Prognostic nomograms and risk-stratifying systems for predicting survival in patients with resected pT2-4aN0M0 esophageal carcinoma

Xiao-Meng Dou[1,2#], Na Zhang[3#], Yan-Yan Fang[1,2#], Bo-Han Zhang[1,2], Jie-Jing Liao[1,2], Jing-Sheng Cai[1,2], Jin-Bo Li[1,2]

[1]Department of Thoracic Surgery, Sun Yat-sen University Cancer Center, Guangzhou, China; [2]State Key Laboratory of Oncology in South China; Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China; [3]Department of Radiotherapy, Sun Yat-Sen Memorial Hospital, Guangzhou, China

Contributions: (I) Conception and design: JB Li, JS Cai; (II) Administrative support: JB Li; (III) Provision of study materials or patients: XM Dou, BH Zhang, JJ Liao; (IV) Collection and assembly of data: N Zhang, YY Fang; (V) Data analysis and interpretation: XM Dou, N Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Dr. Jin-Bo Li; Dr. Jing-Sheng Cai. Department of Thoracic Surgery, Sun Yat-sen University Cancer Center, No. 651, Dongfeng East Road, Guangzhou 510060, China. Email: lijb@sysucc.org.cn; caijs@sysucc.org.cn.

Background: According to the National Comprehensive Cancer Network (NCCN) guidelines, surveillance or adjuvant chemoradiation is recommended for patients with completely resected pT2-4aN0M0 esophageal carcinoma (EC). Due to this population's variant prognosis, we developed novel nomograms to define the high-risk patients who may need closer follow-up or even post-operative therapy.

Methods: Cases with resected pT2-4aN0M0 EC from the Surveillance, Epidemiology, and End Results (SEER) database and the Sun Yat-sen University Cancer Center (SYSUCC) were enrolled in the study. The SEER database cases were randomly assigned into the training cohort (SEER-T) and the internal validation cohort (SEER-V). Cases from the SYSUCC served as the external validation cohort (SYSUCC-V). Overall survival (OS) and cancer specific survival (CSS) were compared between groups. Multivariate analyses were applied to identify the prognostic factors. Nomograms and risk-classifying systems were developed. The nomograms' performances were evaluated by concordance index (C-index), calibration plots and decision curve analysis (DCA).

Results: A total of 2,441 eligible EC cases (SEER-T, n=839; SEER-V, n=279; SYSUCC-V, n=1,323) were included. Age, sex, chemotherapy, lymph node harvested (LNH) and T stage were identified as the independent predictors for CSS. Regarding OS, it also included the prognostic factor of histology. Nomograms were formulated. For CSS, the C-index was 0.68 [95% confidence interval (CI): 0.66–0.71], 0.67 (95% CI: 0.63–0.71) and 0.61 (95% CI: 0.59–0.63) for the SEER-T, SEER-V, and SYSUCC-V, respectively. For OS, the C-index was 0.69 (95% CI: 0.66–0.72), 0.64 (95% CI: 0.59–0.69) and 0.62 (95% CI: 0.61–0.63) for the SEER-T, SEER-V, and SYSUCC-V, respectively. The calibration curves and DCA showed good performances of the nomograms. In further analyses, risk-classification systems stratified pT2-4aN0M0 EC into low-risk and high-risk subgroup. The OS and CSS curves of these 2 subgroups, in the full analysis set or stratified by TNM stage, histology, T stage and LNH categories, showed significant distinctions.

Conclusions: The novel prognostic nomograms and risk-stratifying systems which separated resected pT2-4aN0M0 esophageal carcinoma patients into the low-risk and high-risk prognostic groups were developed. It may help clinicians estimate individual survival and develop individualized treatment strategies.

Keywords: Esophageal carcinoma (EC); nomogram; survival

Introduction

Esophageal carcinoma (EC), the 8th most common cancer and 6th most common cause of cancer-related death worldwide, shows a gradual annual increase incidence (1-3). Although surgery has been considered the preferred first-line treatment for locally advanced EC (4), the survival rate differs substantially between cases. According to the latest version of National Comprehensive Cancer Network (NCCN) guidelines (5), follow-up is recommended for patients with completely resected pT2-4aN0M0 squamous cell carcinoma, and follow-up or adjuvant chemoradiation is recommended for patients with completely resected pT2-4aN0M0 adenocarcinoma. Regarding the variant prognosis of this population, the development of a prognostic model to facilitate the risk stratification and define high-risk cases that may benefit from closer follow-up or adjuvant therapy is imperative.

The 8th edition of the Tumor-Node-Metastasis (TNM) staging system of the American Joint Commission on Cancer (AJCC) is currently applied for predicting the survival of EC patients and guiding clinical practice (6). However, several previous studies have demonstrated the unsatisfactory predictive efficacy of the TNM staging system (7,8). Thus, additional prognostic factors to the TNM staging system, such as age, sex, chemotherapy information and lymph node harvested (LNH), should be considered when evaluated these patients.

In the current study, we analyzed resected cases of pT2-4aN0M0 EC from the Surveillance, Epidemiology, and End Results (SEER) database and Sun Yat-sen University Cancer Center (SYSUCC). The independent prognostic factors of overall survival (OS) and cancer specific survival (CSS) were identified, and prognostic nomograms were formulated. Risk-stratifying systems, based on the nomograms, were established to select the patients with an enhanced likelihood of poor survival. Different statistical methods were carried out to validate the models. Using the models, clinicians could refine treatment strategies and subsequently improve patient prognosis.

We present the following article in accordance with the TRIPOD reporting checklist (available at http://dx.doi.org/10.21037/jtd-20-3393).

Methods

Patient selection

Patients diagnosed with EC from 2004 to 2015 were retrospectively reviewed from the SEER database using SEER*Stat 8.3.5 software (https://seer.cancer.gov/seerstat/), and EC cases from 2001 to 2014 in the SYSUCC were also included. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Informed consent of enrolled patients and SYSUCC review board approval of were obtained in this study. The authenticity of the study has been validated by uploading the raw data onto the Research Data Deposit public platform (www.researchdata.org.cn), with the approval RDD number RDDA2020001859.

The inclusion criteria were as follows: (I) pathologically diagnosed with squamous cell carcinoma or adenocarcinoma; (II) underwent surgery and (III) pT2-4aN0M0 (according to the 8th TNM staging system). The exclusion criteria were as follows: (I) age <18 years; (II) clinicopathological or follow-up information was unavailable and (III) perioperative death.

Figure 1 shows the detailed flowchart of this study. Cases with resected pT2-4aN0M0 EC from the SEER database were randomly divided into the training cohort (SEER-T) and the internal validation cohort (SEER-V) at a ratio of 3:1. Cases with pT2-4aN0M0 EC from the SYSUCC served as the external validation cohort (SYSUCC-V).

Data collection

Clinical and demographic variables including sex, age at diagnosis, surgical type, pTNM stage, histology, tumor grade, tumor size, LNH, tumor location, race, chemotherapy, radiotherapy, follow-up information, and vital status were retrospectively collected. Continuous variables such as age, LNH, and tumor size were transformed into categorical variables based on the cutoff values determined by the X-tile software (9). TNM staging was performed according to the 8th edition of the AJCC TNM staging system (6).
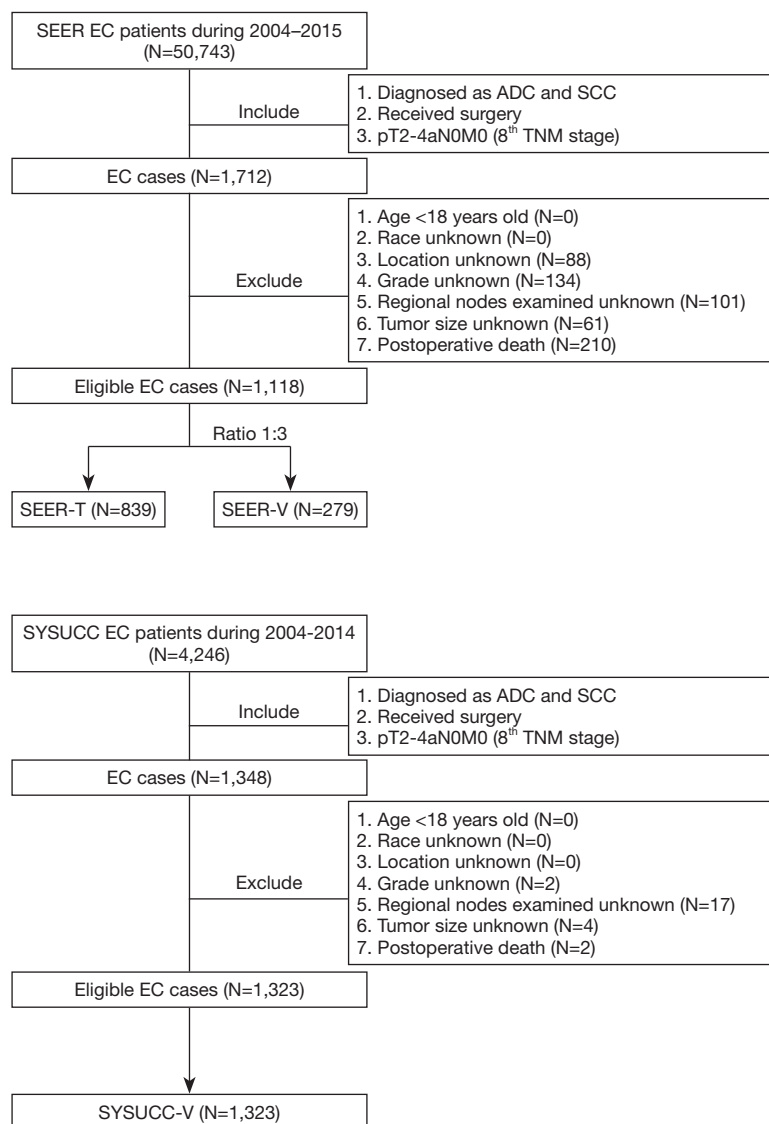
Figure 1 The flow chart of the patient selection in the SEER database and the SYSUCC. SEER, Surveillance, Epidemiology, and End Results; SYSUCC, Sun Yat-sen University Cancer Center; EC, esophageal carcinoma; ADC, adenocarcinoma; SCC, squamous cell carcinoma; SEER-T, SEER training cohort; SEER-V, SEER validation cohort; SYSUCC-V, SYSUCC validation cohort.

Statistical analysis

Statistical analysis was carried out by R version 3.5.2 (http://www.r-project.org), IBM SPSS Statistics (version 25.0; IBM, Armonk, NY, USA), X-tile software (9), and GraphPad Prism 8. OS was defined as the interval from the date of diagnosis to the date of death from any cause or the last follow-up. CSS was defined as the interval from the date of diagnosis to the date of lung cancer-related death or the last follow-up. All time events were estimated using the Kaplan-Meier method with a log-rank test.

Multivariate analyses were applied to identify prognostic factors, and nomograms were developed. Each number/category of the independent prognostic variables in the nomogram was assigned a score on the point scale and finally added up to a total score. After that, a cutoff value, achieved by X-tile software, was applied to transform the continuous score into 2 categories (low-risk and high-risk groups). The concordance index (C-index) (10), calibration curves and decision curve analysis (DCA) (11) were applied to verify the predicted effect of the nomogram.

2366

Dou et al. Prognosis of resected T2-4N0M0 esophageal carcinoma

Pearson's χ^2 test or Fisher's exact test was used to compare categorical variables between groups. Two-sided P<0.05 was considered statistically significant.

Results

Patients characteristics

Between 2004 and 2015, a total of 50,743 EC cases from the SEER database were reviewed. The aforementioned selection criteria yielded a study population of 1,118 eligible EC patients (SEER-T, n=839 cases; SEER-V, n=279

cases). The general clinical features are listed in *Table 1*. The median age was 65 years old (range, 26–91 years). Almost all the patients were white (88.7%). Male patients (79.2%) and adenocarcinoma (70.3%) accounted for a substantial proportion of the entire cohort. There was a slightly high percentage of low-third tumors (75.3%). Over half of patients received radiotherapy (59.9%). The median follow-up time was 30.0 months (range, 1.0–155.0 months). All covariates between the training cohort and the validation cohort were well-balanced (*Table 1*). The clinical characteristics of the SYSUCC-V cohort are summarized in Table S1.

Table 1 Characteristics of the patients at baseline from SEER database

Characteristic	Total (N=1,118), No. of patients (%)	Training cohort (N=839), No. of patients (%)	Validation cohort (N=279), No. of patients (%)	P
Age				0.706
≤63	514 (46.0)	391 (46.6)	123 (44.1)	
63–73	387 (34.6)	285 (34.0)	102 (36.6)	
>73	217 (19.4)	163 (19.4)	54 (19.4)	
Sex				0.240
Male	886 (79.2)	658 (78.4)	228 (81.7)	
Female	232 (20.8)	181 (21.6)	51 (18.3)	
Race				0.836
White	992 (88.7)	746 (88.9)	246 (88.2)	
Black	68 (6.1)	49 (5.8)	19 (6.8)	
Other[‡]	58 (5.2)	44 (5.2)	14 (5.0)	
Marital status				0.878
Un-married	409 (36.6)	308 (36.7)	101 (36.2)	
Married	709 (63.4)	531 (63.3)	178 (63.8)	
Location				0.628
Upper	43 (3.8)	30 (3.6)	13 (4.7)	
Middle	193 (17.3)	149 (17.8)	44 (15.8)	
Low	842 (75.3)	628 (74.9)	214 (76.7)	
Overlapping	32 (3.8)	32 (3.8)	8 (2.9)	
Histology				0.982
Squamous cell carcinoma	332 (29.7)	249 (29.7)	83 (29.7)	
Adenocarcinoma	786 (70.3)	590 (70.3)	196 (70.3)	

Table 1 (*continued*)

Table 1 (*continued*)

Characteristic	Total (N=1,118), No. of patients (%)	Training cohort (N=839), No. of patients (%)	Validation cohort (N=279), No. of patients (%)	P
Grade				0.765[†]
Well	81 (7.2)	57 (6.8)	24 (8.6)	
Moderately	563 (50.4)	425 (50.7)	138 (49.5)	
Poor	466 (41.7)	351 (41.8)	115 (41.2)	
Undifferentiated	8 (0.7)	6 (0.7)	2 (0.7)	
Radiotherapy				0.768
No	448 (40.1)	335 (39.9)	113 (40.5)	
Yes	670 (59.9)	504 (60.1)	166 (59.5)	
Chemotherapy				0.299
No	726 (64.9)	552 (65.8)	174 (62.4)	
Yes	392 (35.1)	287 (34.2)	105 (37.6)	
Esophagectomy				0.232
Total	973 (87.0)	736 (87.7)	237 (84.9)	
Partial	145 (13.0)	103 (12.3)	42 (15.1)	
Lymph node harvested				0.684
≤6	272 (24.3)	209 (24.9)	63 (22.6)	
6–17	526 (47.0)	394 (47.0)	132 (47.3)	
>17	320 (28.6)	236 (28.1)	84 (30.1)	
Tumor size (mm)				0.985
≤45	741 (66.3)	555 (66.2)	186 (66.7)	
45–66	238 (21.3)	179 (21.3)	59 (21.1)	
>66	139 (12.4)	105 (12.5)	34 (12.2)	
T stage				0.494[†]
2	402 (36.0)	294 (35.0)	108 (38.7)	
3	698 (62.4)	532 (63.4)	166 (59.5)	
4a	18 (1.6)	13 (1.5)	5 (1.8)	
TNM stage				0.656[†]
IB	13 (1.2)	9 (1.1)	4 (1.4)	
IC	26 (2.3)	17 (2.0)	9 (3.2)	
IIA	462 (41.3)	344 (41.0)	118 (42.3)	
IIB	599 (53.6)	456 (54.4)	143 (51.3)	
IIIB	18 (1.6)	13 (1.5)	5 (1.8)	

[†]Fisher's exact test; [‡]Other includes American Indian, Alaska native, Asian and pacific islander. SEER, Surveillance, Epidemiology, and End Results; TNM, tumor-node-metastasis.

2368

Dou et al. Prognosis of resected T2-4N0M0 esophageal carcinoma

Cox regression analysis

Regarding OS, a univariate Cox regression analysis suggested that age, sex, location, histology, tumor grade, radiotherapy, chemotherapy, LNH, and T stage were associated with patient's survival (*Table 2*). The multivariate analysis confirmed that age, sex, histology, chemotherapy, LNH, and T stage were independent factors that impacted OS (*Table 2*).

As for CSS, the univariate Cox regression analysis revealed that age, sex, race, location, histology, tumor grade, radiotherapy, chemotherapy, LNH, and T stage were prognostic predictors (*Table 3*). The multivariate analysis confirmed that age, sex, chemotherapy, LNH, and T stage were statistically significant prognostic factors (*Table 3*).

Nomograms and risk-stratifying Systems

Statistically significant factors from the multivariate analysis were entered into the nomogram, which was developed to calculate 3, 5 and 10-year OS and CSS probabilities (*Figure 2*). With regard to OS, the nomogram indicated that LNH was the strongest predictor, followed by age and chemotherapy (*Figure 2A*). Concerning CSS, the graphic demonstrated that T stage was the strongest predicator, followed by LNH and age (*Figure 2B*).

In further analyses, each independent prognostic variable was assigned a risk score according to the point scale using R software. A total score was calculated. The total cutoff values (21.0 for OS and 14.0 for CSS), determined by X-tile software, were used to dichotomize EC cases into low-risk and high-risk subgroups (*Table 4*).

Calibration and validation

For OS, the C-index was 0.69 [95% confidence interval (CI): 0.66–0.72], 0.64 (95% CI: 0.59–0.69), and 0.62 (95% CI: 0.61–0.63) for the SEER-T cohort, the SEER-V cohort, and the SYSUCC-V cohort, respectively. For CSS, the C-index was 0.68 (95% CI: 0.66–0.71), 0.67 (95% CI: 0.63–0.71), and 0.61 (95% CI: 0.59–0.63) for the SEER-T, SEER-V, and SYSUCC-V, respectively. The calibration curves indicated an optimal consistency between predicted and actual observations (Figure S1 for OS and Figure S2 for CSS). The DCA demonstrated that the nomograms had a satisfying clinical net benefit when compared with the TNM staging system (Figure S3A for OS and Figure S3B for CSS).

Table 2 Univariate and multivariate COX proportional hazard model analysis for overall survival

Characteristic	Univariate analysis			Multivariate analysis[†]		
	HR	95% CI	P	HR	95% CI	P
Age			<0.001			<0.001
≤63	Ref			Ref		
63–73	1.211	0.985–1.488		1.212	0.983–1.494	
>73	2.058	1.643–2.579		1.700	1.336–2.163	
Sex			0.058			0.001
Male	Ref			Ref		
Female	0.806	0.646–1.007		0.674	0.529–0.857	
Race			0.169			
White	Ref					
Black	1.304	0.924–1.841				
Other[‡]	0.797	0.519–1.224				
Marital status			0.115			
Un-married	Ref					
Married	0.863	0.719–1.037				

Table 2 (*continued*)

Table 2 (*continued*)

Characteristic	Univariate analysis			Multivariate analysis[†]		
	HR	95% CI	P	HR	95% CI	P
Location			0.004			0.220
Upper	Ref			Ref		
Middle	0.654	0.421–1.016		0.750	0.476–1.182	
Low	0.516	0.344–0.775		0.646	0.410–1.016	
Overlapping	0.503	0.277–0.915		0.563	0.300–1.055	
Histology			0.001			0.035
Squamous cell carcinoma	Ref			Ref		
Adenocarcinoma	0.718	0.595–0.866		0.768	0.602–0.981	
Grade			0.091			0.113
Well	Ref			Ref		
Moderately	1.373	0.928–2.031		1.454	0.976–2.166	
Poor	1.588	1.070–2.356		1.631	1.089–2.444	
Undifferentiated	1.127	0.343–3.710		1.434	0.432–4.753	
Radiotherapy			<0.001			0.509
No	Ref			Ref		
Yes	0.716	0.599–0.857		1.113	0.810–1.530	
Chemotherapy			<0.001			<0.001
No	Ref			Ref		
Yes	1.754	1.466–2.099		1.892	1.371–2.610	
Lymph node harvested			<0.001			<0.001
≤6	Ref			Ref		
6–17	0.755	0.615–0.928		0.742	0.601–0.915	
>17	0.539	0.421–0.691		0.550	0.427–0.707	
Tumor size			0.357			
≤45	Ref					
45–66	0.989	0.790–1.238				
>66	0.812	0.611–1.081				
T stage			0.002			<0.001
2	Ref			Ref		
3	1.410	1.163–1.710		1.527	1.248–1.869	
4a	1.435	0.672–3.065		1.158	0.523–2.564	

[†]Variables with P value less than 0.1 are included in the multivariate analysis; [‡]Other includes American Indian, Alaska native, Asian and pacific islander. HR, hazard ratio; CI, confidence interval.

Table 3 Univariate and multivariate COX proportional hazard model analysis for cancer-specific survival

Characteristic	Univariate analysis			Multivariate analysis[†]		
	HR	95% CI	P	HR	95% CI	P
Age			<0.001			0.008
≤63	Ref			Ref		
63–73	1.185	0.925–1.520		1.193	0.927–1.534	
>73	1.905	1.449–2.504		1.595	1.187–2.143	
Sex			0.085			0.005
Male	Ref			Ref		
Female	0.787	0.599–1.034		0.654	0.486–0.880	
Race			0.036			0.094
White	Ref			Ref		
Black	1.649	1.126–2.417		1.540	1.022–2.322	
Other[‡]	1.001	0.622–1.612		0.909	0.552–1.496	
Marital status			0.166			
Un-married	Ref					
Married	0.855	0.686–1.067				
Location			0.023			0.222
Upper	Ref			Ref		
Middle	0.882	0.497–1.563		0.992	0.543–1.812	
Low	0.615	0.358–1.054		0709	0.388–1.296	
Overlapping	0.777	0.379–1.592		0.807	0.375–1.736	
Histology			0.013			0.582
Squamous cell carcinoma	Ref			Ref		
Adenocarcinoma	0.749	0.596–0.941		0.916	0.671–1.251	
Grade			0.062			0.116
Well	Ref			Ref		
Moderately	1.427	0.876–2.322		1.503	0.913–2.474	
Poor	1.769	1.085–2.884		1.779	1.076–2.940	
Undifferentiated	1.220	0.283–5.260		1.500	0.345–6.519	
Radiotherapy			0.002			0.525
No	Ref			Ref		
Yes	0.710	0.572–0.881		1.132	0.773–1.658	
Chemotherapy			<0.001			<0.001
No	Ref			Ref		
Yes	1.815	1.462–2.254		2.110	1.434–3.105	

Table 3 (*continued*)

Table 3 (*continued*)

Characteristic	Univariate analysis			Multivariate analysis[†]		
	HR	95% CI	P	HR	95% CI	P
Lymph node harvested			0.004			0.013
≤6	Ref			Ref		
6–17	0.806	0.626–1.039		0.791	0.610–1.026	
>17	0.604	0.449–0.814		0.633	0.468–0.858	
Tumor size			0.420			
≤45	Ref					
45–66	1.025	0.785–1.338				
>66	0.799	0.562–1.136				
T stage			<0.001			<0.001
2	Ref			Ref		
3	1.610	1.266–2.049		1.744	1.356–2.244	
4a	1.875	0.821–4.285		3.115	0.698–4.009	

[†]Variables with P value less than 0.1 are included in the multivariate analysis; [‡]Other includes American Indian, Alaska native, Asian and pacific islander. HR, hazard ratio; CI, confidence interval.

Stratification efficacy evaluation

The total score cutoff values were used to dichotomize EC cases into low-risk and high-risk subgroups. Kaplan-Meier curves for OS in the entire cohort, stratified by histology, TNM stage, T stage, and LNH were used to evaluate the risk-stratifying systems' efficacy. The low-risk subgroup's OS was significantly better than the high-risk subgroup for the entire cohort (P<0.001, *Figure 3A*). For histology, both in the squamous cell carcinoma group and adenocarcinoma group, low-risk patients had significantly favorable OS than high-risk patients (adenocarcinoma, P<0.001, *Figure 3B*; squamous cell carcinoma, P<0.001, *Figure 3C*). In the TNM stage subset analysis, low-risk cases had high survival levels compared with high-risk cases in all the three substages (stage I, P=0.003, *Figure 3D*; stage II, P<0.001, *Figure 3E*; stage IIIB, P=0.020, *Figure 3F*). In the T stage subsets analysis, low-risk cases had better OS rate compared with high-risk cases in all the three T substages (T2, P<0.001, *Figure 3G*; T3, P<0.001, *Figure 3H*; T4a, P=0.020, *Figure 3I*). Regarding LNH, the OS of low-risk cases was also superior than that of high-risk cases in all the three categories (LNH ≤6, P<0.001, *Figure 3J*; 6< LNH ≤17, P<0.001, *Figure 3K*; LNH >17, P<0.001, *Figure 3L*).

As for CSS, the survival curves of these two different risk subgroups, in the full analysis set (P<0.001, *Figure 4A*) or stratified by histology(adenocarcinoma, P<0.001, *Figure 4B*; squamous cell carcinoma, P<0.001, *Figure 4C*), TNM stage (stage I, P<0.001, *Figure 4D*; stage II, P<0.001, *Figure 4E*; stage IIIB, P=0.031, *Figure 4F*), T stage (T2, P<0.001, *Figure 4G*; T3, P<0.001, *Figure 4H*; T4a, P=0.031, *Figure 4I*) and LNH categories (LNH ≤6, P=0.001, *Figure 4J*; 6< LNH ≤17, P<0.001, *Figure 4K*; LNH > 17, P<0.001, *Figure 4L*), also showed significant distinctions.

Discussion

In the present study, novel prognostic nomograms, including routinely available factors such as age, sex, histology, chemotherapy, LNH, and T stage, was established to predict OS and CSS. The risk-stratifying systems stratified patients with resected pT2-4aN0M0 EC into two different prognostic groups. The validation of the models using the C-index, DCA, and calibration curves demonstrated its good performances. With the help of the easy-to-use models, clinicians could identify a subset of EC patients with poor survival who may need closer follow-up or intervention therapy.

According to the NCCN guidelines for EC (5), surveillance is recommended for patients with completely
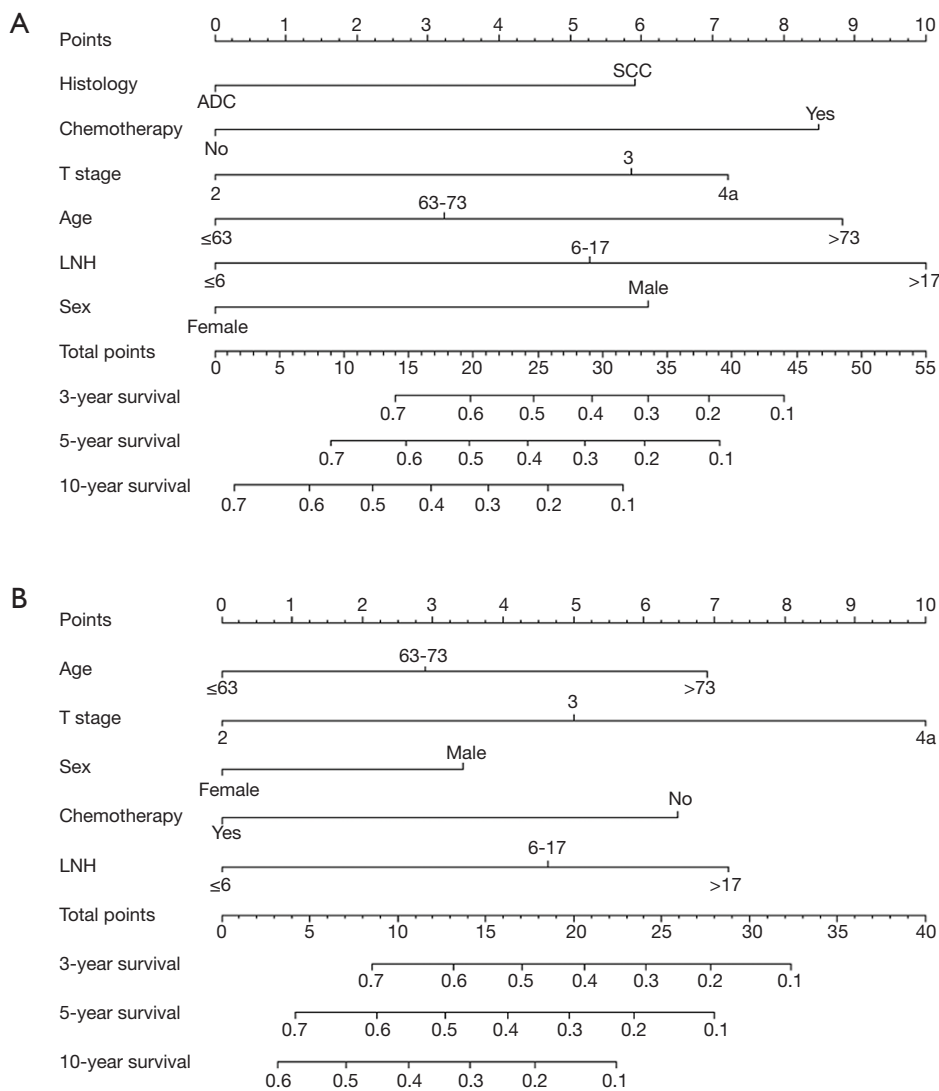
Figure 2 Nomogram for predicting the 3-year, 5-year and 10-year survival rate in resected stage pT2-4aN0M0 EC patients. OS (A) and CSS (B). OS, overall survival; CSS, cancer specific survival; EC, esophageal carcinoma; LNH, lymph node harvested.

resected pT2-4aN0M0 squamous cell carcinoma, and surveillance or adjuvant chemoradiation is recommended for patients with completely resected pT2-4aN0M0 adenocarcinoma. Previous studies have suggested a dismal 5-year survival rate for resectable EC (12). Therefore, it was necessary to identify the prognostic factors of resected pT2-4AN0M0 EC and formulate a nomogram to determine high-risk patients who may be the best adjuvant therapy candidates.

Compared with advanced tumors, the clinical manifestation of relapse or progression of nodal negative tumors might be less obvious. As we all know, there are

still many developing countries in the world. Esophageal carcinoma mostly occurs in the group of people with relatively poor economic situation, and most of them live in rural area far away. The expensive follow-up examinations and long distances form major obstacles for these patients to get scheduled follow-up examinations. Based on our classifying systems, we strongly recommend that the high-risk patients should return to hospital for a scheduled follow-up examination or even postoperative therapy, irrespective of cost.

Table 4 Constructed prognostic score to predict survival in stage pT2-4aN0M0 EC patients

	Category	Score
Covariate of OS		
Histology	ADC	0.0
	SCC	5.0
Chemotherapy	No	0.0
	Yes	7.0
T stage	2	0.0
	3	5.0
	4a	10.0
Age	≤63	0.0
	63–73	3.0
	>73	7.0
LNH	≤6	0.0
	6–17	4.0
	>17	8.0
Sex	Female	0.0
	Male	5.0
Risk-stratifying system for OS (cutoff value =21.0)		
Low risk		≤21.0
High risk		>21.0
Covariate of CSS		
Chemotherapy	No	0.0
	Yes	6.0
T stage	2	0.0
	3	5.0
	4a	10.0
Age	≤63	0.0
	63–73	2.0
	>73	4.0
LNH	≤6	0.0
	6–17	3.0
	>17	5.0
Sex	Female	0.0
	Male	3.0

Table 4 (*continued*)

Table 4 (*continued*)

	Category	Score
Risk-stratifying system for CSS (cutoff value =14.0)		
Low risk		≤14.0
High risk		>14.0

OS, overall survival; CSS, cancer specific survival; EC, esophageal carcinoma; ADC, adenocarcinoma; SCC, squamous cell carcinoma; LNH, lymph node harvested.

According to the current 8th edition of the AJCC TNM staging system (6), the covariates adopted to subdivide pT2-4aN0M0 EC are depth of tumor invasion, location, grade, and histology. Our nomograms confirmed that the TNM stage factors such as depth of tumor invasion and histology were the independent prognostic predictors, but grade and location were not. From our perspective, the relatively small number of cases might interpret, to a certain extent, the no statistically significance of grade. As is widely acknowledged, patients with upper-third EC are unlikely to undergo surgery due to anatomical difficulties and technique challenges (4), and the limited treatment options may impair the survival of these patients. Only 43 (3.8%) upper-third EC patients in our cohort, and all cases included in our study underwent surgical resection. Therefore, the impact of location on long-term survival was not deep enough to reach statistically significance. In our nomograms, 4 other prognostic factors (age, sex, LNH, and chemotherapy) were also included, which contributed to the nomograms' superior predictive value compared with the TNM staging system.

Despite several previously reported predictive models (7,8,13-20), a nomogram for resected pT2-4aN0M0 EC from two large databases with long-term follow-up has not been established. In the study by Du *et al.* (15), the authors reviewed data of 4,566 resected localized EC from the SEER database and developed a predictive nomogram. Unfortunately, there was unavailable clinicopathological information in their research, which could affect the model's predictive efficacy. Wu *et al.* (16) also analyzed 20,623 EC adenocarcinoma cases from the SEER dataset and established a nomogram to predict OS and CSS. Grade, T stage, N stage, M stage, underwent surgery, insurance record and marital status were entered into their nomogram which showed great
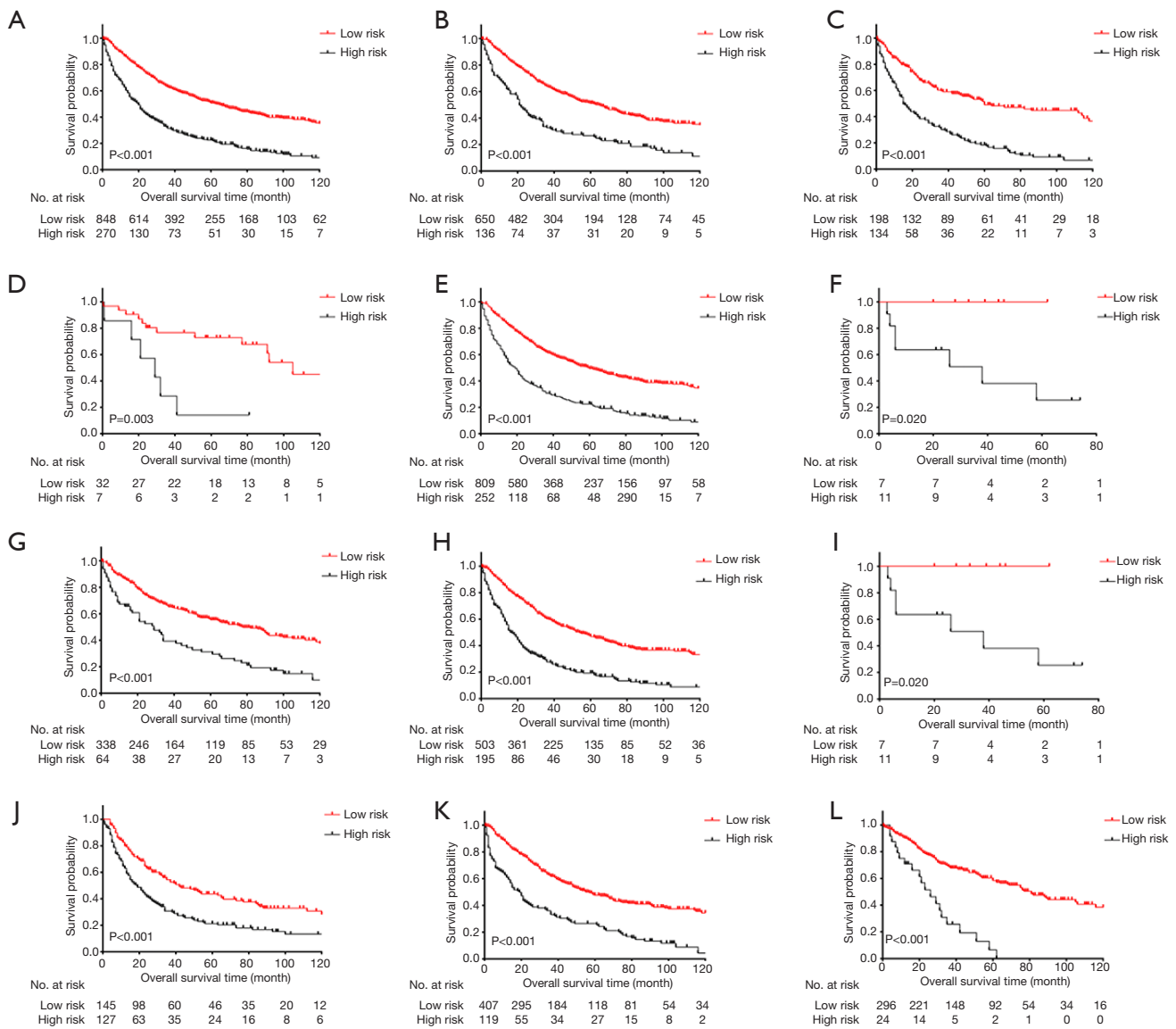
Figure 3 Kaplan-Meier estimates of OS in resected stage pT2-4aN0M0 EC patients. The entire cohort (A), ADC group (B), SCC group (C), Stage I group (D), Stage II group (E), Stage III group (F), T2 stage group (G), T3 stage group (H), T4a stage group (I), LNH ≤6 group (J), 6< LNH ≤17 group (K) and LNH >17 group (L). OS, overall survival; EC, esophageal carcinoma; ADC, adenocarcinoma; SCC, squamous cell carcinoma; LNH, lymph node harvested.

performance (C index: 0.720–0.733) when compared with our nomogram. However, chemotherapy, a critical prognostic factor for survival, was lacking in their study. Moreover, the tumor stage was diagnosed either pathologically or clinically due to the inclusion criteria in their research, which may lead to bias. Molecular indicators such as integrated mRNA-lncRNA signature (21), Linc-PINT (22) and progranulin (23) were also reported in several prognostic nomograms. In our study however, only routinely available variables such as sex, age, histology, chemotherapy, T stage and LNH were used to formulate the nomogram, making the nomogram convenient in clinical practice. Furthermore, the nomogram was externally validated by another large cohort, and the results were generally applicable. Large case numbers with long-term follow-up and the use of both internal and external validation make our nomogram reliable.
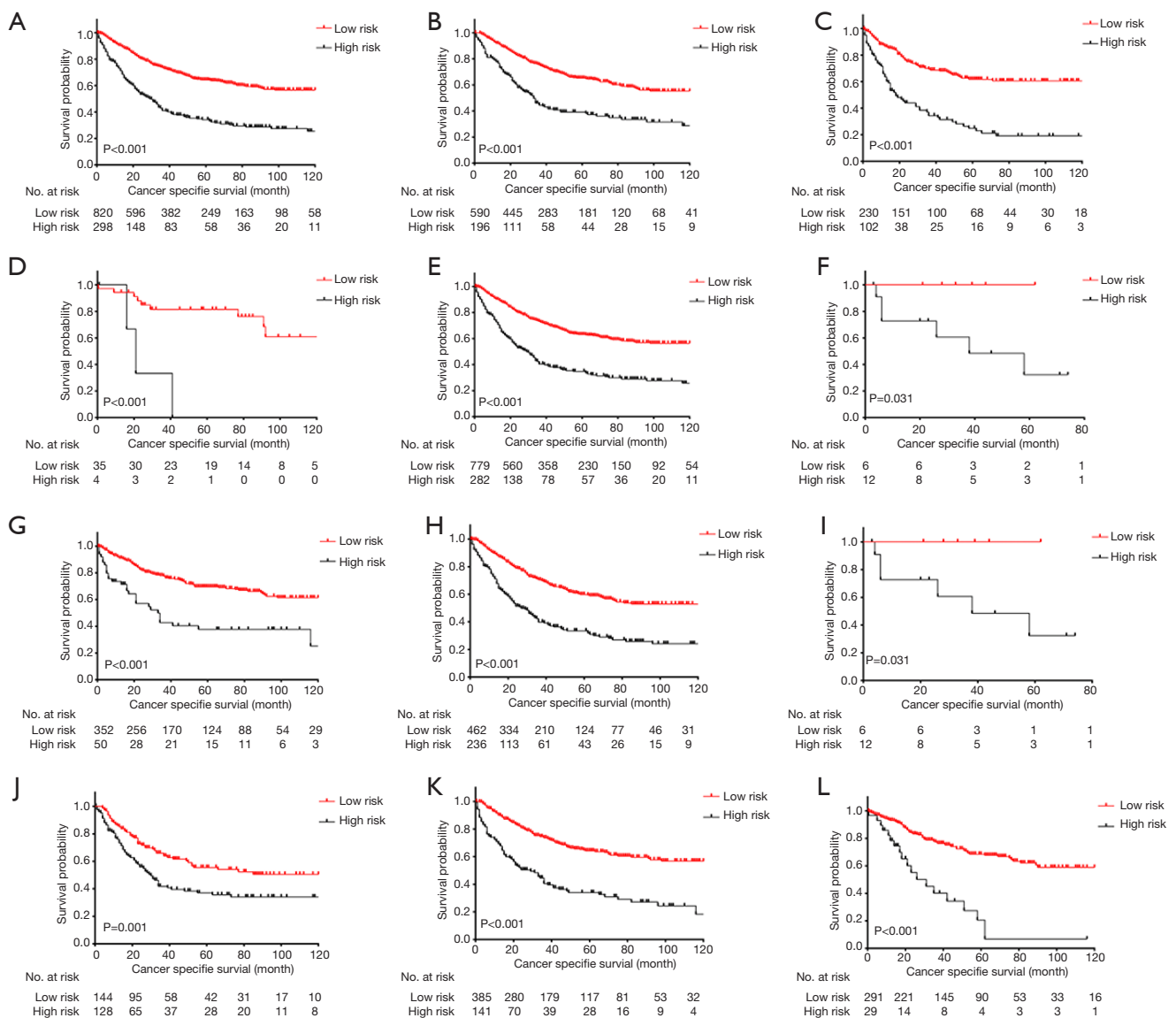
Figure 4 Kaplan-Meier estimates of CSS in resected stage pT2-4aN0M0 EC patients. The entire cohort (A), ADC group (B), SCC group (C), Stage I group (D), Stage II group (E), Stage III group (F), T2 stage group (G), T3 stage group (H), T4a stage group (I), LNH ≤6 group (J), 6< LNH ≤17 group (K) and LNH >17 group (L). CSS, cancer specific survival; EC, esophageal carcinoma; ADC, adenocarcinoma; SCC, squamous cell carcinoma; LNH, lymph node harvested.

Our study also has several limitations. First, the factors used to formulate the nomograms represented only a portion of clinicopathological characteristics owing to the SEER database, which had limited relevant information. Second, the quality of several data in the SEER database is relatively poor, for example, the median number of lymph node harvested in the SEER database is only 12, which is less than the number (at least 15 examined lymph nodes) recommended by the NCCN guideline (24), and it might lead to stage migration. Third, although the developed prognostic score models showed good discrimination capacity, further large-scale validation cohorts are warranted. Finally, the retrospective nature of the study may have contributed to selection bias. Further efforts on prospective study results collection and broader prognostic factors recruitment such as the timing and regimen of

2376

Dou et al. Prognosis of resected T2-4N0M0 esophageal carcinoma

chemotherapy, timing and dosage of radiotherapy, resection margin, tumor markers, and genetic molecule indicators are encouraged to improve this nomogram.

In conclusion, the novel nomograms demonstrated good prognostic prediction for patients with resected pT2-4aN0M0 EC and stratified the population subset into low-risk and high-risk prognostic subgroups, which may aid individual clinical practice.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at http://dx.doi.org/10.21037/jtd-20-3393

Peer Review File: Available at http://dx.doi.org/10.21037/jtd-20-3393

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jtd-20-3393). 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).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020;70:7-30.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.

3. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. CA Cancer J Clin 2015;65:87-108.

4. Uzunoglu FG, Reeh M, Kutup A, et al. Surgery of esophageal cancer. Langenbecks Arch Surg 2013;398:189-93.

5. Milleron B, Westeel V, Gounant V, et al. Pathological complete response: A predictive survival factor after neoadjuvant chemotherapy in lung cancer. Bull Cancer 2016;103:66-72.

6. Rice TW, Gress DM, Patil DT, et al. Cancer of the esophagus and esophagogastric junction-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin 2017;67:304-17.

7. Zheng Y, Fu S, He T, et al. Predicting prognosis in resected esophageal squamous cell carcinoma using a clinical nomogram and recursive partitioning analysis. Eur J Surg Oncol 2018;44:1199-204.

8. Su D, Zhou X, Chen Q, et al. Prognostic Nomogram for Thoracic Esophageal Squamous Cell Carcinoma after Radical Esophagectomy. Plos One 2015;10:e0124437.

9. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 2004;10:7252-9.

10. Harrell FE, Jr., Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982;247:2543-6.

11. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006;26:565-74.

12. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391:1023-75.

13. Zhu C, You Y, Liu S, et al. A Nomogram to Predict Distant Metastasis for Patients with Esophageal Cancer. Oncol Res Treat 2020;43:2-9.

14. Min BH, Yang JW, Min YW, et al. Nomogram for prediction of lymph node metastasis in patients with superficial esophageal squamous cell carcinoma. J Gastroenterol Hepatol 2020;35:1009-15.

15. Du F, Sun Z, Jia J, et al. Development and Validation of an Individualized Nomogram for Predicting Survival in Patients with Esophageal Carcinoma after Resection. J

Cancer 2020;11:4023-9.

16. Wu XX, Chen RP, Chen RC, et al. Nomogram predicting cancer-specific mortality in patients with esophageal adenocarcinoma: a competing risk analysis. J Thorac Dis 2019;11:2990-3003.

17. Tang X, Zhou X, Li Y, et al. A Novel Nomogram and Risk Classification System Predicting the Cancer-Specific Survival of Patients with Initially Diagnosed Metastatic Esophageal Cancer: A SEER-Based Study. Ann Surg Oncol 2019;26:321-8.

18. Deng W, Zhang W, Yang J, et al. Nomogram to Predict Overall Survival for Thoracic Esophageal Squamous Cell Carcinoma Patients After Radical Esophagectomy. Ann Surg Oncol 2019;26:2890-8.

19. Zhang WY, Chen XX, Chen WH, et al. Nomograms for predicting risk of locoregional recurrence and distant metastases for esophageal cancer patients after radical esophagectomy. BMC Cancer 2018;18:879.

20. Xie K, Liu S, Liu J. Nomogram predicts survival benefit

for non- metastatic esophageal cancer patients who underwent preoperative radiotherapy. Cancer Manag Res 2018;10:3657-68.

21. Lan T, Xiao Z, Luo H, et al. Bioinformatics analysis of esophageal cancer unveils an integrated mRNA-lncRNA signature for predicting prognosis. Oncol Lett 2020;19:1434-42.

22. Zhang L, Chen J, Wang L, et al. Linc-PINT acted as a tumor suppressor by sponging miR-543 and miR-576-5p in esophageal cancer. J Cell Biochem 2019;120:19345-57.

23. Li G, Dong T, Yang D, et al. Progranulin promotes lymphangiogenesis through VEGF-C and is an independent risk factor in human esophageal cancers. Hum Pathol 2018;75:116-24.

24. National Comprehensive Cancer Network. Esophageal and Esophagogastric Junction Cancers. Version 1.2019. Available online: https://www.nccn.org/professionals/physician_gls/pdf/esophageal.pdf.

J Thorac Dis 2021;13(4):2363-2377 | http://dx.doi.org/10.21037/jtd-20-3393
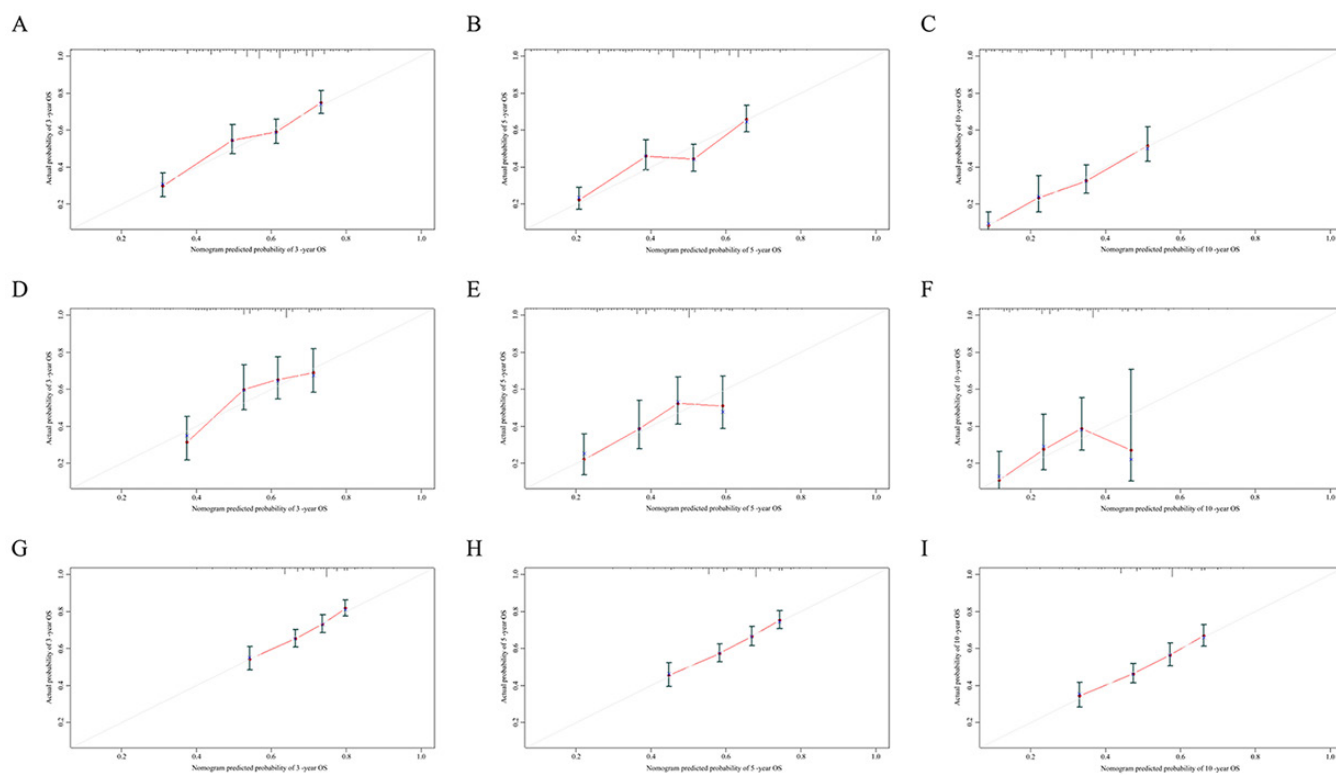
Figure S1 The calibration curves for predicting OS in the SEER training cohort (A, B, C), the SEER validation cohort (D, E, F) and the SYSUCC validation cohort (G, H, I). OS, overall survival; SEER, Surveillance, Epidemiology, and End Results; SYSUCC, Sun Yat-sen University Cancer Center.
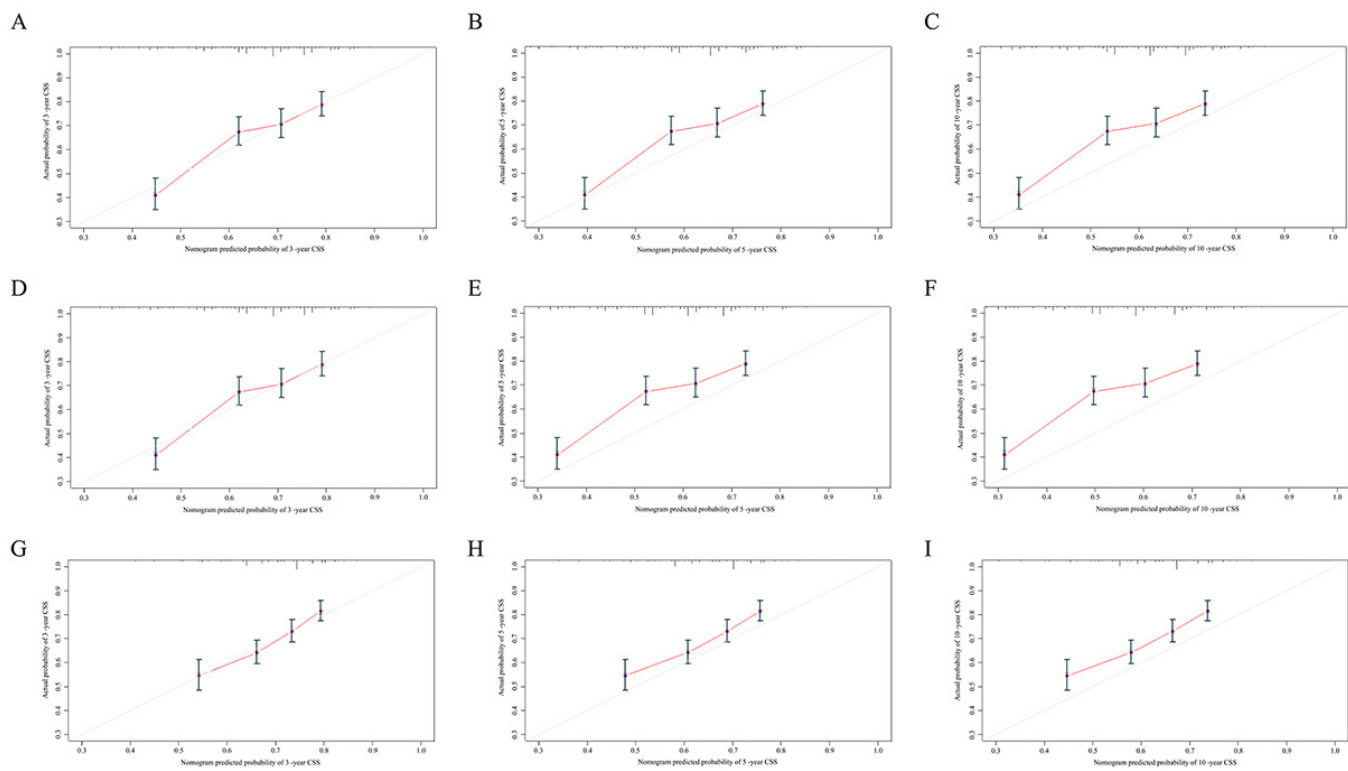
Figure S2 The calibration curves for predicting CSS in the SEER training cohort (A, B, C), the SEER validation cohort (D, E, F) and the SYSUCC validation cohort (G, H, I). CSS, cancer specific survival; SEER, Surveillance, Epidemiology, and End Results; SYSUCC, Sun Yat-sen University Cancer Center.
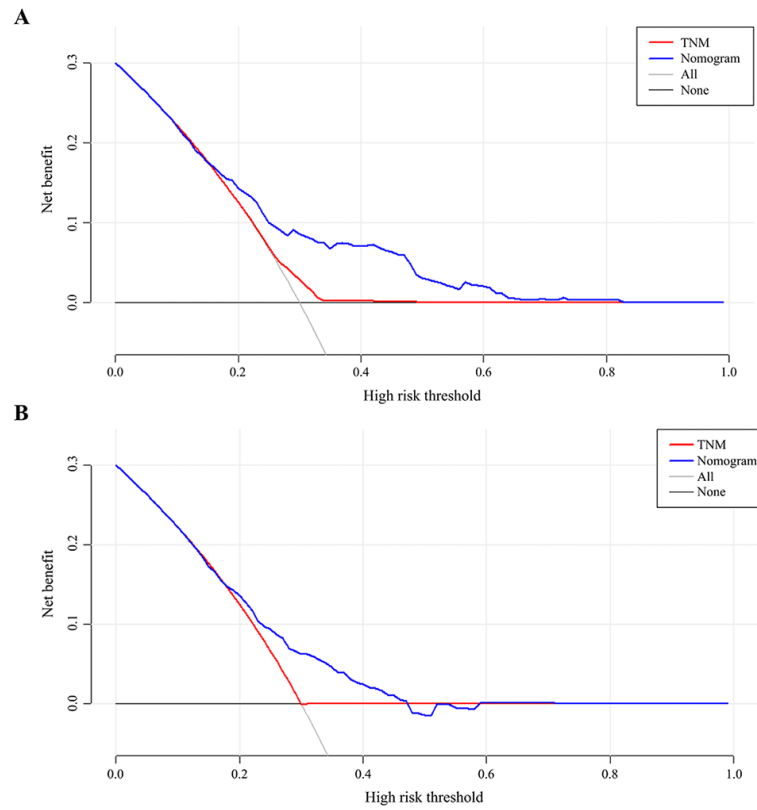
Figure S3 The decision curve analysis of the nomograms and the 8th TNM stage. OS (A) and CSS (B). OS, overall survival; CSS, cancer specific survival.