. Gong L, Wang QI, Zhao L, et al. Factors affecting the risk of brain metastasis in small cell lung cancer with surgery: is prophylactic cranial irradiation necessary for stage I-III disease? Int J Radiat Oncol Biol Phys 2013;85:196-200.
11. Harada H, Asakura H, Ogawa H, et al. Prognostic factors in patients with brain metastasis from non-small cell lung cancer treated with whole-brain radiotherapy. Journal of Cancer Research and Therapeutics 2016;12:267-70.
12. Yu X, Fan Y. Real-World Data on Prognostic Factors for Overall Survival in EGFR-Mutant Non-Small-Cell Lung Cancer Patients with Brain Metastases. J Cancer 2019;10:3486-93.
13. Zheng Y, Wang L, Zhao W, et al. Risk factors for brain metastasis in patients with small cell lung cancer without prophylactic cranial irradiation. Strahlenther Onkol 2018;194:1152-62.
14. Kim YJ, Kim JS, Kim IA. Molecular subtype predicts incidence and prognosis of brain metastasis from breast cancer in SEER database. J Cancer Res Clin Oncol 2018;144:1803-16.
15. Zhang F, Zheng W, Ying L, et al. A Nomogram to Predict Brain Metastases of Resected Non-Small Cell Lung Cancer Patients. Ann Surg Oncol 2016;23:3033-9.
16. Waqar SN, Samson PP, Robinson CG, et al. Non-small-cell Lung Cancer With Brain Metastasis at Presentation. Clin Lung Cancer 2018;19:e373-9.
17. Hubbs JL, Boyd JA, Hollis D, et al. Factors associated with the development of brain metastases: analysis of 975 patients with early stage nonsmall cell lung cancer. Cancer 2010;116:5038-46.
18. Lim JH, Um SW. The risk factors for brain metastases in patients with non-small cell lung cancer. Ann Transl Med 2018;6:S66.
19. Sperduto PW, Yang TJ, Beal K, et al. Estimating Survival in Patients With Lung Cancer and Brain Metastases: An Update of the Graded Prognostic Assessment for Lung Cancer Using Molecular Markers (Lung-molGPA). JAMA Oncol 2017;3:827-31.
20. Yawn BP, Wollan PC, Schroeder C, et al. Temporal and gender-related trends in brain metastases from lung and breast cancer. Minnesota medicine 2003;86:32-7.
21. Goncalves PH, Peterson SL, Vigneau FD, et al. Risk of brain metastases in patients with nonmetastatic lung cancer: Analysis of the Metropolitan Detroit Surveillance, Epidemiology, and End Results (SEER) data. Cancer 2016;122:1921-7.
22. Riihimäki M, Hemminki A, Fallah M, et al. Metastatic sites and survival in lung cancer. Lung Cancer 2014;86:78-84.
23. Andratschke N, Kraft J, Nieder C, et al. Optimal management of brain metastases in oncogenic-driven non-small cell lung cancer (NSCLC). Lung Cancer 2019;129:63-71.
24. Bernhardt D, Adeberg S, Bozorgmehr F, et al. Outcome and prognostic factors in single brain metastases from small-cell lung cancer. Strahlenther Onkol 2018;194:98-106.
25. Jolly K, Chambers R. Improving outcomes for patients with obesity. Practitioner 2014;258:29-31, 3.
26. Zindler JD, Jochems A, Lagerwaard FJ, et al. Individualized early death and long-term survival prediction after stereotactic radiosurgery for brain metastases of non-small cell lung cancer: Two externally validated nomograms. Radiother Oncol 2017;123:189-94.
27. Yang L, Wang S, Gerber DE, et al. Main bronchus

location is a predictor for metastasis and prognosis in lung adenocarcinoma: A large cohort analysis. Lung Cancer 2018;120:22-6.
28. Ando T, Kage H, Saito M, et al. Early stage non-small cell lung cancer patients need brain imaging regardless of symptoms. Int J Clin Oncol 2018;23:641-6.
29. Sperduto PW, Kased N, Roberge D, et al. Summary report on the graded prognostic assessment: an accurate and facile diagnosis-specific tool to estimate survival for patients with brain metastases. J Clin Oncol 2012;30:419-25.

Cite this article as: Zuo C, Liu G, Bai Y, Tian J, Chen H. The construction and validation of the model for predicting the incidence and prognosis of brain metastasis in lung cancer patients. Transl Cancer Res 2021;10(1):22-37. doi: 10.21037/tcr-20-2745

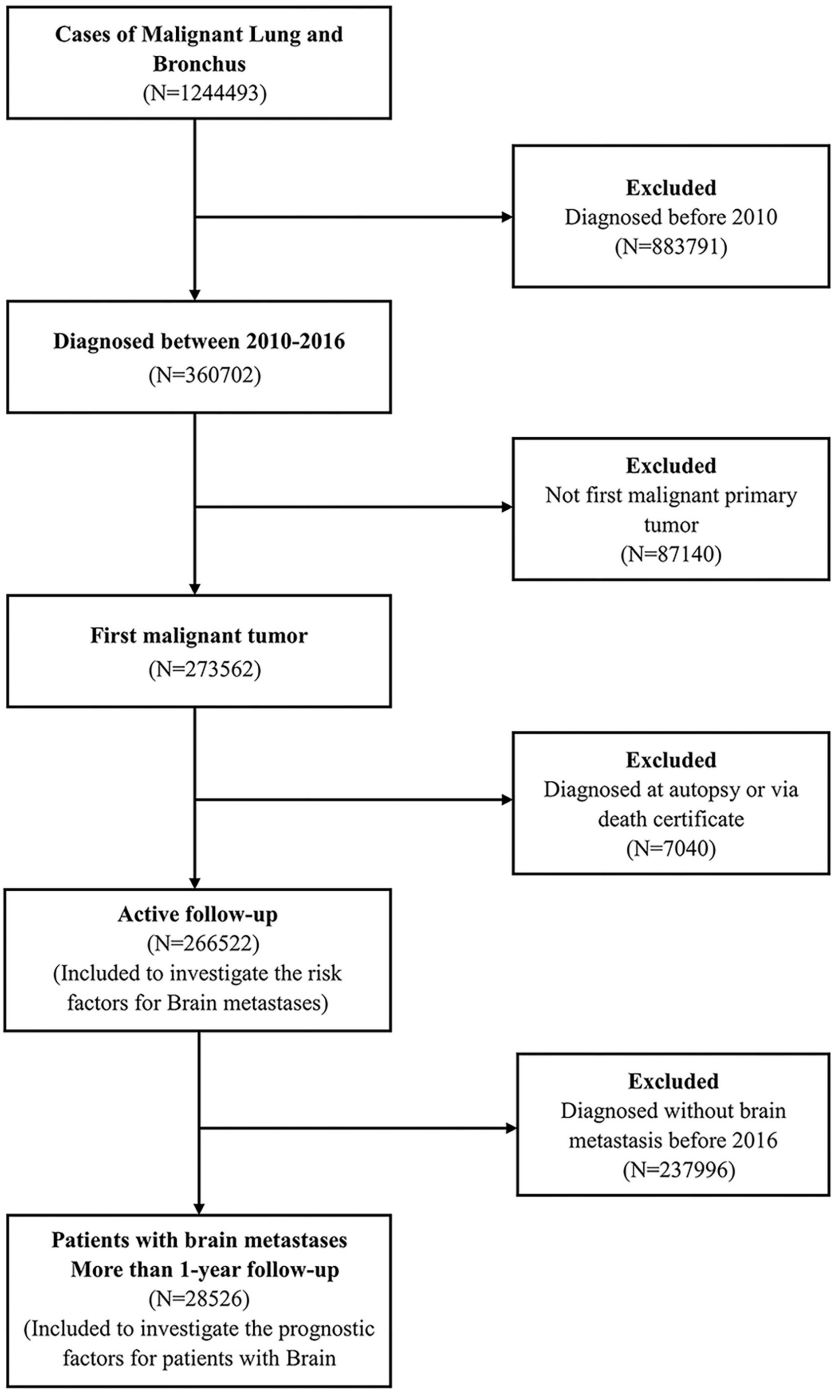

Figure S1 The flow-chart for the population selection.

http://dx.doi.org/10.21037/tcr-20-2745