Am J Transl Res 2022;14(8):5870-5882
www.ajtr.org /ISSN:1943-8141/AJTR0141638

Original Article

Predicting prognosis for patients with ESCC before surgery by SVMs ranking with nomogram analyses

Zhi Sheng Jiang[1*], Meng Qing Xu[2*], Zhuang Zhuang Cong[1], Li Wen Hu[1], Jing Luo[1], Yi Fei Diao[1], Yi Shen[1]

Departments of [1]Cardiothoracic Surgery, [2]Gastroenterology, Jinling Hospital, Nanjing 210002, Jiangsu, China. [*]Equal contributors.

Received January 9, 2022; Accepted May 19, 2022; Epub August 15, 2022; Published August 30, 2022

Abstract: Objective: A SVM predictive model consisting of preoperative tumor markers and inflammatory factors was established to explore its significance in evaluating the prognosis of patients with ESCC. Methods: Clinical data of 311 patients with ESCC who underwent surgery were collected and followed up until October 2019. Statistical software SPSS version 22.0, and R (version 3.6.1) were used to analyze the data. Results: In the Test, Val1 and Val2 groups, the sensitivity of preoperative optimal combination (SVM5) to predict the prognosis of patients with ESCC was 88.89%, 76.92%, and 73.68%, respectively. The specificity was 92.00%, 74.42%, and 78.00%, respectively. The sensitivity and specificity were not statistically different from those of SVM9 ($P > 0.05$), while the sensitivity of SVM9+5 for predicting the prognosis of patients with ESCC was 91.84%, 82.26%, and 80.36%, respectively. The specificity was 97.44%, 75.93%, and 78.00%, respectively. Its sensitivity and specificity were higher than those of SVM9 ($P < 0.001$). Conclusions: We used a nomogram to input the indicators in the SVM5 into the artificial intelligence program for patients with ESCC who have not yet developed an individualized plan. It can predict and evaluate the postoperative outcome of patients with ESCC with a sensitivity of 79.04%, specificity of 81.82%, PPV of 83.54%, NPV of 76.97%, and accuracy of 80.32%. For patients who have undergone surgery, we can enter the indicators in SVM9+5 into the artificial intelligence program.

Keywords: Esophageal squamous cell carcinoma (ESCC), tumor markers, inflammatory markers, SVM nomogram

Introduction

Esophageal cancer is very common in China and poses a threat to public health [1, 2]. Esophageal squamous cell carcinoma (ESCC) is the main histological subtype in China [3]. Many factors affect the prognosis of patients with ESCC, including the patient's performance status (ECOG), clinicopathological characteristics, tumor markers, and inflammatory nutritional indicators. However, the previous related studies were mostly single-factor studies. Therefore, exploring the factors affecting the prognosis and constructing a prognostic predictive model are undoubtedly of great significance for patients with ESCC.

Postoperative TNM staging is the most valuable index for evaluating the prognosis of patients with ESCC [4]. However, it is an index that can only be confirmed after surgery and can only provide a theoretical basis for postoperative treatment strategies. Thus, it is of little significance to formulate individualized treatment strategies before surgery in selected patients, especially for those with a poor physical condition or in whom performing surgery is difficult, since it is difficult to evaluate an accurate TNM staging for these patients. The exploration for indicators that can predict the prognosis before surgery has recently attracted great attentions, specifically regarding finding easy-to-detect indicators from patients' preoperative serum for assessment of tumor-related prognoses [5]. Preoperative tumor markers and tumor-related inflammation indicators are routinely detected and are easily available. Therefore, these markers have become the current research hotspots for evaluating and predicting the prognosis of tumors.

In recent years, there have been increasing numbers of studies on newly combined inflammatory indicators, such as GPS, NLR, LMR, P-

CRP, CPR, and other inflammatory indicators in multiple tumors [6-8]. However, there are few studies on the prognostic value of patients with ESCC, and its postoperative prognostic value needs to be further verified. Studies have found that simple tumor markers and inflammatory indicators have a certain evaluation value for the survival rate of patients with ESCC, but they lack high sensitivity. To date, assessing the prognosis of patients with ESCC after radical surgery by detecting preoperative serum indicators remains controversial [9].

Recently, support vector machines (SVMs), a new data-mining technology, have been used to predict tumor progression and clinical outcomes by integrating molecular markers and/or clinical features [10, 11]. For example, SVM analysis displayed moderately strong power in predicting regional lymph node metastasis preoperatively [12]. Thus, SVMs will likely continue to yield valuable insights into the accurate prediction of prognosis. In this study, we selected preoperative age, Eastern Cooperative Oncology Group (ECOG), tumor marker, inflammatory markers, and clinicopathological features using SVM learning models with a nomogram, to test the hypothesis that these markers may serve as substitutes for clinical features in predicting prognosis for patients with ESCC before surgery. The SVM model can continuously combine multiple detection indicators to evaluate its ability to predict the survival prognosis after surgery, until a model with a small number of combinations, high sensitivity, and specificity would be selected for clinical application. The establishment of an SVM model to assess postoperative risks of patients with ESCC will not only provide advice for the choice of a proper treatment strategy for the patient, but also assist in the formulation of individualized medical treatment after surgery by combining it with a nomogram.

Materials and methods

Patients and follow-up

Clinical data of 311 patients with ESCC who underwent radical resection surgery at Jinling Hospital from June 2014 to November 2016 were collected and followed up until October 2019. The data mainly included basic preoperative information such as sex; age; ECOG; BMI; SCC, CY211, AFP, CEA, CA199, and CA125; inflammation indicators including NLR, CRP, WBC, ALB, PALB, GPS, LMR, P-CRP, and CPR; and postoperative indices including tumor location, tumor size, INV, T, N, and TNM. Overall survival (OS) was defined from the date of surgery until the date of death or the date of the last follow-up.

Statistical analyses

Data analyses were performed using the statistical software package SPSS (version 3.6.1; http://www.R-project.org). Chi-square (χ^2) tests were used to analyze differences between SVM models for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Univariate and multivariate analyses of the relative prognostic importance of the parameters were performed using the Cox proportional hazards model. In addition, an SVM uses an implicit mapping of the input data into a high-dimensional feature space defined by a kernel function [13-15]. Then, the output results of the SVM model were subjected to ROC curve analysis. The Kaplan-Meier method was used to calculate and plot survival curves and re-verify the ability to evaluate the SVM model and identify the high- and low-mortality risk of patients with ESCC after surgery. A p value of < 0.05 was considered statistically significant.

Results

Baseline characteristics of the participants

A total of 311 patients with ESCC after surgery were included in this study. There were 241 males and 70 females, accounting for 77.5% and 22.5% of the total cases, respectively. Their ages ranged from 40 to 83 years, and the median age was 66 years. There were 42 (13.5%) well-differentiated patients, 179 (57.6%) moderately-differentiated, and 90 (28.9%) poorly-differentiated patients. There were 55 patients with INV. Regarding TNM staging, 63 cases (20.3%) were stage I, 115 (37.0%) were stage II, 111 (35.7%) were stage III, and 22 (7.1%) were stage IV (Table S1). Tumor markers (AFP, CEA, CA199, CA125, SCC, and CY211), inflammation-related indicators (CRP, NLR, WBC, ALB, PALB, GPS, LMR, P-CRP, and CPR), and other indicators were grouped by median (see Table S2 for details about specific values). On October 15, 2019, 143 (46.0%) patients with ESCC were still alive,

whereas 168 (54.0%) had died. In this study, the 1-year survival rate of patients with ESCC after surgery was 85.9%, the 3-year survival rate was 52.1%, and the 5-year survival rate was 45.3% (Figure S1).

Correlation analysis of various indices in patients with ESCC and their relationship with survival

Positive correlations were observed between SCC and tumor size (r = 0.339, P < 0.01), T (r = 0.396, P < 0.01), N (r = 0.220, P < 0.05), and TNM staging (r = 0.265, P < 0.01). Positive correlations were observed between CY211 and tumor size and TNM staging (all P < 0.05) (Table S3). Positive correlations were observed between ECOG and NLR (r = 0.297, P < 0.01), GPS, CPR, cell differentiation, tumor size, T, N, and TNM staging. Negative correlations were found between LMR and NLR, CRP, GPS, CPR, tumor size, and tumor invasion. Positive correlations were observed between NLR and CRP, GPS, P-CRP, CPR, cell death, tumor size, T, N, and TNM (r = 0.272, P < 0.01). Positive correlations were observed between CRP and GPS (rs = 0.645, P < 0.01), P-CRP (r = 0.874, P < 0.01), CPR (r = 0.845, P < 0.01), and tumor size (Table S4). Kaplan-Meier survival analysis showed that age, ECOG, SCC, CY211, NLR, CPR, LMR, P-CRP, CRP, GPS, size, INV, cell differentiation, T, N, and TNM staging had an impact on the prognostic survival of patients with ESCC after surgery (P < 0.05) (see Figures S1, S2). The survival curve of patients has been found to decrease, with worsening prognosis. With deeper invasion, metastasis, high ECOG score, low differentiation, age (< 66 years), tumor diameter (≥3 cm), and neurovascular invasion, the prognosis of patients with ESCC after surgery is reported to be poor (P < 0.05). Survival analysis revealed that high expression of NLR, SCC, CY211, CPR, P-CRP, CRP, and GPS correlated with poor prognosis in ESCC (P < 0.05) (Figure S3). Low expression of LMR suggests a poor prognosis for patients with ESCC after surgery (P = 0.015). TNM staging suggests a poor prognosis (P < 0.001).

Risk factors affecting the prognosis of patients with ESCC and their ability to predict survival

Univariate Cox regression analyses demonstrated that preoperative age, ECOG, SCC, CY211, NLR, CRP, GPS, LMR, P-CRP, CPR, and postoperative T, N, TNM, cell differentiation, tumor size, INV, and other indicators all affected the survival and prognosis of patients with ESCC after surgery (P < 0.05) (Table S5). Furthermore, multivariate analyses showed that preoperative age, ECOG, NLR, SCC, CY211, postoperative INV, cell differentiation, and TNM staging were independent factors affecting ESCC patient survival (P < 0.05). Among them, the risk of death in clinical stage III+IV was approximately 2.609 times that of those in stage I+II, with a 95% confidence interval (CI) of 1.862-3.656 (Table S6).

Our results showed that the AUC (0.760) of TNM was the largest, which was consistent with the prognostic value of TNM recognized by previous literature. The AUC of preoperative age predicting patient survival was 0.553, with a 95% CI of 0.489-0.618. The AUC of preoperative ECOG predicting patient survival was 0.738, with a 95% CI of 0.683-0.793. The AUC of preoperative tumor markers SCC and CY211 to predict the survival and prognosis of patients was 0.675 (0.616, 0.734) and 0.649 (0.588, 0.710), respectively. The AUC of pre-inflammatory indicators was as follows: NLR (AUC = 0.672 (0.613, 0.732), P < 0.001), CPR (AUC = 0.634 (0.572, 0.696), P < 0.001), LMR (AUC = 0.618 (0.555, 0.681), P < 0.001), P-CRP (AUC = 0.598 (0.535, 0.662), P = 0.003), CRP (AUC = 0.589 (0.525, 0.625), P = 0.007), and GPS (AUC = 0.573 (0.509, 0.637), P = 0.027) (see **Table 1**).

The SVM combined with ROC models in predicting prognosis of patients with ESCC

The SVM model for ESCC refers to predicting the survival probability of patients with ESCC. The performance of different SVM models for ESCC was evaluated by comparing the size of six indicators, namely, sensitivity, specificity, Youden index, PPV, NPV, and accuracy, of which sensitivity and specificity are more important than the rest. The indices used were sensitivity and specificity. An SVM model was used as an evaluation tool. Through the repeated combination of various conventional detection indicators, the optimal combination with the best predictive ability and the fewest indicators was selected. In the case of slight differences in

Establishment of the SVM model

Table 1. The AUC of ROC curves with preoperative and postoperative clinical markers in predicting postoperative survival prognosis in patients with ESCC

Clinical Indexes	AUC	95% *CI*	*P* value
Reference	0.500	/	/
Age	0.553	(0.489, 0.618)	0.104
ECOG	0.738	(0.683, 0.793)	< 0.001
NLR	0.672	(0.613, 0.732)	< 0.001
SCC	0.675	(0.616, 0.734)	< 0.001
CY211	0.649	(0.588, 0.710)	< 0.001
CRP	0.589	(0.525, 0.625)	0.007
GPS	0.573	(0.509, 0.637)	0.027
LMR	0.618	(0.555, 0.681)	< 0.001
P-CRP	0.598	(0.535, 0.662)	0.003
CPR	0.634	(0.572, 0.696)	< 0.001
Tumor Size	0.646	(0.583, 0.708)	< 0.001
Cell Differentiation	0.640	(0.579, 0.701)	< 0.001
INV	0.561	(0.497, 0.625)	0.065
T	0.686	(0.626, 0.747)	< 0.001
N	0.715	(0.658, 0.772)	< 0.001
TNM	0.760	(0.706, 0.814)	< 0.001

Note: ESCC: esophageal squamous cell carcinoma; ECOG: Eastern Cooperative Oncology Group (performance status); SCC: squamous cell carcinoma antigen; CY211: cytokeratin 19 fragment; GPS: Glasgow Prognostic Score; NLR: nneutrophil to lymphocyte ratio; CRP: C-reactive protein; PALB: prealbumin; LMR: lymphocyte to monocyte ratio; P-CRP: the platelet × C-reactive protein multiplier value; CPR: C-reactive protein to prealbumin ratio; INV: invasion of nerve or (and) vessels; T: tumor invasion; N: lymph node metastasis; TNM: TNM staging; AUC: area under the curve.

specificity, sensitivity was the most valuable evaluation index. As shown in **Table 2**, postoperative SVM9 predicted the survival of patients with ESCC. In the Test, Val1 and Val2 groups, the sensitivity was 81.63%, 79.03%, and 80.36%, and the specificity was 92.31%, 77.78%, and 72.00%, respectively. The AUC was 0.870, 0.784, and 0.762, respectively, and the AUC of Val1+Val2 was 0.773 (0.709, 0.837). The sensitivity of the preoperative combination of all indicators (SVM1) to predict the prognosis of patients with ESCC was 91.07%, 82.46%, and 77.78%, respectively. Compared with SVM9, the difference was not statistically significant (χ^2 = 0.772, P = 0.380). The specificity was 96.08%, 72.22%, and 68.42%, respectively. There was no statistically significant difference compared with SVM9 (χ^2 = 0.335, P = 0.563) (**Table 3**). The AUC was 0.936, 0.773, and 0.731, respectively, and the AUC of Val1+Val2 was 0.754 (0.685, 0.823). The sensitivity of the preoperative optimal combination (SVM5) for predicting the prognosis of patients with ESCC was 88.89%, 76.92%, and 73.68%, respectively. The specificity was 92.00%, 74.42%, 78.00%, respectively. Sensitivity and specificity were not statistically compared with those of SVM9. However, its sensitivity was higher than that of SVM8 (P < 0.001) and lower than that of SVM1 (P < 0.001). Its specificity was higher than that of SVM1 (P = 0.011), but the difference was not statistically different from SVM8 (P = 0.069). The AUCs were 0.904, 0.757, and 0.758, respectively, where the AUC of Val1+Val2 was 0.759 (0.692, 0.826). Based on this, all preoperative and postoperative indicators of patients with ESCC were included in the SVM model (i.e., SVM9+1). In the Val1 and Val2 groups, the sensitivities were 89.80%, 83.87%, and 76.79%, respectively. The specificity was 100%, 83.33%, and 78.00%, respectively. Its sensitivity and specificity were higher than those of SVM9 (P < 0.001), with AUC values of 0.949, 0.836, and 0.774, respectively. The AUC of Val1+Val2 was 0.806 (0.746, 0.867). The sensitivity of SVM9+5 for predicting the prognosis of patients with ESCC was 91.84%, 82.26%, and 80.36%, respectively. The specificity was 97.44%, 75.93%, and 78.00%, respectively. Its sensitivity and specificity were higher than that of SVM9 (P < 0.001); its sensitivity was lower than that of SVM9+ECOG and higher than that of SVM9+1 (all P < 0.001). Its specificity was lower than that of SVM9+1 and higher than that of SVM9+ECOG (all P < 0.001) (**Table 3**). The AUC was 0.946, 0.791, and 0.792, respectively, and the AUC of Val1+Val2 was 0.791 (0.729, 0.854) (see **Table 4**). Our results showed that there was no statistical difference between SVM5 and SVM9 in terms of sensitivity, specificity, PPV, and NPV. In addition, the AUC of SVM5 (Test+Val1+Val2) (0.804 (0.753, 0.855)) and SVM9 (0.800 (0.748, 0.851)) were also equivalent and higher than that of TNM staging. SVM9+5 had a slightly higher sensitivity than that of SVM9+1 and a slightly lower specificity than that of SVM9+1. There were no statistical differences in PVV, NPV, ACC, etc. and the AUC of SVM9+5 (0.835 (0.787, 0.883)) was slightly lower than that of SVM9+1 (0.846 (0.800, 0.893)), which was much higher than that of TNM staging (see **Figure 1**).

Table 2. Comparisons of sensitivity, specificity, Youden's index, PPV, NPV and accuracy among different SVM models

Variable combinations	Sensitivity (%)			Specificity (%)			Yuedens' Index (%)			PPV (%)			NPV (%)			Accuracy (%)		
	Test	Val1	Val2	Test	Val1	Val2	Test	Val1	Val2	Test	Val1	Val2	Test	Val1	Val2	Test	Val1	Val2
Before Surgery																		
SVM1 (107, 111, 92)	91.07	82.46	77.78	96.08	72.22	68.42	87.15	54.68	46.20	96.23	75.81	77.78	90.74	79.59	68.42	93.46	77.48	73.91
SVM2 (95, 108, 107)	86.67	75.38	68.42	94.00	67.44	70.00	80.67	42.82	38.42	92.86	77.78	72.22	88.68	64.44	66.04	90.53	72.22	69.16
SVM3 (95, 108, 107)	88.89	75.38	71.93	90.00	72.09	76.00	78.89	47.47	47.93	88.89	80.33	77.36	90.00	65.96	70.37	89.47	74.07	73.83
SVM4 (106, 103, 101)	77.97	72.22	75.93	91.49	81.63	72.34	69.46	53.85	48.27	92.00	81.25	75.93	76.79	72.73	72.34	83.96	76.69	74.26
SVM5 (95, 108, 107)	88.89	76.92	73.68	92.00	74.42	78.00	80.89	51.34	51.68	90.91	81.97	79.25	90.20	68.09	72.22	90.53	75.93	75.70
SVM6 (122, 84, 104)	84.85	80.00	72.55	91.07	64.71	67.92	75.92	44.71	40.47	91.80	76.92	68.52	83.61	68.75	72.00	87.70	73.81	70.19
SVM7 (122, 84, 104)	86.36	82.00	70.59	94.64	73.53	60.38	81.00	55.53	30.97	95.00	82.00	63.16	85.48	73.53	68.09	90.16	78.57	65.38
SVM8 (108, 90, 112)	78.33	71.11	67.74	95.83	75.56	78.00	74.16	46.67	45.74	95.92	74.42	79.25	77.97	72.34	66.10	86.11	73.33	72.32
After Surgery																		
SVM Model 9 (88, 116, 106)	81.63	79.03	80.36	92.31	77.78	72.00	73.94	56.81	52.36	93.02	80.33	76.27	80.00	76.36	76.60	86.36	78.45	76.42
All markers																		
SVM9+1 (88, 116, 106)	89.80	83.87	76.79	100.00	83.33	78.00	89.80	67.20	54.79	100.00	85.25	79.63	88.64	81.82	75.00	94.32	83.62	77.36
SVM9+2 (88, 116, 106)	89.80	83.87	76.79	100.00	83.33	72.00	89.80	67.20	48.79	100.00	85.25	75.44	88.64	81.82	73.47	94.32	83.62	74.53
SVM9+3 (88, 116, 106)	91.84	82.26	80.36	94.87	72.22	76.00	86.71	54.48	56.36	95.74	77.27	78.95	90.24	78.00	77.55	93.18	77.59	78.30
SVM9+4 (88, 116, 106)	91.84	82.26	80.36	94.87	75.93	78.00	86.71	58.18	58.36	95.74	79.69	80.36	90.24	78.85	78.00	93.18	79.31	79.25
SVM9+5 (88, 116, 106)	91.84	82.26	80.36	97.44	75.93	78.00	89.27	58.18	58.36	97.83	79.69	80.36	90.48	78.85	78.00	94.32	79.31	79.25
SVM9+6 (88, 116, 106)	89.80	82.26	76.79	100.00	79.63	72.00	89.80	61.89	48.79	100.00	82.26	75.44	88.64	79.63	73.47	94.32	81.03	74.53
SVM9+7 (122, 84, 104)	95.45	82.00	78.43	94.64	70.59	69.81	90.10	52.59	48.24	95.45	80.39	71.43	94.64	72.73	77.08	95.08	77.38	74.04
SVM9+8 (88, 116, 106)	91.84	80.65	78.57	92.31	72.22	68.00	84.14	52.87	46.57	93.75	76.92	73.33	90.00	76.47	73.91	92.05	76.72	73.58
SVM9+ECOG, NLR (122, 84, 104)	93.94	76.00	74.51	94.64	70.59	69.81	88.58	46.59	44.32	95.38	79.17	70.37	92.98	66.67	74.00	94.26	73.81	72.12
SVM9+ECOG (88, 116, 106)	93.88	82.26	85.71	87.18	68.52	64.00	81.06	50.78	49.71	90.20	75.00	72.73	91.89	77.08	80.00	90.91	75.86	75.47

Note: test group, validation group 1, validation group 2 were randomly divided into three independent SVM model groups. The test group was also known as the learning group, then two validation groups were used to recognize and verify the precise of predicting the prognosis for ESCC. Data are presented as percentages. PPV: positive predictive value; NPV: negative predictive value; Val: validation; SVM1: all markers before surgery (ECOG, NLR, SCC, CY211, CPR, LMR, P-CRP, CRP, GPS, age); SVM2: SVM Model 1 (omit age); SVM3: SVM Model 2 (omit GPS); SVM4: SVM Model 3 (omit CRP); SVM5: SVM Model 4 (omit P-CRP): ECOG, NLR, SCC, CY211, CPR, LMR; SVM6: SVM Model 5 (omit LMR): ECOG, NLR, SCC, CY211, CPR; SVM7: SVM Model 6 (omit CPR): ECOG, NLR, SCC, CY211; SVM8: SVM Model 7 (omit Cy211): ECOG, NLR, SCC; SVM9: TNM, Diff, Size, INV.

Table 3. Comparisons among different marker combination obtained before surgery, after surgery and all markers according to sensitivity, specificity, PPV, NPV and accuracy in predicting prognoses of patients with ESCC

Variable combinations	Test+Validation group 1+Validation group 2 (n = 310)												Accuracy (%)
	Sensitivity (%)	χ^2	*P* value	Specificity (%)	χ^2	*P* value	PPV (%)	χ^2	*P* value	NPV (%)	χ^2	*P* value	
SVM1	83.83	0.772[#]	0.380	79.72	0.335[#]	0.563	82.84	0.023[#]	0.880	80.85	0.475[#]	0.491	81.94
		34.304[$1]	< 0.001		6.498[$1]	0.011		0.029[$1]	0.865		0.659[$1]	0.417	
SVM5	79.04	0.002[#]	0.968	81.82	0.471[#]	0.493	83.54	0.101[#]	0.751	76.97	0.014[#]	0.905	80.32
		37.345[$2]	< 0.001		3.311[$2]	0.069		0.001[$2]	0.982		0.979[$2]	0.322	
SVM8	72.46	0.690[#]	0.406	83.22	0.005[#]	0.941	83.45	0.083[#]	0.774	72.12	1.212[#]	0.271	77.42
		9.535[$3]	0.002		4.087[$3]	0.043		0.021[$3]	0.886		3.190[$3]	0.074	

SVM9	80.24#			79.72#			82.21#			77.55#			80.00#
SVM9+1	83.23	19.401#	< 0.001	86.01	19.925#	< 0.001	87.42	1.695#	0.193	81.46	0.698#	0.403	84.52
		48.159$1	< 0.001		48.788$1	< 0.001		0.419$1	0.517		0.068$1	0.795	
SVM9+5	84.43	17.760#	< 0.001	82.52	18.855#	< 0.001	84.94	0.447#	0.504	81.94	0.327#	0.567	83.55
		40.424$2	< 0.001		47.204$2	< 0.001		2.496$2	0.114		0.201$2	0.654	
SVM9+ECOG	86.63	34.033#	< 0.001	72.03	61.231#	< 0.001	78.38	0.800#	0.371	82.40	0.985#	0.321	80.00
		12.669$3	< 0.001		37.533$3	< 0.001		4.856$3	0.028		0.068$3	0.795	

SVM1: all markers before surgery (ECOG, NLR, SCC, Cy211, CPR, LMR, P-CRP, CRP, GPS, age); SVM5: ECOG, NLR, SCC, Cy211, CPR, LMR; SVM8: ECOG, NLR, SCC; SVM9: TNM, diff, size, INV; #: χ^2 test was used in comparisons among markers obtained SVM1, SVM5, SVM8 and SVM9, and among markers of SVM9+1, SVM9+5, SVM9+ECOG and SVM9, respectively; $1: χ^2 test was used in comparisons between markers of SVM5 and SVM1, and between SVM9+5 and SVM9+1; $2: χ^2 test was used in comparisons between markers of SVM8 and SVM5, and between SVM9+ECOG and SVM9+5; $3: χ^2 test was used in comparisons between markers of SVM8 and SVM1, and between SVM9+ECOG and SVM9+1; *P* value: corresponding comparisons.

Table 4. The SVM combined with ROC models in predicting prognosis of patients with ESCC by using testing data, validation data separately

Combinations	Test			Val1			Val2			Val1+2		
	AUC	95% CI	*P* value	AUC	95% CI	*P* value	AUC	95% CI	*P* value	AUC	95% CI	*P* value
Before Surgery												
SVM Model 1	0.936	(0.882, 0.989)	< 0.001	0.773	(0.683, 0.864)	< 0.001	0.731	(0.623, 0.839)	< 0.001	0.754	(0.685, 0.823)	< 0.001
SVM Model 2	0.903	(0.834, 0.973)	< 0.001	0.714	(0.612, 0.816)	< 0.001	0.692	(0.590, 0.794)	< 0.001	0.705	(0.633, 0.776)	< 0.001
SVM Model 3	0.894	(0.823, 0.966)	< 0.001	0.737	(0.639, 0.836)	< 0.001	0.740	(0.643, 0.836)	< 0.001	0.740	(0.671, 0.808)	< 0.001
SVM Model 4	0.847	(0.769, 0.926)	< 0.001	0.769	(0.675, 0.863)	< 0.001	0.741	(0.642, 0.841)	< 0.001	0.756	(0.688, 0.824)	< 0.001
SVM Model 5	0.904	(0.835, 0.973)	< 0.001	0.757	(0.660, 0.853)	< 0.001	0.758	(0.664, 0.853)	< 0.001	0.759	(0.692, 0.826)	< 0.001
SVM Model 6	0.880	(0.813, 0.946)	< 0.001	0.724	(0.609, 0.839)	< 0.001	0.702	(0.600, 0.804)	< 0.001	0.715	(0.639, 0.790)	< 0.001
SVM Model 7	0.905	(0.845, 0.965)	< 0.001	0.778	(0.671, 0.884)	< 0.001	0.655	(0.549, 0.761)	0.007	0.709	(0.633, 0.784)	0.007
SVM Model 8	0.871	(0.799, 0.942)	< 0.001	0.733	(0.627, 0.839)	< 0.001	0.729	(0.633, 0.824)	< 0.001	0.730	(0.659, 0.801)	< 0.001
After Surgery												
SVM Model 9	0.870	(0.789, 0.950)	< 0.001	0.784	(0.697, 0.871)	< 0.001	0.762	(0.667, 0.856)	< 0.001	0.773	(0.709, 0.837)	< 0.001
Preoperative and postoperative markers												
SVM Model 9+1	0.949	(0.898, 1.000)	< 0.001	0.836	(0.758, 0.914)	< 0.001	0.774	(0.681, 0.866)	< 0.001	0.806	(0.746, 0.867)	< 0.001
SVM Model 9+2	0.949	(0.898, 1.000)	< 0.001	0.836	(0.758, 0.914)	< 0.001	0.744	(0.647, 0.841)	< 0.001	0.792	(0.730, 0.854)	< 0.001
SVM Model 9+3	0.934	(0.873, 0.994)	< 0.001	0.772	(0.683, 0.862)	< 0.001	0.782	(0.690, 0.873)	< 0.001	0.777	(0.713, 0.841)	< 0.001
SVM Model 9+4	0.934	(0.873, 0.994)	< 0.001	0.791	(0.705, 0.877)	< 0.001	0.792	(0.702, 0.882)	< 0.001	0.791	(0.729, 0.854)	< 0.001
SVM Model 9+5	0.946	(0.893, 1.000)	< 0.001	0.791	(0.705, 0.877)	< 0.001	0.792	(0.702, 0.882)	< 0.001	0.791	(0.729, 0.854)	< 0.001
SVM Model 9+6	0.949	(0.898, 1.000)	< 0.001	0.809	(0.726, 0.893)	< 0.001	0.744	(0.647, 0.841)	< 0.001	0.778	(0.715, 0.842)	< 0.001
SVM Model 9+7	0.950	(0.906, 0.995)	< 0.001	0.763	(0.654, 0.872)	< 0.001	0.741	(0.644, 0.839)	< 0.001	0.752	(0.679, 0.824)	< 0.001
SVM Model 9+8	0.921	(0.855, 0.987)	< 0.001	0.764	(0.674, 0.854)	< 0.001	0.733	(0.635, 0.831)	< 0.001	0.749	(0.683, 0.816)	< 0.001
SVM Model 9+ECOG, NLR	0.943	(0.895, 0.991)	< 0.001	0.733	(0.620, 0.846)	< 0.001	0.722	(0.622, 0.822)	< 0.001	0.727	(0.653, 0.801)	< 0.001
SVM Model 9+ECOG	0.905	(0.833, 0.978)	< 0.001	0.754	(0.662, 0.846)	< 0.001	0.749	(0.652, 0.845)	< 0.001	0.751	(0.685, 0.818)	< 0.001

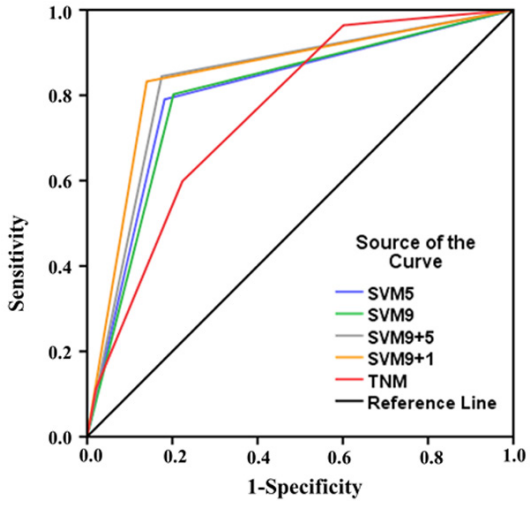

Figure 1. The ROC curve and the AUC to quantify the impact weights (or powers) of the SVM model. According to the move of cut-off point, the sensitivity and false negative rate (1-specificity) being studied were obtained, then we plotted each point by vertical axis (sensitivity) and horizontal axis (1-specificity) to generate a ROC curve and calculated the area (the larger area, the higher diagnostic value) under the curve, in order to gain further prediction analysis through SVM model. Abbreviations: SVM5: ECOG, NLR, SCC, CY211, CPR, LMR; SVM9: TNM, Diff, Size, INV; SVM1: All markers before surgery (ECOG, NLR, SCC, CY211, CPR, LMR, P-CRP, CRP, GPS, age).

Multi-factor combined detection predicts the advantages of postoperative survival in ESCC

This study proposed for the first time the establishment and verification of an early warning model for the survival of patients with ESCC before and after surgery. It was found that the combination of preoperative routine clinical indicators could be used to evaluate the postoperative mortality risk of patients with ESCC using the SVM model. The mortality risk of patients with ESCC who had undergone surgery could be evaluated using the SVM model combining various indicators before and after surgery. As shown in **Figure 2**, the predicted low-risk group (PLR) represented the low-mortality risk group; that is, this group of patients had a long survival time after surgery. The predicted high-risk group (PHR) represented the high mortality risk group; that is, this group of patients had a short survival time after surgery. Kaplan-Meier survival analysis showed that in the preoperative SVM5 test group, the average postoperative survival time of patients with ESCC in the low-risk group (51 cases, 53.5 months) was longer than that of the high-risk group (44 cases, 25 months), and their 3-year cumulative survival rates were 92.2% and 20.5%, respectively ($P < 0.001$), with a difference of 71.7%. Verification groups 1 and 2 were completely independent of the test group. In Validation group 1, the average postoperative survival time of patients with ESCC in PLR (47 cases, 49 months) was longer than that of PHR (61 cases, 19 months), their 3-year cumulative survival rates were 76.6% and 24.6%, respectively ($P < 0.001$), and the difference was 52.0%. In Validation group 2, the average postoperative survival time of patients with ESCC in PLR (53 cases, 50.5 months) was longer than that of PHR (54 cases, 22 months), and their 3-year cumulative survival rates were 75.9% and 26.4%, respectively ($P < 0.001$), with a difference of 49.5%. In the SVM9 test group, the average postoperative survival time of patients with ESCC in PLR (45 cases, 49 months) was longer than that of PHR (43 cases, 17 months), and the 3-year cumulative survival rates were 82.2% and 23.3%, respectively ($P < 0.001$), with a difference of 58.9%. In the SVM9+5 test group, the average postoperative survival time of patients with ESCC in PLR (42 cases, 50.5 months) was longer than that of PHR (46 cases, 17.3 months), and their 3-year cumulative survival rates were 92.9% and 17.4%, respectively ($P < 0.001$), with a difference of 75.5%. The same trend was shown in the validation group (**Figure 2**). In addition, there were significant differences in the distribution of risk factors obtained by the SVM model between the PHR and PLR group through the Heatmap, and the result again indicated that the model exhibited good predictive ability (**Figure 3**).

Establishment of a nomogram for predicting OS

We introduced a nomogram based on patients' basic clinical features, preoperative tumor markers, and inflammation indicators. The nomogram (SVM5) predicted the OS among ESCC patients (**Figure 4**) with a reliable performance (AUC of 0.804). It can predict and evaluate the postoperative outcome of patients with ESCC with a sensitivity of 79.04%, specificity of 81.82%, PPV of 83.54%, NPV of 76.97%, and accuracy of 80.32%. The nomogram (SVM9+5) predicts OS among patients with ESCC (**Figure 5**) with a reliable performance (AUC of 0.835).

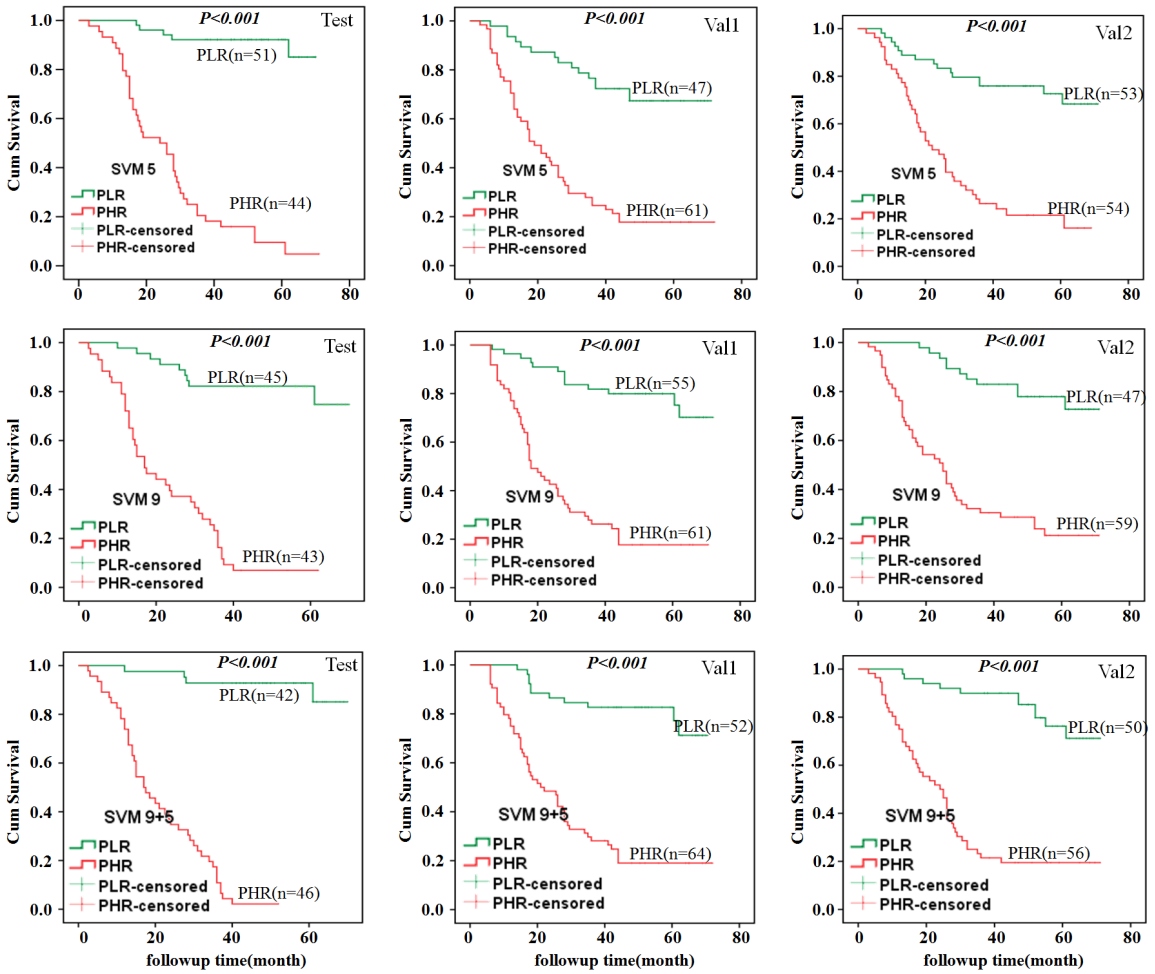

Figure 2. Application of the support vector machines (SVM) 5, SVM9, and SVM9+5 models to refine the assessment of risk in patients with ESCC. The predicted low-risk group (PLR) represented the low-mortality risk group; that is, this group of patients had a long survival time after surgery. The predicted high-risk group (PHR) represented the high mortality risk group; that is, this group of patients had a short survival time after surgery. We randomly divided the 311 patients into test, Val1, and Val2 groups, and then each group was divided into a high-risk group and a low-risk group. The test group was used to predict the high-risk (PHR) and low-risk groups (PLR) of patients with ESCC by learning different combination indices. In the testing set, the features of the selected variables in each patient were input into the SVM model. After the completion of the test process, the independent Val1 and Val2 groups were given the same indexes as the test group to predict the high- and low-risk groups for patients. Abbreviations: PLR, predicted low-risk group; PHR, predicted high-risk group.

It can predict and evaluate the postoperative outcome of patients with ESCC with sufficient sensitivity (84.43%), specificity (82.52%), PPV (84.94%), NPV (81.94%), and accuracy (83.55%). This nomogram may be used to comprehend treatment and follow-up plans for patients with ESCC.

Discussion

Postoperative survival time had a significant effect on the prognosis evaluation system. In recent years, although the incidence of ESCC has decreased, the survival rate remains poor. TNM staging is currently a well-recognized evaluation index for predicting ESCC prognosis [16]. However, TNM staging is an indicator that can only be confirmed after surgery, which is only suitable for postoperative patients and has little significance for preoperative planning.

Systemic inflammation reactions and malnutrition are some reasons for the poor prognosis of patients with ESCC [17, 18]. In recent years,

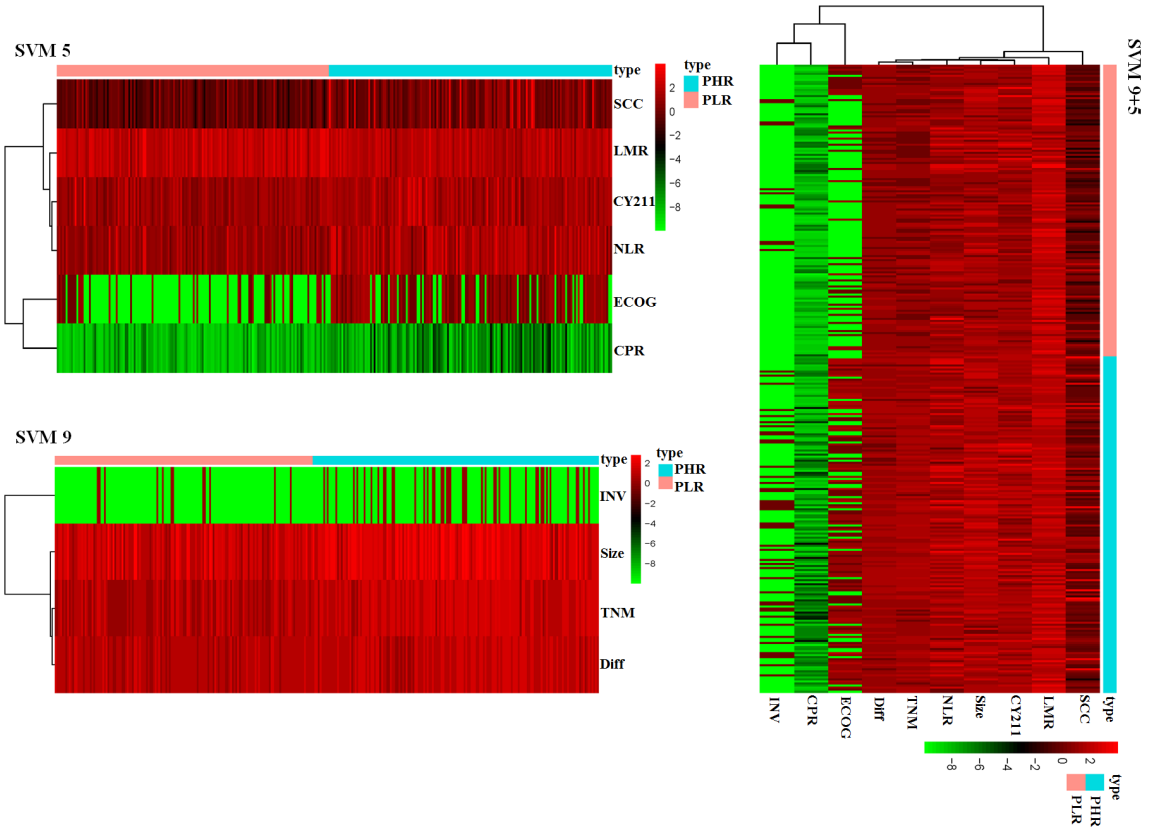

Figure 3. Heatmap of high and low distribution profiles of risk factors affecting the prognosis of patients with ESCC.

many studies have been conducted on newly combined inflammation indicators, such as GPS, P-CRP, and CPR [6-8]. The inflammation index GPS is the ratio of CRP to ALB, and CPR is the ratio of CRP and PALB [19]. It can simultaneously reflect the inflammatory response and nutritional status. It is an easy-to-measure and valuable preoperative prognostic indicator [20, 21]. However, it is rarely found in prognosis studies of patients with ESCC after surgery, and its prognostic value for patients needs to be further verified. Our study found that the inflammation indicators NLR, LMR, P-CRP, CPR, and GPS are risk factors that affect the prognosis of patients with ESCC, in which NLR is an independent risk factor.

Previous studies have shown that a higher SCC level in patients with ESCC before treatment indicates a relatively late stage of the tumor and a poor prognosis [22, 23], which is consistent with our research results. CY211 is a member of the keratin family [24]. It is overexpressed in many malignant tumors [25-27]. Our study found that elevated levels of SCC and CY211 before surgery indicate a poor prognosis for patients with ESCC. Many studies support these results [28, 29].

The ECOG performance status score calculated for the same patients was compared in terms of survival prognosis. ECOG has almost the vital power for predicting the postoperative prognosis of patients with ESCC [30]. At the same time, it also has an impact on the survival of lung cancer [20], liver cancer, and other cancers. We screened out the prognostic factors of ESCC based on patient age, physical fitness score, BMI, preoperative tumor markers and inflammatory factors, postoperative TNM staging, and other indicators, which were included in the analysis. Furthermore, we established an applicable convenient artificial intelligence model for clinical application by SVM combined with nomogram analysis.

Due to the lack of high sensitivity, it may be difficult to use only a single positive tumor marker

Establishment of the SVM model

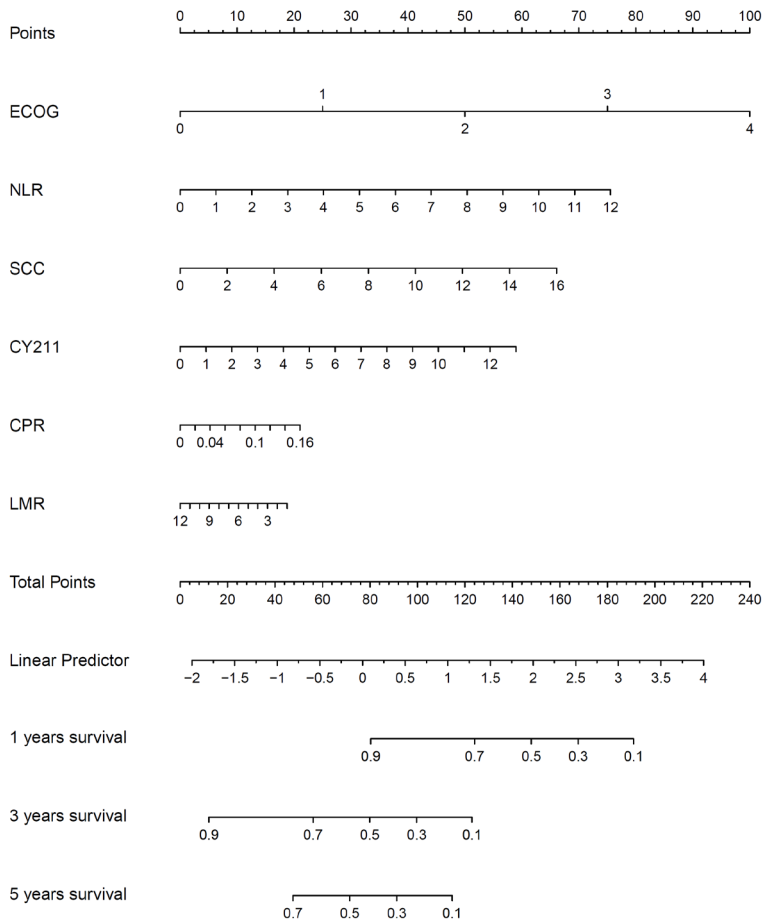

Figure 4. The nomogram (SVM5) predicted individual patient-level 1, 3, 5-year overall survival based on preoperative clinical index. Vertical lines were drawn from the correct status of each prognostic factor to the top axis (points). After the addition of all the points, a vertical line was drawn from the "total points" axis to the bottom axes. This helps in the conversion into a 1-, 3-, and 5-year survival probability. Abbreviations: ECOG, Eastern Cooperative Oncology Group; SCC, squamous cell carcinoma antigen; CY211, cytokeratin 19 fragment; NLR, neutrophil-to-lymphocyte ratio; LMR, lymphocyte to monocyte ratio; CPR, C-reactive protein to prealbumin ratio.

or inflammatory index to evaluate prognosis in the clinic [28]. In order to assess the prognosis of patients more accurately with ESCC, we combined basic patient characteristics (age, BMI, and ECOG score), serum tumor markers, and inflammation indices to improve the sensitivity, specificity, and accuracy of prediction and judgment. The area under the ROC curve reflects the value of various factors in the prognosis of patients with ESCC. TNM staging is a well-recognized standard [4]. The results of our study showed the AUC of the single index, whether tumor markers or inflammatory factors, is smaller than the AUC of TNM staging, indicating that a single index has a certain predictive value, but lacks sensitivity, and the predictive ability is not ideal.

The current results show that combining multiple indices that have an impact on the prognosis of ESCC can greatly improve the sensitivity and specificity of the SVM and enhance its predictive ability. The preoperative best combination (SVM5) obtained by combining preoperative indicators that have an impact on the prognosis of ESCC has an area under the line of 0.804, which far exceeds the predictive power of TNM staging [4] (AUC = 0.760), which is equivalent to the AUC of the postoperative combination SVM9 (AUC = 0.800), and its sensitivity and specificity for predicting postoperative survival of ESCC have no statistical difference compared with those of SVM9 ($P > 0.05$). This also confirms the predicted value of the SVM model from the side. For patients who have already undergone surgery, we can combine preoperative and postoperative indicators that have an impact on the prognosis of ESCC (SVM9+5) to predict the survival time of patients after surgery, with an area under the line of 0.835, which is only slightly lower than the combination of all indicators before and after surgery (SVM9+1), far more than the postoperative combination SVM9. Moreover, its sensitivity and specificity for predicting postoperative survival of ESCC were higher than those of SVM9 ($P < 0.001$); at the same time, its sensitivity was higher than that of SVM9+1, and specificity was lower than that of SVM9+1 (all $P < 0.001$). Moreover, through survival analysis, we performed survivability discrimination verification on the PLR and PHR groups obtained by the SVM model and found that the survival time of the PLR group of both SVM5 and SVM9+5 was much longer than that of the PHR

Establishment of the SVM model

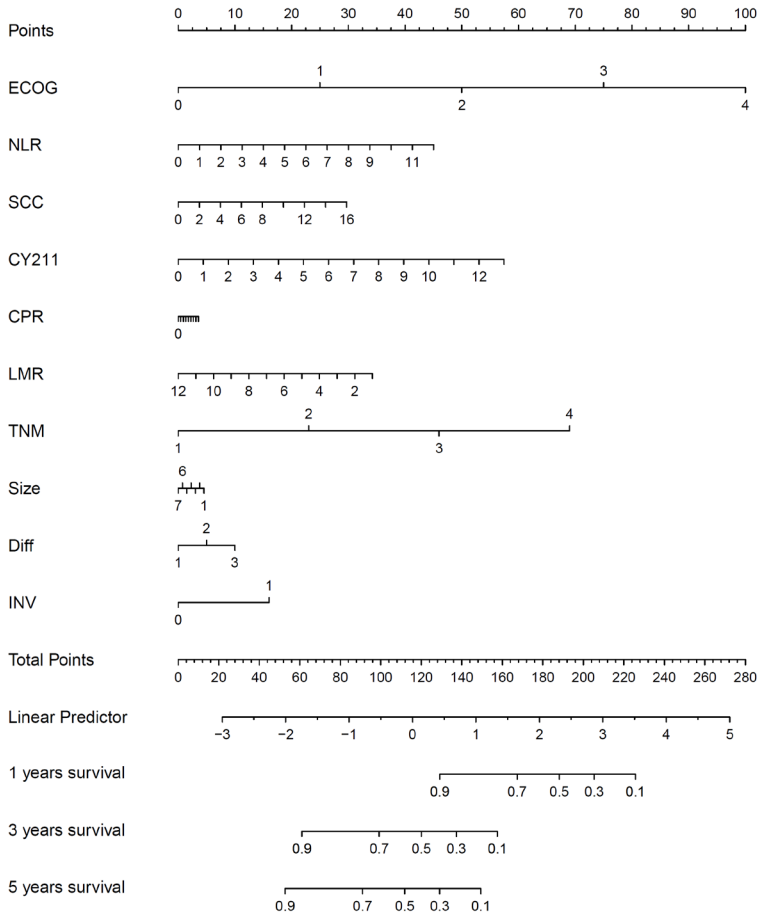

Figure 5. The nomogram (SVM9+5) predicted individual patient-level 1, 3, 5-year overall survival based on preoperative and postoperative clinical index. Vertical lines were drawn from the correct status of each prognostic factor to the top axis (points). After the addition of all the points, a vertical line was drawn from the "total points" axis to the bottom axes. This helps in the conversion into a 1-, 3-, and 5-year survival probability. Abbreviations: TNM, tumor node metastasis; INV, invasion of nerve or (and) vessels; Size, tumor size; Diff, cell differentiation.

group, and the difference in cumulative survival rate was large, so we verified the reliability of the SVM model again. We used a nomogram to input the indicators in the SVM5 combination into the artificial intelligence program for patients with ESCC who have not yet developed an individualized plan. It can predict and evaluate the postoperative outcome of patients with ESCC with a sensitivity of 79.04%, specificity of 81.82%, PPV of 83.54%, NPV of 76.97%, and accuracy of 80.32%. For patients who have undergone surgery, we can enter the indicators in SVM9+5 into the artificial intelligence program, which can predict and evaluate the postoperative outcomes with a sensitivity of 84.43%, specificity 82.52%, PPV of 84.94%, NPV of 81.94%, and accuracy of 83.55%. In today's information age, through the establishment of new predictive models, medical research and clinical data are combined at the level of patient diagnosis and treatment to promote the progress and development of precision medicine research [11].

The data for this study were collected from a single institution database, and there is a certain degree of selection bias, but the SVM model [13-15] can randomly separate them into a test group and two independent verification groups, which can reduce the selection bias and ensure the reliability of the data. In addition, we used the SVM model and ROC curve to qualitatively and quantitatively evaluate the prediction model and used the nomogram to evaluate the survival outcome of patients with ESCC. Additionally, different treatment plans were adopted according to the different predicted high- and low-risk groups. Based on the differences in the high- and low-risk groups of patients with ESCC, we selected individualized medical treatments such as personalized surgical planning (or appropriate surgery), the best dose and time of radiotherapy and chemotherapy, and appropriate follow-up intervals. It is worth noting that overtreatment should be avoided in the high-risk group, and undertreatment should be avoided in the low-risk group. The establishment of this SVM model not only has prompt significance for the surgical strategies for patients, but also assists in the formulation of individualized medical programs such as the best postoperative radiotherapy and chemotherapy and a reasonable period of postoperative follow-up [10-12]. In the future, we will conduct multi-center, large-sample, prospective studies to validate our results.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 8200-2454 and 81702444) and the Natural Science Foundation of Jiangsu Province (BK201812-39). All data generated or analyzed during this study are included in this article and its supplementary information files.

Written informed consent was obtained from all patients, and protocols for this study were approved by the Ethics Committee of Jinling Hospital (The certificate No. 2018NZKY-021-03).

Disclosure of conflict of interest

None.

Abbreviations

AJCC, American Joint Committee on Cancer; TNM, tumor node metastasis; T, tumor invasion; N, lymph node metastasis; M, distant metastasis; ESCC, esophageal squamous cell carcinoma; SVM, Support Vector Machine; ECOG, Eastern Cooperative Oncology Group; EAC, esophageal adenocarcinoma; SCC, squamous cell carcinoma antigen; CY211, cytokeratin 19 fragment; AFP, alpha-fetoprotein; CEA, carcinoma embryonic antigen; CA199, carbohydrate antigen 19-9; CA125, carbohydrate antigen 12-5; NLR, Neutrophil-to-Lymphocyte ratio; CRP, C-reactive protein; WBC, white blood cell; ALB, albumin; PALB, prealbumin; GPS, Glasgow prognostic score; LMR, lymphocyte to monocyte ratio; P-CRP, the platelet × C-reactive protein multiplier value; CPR, C-reactive protein to prealbumin ratio; INV, invasion of nerve or (and) vessels; OS, overall survival; r, coefficient of correlation; PLR, predicted low-risk group; PHR, predicted high-risk group; HR, hazard ratio; CI, confidence interval; AUC, area under the curve; ROC, receiver operating characteristic.

Address correspondence to: Dr. Yi Shen, Department of Cardiothoracic Surgery, Jinling Hospital, Nanjing 210002, Jiangsu, China. E-mail: dryishen@nju.edu.cn

References

[1] International Agency for Research on Cancer (IARC), GLOBOCAN 2020: https://gco.iarc.fr/today/online-analysis-table.

[2] Hou H, Meng Z, Zhao X, Ding G, Sun M, Wang W and Wang Y. Survival of esophageal cancer in China: a pooled analysis on hospital-based studies from 2000 to 2018. Front Oncol 2019; 9: 548.

[3] Mao YS, He J and Cheng GY. Current status of surgical management of esophageal cancer in China and the future strategy. Zhonghua Zhong Liu Za Zhi 2010; 32: 401-4.

[4] Nomura M, Shitara K, Kodaira T, Hatooka S, Mizota A, Kondoh C, Yokota T, Takahari D, Ura T and Muro K. Prognostic impact of the 6th and 7th American Joint Committee on Cancer TNM staging systems on esophageal cancer patients treated with chemoradiotherapy. Int J Radiat Oncol Biol Phys 2012; 82: 946-52.

[5] Matsunaga T, Miyata H, Sugimura K, Motoori M, Asukai K, Yanagimoto Y, Yamamoto K, Akita H, Nishimura J, Wada H, Takahashi H, Yasui M, Omori T, Ohue M, Fujiwara Y and Yano M. Prognostic significance of C-reactive protein-to-prealbumin ratio in patients with esophageal cancer. Yonago Acta Med 2019; 63: 8-19.

[6] Yodying H, Matsuda A, Miyashita M, Matsumoto S, Sakurazawa N, Yamada M and Uchida E. Prognostic significance of neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio in oncologic outcomes of esophageal cancer: a systematic review and meta-analysis. Ann Surg Oncol 2016; 23: 646-54.

[7] Hirahara N, Tajima Y, Fujii Y, Kaji S, Yamamoto T, Hyakudomi R, Taniura T, Miyazaki Y, Kishi T and Kawabata Y. Preoperative prognostic nutritional index predicts long-term surgical outcomes in patients with esophageal squamous cell carcinoma. World J Surg 2018; 42: 2199-2208.

[8] Lu J, Xu BB, Zheng ZF, Xie JW, Wang JB, Lin JX, Chen QY, Cao LL, Lin M, Tu RH, Huang ZN, Zheng CH, Huang CM and Li P. CRP/prealbumin, a novel inflammatory index for predicting recurrence after radical resection in gastric cancer patients: post hoc analysis of a randomized phase III trial. Gastric Cancer 2019; 22: 536-545.

[9] Ma L, Xie XW, Wang HY, Ma LY and Wen ZG. Clinical evaluation of tumor markers for diagnosis in patients with non-small cell lung cancer in China. Asian Pac J Cancer Prev 2015; 16: 4891-4.

[10] Zhu ZH, Sun BY, Ma Y, Shao JY, Long H, Zhang X, Fu JH, Zhang LJ, Su XD, Wu QL, Ling P, Chen M, Xie ZM, Hu Y and Rong TH. Three immunomarker support vector machines-based prognostic classifiers for stage IB non-small-cell lung cancer. J Clin Oncol 2009; 27: 1091-9.

[11] Sheng-Hua YU and Gong SH. Combination forecast and its application based on lasso

and support vector machine. Journal of Quantitative Economics 2016; 33: 46-49.
[12] Fan XJ, Wan XB, Huang Y, Cai HM, Fu XH, Yang ZL, Chen DK, Song SX, Wu PH, Liu Q, Wang L and Wang JP. Epithelial-mesenchymal transition biomarkers and support vector machine guided model in preoperatively predicting regional lymph node metastasis for rectal cancer. Br J Cancer 2012; 106: 1735-41.
[13] Cherkassky V. The nature of statistical learning theory~. IEEE Trans Neural Netw 1997; 8: 1564.
[14] Scholkopf B and Smola A. Learning with Kernels. Jouran of The Korean Society of Civil Engineers 2011; 59: 86-86.
[15] Alexandros Karatzoglou, Alex Smola, Kurt Hornik and Achim Zeileis. kernlab - an S4 package for kernel methods in R. Journal of Statistical Software 2014; 11: 1-20.
[16] Zhang D, Zheng Y, Wang Z, Huang Q, Cao X, Wang F and Liu S. Comparison of the 7th and proposed 8th editions of the AJCC/UICC TNM staging system for esophageal squamous cell carcinoma underwent radical surgery. Eur J Surg Oncol 2017; 43: 1949-1955.
[17] Singh R, Mishra MK and Aggarwal H. Inflammation, immunity, and cancer. Mediators Inflamm 2017; 2017: 6027305.
[18] Yang PW, Huang PM, Yong LS, Chang YH, Wu CW, Hua KT, Hsieh MS and Lee JM. Circulating interleukin-6 is associated with prognosis and genetic polymorphisms of miR608 in patients with esophageal squamous cell carcinoma. Ann Surg Oncol 2018; 25: 2449-2456.
[19] Lindenmann J, Fink-Neuboeck N, Avian A, Pichler M, Habitzruther M, Maier A and Smolle-Juettner FM. Preoperative Glasgow Prognostic Score as additional independent prognostic parameter for patients with esophageal cancer after curative esophagectomy. Eur J Surg Oncol 2017; 43: 445-453.
[20] Forrest LM, McMillan DC, McArdle CS, Angerson WJ and Dunlop DJ. Comparison of an inflammation-based prognostic score (GPS) with performance status (ECOG) in patients receiving platinum-based chemotherapy for inoperable non-small-cell lung cancer. Br J Cancer 2004; 90: 1704-6.
[21] Ma Q, Liu W, Jia R, Jiang F, Duan H, Lin P, Zhang L, Long H, Zhao H and Ma G. Inflammation-based prognostic system predicts postoperative survival of esophageal carcinoma patients with normal preoperative serum carcinoembryonic antigen and squamous cell carcinoma antigen levels. World J Surg Oncol 2016; 14: 141.
[22] Kubik S, Moszynska-Zielinska M, Fijuth J, Tomalczyk A, Jesionek-Kupnicka D, Ura L and Gottwald LM. Assessment of the relationship between serum squamous cell carcinoma antigen (SCC-Ag) concentration in patients with locally advanced squamous cell carcinoma of the uterine cervix and the risk of relapse. Prz Menopauzalny 2019; 18: 23-26.
[23] Sunpaweravong S, Puttawibul P, Sunpaweravong P, Nitiruangjaras A, Boonyaphiphat P and Kemapanmanus M. Correlation between serum SCCA and CYFRA 2 1-1, tissue Ki-67, and clinicopathological factors in patients with esophageal squamous cell carcinoma. J Med Assoc Thai 2016; 99: 331-7.
[24] Lacroix M. Significance, detection and markers of disseminated breast cancer cells. Endocr Relat Cancer 2006; 13: 1033-67.
[25] Holdenrieder S, Wehnl B, Hettwer K, Simon K, Uhlig S and Dayyani F. Carcinoembryonic antigen and cytokeratin-19 fragments for assessment of therapy response in non-small cell lung cancer: a systematic review and meta-analysis. Br J Cancer 2017; 116: 1037-1045.
[26] Di Gioia D, Blankenburg I, Nagel D, Heinemann V and Stieber P. Tumor markers in the early detection of tumor recurrence in breast cancer patients: CA 125, CYFRA 21-1, HER2 shed antigen, LDH and CRP in combination with CEA and CA 15-3. Clin Chim Acta 2016; 461: 1-7.
[27] Chang CC, Yang SH, Chien CC, Chen SH, Pan S, Lee CL, Lin CM, Sun HL, Huang CC, Wu YY, Yang RN and Huang CJ. Clinical meaning of age-related expression of fecal cytokeratin 19 in colorectal malignancy. BMC Cancer 2009; 9: 376.
[28] Cao X, Zhang L, Feng GR, Yang J, Wang RY, Li J, Zheng XM and Han YJ. Preoperative Cyfra21-1 and SCC-Ag serum titers predict survival in patients with stage II esophageal squamous cell carcinoma. J Transl Med 2012; 10: 197.
[29] Kosugi S, Nishimaki T, Kanda T, Nakagawa S, Ohashi M and Hatakeyama K. Clinical significance of serum carcinoembryonic antigen, carbohydrate antigen 19-9, and squamous cell carcinoma antigen levels in esophageal cancer patients. World J Surg 2004; 28: 680-5.
[30] Liu G, Wang Y, Wang C, He Y and E M. Clinical efficacy and safety of apatinib as maintenance treatment in patients with advanced esophageal squamous cell carcinoma. Expert Rev Clin Pharmacol 2020; 13: 1423-1430.

Table S1. Baseline characteristics of the participants

Clinical Indicators	Patients Cohort		Clinical Indicators	Patients Cohort	
	n	%		n	%
Gender (311)			T (311)		
Male	241	77.5	T1	63	20.3
Female	70	22.5	T2	85	27.3
Age (years) (311)			T3	159	51.1
Range	40-83		T4	4	1.3
Median	66		N (311)		
ECOG (311)			N0	173	55.6
Range	0-4		N1	86	27.7
Median	1		N2	35	11.3
BMI (kg/m^2) (310)			N3	17	5.5
Range	16.0-30.8		TNM (311)		
≤ 18.5	33	10.6	I	63	20.3
18.6-23.9	160	51.6	II	115	37.0
24-27.9	95	30.6	III	111	35.7
≥ 28	22	7.1	IV	22	7.1
Tumor Location (311)			Cell Differentiation (311)		
Upper	17	5.5	Well	42	13.5
Middle	217	70.2	Moderate	179	57.6
Lower	75	24.3	Poorly	90	28.9
Tumor Size (cm) (311)			State (311)		
Range	0.5-7.0		Death	143	46.0
Median	3.0		Survival	168	54.0
< 3.0	109	35.0	Follow-up (months) (311)		
≥ 3.0	202	65.0	Range	3-72	
INV (311)			Median	38	
No	256	82.3	Mean ± SD	37.04±20.31	
Yes	55	17.7			

Note: ECOG = Eastern Cooperative Oncology Group (performance status); Tumor Location: Upper part = 20-25 cm from incisor; Middle part = 25-30 cm from incisor; Lower part = 30-40 cm from incisor; INV = Invasion of Nerve or (and) Vessels; T = Tumor invasion; N = Lymph node metastasis; TNM = TNM staging.

Table S2. Basic characteristics of preoperative inflammatory and tumor markers for 311 patients with ESCC

Markers	ESCC patients		Markers	ESCC patients	
	n	%		n	%
AFP (ug/L) (311)			WBC (×10^9/L) (311)		
< 2.68	154	49.5	< 6.00	154	49.5
≥ 2.68	157	50.5	≥ 6.00	157	50.5
CEA (ug/L) (311)			ALB (g/L) (311)		
< 1.90	154	49.5	< 41.00	154	49.5
≥ 1.90	157	50.5	≥ 41.00	157	50.5
CA19-9 (IU/ml) (310)			PALB (g/L) (311)		
< 8.47	154	49.7	< 286.00	155	50.2
≥ 8.47	156	50.3	≥ 286.00	156	49.8
CA125 (IU/ml) (311)			GPS (311)		
< 8.30	155	49.8	< 1	153	49.2
≥ 8.30	156	50.2	≥ 1	158	50.8

SCC (ng/ml) (311)			LMR (311)		
< 0.90	143	46.0	< 4.28	155	49.8
≥ 0.90	168	53.4	≥ 4.28	156	50.2
CY211 (ng/ml) (310)			P-CRP (311)		
< 2.65	154	49.7	< 152.40	154	49.5
≥ 2.65	156	50.3	≥ 152.40	157	50.5
CRP (mg/dL) (311)			CPR (147)		
< 0.90	152	48.9	< 0.0037	157	50.5
≥ 0.90	159	51.1	≥ 0.0037	154	49.5
NLR (311)					
< 2.43	154	49.4			
≥ 2.43	157	50.3			

Note: ESCC = Esophageal squamous cell carcinoma; SCC = Squamous Cell Carcinoma Antigen; CY211 = cytokeratin 19 fragment; NLR = Neutrophil To Lymphocyte Ratio; CRP = C-reactive protein; PALB = prealbumin; GPS = Glasgow Prognostic Score; LMR, Lymphocyte to monocyte ratio; P-CRP, the platelet × C-reactive protein multiplier value; CPR, C-reactive protein to prealbumin ratio; These indicators were grouped by median.

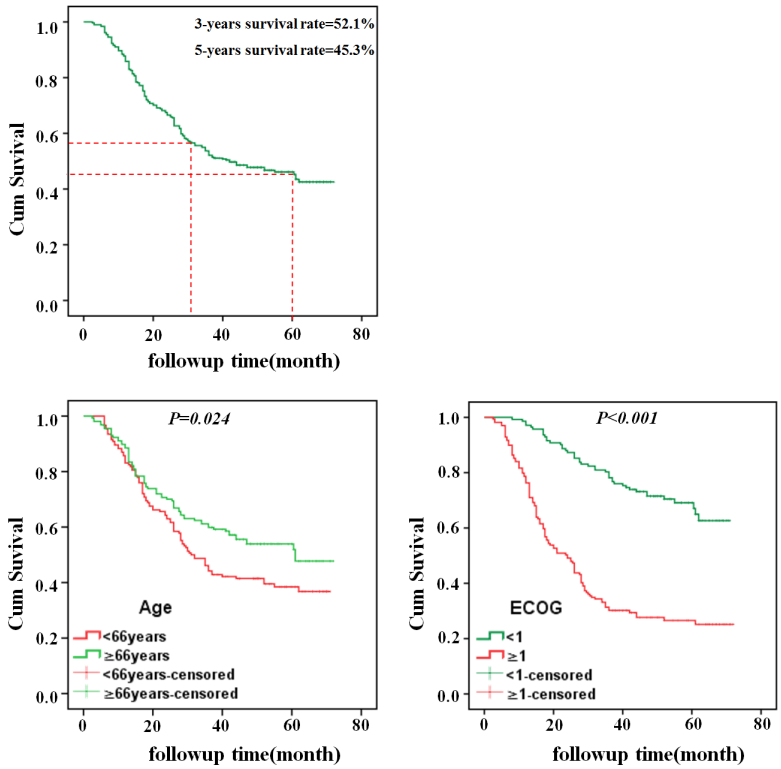

Figure S1. Distribution map of preoperative age, ECOG impact on survival and overall survival trend of ESCC patients. Overall survival (OS) was defined from the date of surgery until the date of death or the date of the last follow-up. Follow-up ranged from 0 to 72 months after surgery until October 2019, and the median follow-up time was 38 months. It can be seen from the figure: the 1-year survival rate of patients with ESCC after surgery was 85.9%, the 3-year survival rate was 52.1%, and the 5-year survival rate was 45.3%. And Kaplan-Meier survival analysis showed: Age and ECOG have an impact on the prognostic survival of ESCC patients after surgery ($P < 0.05$).

Table S3. Quantitative correlation analysis between preoperative tumor markers and postoperative clinical indicators in patients with ESCC

Markers	AFP	CEA	CA19-9	CA125	SCC	CY211	Diff	INV	Size	T	N	TNM
AFP	1											
CEA	-0.096	1										
CA199	0.036	0.309**	1									
CA125	0.114*	0.125*	0.276**	1								
SCC	-0.025	0.112*	0.022	0.107	1							
CY211	0.017	0.004	0.053	0.087	0.204**	1						
Diff	0.035	0.104	0.143*	0.039	0.004	0.006	1					
INV	-0.167**	0.093	0.034	-0.036	0.019	0.022	0.245**	1				
Size	0.013	-0.034	0.024	0.041	0.339**	0.124*	0.094	0.154**	1			
T	0.052	0.006	-0.083	0.110	0.396**	0.102	0.257**	0.171**	0.481**	1		
N	0.005	0.072	-0.028	0.069	0.220**	0.065	0.417**	0.346**	0.266**	0.413**	1	
TNM	-0.011	0.064	-0.06	0.077	0.265**	0.120*	0.448**	0.355**	0.365**	0.648**	0.884**	1

Note: Size = Tumor Size; Diff = Cell Differentiation; INV = Invasion of Nerve or (and) Vessels; T = Invasion depth; N = Lymph node metastasis; TNM = TNM staging. *p < 0.05; **p < 0.01.

Table S4. Quantitative correlation analysis between preoperative inflammatory markers and postoperative clinical indicators in patients with ESCC

	Sex	Age	BMI	ECOG	NLR	CRP	GPS	LMR	PCRP	CPR	Diff	INV	Size	T	N	TNM
Sex	1															
Age	0.184**	1														
BMI	-0.027	-0.035	1													
ECOG	-0.094	0.164**	0.013	1												
NLR	-0.179**	-0.077	-0.046	0.297**	1											
CRP	0.003	-0.041	0.019	0.102	0.198**	1										
GPS	-0.029	-0.024	0.063	0.174**	0.265**	0.645**	1									
LMR	0.224**	0.064	0.184**	-0.098	-.479**	-0.124*	-.145*	1								
PCRP	0.028	-0.039	-0.011	0.108	0.152**	0.874**	0.589**	-0.106	1							
CPR	0.014	0.022	0.013	0.131*	0.239**	0.845**	0.624**	-0.19**	0.771**	1						
Diff	-0.005	0.074	-0.126*	0.240**	0.167**	0.094	0.143*	0.045	0.143*	0.106	1					
INV	0.033	0.047	0.011	0.018	0.086	0.032	0.055	0.035	0.010	0.066	0.245**	1				
Size	-0.023	0.077	-0.015	0.127*	0.180**	0.185**	0.167**	-.127*	0.204**	0.198**	0.094	0.154**	1			
T	-.165**	-0.036	-0.03	0.128*	0.246**	0.101	0.122*	-.178**	0.141*	0.194**	0.257**	0.171**	0.481**	1		
N	-0.012	-0.071	-0.043	0.167**	0.256**	0.071	0.065	-0.077	0.041	0.103	0.417**	0.346**	0.266**	0.413**	1	
TNM	-0.018	-0.063	-0.05	0.163**	0.272**	0.080	0.083	-0.099	0.092	0.125*	0.448**	0.355**	0.365**	0.648**	0.884**	1

Note: Size = Tumor Size; Diff = Cell Differentiation; INV = Invasion of Nerve or (and) Vessels; T = Invasion depth; N = Lymph node metastasis; TNM = TNM staging. *p < 0.05; **p < 0.01.

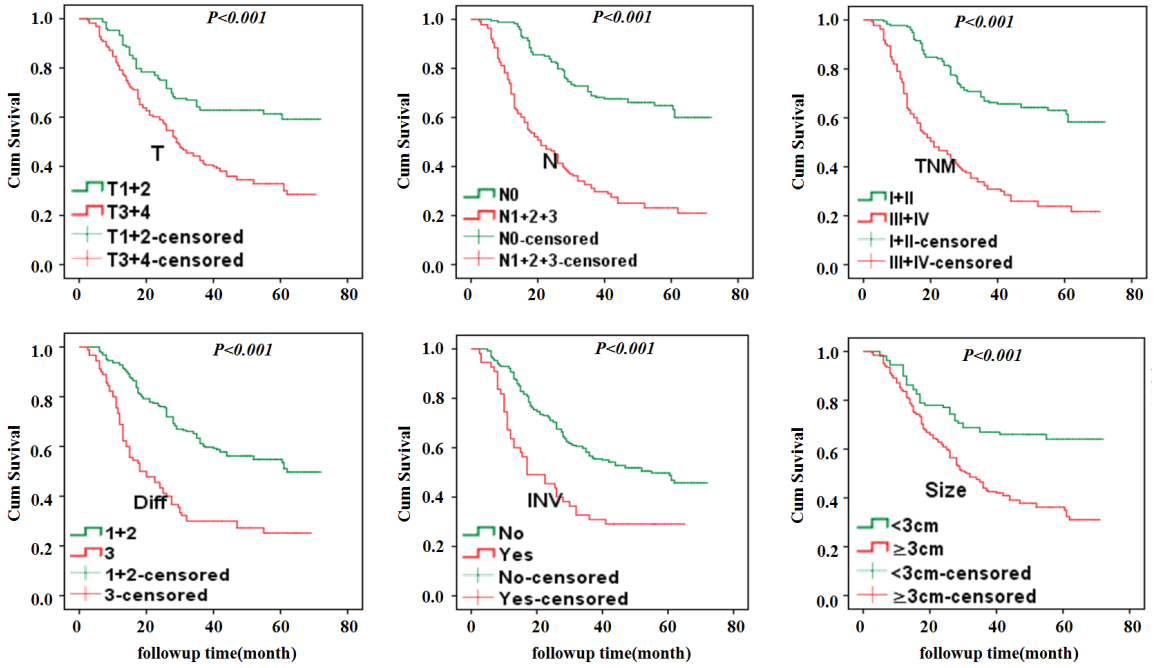

Figure S2. Post-operative indicators correlate with survival in patients with ESCC. The survival curve of patients is getting lower and lower, and the prognosis is getting worse. Mainly include the following indicators that are meaningful for survival and prognosis: Size, INV, cell differentiation, T, N, and TNM staging had an impact on the prognostic survival of patients with ESCC after surgery.

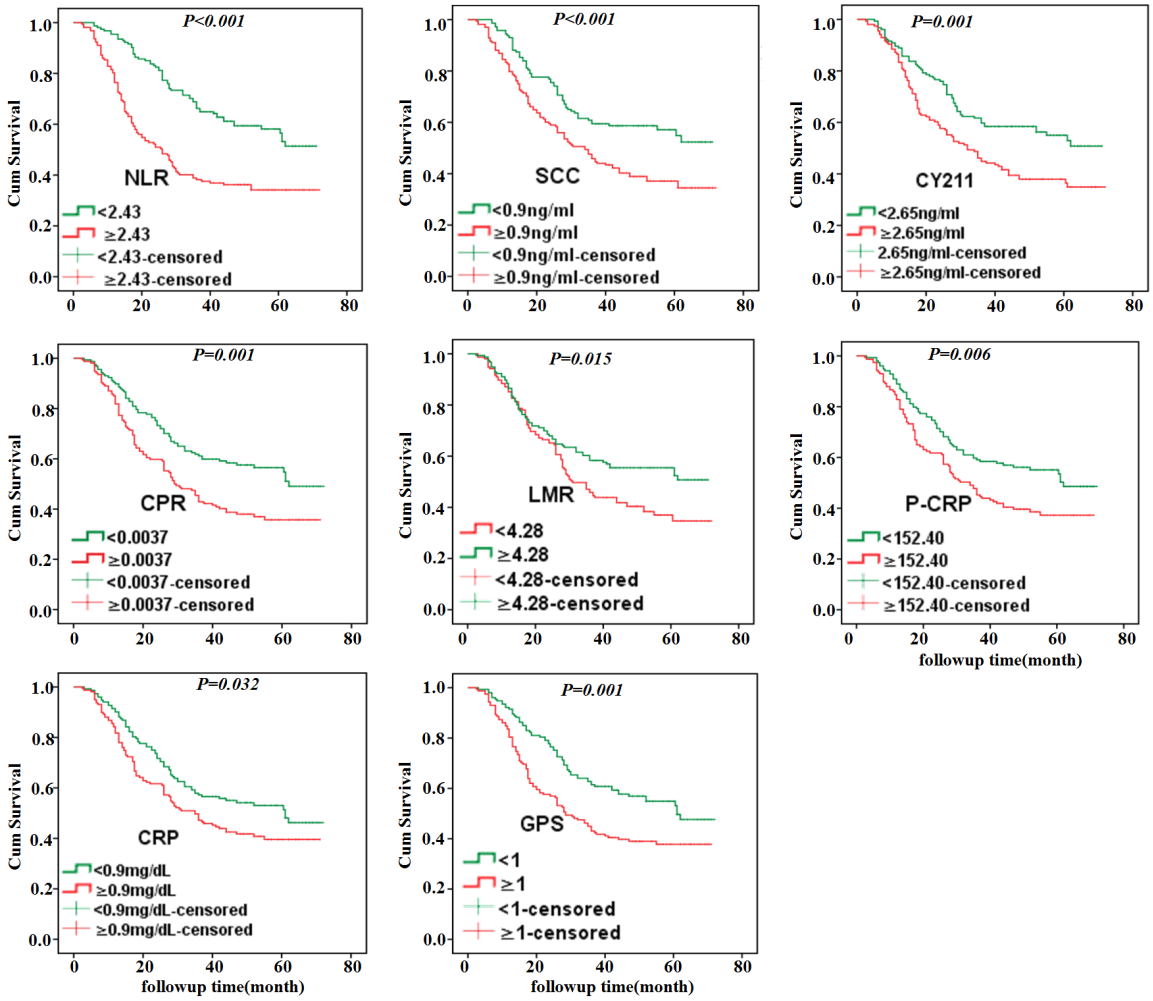

Figure S3. Preoperative related indicators correlate with survival in patients with ESCC. Mainly include the following indicators that are meaningful for survival and prognosis: Survival analysis revealed that high expression of NLR, SCC, CY211, CPR, P-CRP, CRP, and GPS correlated with poor prognosis in ESCC ($P < 0.05$). Low expression of LMR suggests a poor prognosis for patients with ESCC after surgery.

Table S5. Risk factors affecting the prognosis of patients with ESCC by Cox single factor analysis

Clinical Parameters	B	SE	Wald	df	Cox single-factor regression model analysis HR (95% CI)	*P* value
Markers before surgery						
Gender (Female vs Male)	-0.305	0.197	2.408	1	0.737 (0.501-1.084)	0.121
Age (≥ 66 vs < 66 years)	-0.348	0.156	4.993	1	0.706 (0.521-0.958)	0.025
BMI			2.705	3		0.439
(18.6-23.9 vs ≤ 18.5 kg/m^2)	-0.083	0.243	0.117	1	0.920 (0.572-1.481)	0.732
(24-27.9 vs ≤ 18.5 kg/m^2)	-0.332	0.265	1.571	1	0.718 (0.427-1.206)	0.210
(≥ 28 vs ≤ 18.5 kg/m^2)	-0.329	0.372	0.781	1	0.720 (0.347-1.493)	0.377
ECOG (≥ 1 vs < 1)	1.351	0.176	58.918	1	3.863 (2.736, 5.455)	< 0.001
NLR (≥ 2.43 vs < 2.43)	0.767	0.159	23.363	1	2.152 (1.577, 2.937)	< 0.001
AFP (≥ 2.68 vs < 2.68 ug/L)	0.118	0.155	0.583	1	1.125 (0.831-1.524)	0.445
CEA (≥ 1.90 vs < 1.90 ug/mL)	0.004	0.154	0.001	1	1.004 (0.742-1.359)	0.978

CA19-9 (≥ 8.47 vs < 8.47 IU/mL)	0.302	0.156	3.731	1	1.352 (0.996-1.837)	0.053
CA125 (≥ 8.30 vs < 8.30 IU/mL)	0.238	0.155	2.363	1	1.269 (0.937-1.720)	0.124
SCC (≥ 0.90 vs < 0.90 ng/mL)	0.537	0.160	11.277	1	1.711 (1.251-2.341)	0.001
CY211 (≥ 2.65 vs < 2.65 ng/mL)	0.496	0.157	9.943	1	1.642 (1.207-2.236)	0.002
CRP (≥ 0.90 vs < 0.90 mg/dL)	0.033	0.012	8.183	1	1.034 (1.011-1.058)	0.004
WBC (≥ 6.00 vs < 6.00 × 10^9/L)	-0.154	0.155	0.986	1	0.858 (0.633-1.161)	0.321
ALB (≥ 41.00 vs < 41.00 g/L)	-0.064	0.155	0.170	1	0.938 (0.692-1.271)	0.680
PALB (≥ 286.00 vs < 286.00 g/L)	-0.261	0.155	2.834	1	0.770 (0.568-1.044)	0.092
GPS (≥ 1 vs < 1)	0.495	0.156	10.000	1	1.640 (1.207-2.229)	0.002
LMR (≥ 4.28 vs < 4.28)	-0.376	0.156	5.786	1	0.687 (0.505-0.933)	0.016
P-CRP (≥ 152.40 vs < 152.40)	0.420	0.156	7.249	1	1.523 (1.121-2.068)	0.007
CPR (≥ 0.0037 vs < 0.0037)	0.504	0.156	10.374	1	1.655 (1.218-2.248)	0.001
Tumor Location						0.795
(Middle vs ≤ Upper)	0.142	0.366	0.151	1	1.153 (0.563-2.360)	0.698
(Lower vs ≤ Upper)	0.232	0.386	0.361	1	1.261 (0.592-2.686)	0.548
Markers after surgery						
T (T3+T4 vs T1+T2)	0.779	0.163	22.814	1	2.180 (1.583-3.002)	< 0.001
N (N1+N2+N3 vs N0)	1.251	0.161	60.371	1	3.493 (2.548-4.789)	< 0.001
TNM (III+IV vs I+II)	1.161	0.159	53.672	1	3.194 (2.341-4.358)	< 0.001
Cell Differentiation (3 vs 1+2)	0.898	0.159	31.892	1	2.455 (1.797-3.353)	< 0.001
Tumor Size (≥ 3.00 vs < 3.00 cm)	0.810	0.185	19.214	1	2.248 (1.565-3.230)	< 0.001
INV (yes vs no)	0.731	0.184	15.844	1	2.077 (1.449-2.977)	< 0.001

Note: HR, hazard ratio; CI, confidence interval; vs., versus; GPS=Glasgow Prognostic Score; Tumor Location: Upper part = 20-25 cm from incisor; Middle part = 25-30 cm from incisor; Lower part = 30-40 cm from incisor; CY211 = Cyfra21-1; INV = Invasion of Nerve or (and) Vessels; Cell Differentiation: 1 = Well; 2 = Moderate; 3 = Poorly.

Table S6. Risk factors affecting the prognosis of patients with ESCC by cox multiple factor regression analysis

Clinical Markers	B	SE	Wald	df	Cox multi-factor regression model analysis	*P* value
					HR (95% CI)	
Age (≥ 66 vs < 66 years)	-0.500	0.161	9.703	1	0.606 (0.443, 0.831)	0.002
ECOG (≥ 1 vs < 1)	1.527	0.188	66.139	1	4.603 (3.186, 6.649)	< 0.001
NLR (≥ 2.43 vs < 2.43)	0.631	0.169	14.014	1	1.880 (1.351, 2.617)	< 0.001
SCC (≥ 0.90 vs < 0.90 ng/mL)	0.385	0.169	5.210	1	1.469 (1.056, 2.044)	0.022
CY211 (≥ 2.65 vs < 2.65 ng/mL)	0.454	0.159	8.142	1	1.574 (1.153, 2.150)	0.004
INV (yes vs no)	0.578	0.199	8.400	1	1.783 (1.206, 2.636)	0.004
Cell Differentiation (3 vs 1+2)	0.399	0.175	5.204	1	1.491 (1.058, 2.101)	0.023
TNM (III+IV vs I+II)	0.959	0.172	31.022	1	2.609 (1.862, 3.656)	< 0.001

Note: HR, hazard ratio; CI, confidence interval; vs., versus; INV = Invasion of Nerve or (and) Vessels; Cell Differentiation: 1 = Well; 2 = Moderate; 3 = Poorly.