International Journal of Surgery 66 (2019) 53–61

Contents lists available at ScienceDirect

International Journal of Surgery

journal homepage: www.elsevier.com/locate/ijsu

Original Research

Prediction of lymph node metastatic status in superficial esophageal squamous cell carcinoma using an assessment model combining clinical characteristics and pathologic results: A retrospective cohort study

Yue Zhou[a,1], Junjie Du[a,1], Yiqiu Wang[b,1], Hai Li[c], Guoqiang Ping[c], Jinhua Luo[a], Liang Chen[a], Shijiang Zhang[a,*], Wei Wang[a,**]

[a] Department of Thoracic and Cardiovascular Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China
[b] Department of Surgical Oncology, The Affiliated Xuzhou Hospital of Medical College of Southeast University, Xuzhou, 221009, China
[c] Department of Pathology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China

Keywords:
Lymph node metastasis
Prediction model
Superficial esophageal squamous cell carcinoma

A B S T R A C T

Background: It is important to identify the risk of lymph node metastasis (LNM) in patients with superficial esophageal squamous carcinoma (SESC) who have received endoscopic resection (ER). We aimed to develop a risk-predicting model for metastasis of SESC to lymph nodes using clinicopathological features and pathological results.

Methods: Clinical data on 539 consecutive patients who underwent esophagectomy for SESC in our hospital were collected. Their post-surgical pathological results were assessed and analyzed. Multivariate logistic regression was used to identify all independent risk factors associated with LNM that then were incorporated into the prediction model.

Results: LNM was identified in 53 of 366 patients and 30 of 173 patients by positive histopathological results in the training and validation cohorts. The risk factors associated with LNM were large tumor size, poor tumor grade, deep invasion, and presence of angiolymphatic invasion. The model achieved good discriminatory ability of 0.80 (95%CI, 0.74–0.86) and 0.81 (95%CI, 0.75–0.86) in predicting LNM in the training and validation cohorts respectively. A LNM-predicting nomogram was formed with an area under curve of 0.80 (95% CI, 0.74–0.86), which had well-fitted calibration curves.

Conclusions: A prediction model was constructed to generates 3 categories for estimated LNM risk in SESC patients. It provides a practical way of estimation of LNM risk in SESC patients who had received ER.

1. Introduction

With the development of fiber endoscopy and endoscopic ultrasonography, the accurate initial anatomic staging of esophageal cancer has become available. As a result, more patients with superficial esophageal squamous carcinoma (SESC) have been diagnosed [1]. Compared with the radical resection of esophageal cancer via traditional thoracotomy, the advantages of endoscopic resection (ER) lie in less procedure-associated trauma and quicker recovery time. Consequently, patients with early stages (tumor confined to the mucosa and submucosal layer) are favorably treated by ER, including endoscopic mucosal resection and endoscopic submucosal dissection. But, if SESC patients have positive lymph node metastasis (LNM) when receiving the procedure, ER will leave residual cancer behind and result in worse prognosis. Therefore, ER is not a treatment option for SESC patients with confirmed LNM.

Unfortunately, the diagnosis of LNM is not feasible for SESC patients who undergo ER due to the fact that it is determined by

Abbreviations: SESC, superficial esophageal squamous carcinoma; ER, endoscopic resection; LNM, lymph node metastasis; ALI, angiolymphatic invasion; ROC, receiver operationg characteristic; AUC, area under curve

* Corresponding author. Department of Thoracic and Cardiovascular Surgery, First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, 210029, China.
** Corresponding author. Department of Thoracic Surgery, First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, 210029, China.
E-mail addresses: drzhangshijiang@163.com (S. Zhang), drwangwei4445@163.com (W. Wang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ijsu.2019.04.014
Received 8 December 2018; Received in revised form 13 March 2019; Accepted 22 April 2019
Available online 25 April 2019
1743-9191/

histopathological examination of lymph nodes obtained from surgical resection. Therefore, it is crucial to accurately estimate the risk of LNM of SESC patients after receiving ER. If a patient is considered with a high risk of LNM, a salvage surgical lymph node resection will be needed for curative intent. On the other hand, for those who are ruled out LNM, lymph node dissection with open surgery may increase procedure-associated trauma over potential benefit to those patients. Thus, it is important to estimate the risk of LNM for optimal therapeutic choice.

Despite the fact that the tumor of early SESC patients is confined to the mucosal or submucosal layers, the incidence of LNM in these patients can be as high as 14.4–44.61% [2,3]. Our group and others previously found that LNM in SESC patients are closely associated with the tumor size, the tumor grade, the depth of the original lesion, and whether there are angiolymphatic invasions (ALIs) [4,5]. However, there are no practical methods to assess and predict the risk of LNM in SESC patients. Therefore, it is desirable to develope a predictive model that incorporates factors associated with LNM based on clinicopathological data. The aim of this study was to establish a mathematical model of prediction between the clinicopathological factors and the risk of LNM to help doctors facilitate management-related decision making for SESC patients who received ER.

2. Methods

2.1. Study population

The institutional ethics committee approved the study with waiver of consent (approval no.2017-SR-096). The work has been reported in line with the STROCSS criteria [6]. It was registered in www.researchregistry.com with a Research Registry UIN,4561.

A total of 3026 consecutive patients who underwent surgical esophagectomy for esophageal cancer at our hospital from January 2010 to September 2016 were reviewed retrospectively. Patients met the following inclusion criteria were recruited to this study: (I) preoperative clinical N0 according to CT imaging; (II) tumor invasion was confined to mucosa or submucosa according to the postoperative histological findings. We excluded patients who met any one of the following conditions from this study: (I) multiple primary cancers of different pathologic types; (II) ER history; (III) received neoadjuvant therapy before surgery; (IV) incomplete lymph node dissection including not enough for total number of dissected lymph nodes or not enough for the number of dissected lymph node fields; (V) incomplete clinical data.and incomplete systemic lymph node dissection during surgery. Fig. 1 depicts the recruitment of participants. Three hundred and sixty-six eligible patients who underwent surgery between January 2010 and December 2014 were included into the training cohort to establish the model, and 173 patients who underwent surgery after December 2014 were entered into the validation cohort.

2.2. Surgical procedures

Surgical resection was performed via the left thoracic (sweet); abdominal right-thoracic (Ivor-Lewis) or right thoracic-abdominal-cervical (McKeown) approach. The required proximal and distal margins were 6–8 cm or more from the gross tumor. Lymph node dissection in all surgical approaches met the requirement for strict radical resection of esophageal cancer, including the periesophageal, infracardial, posterior mediastinal and paracardial lymph nodes, as well as those located along the lesser gastric curvature and at the origin of the left gastric artery, the celiac trunk, the common hepatic artery and the splenic artery. The whole stomach was used as an esophageal substitute. Esophagogastrostomy was performed with a circular stapler at the thoracic inlet level.

2.3. Clinicopathological variables

The clinicopathological variables in this study included the patient's age, sex, body weight index, blood type, tumor location, tumor size, tumor grade, depth of invasion, the presence of multicentric tumor, the number of resected nodes, the presence of ALI, and the presence of LNM. For histopathologic assessment, surgical specimens were cut into 2-mm slices, and further cut into 4-μm slides for standard hematoxylin and eosin staining, then assessed independently by two experienced pathologists and the final classification was based on the consensus diagnosis of these two pathologists. The classification or measurement criteria of pathologic features were summarized in Table 1.

2.4. Statistical analysis

Variables were characterized descriptively using medians and the interquartile range for continuous variables or categories and percentages for discrete variables. Significance of differences was assessed with the student's t-test for continuous variables and chi-square test for categorical variables (Fish-test was used when the number of variables was less than five). Significance of associations with the outcome of LNM in the training cohort was first evaluated using a univariate logistic model, then by a multivariable analysis. Clinically relevant pathologic variables were included in the multivariable regression model. According to the independent risk factors of multivariable analysis and the resulting model coefficients, a mathematical equation ($\hat{y} = 1/[1 + \exp(-x\beta)]$) was established. Receiver operating characteristic (ROC) curve analysis was used to calculate the optimal cut-off values that were determined by maximizing the Youden index (i.e., sensitivity – (1 − specificity)) and the area under ROC curve was obtained to calculate predicted values.

For clinical use of the equation, a nomogram was formulated based on proportionally converting each regression coefficient in multivariate logistic regression to a 0- to 100-point scale by using the *rms* package of R, version 3.2 (http://www.r-project.org/). The effect of the variable with the highest β coefficient (absolute value) is assigned 100 points. The points are added across independent variables to derive total points, which are converted to predicted probabilities. The predictive performance of the nomogram was measured by area under curve (AUC) and calibration with 1000 bootstrap samples to decrease the overfit bias.

In all analyses, $P < 0.05$ was considered to indicate statistical significance. All analyses were performed using SPSS, version 18.0 (IBM, Armonk, NY) and R programming language, version 3.2.

3. Results

3.1. Characteristics of the study population

During the study period, 3026 consecutive patients with esophageal cancer underwent esophagectomy. Of these, 539 patients who met the inclusion criteria were enrolled. 366 and 173 patients were divided into the training and validation cohorts respectively. The clinicopathologic characteristics of the patients are listed in Table 2. The baseline clinicopathologic data were similar between the training and validation cohorts. The rate of LNM was 14.48% (53/366) and 17.34% (30/173) in the two cohorts, respectively.

3.2. Association of individual pathologic characteristics with nodal metastasis in training cohort

The clinicopathological factors and the presence of LNM of the patients from the training cohort were evaluated (Table 3). Tumor size, tumor grade, depth of invasion, and presence of ALI were the significant factors associated with LNM, while there was no significant difference in other variables. We determined the optimal clinically applicable

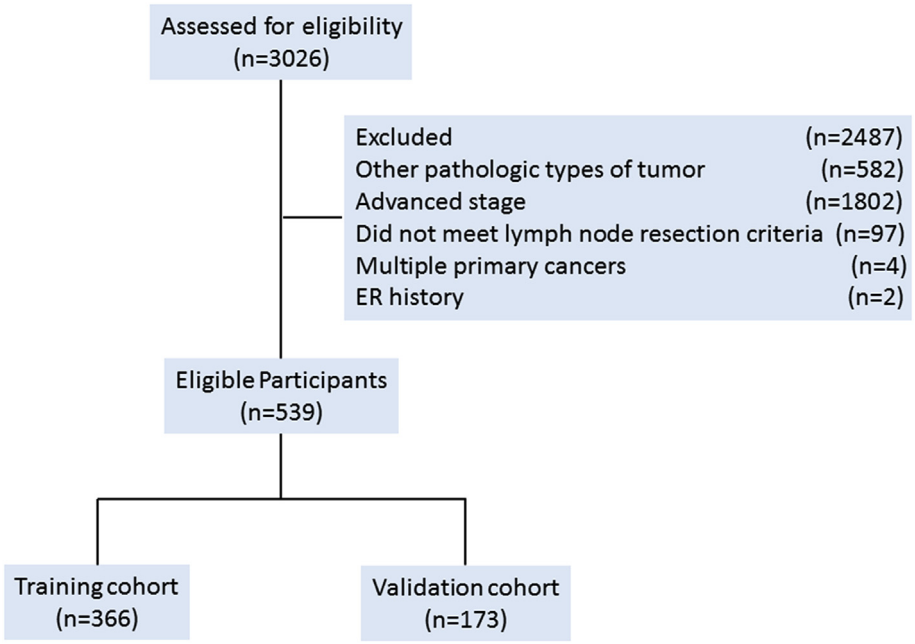

Fig. 1. CONSORT diagram of participant recruitment.

Table 1
Definition of histopathologic risk factors for nodal metastasis.

Risk factor	Classification or measurement criteria
Tumor location	(primary tumor site) is defined by position of upper (proximal) edge of tumor in esophagus[a]
Cervical esophagus	bounded superiorly by the cricopharyngeus and inferiorly by the sternal notch, is typically 15–20 cm from the incisors using esophagoscopy
Upper thoracic	bounded superiorly by the sternal notch and inferiorly by the azygos arch, is typically 20–25 cm from the incisors using esophagoscopy
Middle thoracic	bounded superiorly by the azygos arch and inferiorly by the inferior pulmonary vein, is typically 25–30 cm from the incisors using esophagoscopy
Lower thoracic	bounded superiorly by the inferior pulmonary vein and inferiorly by the lower esophageal sphincter, is typically 30–40 cm from the incisors using esophagoscopy; this location includes cancers whose epicenter is within the proximal 5 cm of the stomach that extend into the EGJ or lower thoracic esophagus. EGJ, esophagogastric junction
Tumor size	Tumor size (in cm) based on maximal cross-sectional dimension on histologic sections or maximal gross tumor size measurement; for Multicentric tumors, size of the largest focus was used
Histologic grade	
I	Well differentiated tumor
II	Moderately differentiated tumor
III	Poorly differentiated tumor
Depth of invasion	
m1	Intraepithelial tumor
m2	Tumor invading the lamina propria
m3	Tumor invading the muscularis mucosa
sm1	Tumor invading the most superficial one third of the submucosa
sm2	Tumor invading the middle one third of the submucosa
sm3	Tumor invading deeper than sm2 level
Multicentric tumor	
Present	Synchronous multiple primary squamous cell carcinoma
Absent	Above criterion not met
Angiolymphatic invasion	
Present	Tumor invading the lymphatic vessel
Absent	Above criterion not met

[a] , American Joint Committee on Cancer; American Cancer Society; American College of Surgeons. AJCC cancer staging manual. 7th edition. New York: Springer; 2010.

tumor size cutoff value was 1.75 cm by maximizing the sum of sensitivity and specificity, i.e. Youden's index (Table 4). Therefore, the tumors were stratified as > 1.75 cm or ≤ 1.75 cm. Using tumor size alone, the area under the ROC curve was 0.65 (95% confidence interval [CI], 0.58–0.72) (Fig. 2).

Univariate logistic regression included tumor size, tumor grade, depth of invasion, and ALI as the candidate predictors of LNM. Multivariable analysis confirmed that tumor size (> 1.75 cm, OR = 2.90; 95% CI, 1.45–5.81), tumor grade (II + III, OR = 2.82; 95% CI, 1.17–6.82), depth of invasion (sm1-3, OR = 2.78; 95% CI, 1.01–7.66), presence of ALI (OR = 6.73; 95% CI, 2.51–18.04) were significant predictors of LNM (Table 5).

3.3. Estimating the possibility of LNM using a mathematic model

Abovementioned four independent risk factors associated with LNM were used to form a mathematic model. The resulting beta coefficients in multivariate analysis were applied to calculate predicted values from the logistic equation by using the following weighted sum: $x\beta = -4.339 + (1.211 \times \text{Tumor size}) + (1.078 \times \text{Tumor grade}) + (1.036 \times \text{Depth of invasion}) + (2.661 \times \text{Angiolymphatic invasion})$. The estimated possibility of LNM was calculated for each

Table 2
Characteristics of the study population.

Variables	Cohort, No. (%)	
	Training (n = 366)	Validation (n = 173)
Age (years)		
Mean ± SD	60.61 ± 7.04	63.27 ± 6.92
Gender, n (%)		
Male	268 (73.22)	120 (69.36)
Female	98 (26.78)	53 (30.64)
BMI (weight/high^2)		
Mean ± SD	23.07 ± 2.96	23.40 ± 3.00
Blood-type, n (%)		
A	120 (32.79)	43 (24.86)
B	101 (27.60)	32 (18.50)
O	90 (24.59)	80 (46.24)
AB	55 (15.03)	18 (10.40)
Tumor location, n (%)		
Cervical	15 (4.10)	8 (4.62)
Upper thoracic	64 (17.49)	30 (17.34)
Middle thoracic	183 (50)	88 (50.87)
Lower thoracic	104 (28.41)	47 (21.17)
Tumor size (cm)		
Mean ± SD	1.80 ± 1.01	1.86 ± 0.88
Differentiation, n (%)		
I	143 (39.07)	64 (36.99)
II	199 (54.37)	101 (58.38)
III	24 (6.56)	8 (4.62)
Depth of invasion, n (%)		
m1	23 (6.28)	16 (9.25)
m2	49 (13.39)	22 (12.72)
m3	60 (16.39)	32 (18.50)
sm1	45 (12.30)	20 (11.56)
sm2	99 (27.05)	54 (31.21)
sm3	90 (24.59)	29 (16.76)
Multicentric invasion, n (%)		
Absent	353 (96.45)	165 (95.38)
Present	13 (3.55)	8 (4.62)
Angiolymphatic invasion, n (%)		
Absent	349 (95.36)	161 (93.06)
Present	17 (4.64)	12 (6.94)
Surgical procedure, n (%)		
Traditional	215 (58.74)	46 (26.59)
Expanded	151 (41.26)	127 (73.41)
Number of resected lymph nodes		
Mean ± SD	12.04 ± 4.94	17.27 ± 7.07

Table 3
Characteristics of patients in training cohort.

Variables	Group		P value
	No LNM	LNM	
Number	313	53	
Age (years)			0.086
Mean ± SD	60.87 ± 7.02	59.08 ± 7.01	
Gender			0.462
Male	227	41	
Female	86	12	
BMI (weight/high^2)			0.442
Mean ± SD	23.12 ± 3.01	22.78 ± 2.72	
Blood-type			0.918
A	104	16	
B	87	14	
O	75	15	
AB	47	8	
Tumor location			0.526
Cervical thoracic	14	1	
Upper thoracic	55	10	
Middle thoracic	159	24	
Lower thoracic	85	19	
Tumor size (cm)			0.003
Mean ± SD	1.73 ± 1.01	2.18 ± 0.95	
Differentiation			< 0.001
I	136	7	
II	157	42	
III	20	4	
Depth of invasion			< 0.001
m1	23	0	
m2	49	0	
m3	55	5	
sm1	36	9	
sm2	88	11	
sm3	62	28	
Multicentric invasion			0.413
Absent	303	50	
Present	10	3	
ALI			< 0.001
Absent	307	42	
Present	6	11	
Surgical procedure			0.968
Traditional	184	31	
Expanded	129	22	
Number of resected lymph nodes			
Mean ± SD	11.88 ± 4.89	13.04 ± 5.21	0.114

LNM, lymph node metastasis.
ALI: angiolymphatic invasion.

patient, ranging from 3.33% (≤1.75 cm, well-differentiated, m1-m3, and no ALI) up to 86.67% (> 1.75 cm, poorly-differentiated, sm1-sm3, with ALI). The ROC curve of the estimated values was generated and the area under the ROC curve (AUROC) which estimates the model's discriminatory ability was 0.80 (95% CI, 0.737–0.862) (Fig. 3A).

According to the maximized Youden's index, i.e., the sum of sensitivity and specificity, the optimal clinically applicable cutoff value of estimated risks was 0.2 (20%) (Table 6). Therefore, we defined the predicted risk groups using a priori based on cutoff values at 5% and 20% of estimated possibility of LNM (Table 7). As listed in Table 7, LNM were observed in 2.84% (4/141), 8.49% (9/106), and 33.61% (40/119) in the low-risk, intermediate-risk, and high-risk groups of prediction, respectively.

3.4. Development of an LNM-Predicting nomogram

The independently associated risk factors were used to form a LNM-predicting nomogram to facilitate clinicians to use abovementioned prediction model (Fig. 4A). The nomogram demonstrated good accuracy in estimating the risk of LNM with an AUC of 0.80 (95% CI, 0.739–0.856). In addition, calibration plots graphically showed good agreement on the presence of LNM between the risk estimation by the nomogram and histopathologic confirmation on surgical specimens (Fig. 4B).

3.5. Validation of the prediction model

For model validation, the fitted model derived from the training cohort was applied to the validation cohort to produce an estimated value of possibility of LNM risk for each patient. All patients in the validation cohort were assigned to a risk group based on their resulting value and LNM was identified in 2 of 70 cases (2.86%) in low-risk group, in 4 of 39 cases (10.26%) in intermediate-risk group, and 24 of 64 cases (37.50%) in high-risk group (Table 7). A ROC curve of estimated value in validation cohort (Fig. 3B) was generated and the resulting area under ROC was 0.814 (95% CI, 0.724–0.905), demonstrating a good discriminatory power for this model. Statistical analysis also showed high uniformity of ROC curves for risk scores between training and validation cohorts (Table 8).

The performance of the model was also examined in the nomogram with the validation cohort. The nomogram displayed an AUC of 0.814 (95% CI, 0.725–0.900) for the estimation of LNM risk in the validation cohort. There was a good calibration curve for the risk estimation (Fig. 4C).

Table 4
Tumor size cutoff values with their sensitivity, specificity, and Youden's index.

Cutoff value [a]	Sensitivity	1 - Specificity	Youden index
−0.900	1.000	1.000	0.000
0.150	1.000	0.997	0.003
0.250	1.000	0.990	0.010
0.350	1.000	0.974	0.026
0.450	1.000	0.965	0.035
0.550	1.000	0.904	0.096
0.650	0.981	0.875	0.106
0.750	0.981	0.856	0.125
0.900	0.981	0.815	0.166
1.050	0.868	0.684	0.184
1.150	0.868	0.681	0.187
1.250	0.811	0.633	0.179
1.350	0.792	0.623	0.169
1.450	0.792	0.620	0.173
1.550	0.717	0.473	0.244
1.650	0.717	0.470	0.247
1.750	**0.717**	**0.454**	**0.263**
1.900	0.660	0.428	0.232
2.100	0.415	0.230	0.185
2.300	0.415	0.217	0.198
2.450	0.415	0.211	0.204
2.600	0.189	0.141	0.048
2.750	0.189	0.137	0.051
2.900	0.189	0.134	0.054
3.100	0.075	0.064	0.012
3.250	0.075	0.061	0.015
3.400	0.075	0.058	0.018
3.550	0.075	0.051	0.024
3.800	0.075	0.045	0.031
4.250	0.057	0.026	0.031
4.550	0.038	0.019	0.019
4.800	0.038	0.016	0.022
5.500	0.000	0.006	−0.006
6.500	0.000	0.003	−0.003
8.000	0.000	0.000	0.000

[a] , The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

Table 5
Univariate and multivariate predictors of lymph node metastasis in training cohort.

Predictor	Univariate analysis		Multivariate analysis		
	odds ratio (95% CI)	*P* value	β	odds ratio (95% CI)	*P* value
Tumor size					
≤1.75 cm	Reference			Reference	
> 1.75 cm	3.05 (1.61–5.77)	0.001	1.211	3.36 (1.57–7.16)	0.002
Tumor grade					
I	Reference			Reference	
II + III	5.05 (2.21–11.53)	< 0.001	1.078	2.94 (1.19–7.26)	0.019
Depth of invasion					
m1-3	Reference			Reference	
sm1-3	6.56 (2.54–16.92)	< 0.001	1.036	2.82 (1.01–7.86)	0.044
ALI					
Absent	Reference			Reference	
Present	13.40 (4.71–38.13)	< 0.001	2.661	14.32 (4.45–46.03)	< 0.001

4. Discussion

The presence of LNM is an independent risk factor for worse prognosis of patients with esophageal squamous carcinoma [7–10]. Compared with the high LNM rates (32.5–69%) of average (all stages) esophageal squamous carcinoma [11–14], the LNM rate of SESC is much lower. However, despite the fact that tumor is confined to the mucosal or submucosal layers, up to 44.61% of patients with SESC still have nodal metastasis [3]. Our data showed LNM rates of SESC patients with M1, M2, M3, SM1, SM2, and SM3 lesions are 0%, 0%, 9.78%, 16.92%, 13.73%, and 35.29% respectively [2,3,15,16], which are in line with the literature reports (0%, 0–5.6%, 4.54–18%, 11–53.1%, 20.5–44.0%, and 18.2–61%). As LNM rates of submucosal tumors are ranging from 25.5% to 53.85% (our data is 21.96%) [2,3], its consequence should not be overlooked.

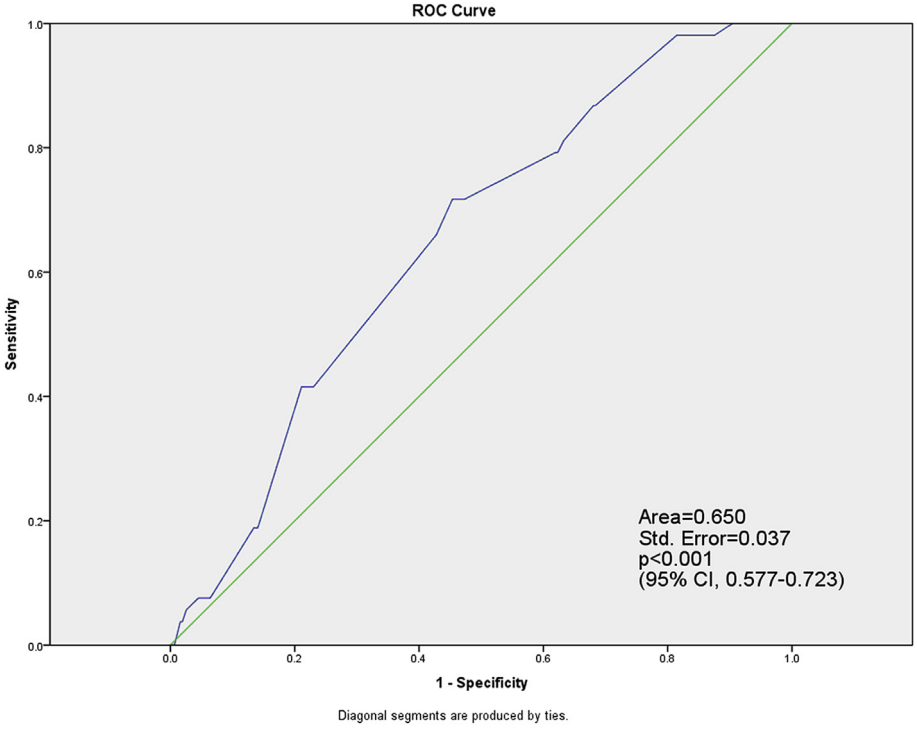

Fig. 2. Receiver operator curve (ROC) of tumor size cutoff values in training cohort.

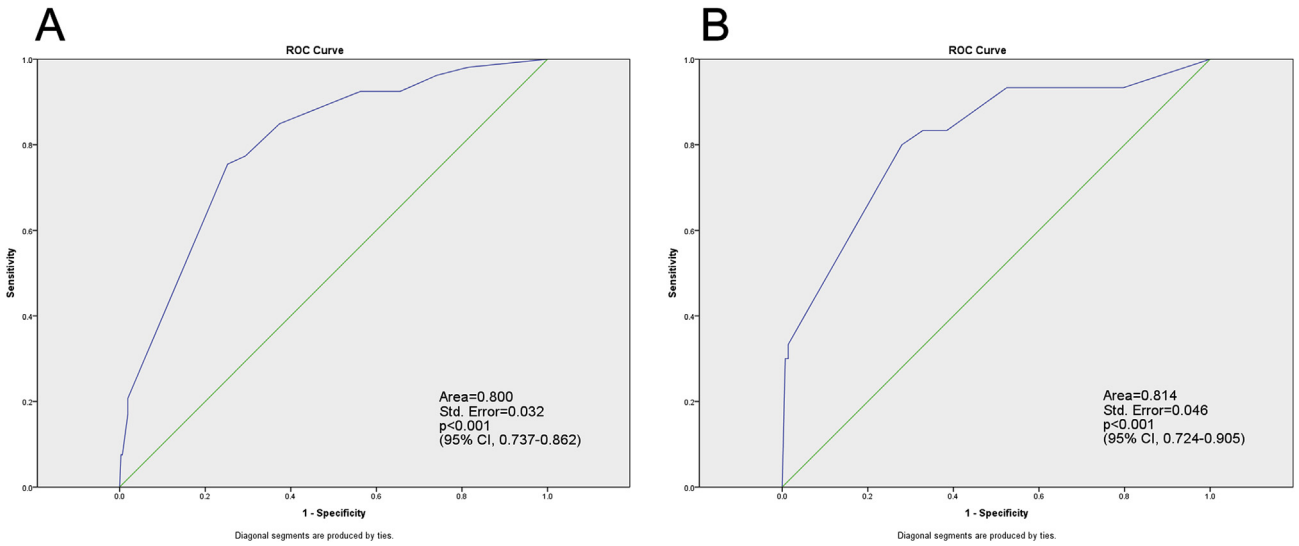

Fig. 3. Receiver operator curve (ROC) of model estimated probabilities of lymph node metastasis in training and validation cohorts. A, the training cohorts, the area under ROC was 0.80 (95% CI, 0.737–0.862). B, the validation cohorts, the area under ROC was 0.814 (95% CI, 0.724–0.905).

Table 6
Cutoff values of estimated LNM risks with their sensitivity, specificity, and Youden's index.

Cutoff value ([a])	Sensitivity	1 - Specificity	Youden index
0.000	1.000	1.000	0.000
0.024	0.981	0.815	0.166
0.036	0.962	0.741	0.221
0.039	0.925	0.655	0.270
0.070	0.925	0.562	0.362
0.104	0.849	0.374	0.475
0.112	0.774	0.294	0.480
0.190	**0.755**	**0.252**	**0.502**
0.305	0.208	0.019	0.188
0.350	0.189	0.019	0.170
0.481	0.170	0.019	0.151
0.623	0.075	0.006	0.069
0.739	0.075	0.003	0.072
1.000	0.000	0.000	0.000

[a] , The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All the other cutoff values are the averages of two consecutive ordered observed test values.

Table 7
Predicted risk groups based on estimated possibility of LNM with observed LNM rate in training and validation cohorts.

Predicted risk group	Estimated possibility of LNM	Observed LNM rate	
		Training cohort	Validation cohort
Low	≤0.05	4/141 (2.84)	2/70 (2.86)
Intermediate	0.05–0.2	9/106 (8.49)	4/39 (10.26)
High	> 0.2	40/119 (33.61)	24/64 (37.50)

LNM, lymph node metastasis.

Due to the substantial esophagectomy-related morbidity and mortality [17], ER is now more widely applied for the management of SESC as it maintains the integrity of the esophagus. Several studies have demonstrated favourable outcomes of ER on SESC patients with m1 and m2 lesions [18–20]. However, no specific guidelines have yet been developed for the use of ER in variety of SESC cases. Although ER is proven to be sufficient for SESC patients without LNM, the oncological safety of ER for SESC patients with potential LNM is still open to debate, mainly due to the fact that the evidence of LNM is not available for patients who undergo ER as it is determined by histologic examination of lymph nodes obtained during surgical resection. In this study, we developed an accurate mathematic model to predict LNM in SESC using clinicopathologic data that can be obtained at ER, which is important for final decision making after ER.

Previous studies showed that, in Early SESC, the rate of lymph node metastasis is significantly associated with the characteristics of local tumors, including tumor size (length, area, and thickness), tumor morphology, tumor grade, depth of invasion, presence of vascular invasion and nerve invasion [3,4,21–24]. Some studies pointed out that tumor size, tumor grade, depth of invasion, and vascular invasion are the main independent risk factors of LNM in SESC [4]. Our findings are consistent with these results.

Although confirming the presentation of a single or multiple independent risk factors is correlated with relatively high risk of LNM, it is far away from estimating the potential incidence of LNM or predicting its outcomes after ER treatment. This study retrospectively collected clinical pathologic data from the training cohort as constructing information and prospectively collected clinical pathologic data from the validation cohort as confirming information. Having analyzed these data of local tumor characteristics using univariate and multivariate regression model, we established a prediction formula to estimate the potential LNM rates by applying mathematic weighting (incorporated) approach to these risk factors and presented the relationships in a form of nomogram for visualization. This prediction model and nomogram incorporate four comprehensive and easily available risk factors/variables which are independently associated with LNM. Although these variables were obtained by pathological examination of samples from surgical resection in this study, they are actually available after any ER procedure. Therefore, the predictive indications of the model are applicable to patients who have undergone ER.

With this prediction model and nomogram, we calculated a score of LNM risk for each patient from training cohort who had individual characteristics of local tumor. ROC analysis of these risk scores was applied to gain optimal cut-off values that were determined by maximizing the Youden index and area under ROC, which demonstrate the model's discriminatory ability between the estimation and observed results of LNM. The predictive model and nomogram also demonstrated good accuracy in estimating the risk of LNM in validation cohort. Based on the cut-off values, patients were divided into three categories, i.e. low risk, intermediate risk, and high risk. According to this

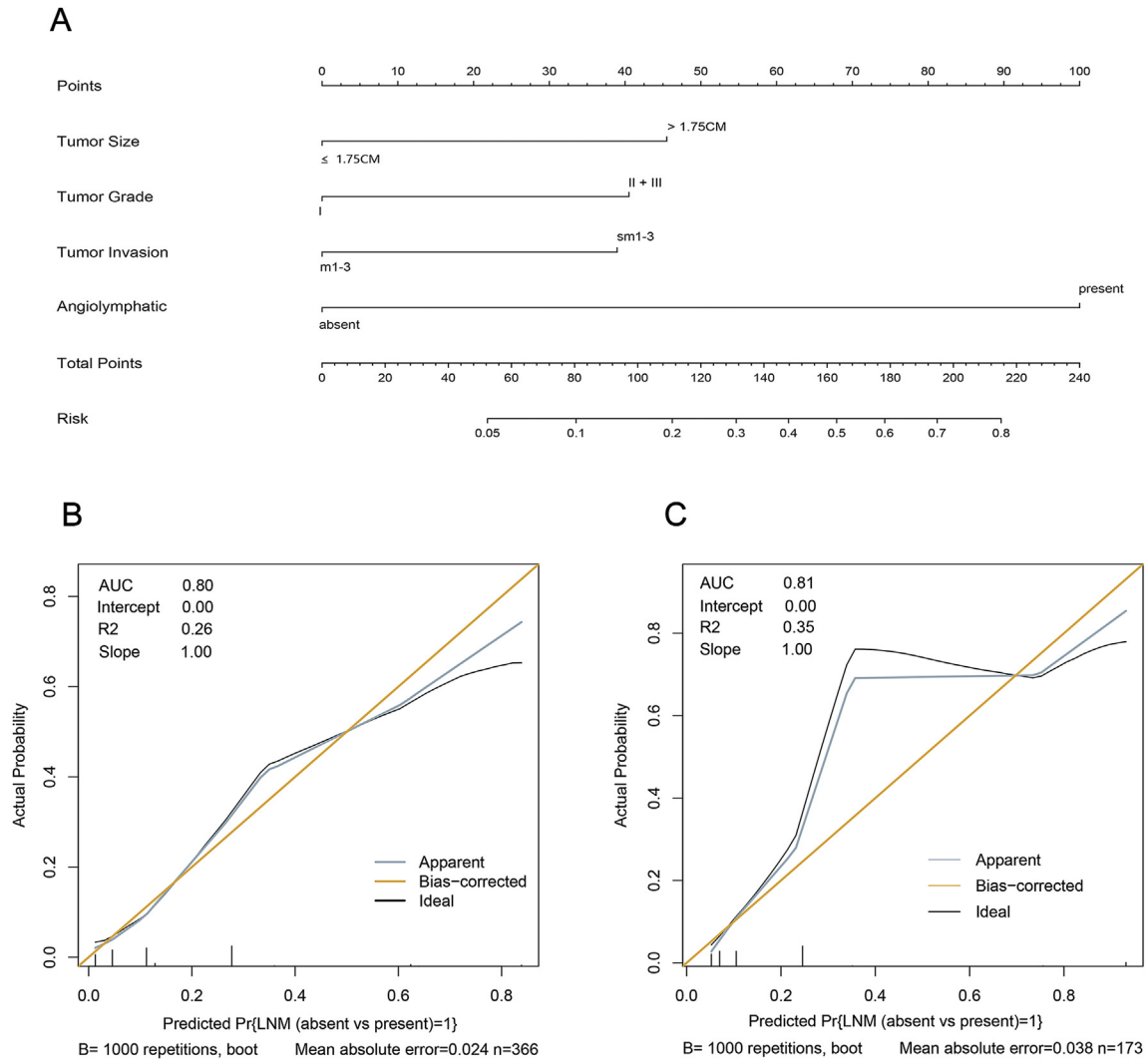

Fig. 4. Nomogram for estimation of LNM risk and its predictive performance. A, the nomogram was formulated based on proportionally converting each regression coefficient in multivariate logistic regression to a 0- to 100-point scale by using the *rms* package of R, version 3.2 (http://www.r-project.org/). B, calibration plots graphically showed good agreement on the presence of LNM between the risk estimation by the nomogram and histopathologic confirmation on surgical specimens in the training cohorts. The AUC value was 0.80 (95% CI, 0.739–0.856). C, calibration plots graphically showed good agreement on the presence of LNM between the risk estimation by the nomogram and histopathologic confirmation on surgical specimens in the validation cohorts. The AUC value was 0.814 (95% CI, 0.725–0.900).

Table 8
Statistical analysis of ROC curves of risk scores between training and validation cohorts.

Variables	ROC	
	Training group	Validation group
Area	0.800	0.814
Standard Error	0.032	0.046
95% CI	0.74–0.86	0.72–0.91
Youden index	0.19	0.19
Difference	−0.014	
Standard Error	0.056	
Z statistic	−0.25	
Significance level	P = 0.8027	

classification, the rates of LNM in these three groups were 2.48%, 8.49%, and 33.61% in training cohort and 2.86%, 10.26%, and 37.50% in validation cohort. Statistical differences were found among low-, intermediate-, and high-risk groups both in training and validation cohorts while there was no significant difference when comparing same risk groups between two cohorts. Also, the ROC curves from these two cohorts showed no statistic difference (Table 2). These results further illustrated that the prediction formula and nomogram established from the training cohort can be applied to the validation cohort. In addition, we found that, although different forms of results (prediction formulas and nomogram) were generated by two distinct statistical software (Spss18.0 and the R Programming Language), the indications are consistent because they were all based on univariate and multivariate logistic regression statistics.

For clinical use of the nomogram, we used 111 and 52 as the cutoff values in estimating the risk of LNM (Fig. 4). Patients with a score of 111 or more are a high-risk subgroup of LNM with an estimated LNM risk of 20% or higher. Patients with a score of 52 or less are a low-risk subgroup of LNM with an estimated risk of 5% or lower. Patients with a score between 52 and 111 are in intermediate-risk subgroup. Based on these predictions, doctors could make further therapeutic decision/choice for SESC patients who receive ER as initial treatment to achieve better long-term outcomes. For patients with a score > 111, a salvage surgical lymph node resection is recommended for curative reasons. For patients with a score ≤52, a regular follow-up after ER might be safe and sufficient. For patient with a score between 52 and 111, they should be kept under close surveillance for LNM. However, perspective

clinical trials with long-term outcomes might be needed for further evaluation of the accuracy and efficacy of this prediction model in these specific patients.

This study had limitations. First, there might be statistical bias associated with the small sample size. Second, the analysis was based on data from our single institution and the results would be more generalizable if an external validation was performed. Third, the clinical outcomes of these patients are lacking and no long-term survival data associated with LNM are available in this study. Therefore, the accuracy of this prediction model could not be further confirmed. Last but not least, although LNM is crucial for the decision to choose ER or surgical procedures for SESC patients, it is not the only factor that we should consider. The depth of invasion of the tumor does bring technical difficulties to ER to keep the cancer negative vertical margin. For those who received ER with positive vertical margin, salvage surgery must be done immediately. For those who received ER with negative vertical margin, our prediction model provides extra evidence to decide whether further surgery is necessary.

5. Conclusions

A prediction model was constructed by combining four independent risk factors of LNM in patients with early SESC. This model provides a practical way of estimation of LNM risk in SESC patients who had received ER treatment.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Availability of data and materials

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

Consent for publication

Informed consent for the publication of any associated data was obtained from all individuals.

Data statement

Data will be made available on request.

Ethical approval

This study was approved by the Ethics Committee of the First Affiliated Hospital of Nanjing Medical University (approval no.2017-SR-096). Written informed consent was obtained from all patients and/or their guardians.

Sources of funding

This research was supported by the Natural Science Foundation of Jiangsu Province (Grant No. BK20151589).

Author contribution

Study concept and design: YZ, JD, YW, SZ, WW. Acquisition of Data: YZ, JD, HL, GP, JL, SZ. Drafting of the article: YZ, JD, YW. Critical revision for important intellectual content: YZ, JD, YW, HL, GP, JL, LC, SZ, WW. Final approval of the manuscript: YZ, JD, YW, HL, GP, JL, LC, SZ, WW.

Conflicts of interest

The authors declare that they have no competing interests.

Research registration number

Researchregistry4561.

Guarantor

Shijiang Zhang, M.D., Ph.D., Department of Thoracic Surgery, The First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing, China.

Acknowledgment

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijsu.2019.04.014.

References

[1] C.J. Lightdale, K.G. Kulkarni, Role of endoscopic ultrasonography in the staging and follow-up of esophageal cancer, J. Clin. Oncol. 23 (2005) 4483–4489.
[2] J.Y. Moon, G.H. Kim, J.H. Kim, et al., Clinicopathologic factors predicting lymph node metastasis in superficial esophageal squamous cell carcinoma, Scand. J. Gastroenterol. 49 (2014) 589–594.
[3] T. Eguchi, Y. Nakanishi, T. Shimoda, et al., Histopathological criteria for additional treatment after endoscopic mucosal resection for esophageal cancer: analysis of 464 surgically resected cases, Mod. Pathol. 19 (2006) 475–480.
[4] B. Li, H. Chen, J. Xiang, et al., Pattern of lymphatic spread in thoracic esophageal squamous cell carcinoma: a single-institution experience, J. Thorac. Cardiovasc. Surg. 144 (2012) 778–786.
[5] Y. Zhou, J. Du, H. Li, et al., Clinicopathologic analysis of lymph node status in superficial esophageal squamous carcinoma, World J. Surg. Oncol. 14 (2016) 259.
[6] R.A. Agha, M.R. Borrelli, M. Vella-Baldacchino, R. Thavayogan, D.P. Orgillfor the STROCSS Group, The STROCSS statement: strengthening the reporting of cohort studies in surgery, Int. J. Surg. 46 (2017) 198–202.
[7] N. Rizk, E. Venkatraman, B. Park, et al., The prognostic importance of the number of involved lymph nodes in esophageal cancer: implications for revisions of the American Joint Committee on Cancer staging system, J. Thorac. Cardiovasc. Surg. 132 (2006) 1374–1381.
[8] A.J. Greenstein, V.R. Litle, S.J. Swanson, et al., Prognostic significance of the number of lymph node metastases in esophageal cancer, J. Am. Coll. Surg. 206 (2008) 239–246.
[9] T.W. Rice, E.H. Blackstone, L.A. Rybicki, et al., Refining esophageal cancer staging, J. Thorac. Cardiovasc. Surg. 125 (2003) 1103–1113.
[10] M. Wilson, E.L. Rosato, K.A. Chojnacki, et al., Prognostic significance of lymph node metastases and ratio in esophageal cancer, J. Surg. Res. 146 (2008) 11–15.
[11] Y.P. Liu, L. Ma, S.J. Wang, et al., Prognostic value of lymph node metastases and lymph node ratio in esophageal squamous cell carcinoma, Eur. J. Surg. Oncol. 36 (2010) 155–159.
[12] J.R. Siewert, H.J. Stein, M. Feith, et al., Histologic tumor type is an independent prognostic parameter in esophageal cancer: lessons from more than 1,000 consecutive resections at a single center in the Western world, Ann. Surg. 234 (2001) 360–369.
[13] K.N. Wang, Q.R. Xu, L.Q. Chen, Impact of lymph-node metastasis site in patients with thoracic esophageal cancer: does the lymph node metastasis site matter? J. Surg. Oncol. 101 (2010) 334.
[14] N. Altorki, M. Kent, C. Ferrara, et al., Three-field lymph node dissection for squamous cell and adenocarcinoma of the esophagus, Ann. Surg. 236 (2002) 177–183.
[15] B. Li, H. Chen, J. Xiang, et al., Prevalence of lymph node metastases in superficial esophageal squamous cell carcinoma, J. Thorac. Cardiovasc. Surg. 146 (2013) 1198–1203.
[16] H. Makuuchi, H. Shimada, K. Mizutani, et al., Clinical pathological analysis of surgically resected superficial esophageal carcinoma to determine criteria for deciding on treatment strategy, Diagn. Ther. Endosc. 3 (1997) 211–220.
[17] E.W. Steyerberg, B.A. Neville, L.B. Koppert, et al., Surgical mortality in patients with esophageal cancer: development and validation of a simple risk score, J. Clin. Oncol. 24 (2006) 4277–4284.
[18] Y. Tajima, Y. Nakanishi, A. Ochiai, et al., Histopathologic findings predicting lymph node metastasis and prognosis of patients with superficial esophageal carcinoma: analysis of 240 surgically resected tumors, Cancer 88 (2000) 1285–1293.
[19] H. Fujita, S. Sueyoshi, H. Yamana, et al., Optimum treatment strategy for superficial esophageal cancer: endoscopic mucosal resection vs radical esophagectomy, World

J. Surg. 25 (2001) 424–431.
[20] M. Endo, K. Yoshino, T. Kawano, et al., Clinicopathologic analysis of lymph node metastasis in surgically resected superficial cancer of the thoracic esophagus, Dis. Esophagus 13 (2000) 125–129.
[21] H. Li, Y. Zhang, H. Cai, J. Xiang, Pattern of lymph node metastases in patients with squamous cell carcinoma of the thoracic esophagus who underwent three-field lymphadenectomy, Eur. Surg. Res. 39 (2007) 1–6.
[22] K. Araki, S. Ohno, A. Egashira, et al., Pathologic features of superficial esophageal squamous cell carcinoma with lymph node and distal metastasis, Cancer 94 (2002) 570–575.
[23] L. Xue, L. Ren, S. Zou, et al., Parameters predicting lymph node metastasis in patients with superficial esophageal squamous cell carcinoma, Mod. Pathol. 25 (2012) 1364–1377.
[24] E. Ancona, S. Rampado, M. Cassaro, et al., Prediction of lymph node status in superficial esophageal carcinoma, Ann. Surg. Oncol. 15 (2008) 3278–3288.