Construction and verification of prognostic nomogram for early-onset esophageal cancer

Xiaoxiao Liu, Wei Guo, Xiaobo Shi, Yue Ke, Yuxing Li, Shupei Pan, Yingying Jin, Yuchen Wang, Qinli Ruan, Hongbing Ma*

ABSTRACT

This study aimed to build up nomogram models to evaluate overall survival (OS) and cancer-specific survival (CSS) in early-onset esophageal cancer (EOEC). Patients diagnosed with esophageal cancer (EC) from 2004 to 2015 were extracted from the Surveillance Epidemiology and End Results (SEER) database. Clinicopathological characteristics of younger versus older patients were compared, and survival analysis was performed in both groups. Independent related factors influencing the prognosis of EOEC were identified by univariate and multivariate Cox analysis, which were incorporated to construct a nomogram. The predictive capability of the nomogram was estimated by the concordance index (C-index), calibration plot, receiver operating characteristic (ROC) curve, and decision curve analysis (DCA). A total of 534 younger and 17,243 older patients were available from the SEER database. Younger patients were randomly segmented into a training set (n = 266) and a validation set (n = 268). In terms of the training set, the C-index of the OS nomogram was 0.740 (95% CI: 0.707-0.773), and that of the CSS nomogram was 0.752 (95% CI: 0.719-0.785). In view of the validation set, the C-index of OS and CSS were 0.706 (95% CI: 0.671-0.741) and 0.723 (95% CI: 0.690-0.756), respectively. Calibration curves demonstrated the consistent degree of fit between actual and predicted values in nomogram models. From the perspective of DCA, the nomogram models were more beneficial than the tumor-node-metastasis (TNM) and the SEER stage for EOEC. In brief, the nomogram models can be considered as an individualized quantitative tool to predict the prognosis of EOEC patients to assist clinicians in making treatment decisions.

KEYWORDS: Early-onset esophageal cancer (EOEC); nomogram; overall survival (OS); cancer-specific survival (CSS); surveillance, epidemiology, and end results (SEER)

INTRODUCTION

Esophageal cancer (EC) is one of the most aggressive gastrointestinal tumors [1]. EC is the seventh leading cancer type for males in the United States. According to cancer statistics in 2020 manifested that there were approximately 18,440 new cases of EC and 16,170 death from EC, while in 2019 there were about 17,650 new cases of EC [2,3]. It is evaluated that the global incidence and mortality rate of the EC will increase in the future years, particularly in Asia [4]. What's more, the morbidity of esophageal adenocarcinoma has increased significantly in all ages, especially among young people [5]. Besides, some studies have reported that there was no distinction in the survival rate between young and elderly patients with EC, but young patients with EC have more advanced tumors and more malignant potential and invasiveness than older patients [6]. A relevant study had manifested that compared with elderly patients, while patients with early-onset esophageal adenocarcinoma (younger than 50 years) were a high proportion of advanced stages; they had a superior survival rate [7]. Furthermore, due to being often diagnosed at an advanced stage at the time of consultation, esophageal cancer has a poor prognosis [8]. It is still controversial whether younger patients with EC have a better or worse prognosis than older patients, which draws our attention to which clinicopathological factors affect the prognosis of young patients with esophageal cancer [9-15].

According to the age group determined by the World Health Organization (WHO) in 2017, the upper age limit for young people has been raised to 44 years old. For the time being, there is no acknowledged explicit definition of early-onset EC. Therefore, in this study, early-onset esophageal cancer (EOEC) referred to patients with age ≤44 years old.

The TNM staging system is considered to be currently the most extensively used system for prognostic evaluation and clinical treatment of cancer patients. It contains tumor invasive depth, regional lymph node involvement, and distant metastasis, but does not contain demographic information

Department of Radiation Oncology, Xi'an Jiaotong University Second Affiliated Hospital, Xi'an, China

*Corresponding author: Hongbing Ma, Department of Radiation Oncology, Xi'an Jiaotong University Second Affiliated Hospital, No.157, Xi Wu Road, Xi'an 710004, Shaanxi, China. E-mail: mhbxian@126.com

about patients, such as age, gender, race, food habit, and marital status, which are also concerned with the prognosis of cancer patients, resulting in that the TNM staging system cannot thoroughly predict the prognosis of EC patients [16]. Consequently, the main purpose of this study was to develop more plentiful and accurate prognostic models to guide survival. Nomogram can evaluate and analyze the risk factors of prognosis visually and individually [16-18]. However, the nomogram of the prognosis of EOEC has not been completely determined.

In this research, we made use of data from the SEER (Surveillance, Epidemiology, and End Results) database to first probe into the differences in clinicopathological characteristics influencing the prognosis of younger and older patients. We then focused on analyzing the clinicopathological features of EOEC and thoroughly investigating the variables associated with prognosis. Ultimately, we further constructed nomogram models to preferably forecast the prognosis of EOEC.

MATERIALS AND METHODS

Patients

The National Cancer Institute's SEER database (http://seer.cancer.gov/seerstat/) covers 28% of the population of the United States. This study used SEER database data from 1975 to 2016 ("Incidence - SEER 18 Regs Custom Data (with additional treatment fields), Nov 2018 Sub (1975-2016 varying)"). Data were extracted, downloaded, and analyzed using SEER*Stat Software (Version 8.3.6).

Patients diagnosed with EC from 2004 to 2015 were retrieved. Inclusion criteria included the following: (I) EC patients proved by pathology; (II) only one primary tumor; and (III) complete clinicopathological data and follow-up information. Exclusion criteria were as follows: (I) The patient's disease-related information is missing, such as unknown age, sex, race, grade, primary site, TNM stage, SEER stage, and other information; and (II) unknown surgery, radiotherapy, and chemotherapy.

Patients were staged using the American Joint Committee on Cancer (AJCC) version 6th. In this study, the upper esophagus was set up with primary site codes C15.0 (cervical esophagus) and C15.3 (upper third of esophagus). Code C15.1 (thoracic esophagus) and C15.4 (middle third of esophagus) were considered to be the middle esophagus. The lower esophagus was defined with codes C15.2 (abdominal esophagus) and C15.5 (lower third of esophagus) [19].

In the end, a total of 17777 patients with EC met the inclusion and exclusion criteria, which included 17,243 elderly patients older than 45 years and 534 patients with EOEC. For further analysis, the selected patients with EOEC were divided into the training set (n = 266) and a validation set (n = 268) at random (Figure 1).

Variables

Clinical variables captured from the SEER database encompassed age, sex, race, grade, histology, primary site, TNM 6th stage, SEER stage, surgery, radiotherapy, and chemotherapy. Follow-up variables involved survival status, survival time, and cause of death. The principal terminal point was overall survival (OS), which referred to the period from diagnosis to death for any reason. Nevertheless, cancer-specific survival (CSS) was the subordinate terminal point, which was regarded as the period from diagnosis to death owing to the EC. All patients were followed up for at least 3 months.

Construction and verification of nomogram

Significant variables screened by univariate analysis were further screened out independent prognostic factors by multivariate analysis and then the nomogram was constructed. The concordance index (C-index), receiver operating characteristics (ROC) curves, and calibration curves were used to evaluate the performance and accuracy of nomogram [20]. Moreover, the nomogram was compared with the TNM stage and SEER stage utilizing decision curve analysis (DCA) that was a novel way for evaluating predictive models [21,22].

Ethical statement

Data in the present study could be freely obtained from the SEER database, which was utilized and analyzed by the public (http://seer.cancer.gov/seerstat/). Therefore, the study exempted the institutional review board approval.

Statistical analysis

Chi-square test was used for comparison of categorical variables, and the comparison of ordinal variables was based on the Wilcoxon rank-sum test. The Kaplan–Meier method was utilized for demonstrating the effect of clinicopathological variables on the survival rate of patients. The inspection level was defined as $p < 0.05$, which represented that the difference was considered dramatically significant. All statistical analyses and drawings were implemented utilizing R software version 3.6.2 (http://www.R-project.org).

RESULTS

Comparison of clinical characteristics between younger and older of EC

As indicated in Table 1, we enrolled a total of 534 younger and 17,243 older patients with EC after screening. Older patients accounted for the majority of the entire cohort. Nevertheless, the number of younger patients represented only 3.1% of the older patients. Statistical analysis revealed significant differences in,
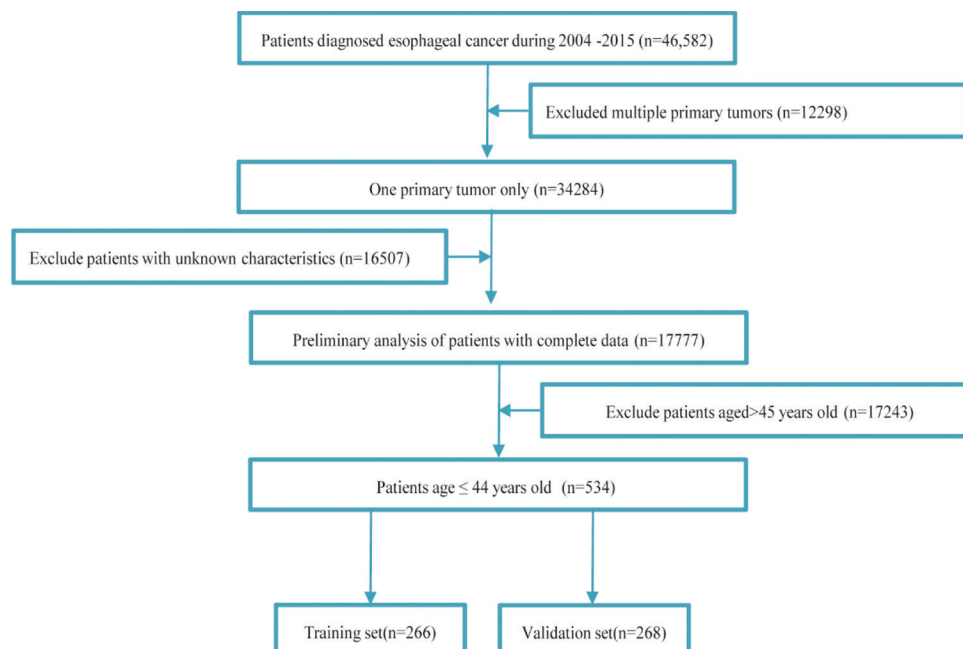
FIGURE 1. Flow diagram of EOEC patient screening.

race ($p = 0.043$), histology ($p < 0.001$), primary site ($p < 0.001$), T ($p < 0.001$), N ($p < 0.001$), M ($p < 0.001$), SEER stage ($p < 0.001$), surgery ($p < 0.001$), and chemotherapy ($p < 0.001$) between patients less than or equal to 44 years of age and older than 45 years. Regardless of age, male patients were more likely than female patients. However, there were more white people in the younger group (81.46%), while the older group was more black people (84.99%). The depth of tumor immersion was mainly T3, which accounted for 41.39% and 40.06%, respectively, in the younger and older groups. Moreover, it was shown that patients with EC, young or not, were more inclined to receive adjuvant radiotherapy and chemotherapy. However, radiotherapy was not statistically significant in both age groups of patients ($p = 0.859$), which might be due to the lack of detailed radiotherapy information in the SEER database. Survival analysis illustrated that the medium survival time of young and old patients was 15.00 (6.00-33.00) and 12.00 (5.00-27.00) months, respectively ($p < 0.001$). The survival curves suggested that the OS and CSS of younger patients were higher than those of older patients, and the differences were statistically significant ($p < 0.001$, $p = 0.043$) (Figure 2). Although our study found that the younger group had a better prognosis than the older group, the study of EC patients aged ≤44 years has aroused widespread concern with the speedy ascent in the incidence. Therefore, we next investigated the prognostic evaluation models in the younger patients.

Confirmation of the cut-off age

X-tile software analysis result displayed that the optimal cut-off age in EOEC patients was 37 years; therefore, the patients were fell into two groups (≤36, 37-44, Figure 3) [23].

Characteristics of patients with EOEC

Overall, 534 EOEC patients qualified were involved in this study, of which 266 patients were randomly allocated to the training set and 268 patients into the validation set (Figure 1). Notably, there was no significant difference between the two sets. Amid all EOGC patients, 425 (79.59%) were between 37 and 44 years of age, while 109 (20.41%) were under 37 years of age. Similar to previous reports, males (85.21%) accounted for significantly more than females (14.79%). The race was predominantly white (81.46%). Grades were dominated by Grade II (43.26%) and Grade III (48.69%). The most common histology type was adenocarcinoma (78.09%). Moreover, the most frequent tumor location of EOGC was the lower esophagus (80.71%), afterward the middle esophagus (12.17%) and the upper esophagus (4.12%). For all the patients, T3 and N0 account for 41.39% and 32.96%, respectively, while 61.80% is M0 in the AJCC stage system. Concerning the SEER stage, there were 239 patients (44.76%) of regional and 236 patients of distant that accounted for 44.19% (Tables 1 and S1). Our analysis resulted in medium survival time of 16.00 (7.00-40.00) and 14.00 (6.00-31.00) months for patients ≤36 years and 37-44 years, respectively (p = 0.304). Kaplan–Meier survival curves demonstrated that patients aged 37-44 years had lower OS and CSS than patients aged ≤36 years. However, the differences were not statistically significant ($p = 0.082$, $p = 0.11$) (Figure 4).

Exploration of prognostic factors associated with OS and CSS

Univariate analysis demonstrated that age, race, grade, histology, primary site, TNM stage, SEER stage, surgery, and

TABLE 1. Comparison of demographic and clinicopathological characteristics of EC in younger and older patients

Variables	Age ≤44 years old (n=534)		Age >45 years old (n=17243)		*p* value
	n	%	n	%	
Sex					0.005
Female	79	14.79	3404	19.74	
Male	455	85.21	13839	80.26	
Race					0.043
White	435	81.46	14655	9.84	
Black	60	11.24	1696	84.99	
Other	39	7.30	892	5.17	
Grade					0.694
Grade I	35	6.55	1034	6.00	
Grade II	231	43.26	7125	41.32	
Grade III	260	48.69	8780	50.92	
Grade IV	8	1.50	304	1.76	
Histology					<0.001
Squamous cell carcinoma	104	19.48	5311	30.80	
Adenocarcinoma	417	78.09	11488	66.62	
Other	13	2.43	444	2.58	
Primary site					<0.001
Upper	22	4.12	1127	6.54	
Middle	65	12.17	3374	19.56	
Lower	431	80.71	11975	69.45	
Overlapping area	16	3.00	767	4.45	
T stage					<0.001
T1	128	23.97	5529	32.07	
T2	61	11.42	2089	12.11	
T3	221	41.39	6908	40.06	
T4	124	23.22	2717	15.76	
N stage					<0.001
N0	176	32.96	7881	45.71	
N1	358	67.04	9362	54.29	
M stage					<0.001
M0	330	61.80	12123	70.31	
M1	204	38.20	5120	29.69	
SEER stage					<0.001
Localized	59	11.05	3976	34.60	
Regional	239	44.76	7300	42.34	
Distant	236	44.19	5967	23.06	
Surgery					<0.001
No	298	55.81	11217	65.05	
Yes	236	44.19	6026	34.95	
Radiotherapy					0.859
No	189	35.39	6184	35.86	
Yes	345	64.61	11059	64.14	
Chemotherapy					<0.001
No	106	19.85	5550	32.19	
Yes	428	80.15	11693	67.81	

EOEC: Early-onset esophageal cancer; SEER: Surveillance, epidemiology and end results; Grade I: Well-differentiated; Grade II: Moderately differentiated; Grade III: Poorly differentiated; Grade IV: Undifferentiated; TNM: Tumor-node-metastasis

radiotherapy were significantly correlated with OS (Table 2). Whereas, among the multivariate analysis, age, grade, histology, primary site, SEER stage, and surgery were regarded as independent hazard factors for OS. The result of CSS univariate analysis manifested that age, race, grade, histology, primary site, TNM stage, SEER stage, surgery, and radiotherapy were prognostic risk factors for EOEC patients in the training set (Table 3). The results of our analysis also manifested that although the *p*-value of the primary site was 0.05, as we know, it also affected the prognosis of patients, so we also included this factor of the primary site in the multivariate analysis. In multivariate analysis, age, race, grade, histology, primary site, SEER stage, and surgery were significantly correlated with CSS.

Construction of the nomogram

The nomogram of OS was constructed by incorporating prognostic risk factors that contained age, grade, histology, primary site, SEER stage, and surgery given by performing the multivariate analysis (Figure 5A). Simultaneously, variables were significantly correlated with CSS, which was further included in the CSS nomogram (Figure 5B). The scores corresponding to independent risk factors were summed to calculate the aggregate score. According to the total score value, the predicted probabilities of OS and CSS in 3 and 5 years could be obtained.

Verification of the nomogram

The OS and CSS nomogram models were verified internally through the training set and externally by the validation set, respectively. According to internal verification, the C-index of OS nomogram was 0.740 (95% CI: 0.707-0.773), and that of CSS nomogram was 0.752 (95% CI: 0.719-0.785). In terms of the external verification, the C-index of OS and CSS were 0.706 (95% CI: 0.671-0.741) and 0.723 (95% CI: 0.690-0.756), respectively. The area under the ROC curves (area under curve, AUC) that were applied to evaluate the discernment of the nomogram models were both high in the training set and verification set (Figure 6). Additionally, the calibration curves were relatively close to the ideal curve, which indicated the probabilities of 3-year and 5-year OS and CSS forecasted by nomogram models were in accordance with the actual survival of patients (Figure 7).

What is more, we also performed a contrast of the nomogram with the AJCC TNM stage and SEER stage. First of all, the C-index of OS nomogram in training set was 0.740, which was obviously superior to the TNM 6th stage (0.692, 95% CI: 0.653-0.731; *p* = 0.006) and SEER stage (0.667, 95% CI: 0.631-0.702; *P* < 0.001). Likewise, the C-index of the CSS nomogram in the training set was 0.752, which was significantly better than the TNM 6th stage (0.689, 95%CI: 0.652-0.726; *p* < 0.001) and SEER stage (0.667, 95% CI: 0.632-0.702; *p* < 0.001). Compared with
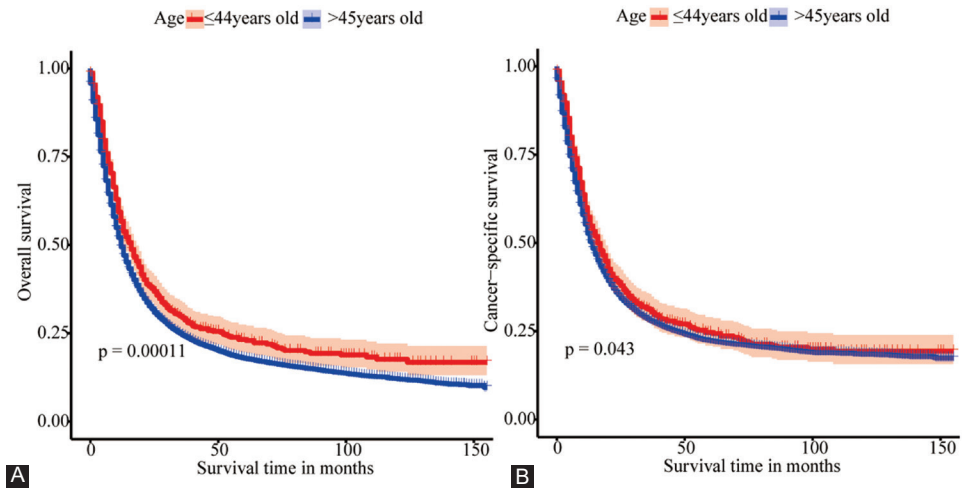
FIGURE 2. Kaplan-Meier survival curves of different age groups of esophageal cancer. (A) The survival curve for overall survival (OS) was compared between patients ≤44 and >45 years of age. (B) The survival curve for cancer-specific survival (CSS) was compared between patients ≤ 44 and >45 years of age.
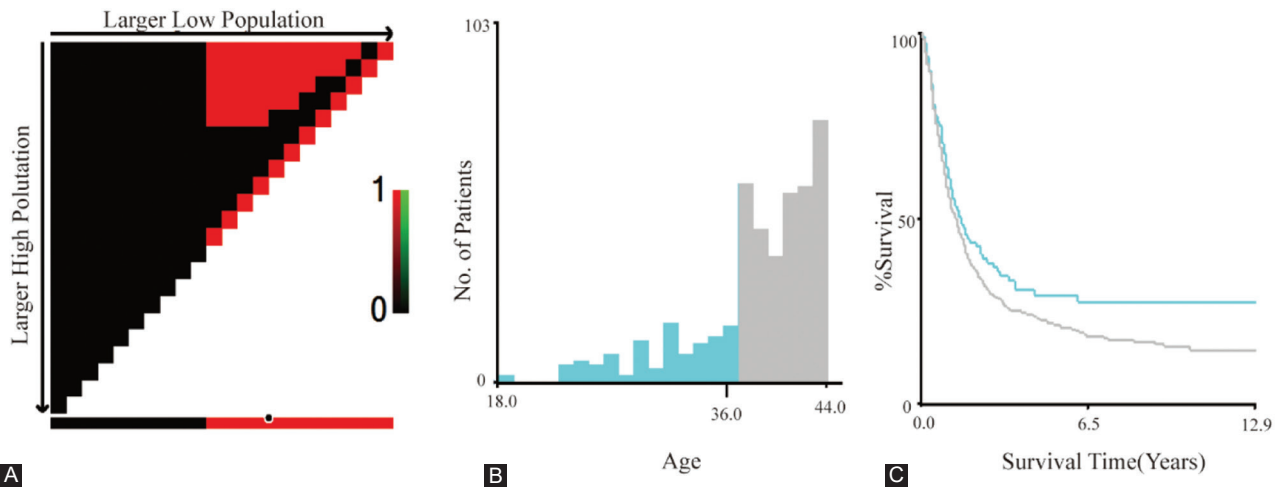
FIGURE 3. The X-tile analysis was implemented on the survival data of EOEC patients to ascertain the optimal threshold for age. The best cutting point was highlighted with a black circle in the left figure 3A. Histogram of the entire group in middle Figure 3B, and Kaplan-Meier plot (C). The figure manifested that the optimal age cut-off point for EOEC patients was 37 years (≤ 36, 37 – 44, $p < 0.001$).
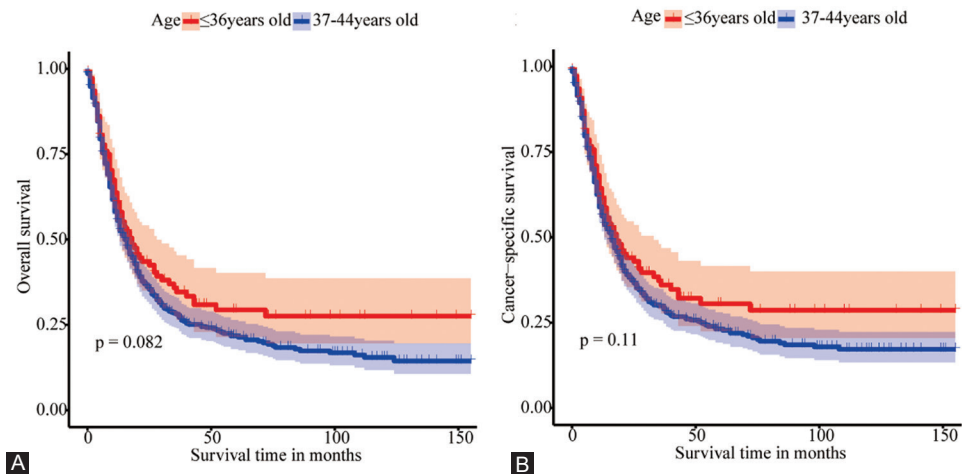
FIGURE 4. Kaplan-Meier survival curves of different age groups of EOEC. (A) The survival curve for OS was compared between patients≤36 and 37-44 years of age. (B) The survival curve for CSS was compared between patients≤36 and 37-44 years of age.

TABLE 2. Univariate and multivariate analysis of variables related to OS in the training set (n=266)

Variables	Univariable analysis			Multivariable analysis		
	HR	95%C	*p*-value	HR	95%CI	*p*-value
Age (years)			0.010			
≤36	Ref			Ref		
37-44	1.58	1.09-2.30	0.016	1.95	1.30-2.92	0.001
Sex			0.300			
Female						
Male	0.83	0.57-1.20	0.309			
Race			0.030			
White	Ref			Ref		
Black	1.67	1.10-2.52	0.016	1.43	0.85-2.40	0.182
Other	1.60	0.89-2.88	0.120	1.79	0.95-3.35	0.070
Grade			<0.001			
Grade I	Ref			Ref		
Grade II	2.87	1.17-7.09	0.022	2.28	0.89-5.83	0.087
Grade III	4.74	1.93-11.63	<0.001	2.71	1.07-6.89	0.036
Grade IV	10.47	2.47-44.41	0.001	9.87	2.22-43.81	0.003
Histology			0.020			
Squamous cell carcinoma	Ref			Ref		
Adenocarcinoma	0.96	0.68-1.35	0.810	1.53	0.92-2.52	0.101
Other	4.27	1.77-10.29	0.001	5.27	1.99-13.97	<0.001
Primary site			0.030			
Upper	Ref			Ref		
Middle	1.98	0.87-4.52	0.104	2.24	0.94-5.35	0.069
Lower	1.55	0.73-3.31	0.257	1.66	0.69-3.96	0.256
Overlapping area	5.96	1.87-19.05	0.003	7.69	2.29-25.87	<0.001
T stage			<0.001			
T1	Ref			Ref		
T2	0.52	0.30-0.92	0.025	0.96	0.52-1.78	0.895
T3	1.02	0.70-1.50	0.904	1.09	0.71-1.68	0.689
T4	1.78	1.20-2.65	0.004	1.12	0.73-1.72	0.603
N stage			0.002			
N0	Ref			Ref		
N1	1.62	1.19-2.21	0.002	0.96	0.68-1.35	0.809
M stage			<0.001			
M0	Ref			Ref		
M1	2.80	2.11-3.72	<0.001	1.53	0.87-2.71	0.142
SEER stage			<0.001			
Localized	Ref			Ref		
Regional	3.65	1.76-7.58	<0.001	2.99	1.27-7.02	0.012
Distant	8.26	4.01-17.02	<0.001	3.07	1.11-8.49	0.031
Surgery			<0.001			
No	Ref			Ref		
Yes	0.29	0.21-0.39	<0.001	0.38	0.26-0.55	<0.001
Radiotherapy			0.007			
No	Ref			Ref		
Yes	0.67	0.50-0.89	0.006	0.73	0.52-1.03	0.074
Chemotherapy			0.700			
No	Ref					
Yes	0.93	0.66-1.33	0.695			

SEER: Surveillance, epidemiology and end results; OS: Overall survival; HR: Hazard ratio; CI: Confidence interval; Grade I: Well-differentiated; Grade II: Moderately differentiated; Grade III: Poorly differentiated; Grade IV: Undifferentiated; TNM: Tumor-node-metastasis

TABLE 3. Univariate and multivariate analysis of variables related to CSS in the training set (n=266)

Variables	Univariable analysis			Multivariable analysis		
	HR	95%CI	*p*-value	HR	95%CI	*p*-value
Age (years)			0.020			
≤36	Ref			Ref		
37-44	1.53	1.05-2.23	0.026	1.91	1.27-2.87	0.002
Sex			0.400			
Female						
Male	0.84	0.57-1.22	0.351			
Race			0.040			
White	Ref			Ref		
Black	1.63	1.06-2.48	0.025	1.48	0.87-2.51	0.149
Other	1.62	0.90-2.93	0.108	1.93	1.03-3.62	0.041
Grade			<0.001			
Grade I	Ref			Ref		
Grade II	2.69	1.09-6.64	0.032	2.01	0.78-5.16	0.147
Grade III	4.69	1.91-11.51	<0.001	2.59	1.02-6.59	0.046
Grade IV	10.23	2.41-43.38	0.002	9.08	2.04-40.41	0.004
Histology			0.020			
Squamous cell carcinoma	Ref			Ref		
Adenocarcinoma	1.02	0.71-1.45	0.923	1.627	0.97-2.74	0.068
Other	4.53	1.87-10.95	<0.001	5.23	1.96-13.97	<0.001
Primary site			0.050			
Upper	Ref			Ref		
Middle	1.78	0.77-4.10	0.175	2.03	0.84-4.89	0.115
Lower	1.53	0.72-3.27	0.270	1.57	0.65-3.79	0.314
Overlapping area	5.85	1.83-18.68	0.003	7.44	2.21-24.98	0.001
T stage			<0.001			
T1	Ref			Ref		
T2	0.54	0.30-0.95	0.033	0.97	0.52-1.81	0.934
T3	1.01	0.68-1.48	0.973	1.05	0.68-1.62	0.840
T4	1.78	1.19-2.66	0.005	1.12	0.73-1.73	0.595
N stage			0.001			
N0	Ref			Ref		
N1	1.67	1.22-2.29	0.002	1.00	0.70-1.42	0.989
M stage			<0.001			
M0	Ref			Ref		
M1	2.91	2.18-3.88	<0.001	1.66	0.93-2.96	0.087
SEER stage			<0.001			
Localized	Ref			Ref		
Regional	3.45	1.66-7.16	<0.001	2.66	1.12-6.28	0.026
Distant	8.07	3.91-16.64	<0.001	2.61	0.93-7.30	0.067
Surgery			<0.001			
No	Ref			Ref		
Yes	0.27	0.20-0.37	<0.001	0.37	0.25-0.54	<0.001
Radiotherapy			0.020			
No	Ref			Ref		
Yes	0.70	0.53-0.94	0.018	0.82	0.58-1.17	0.271
Chemotherapy			0.948			
No	Ref					
Yes	0.99	0.69-1.42	0.900			

SEER: Surveillance, epidemiology and end results; CSS: Cancer-specific survival; HR: Hazard ratio; CI: Confidence interval; Grade I: Well-differentiated; Grade II: Moderately differentiated; Grade III: Poorly differentiated; Grade IV: Undifferentiated; TNM: Tumor-node-metastasis

TNM 6th stage and SEER stage, the DCA results of nomogram models had advanced net benefits, which demonstrated that nomogram models had superior clinical prognostic worth than TNM 6th stage and SEER stage (Figure 8).
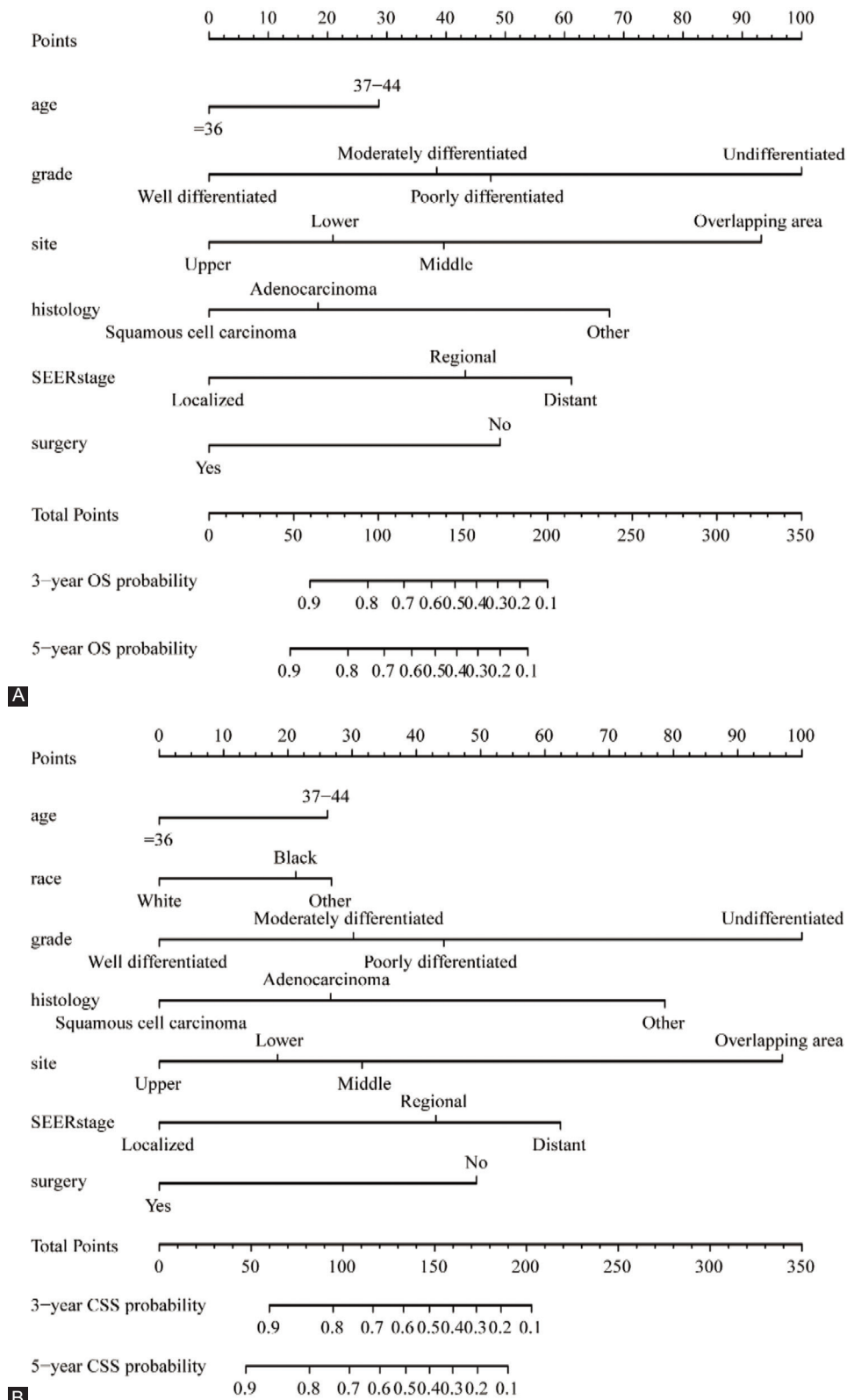
FIGURE 5. Nomogram models of OS and CSS for EOEC patients in the training set. (A) Nomogram models of 3-year and 5-year OS for patients with EOEC. (B) Nomogram models of 3- and 5-year CSS for EOEC patients.

DISCUSSION

In this research, we extracted a total of 17777 EC patients with complete information from the SEER database, of which contained 17,243 elderly patients older than 45 years and 534 patients with EOEC. Subsequently, we explored the differences in clinicopathological factors affecting the prognosis of younger and older patients and carried out survival analysis, which found that the survival rate of younger was higher than that of older patients. As incidence increased in young
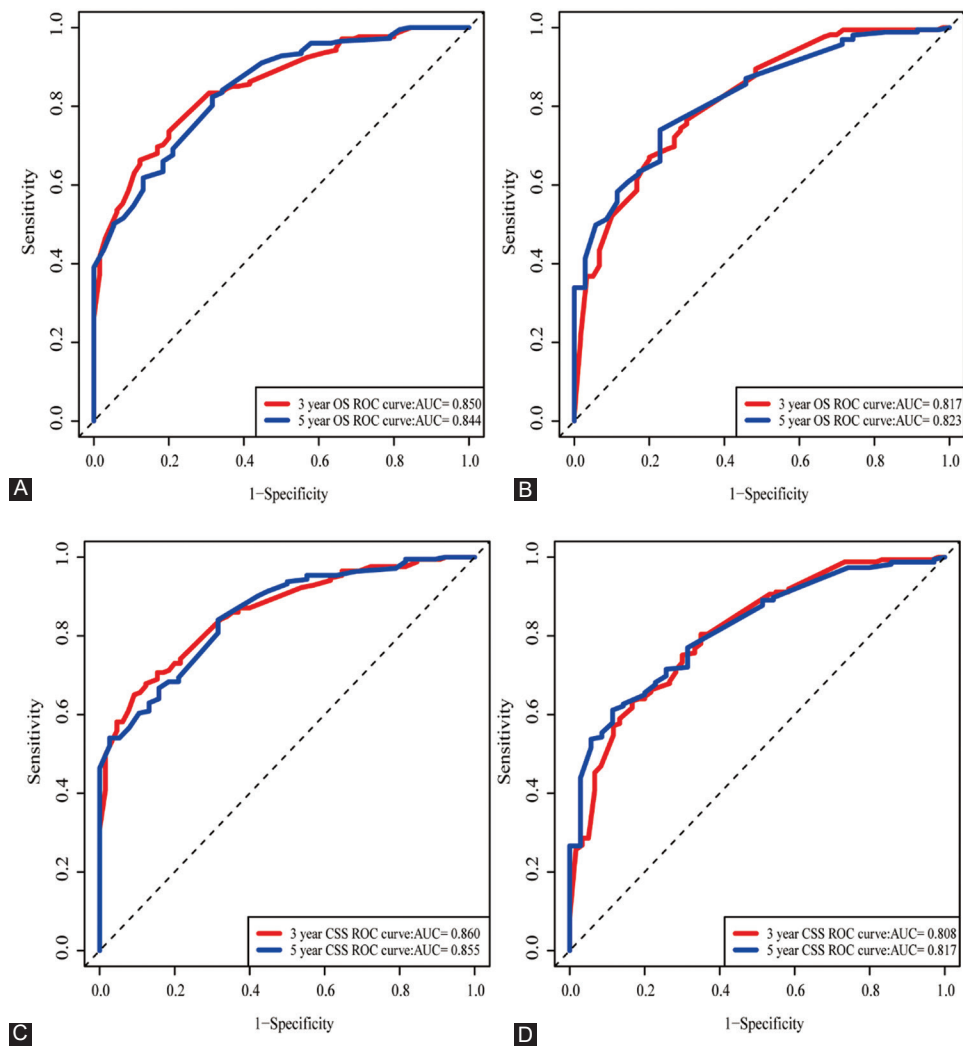
FIGURE 6. The area under the receiver operating characteristic (ROC) curve (AUC) was utilized to weigh up the performance of OS and CSS nomogram models. (A, B) ROC curves for 3- and 5-year OS in the training set and validation set; (C, D) ROC curves for 3- and 5-year CSS in the training set and validation set.

people worldwide, we consequently focused on investigating prognostic risk factors for EOEC in this study. We randomly divided 534 patients with EOEC into two groups, of which 266 patients were regarded as the training group and 268 patients were considered to be the external validation group. The independent hazard factors relevant to OS and CSS were identified by univariate and multivariate analysis. We further integrated these factors into the nomogram to forecast probabilities of OS and CSS in 3 and 5 years, which demonstrated high accuracy of this nomogram through internal and external validation. Compared with the TNM stage and SEER stage, the nomogram was more predictive, guiding the prognosis evaluation of EOEC patients in terms of visualization and individualization.

EC is one of the most mortality cancers worldwide [3]. According to the Global Burden of Disease (GBD) analysis, the topmost burden of EC (from the standpoint of disability-adjusted life-years) was in East Asia, notably China [24]. In addition, the relevant GBD study found that the age-standardized

mortality and incidence rates of EC in China were 2.1 times higher than the global average [25]. A relevant study had reported that risk factors for EC included poor dietary habits such as smoking, alcohol consumption, low fruit intake, and eating spicy foods [26]. As with previous reports, our research also found that EOEC was predominantly male and adenocarcinoma was the dominant histological type of EOEC. Relevant study reported androgen receptor regulated the growth of esophageal adenocarcinoma in a paracrine manner [27]. The incidence of adenocarcinoma of EC in developed countries in Europe and the United States had exceeded that of squamous cell carcinoma that was the predominant histological subtype of esophageal cancer in Central Asia [24,28,29]. Moreover, our findings manifested that the primary site of EOEC was mainly concentrated in the abdominal esophagus and the lower third of the esophagus, which might be adjacent to the stomach where adenocarcinoma was prone to occur. Moreover, it was worth noting that gender was not an independent risk factor for EOEC patients, which was consistent with the findings of
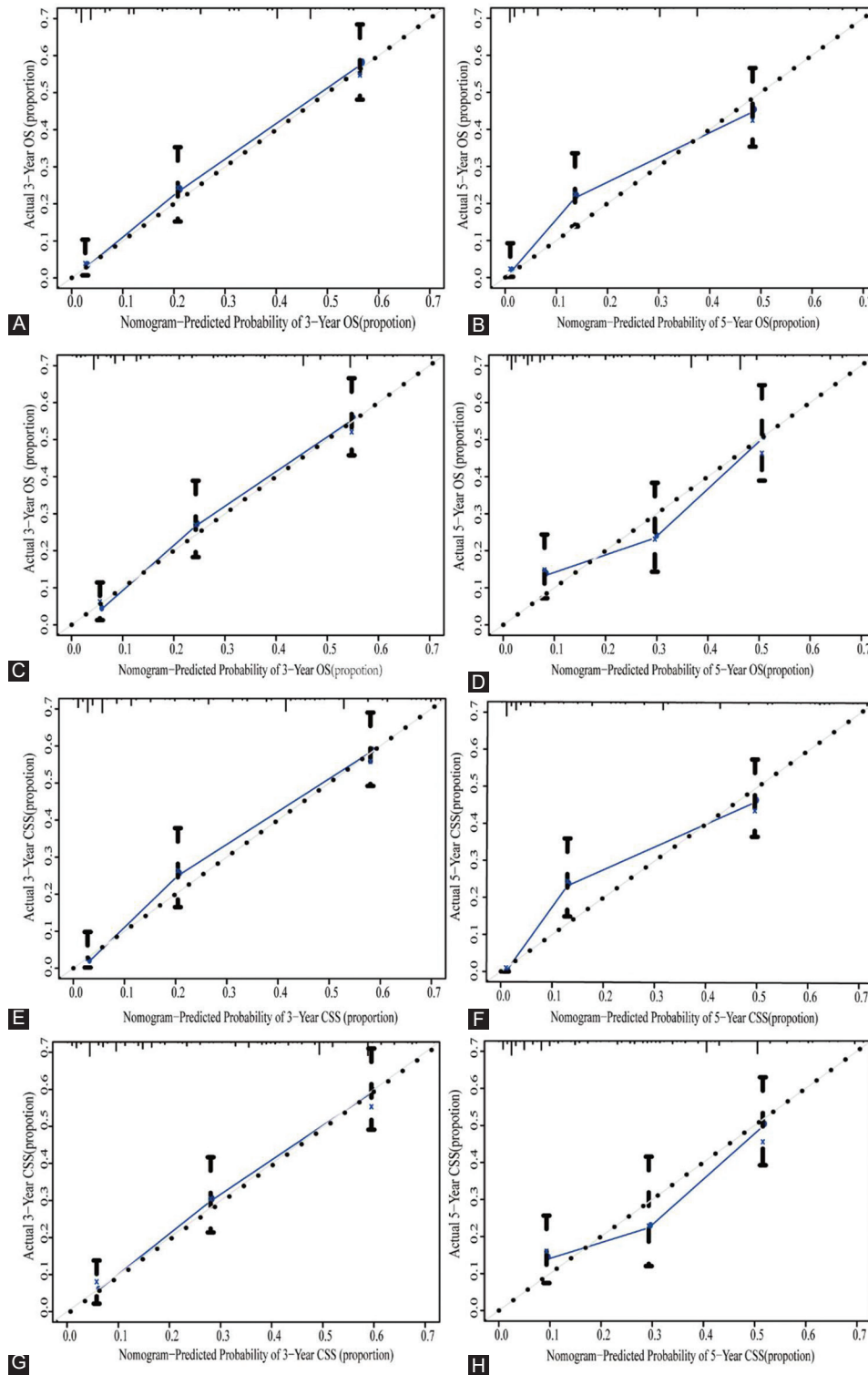
FIGURE 7. Calibration curves for nomogram models related to OS and CSS. (A, B) Calibration curves for 3- and 5-year OS in the training set; (C, D) Calibration curves for 3- and 5-year OS in validation set; (E, F) Calibration curves for 3- and 5-year CSS in the training set; (G, H) Calibration curves for 3- and 5-year CSS in the validation set.

Zeng et al. [30]. Multivariate analysis indicated that age, grade, histology, primary site, SEER stage, and surgery were significantly relevant to OS, which was considered as independent predictors for EOEC. In addition, the independent predictors of CSS for EOEC also included race and radiotherapy.

As shown in the nomogram models, the grade had the greatest impact on prognosis in this study. The more severe the

differentiation of EOEC, the worse the prognosis. As the SEER stage increased, the tumor gradually progressed, the survival time gradually decreased, and the prognosis became worse, which was the same trend obtained from the nomogram models. It had attracted our attention that nomogram demonstrated that the primary site was also a weighty variable in the prognosis of EOEC patients, especially lower esophageal that
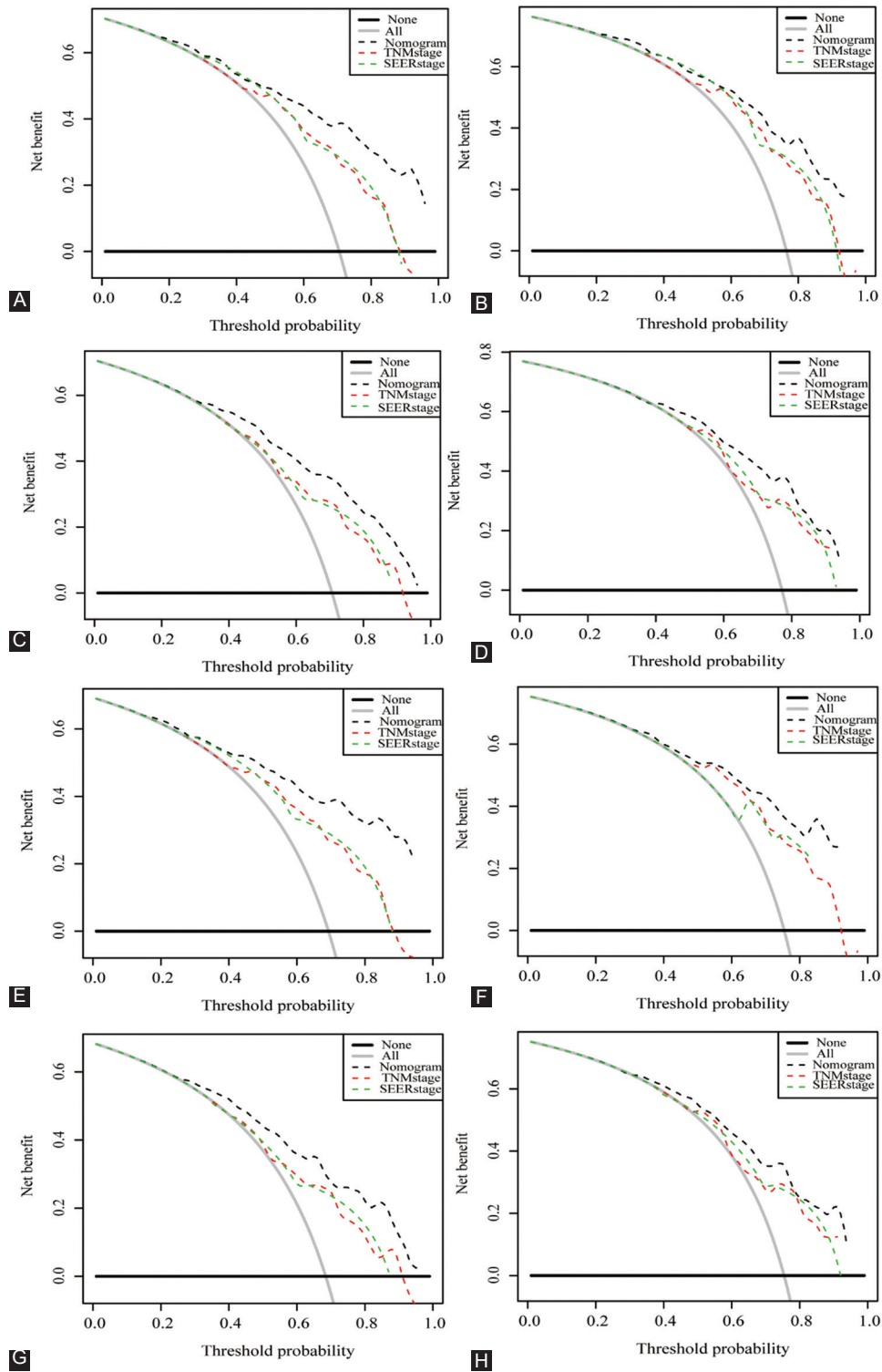
FIGURE 8. Decision curve analysis (DCA) for nomogram, TNM 6th stage, and SEER stage. (A, B) DCA for 3- and 5-year OS in the training set; (C, D) DCA for 3- and 5-year OS in validation set; (E, F) DCA for 3- and 5-year CSS in the training set; (G, H) DCA for 3- and 5-year CSS in the validation set.

was prone to Barrett's esophagus, which was closely related to gastroesophageal reflux disease (GERD) that was regarded as a precancerous lesion [31]. It is well-known that the main treatment of cancer was surgery, chemotherapy, and radiotherapy. Endoscopic resection (ER) was used to treat the early-stage of EC, while chemoradiotherapy was commonly used for locally advanced EC [32]. Our results also validated the importance of

surgery, which improved the prognosis of patients with EOEC from the nomogram. It is generally considered that different chemoradiotherapy regimens have different efficacy and may also have different outcomes. What is more, radiotherapy and chemotherapy variables were not included in our prognostic nomogram models. It might be due to the lack of detailed radiotherapy and chemotherapy protocol information in the

SEER database. In recent years, targeted therapy has become increasingly hot, and it refers to novel treatments developed by blocking immune checkpoints. Clinical trials of programmed death-1 (PD-1) and its ligand (PD-L1) inhibitors in the treatment of EC had manifested that compared with monotherapy, the combination therapy improved the survival rate, and significantly benefit patients [33,34].

The limitations of this study were several aspects. First and foremost, the SEER database provided limited information on chemoradiotherapy specific treatment regimens, targeted therapies, and genomic status, which affected prognosis. The genes with the highest frequency of mutations in early-onset esophageal adenocarcinoma were TP53 (73%) and P16 (16%), and other mutations occurred only in: APC, CDH1, CTNNB1 FGFR2, and STK11 [35]. Obesity was concerned with the early-onset of gastrointestinal cancer [36]. Furthermore, the data of this study were all from the United States, and the study would be more meaningful if data from China were further verified. What is more, our findings might be influenced by patients' willingness to treat. Most important of all, this study was a retrospective study based on the SEER database and required further verification with a prospective cohort study to have sufficient evidence to verify the findings.

Nevertheless, the C-index, AUC, and calibration curves were usually utilized to weigh up the nomogram, which proved that the nomogram had advanced accuracy. Compared with the traditional staging, DCA illuminated that the nomogram had better practicability.

CONCLUSIONS

In conclusion, the nomogram models were available for an individualized quantitative implement to predict the prognosis of EOEC patients to assist clinicians in making treatment decisions.

ACKNOWLEDGMENTS

REFERENCES

[1]   Feng RM, Zong YN, Cao SM, Xu RH. Current cancer situation in China: Good or bad news from the 2018 global cancer statistics? Cancer communications (London, England) 2019;39(1):22. https://doi.org/10.1186/s40880-019-0368-6.

[2]   Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69(1):7-34.

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

[4]   Malhotra GK, Yanala U, Ravipati A, Follet M, Vijayakumar M, Are C. Global trends in esophageal cancer. J Surg Oncol 2017;115(5):564-79. https://doi.org/10.1002/jso.24592.

[5]   Brown LM, Devesa SS, Chow WH. Incidence of adenocarcinoma of the esophagus among white Americans by sex, stage, and age. J Natl Cancer Inst 2008;100(16):1184-7. https://doi.org/10.1093/jnci/djn211.

[6]   Nozoe T, Saeki H, Ohga T, Sugimachi K. Clinicopathologic characterisics of esophageal squamous cell carcinoma in younger patients. Ann Thorac Surg 2001;72(6):1914-7. https://doi.org/10.1016/s0003-4975(01)03091-0.

[7]   Kolb JM, Han S, Scott FI, Murphy CC, Hosokawa P, Wani S. Early onset esophageal adenocarcinoma presents with advanced stage disease but has improved survival compared with older individuals. Gastroenterology 2020;159(6):2238-40.e4. https://doi.org/10.1053/j.gastro.2020.08.002.

[8]   Mansour NM, Groth SS, Anandasabapathy S. Esophageal adenocarcinoma: Screening, surveillance, and management. Ann Rev Med 2017;68:213-27. https://doi.org/10.1146/annurev-med-050715-104218.

[9]   Adam DJ, Craig SR, Sang CT, Walker WS, Cameron EW. Oesophagogastrectomy for carcinoma in patients under 50 years of age. J R Coll Surg Edinburgh 1996;41(6):371-3.

[10] Markar SR, Karthikesalingam A, Low DE. Outcomes assessment of the surgical management of esophageal cancer in younger and older patients. Ann Thorac Surg 2012;94(5):1652-8. https://doi.org/10.1016/j.athoracsur.2012.06.067.

[11] Mori M, Ohno S, Tsutsui S, Matsuura H, Kuwano H, Sugimachi K. Esophageal carcinoma in young patients. Ann Thorac Surg 1990;49(2):284-6. https://doi.org/10.1016/0003-4975(90)90152-v.

[12] Oezcelik A, Ayazi S, DeMeester SR, Zehetner J, Abate E, Dunn J, et al. Adenocarcinoma of the esophagus in the young. J Gastrointest Surg 2013;17(6):1032-5. https://doi.org/10.1007/s11605-013-2177-6.

[13] Patil PK, Patel SG, Mistry RC, Deshpande RK, Desai PB. Cancer of the esophagus in young adults. J Surg Oncol 1992;50(3):179-82. https://doi.org/10.1002/jso.2930500310.

[14] Vallböhmer D, Hölscher AH, Brabender J, Prenzel K, Gutschow C, Schröder W, et al. Clinicopathologic and prognostic factors of young and elderly patients with esophageal adenocarcinoma: Is there really a difference? Dis Esophagus 2008;21(7):596-600. https://doi.org/10.1111/j.1442-2050.2008.00817.x.

[15] Yang S, Li H, Jia C, Ma X, Guo W, Li H. Clinicopathological features and prognosis of patients <45 years old with esophageal adenocarcinoma comparing to other age groups. J Thorac Dis 2016;8(10):2724-9. https://doi.org/10.21037/jtd.2016.09.61.

[16] Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26(8):1364-70. https://doi.org/10.1200/jco.2007.12.9791.

[17] Park SY. Nomogram: An analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg 2018;155(4):1793.

[18] Bianco FJ Jr. Nomograms and medicine. Eur Urol 2006;50(5):884-6.

[19] Ai D, Chen Y, Liu Q, Deng J, Zhao K. The effect of tumor locations of esophageal cancer on the metastasis to liver or lung. J Thorac Dis 2019;11(10):4205-10. https://doi.org/10.21037/jtd.2019.09.67.

[20] Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak 2008;8:53. https://doi.org/10.1186/1472-6947-8-53.

[21] Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ 2016;352:i6. https://doi.org/10.1136/bmj.i6.

[22] Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. Med Decis Mak 2006;26(6):565-74. https://doi.org/10.1177/0272989x06295361.

[23] 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(21):7252-9.
https://doi.org/10.1158/1078-0432.ccr-04-0713.

[24] GBD 2017 Oesophageal Cancer Collaborators. The global, regional, and national burden of oesophageal cancer and its attributable risk factors in 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. Lancet Gastroenterol Hepatol 2020;5(6):582-97.
https://doi.org/10.1002/cncr.33476.

[25] Liu X, Zhou M, Wang F, Mubarik S, Wang Y, Meng R, et al. Secular trend of cancer death and incidence in 29 cancer groups in China, 1990-2017: A joinpoint and age-period-cohort analysis. Cancer Manag Res 2020;12:6221-38.
https://doi.org/10.2147/cmar.s247648.

[26] GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. Lancet (London, England) 2018;392(10159):1923-94.
https://doi.org/10.1038/sj.bdj.2015.751.

[27] Awan AK, Iftikhar SY, Morris TM, Clarke PA, Grabowska AM, Waraich N, et al. Androgen receptors may act in a paracrine manner to regulate oesophageal adenocarcinoma growth. Eur J Surg Oncol 2007;33(5):561-8.
https://doi.org/10.1016/j.ejso.2006.12.001.

[28] Abbas G, Krasna M. Overview of esophageal cancer. Ann Cardiothorac Surg 2017;6(2):131-6.

[29] Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut 2015;64(3):381-7.
https://doi.org/10.1136/gutjnl-2014-308124.

[30] Zeng Y, Ruan W, Liu J, Liang W, He J, Cui F, et al. Esophageal cancer in patients under 50: A SEER analysis. J Thorac Dis 2018;10(5):2542-50.
https://doi.org/10.21037/jtd.2018.05.72.

[31] Labenz J. Barrett's esophagus. Internist 2016;57(11):1079-92.

[32] Watanabe M, Otake R, Kozuki R, Toihata T, Takahashi K, Okamura A, et al. Recent progress in multidisciplinary treatment for patients with esophageal cancer. Surg Today 2020;50(1):12-20.
https://doi.org/10.1007/s00595-019-01878-7.

[33] Hong Y, Ding ZY. PD-1 Inhibitors in the advanced esophageal cancer. Front Pharmacol 2019;10:1418.

[34] Yang H, Wang K, Wang T, Li M, Li B, Li S, et al. The combination options and predictive biomarkers of PD-1/PD-L1 inhibitors in esophageal cancer. Front Oncol 2020;10:300.
https://doi.org/10.3389/fonc.2020.00300.

[35] van Nistelrooij AM, van Marion R, Biermann K, Spaander MC, van Lanschot JJ, Wijnhoven BP, et al. Early onset esophageal adenocarcinoma: A distinct molecular entity? Oncoscience 2016;3(1):42-8.
https://doi.org/10.18632/oncoscience.290.

[36] Juo YY, Gibbons MA, Dutson E, Lin AY, Yanagawa J, Hines OJ, et al. Obesity is associated with early onset of gastrointestinal cancers in California. J Obes 2018;2018:7014073.
https://doi.org/10.1155/2018/7014073.

Related articles published in BJBMS

1. Development and validation of nomograms for predicting survival of elderly patients with stage I small-cell lung cancer
   Yaji Yang et al., BJBMS, 2020

2. Trends of incidence and prognosis of upper tract urothelial carcinoma
   Ming Chen et al., BJBMS, 2020

SUPPLEMENTRY TABLE

TABLE S1. Demographic and clinicopathologic characteristics of EOEC patients

Variables	Training set (n=266)		Validation set (n=268)	
	n	%	n	%
Age (years)				
≤36	54	20.30	55	20.52
37-44	212	79.70	213	79.48
Sex				
Female	44	16.54	35	13.06
Male	222	83.46	233	86.94
Race				
White	219	82.33	216	80.60
Black	30	11.28	30	11.19
Other	17	6.39	22	8.21
Grade				
Grade I	16	6.02	19	7.09
Grade II	120	45.11	111	41.42
Grade III	126	47.37	134	50.00
Grade IV	4	1.50	4	1.49
Histology				
Squamous cell carcinoma	54	20.30	50	18.66
Adenocarcinoma	206	77.44	211	78.73
Other	6	2.26	7	2.61
Primary site				
Upper	12	4.51	10	3.73
Middle	36	13.53	29	10.82
Lower	213	80.08	218	81.34
Overlapping area	5	1.88	11	4.10
T stage				
T1	58	21.80	70	26.12
T2	34	12.78	27	10.07
T3	103	38.72	118	44.03
T4	71	26.69	53	19.78
N stage				
N0	54	20.30	84	31.34
N1	212	79.70	184	68.66
M stage				
M0	44	16.54	172	64.18
M1	222	83.46	96	35.82
SEER stage				
Localized	219	82.33	31	11.57
Regional	30	11.28	128	47.76
Distant	17	6.39	109	40.67
Surgery				
No	148	55.64	150	55.97
Yes	118	44.36	118	44.03
Radiotherapy				
No	92	34.59	97	36.19
Yes	174	65.41	171	63.81
Chemotherapy				
No	56	21.05	50	18.66
Yes	210	78.95	218	81.34

EOEC: Early-onset esophageal cancer; SEER: Surveillance, epidemiology and end results; Grade I: Well-differentiated; Grade II: Moderately differentiated; Grade III: Poorly differentiated; Grade IV: Undifferentiated; TNM: Tumor-node-metastasis