Increases prognostic value of clinical-pathological nomogram in patients with esophageal squamous cell carcinoma

Jing Feng Hu[1,2], Xin Song[1,2], Kan Zhong[1,2], Xue Ke Zhao[1,2], Fu You Zhou[3], Rui Hua Xu[1,2], Ji Lin Li[4], Xian Zeng Wang[5], Xue Min Li[6], Pan Pan Wang[1,2], Ling Ling Lei[1,2], Meng Xia Wei[1,2], Ran Wang[1,2], Zong Min Fan[1,2], Xue Na Han[1,2], Yao Chen[1,2], Liu Yu Li[1,2], Jia Jia Ji[1,2], Yuan Ze Yang[1,2], Bei Li[1,2], Miao Miao Yang[1,2], Hai Jun Yang[3], Fu Bao Chang[7], Jing Li Ren[8], Sheng Li Zhou[9] and Li Dong Wang[1,2]*

[1]State Key Laboratory of Esophageal Cancer Prevention & Treatment of The First Affiliated Hospital, Zhengzhou University, Zhengzhou, Henan, China, [2]Henan Key Laboratory for Esophageal Cancer Research of The First Affiliated Hospital, Zhengzhou University, Zhengzhou, Henan, China, [3]Department of Thoracic Surgery, Anyang Tumor Hospital, Anyang, Henan, China, [4]Department of Pathology, Linzhou Esophageal Cancer Hospital, Linzhou, Henan, China, [5]Department of Thoracic Surgery, Linzhou People's Hospital, Linzhou, Henan, China, [6]Department of Pathology, Hebei Provincial Cixian People's Hospital, Cixian, Hebei, China, [7]Department of Surgery, Central Hospital of Linzhou City, Linzhou, Henan, China, [8]Department of Pathology, The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China, [9]Department of Pathology, Henan Provincial People's Hospital, People's Hospital of Zhengzhou University, People's Hospital of Henan University, Zhengzhou, Henan, China

Background: This study was intended to construct a brand new prognostic nomogram after combine clinical and pathological characteristics to increases prognostic value in patients with esophageal squamous cell carcinoma.

Methods: A total of 1,634 patients were included. Subsequently, the tumor tissues of all patients were prepared into tissue microarrays. AIPATHWELL software was employed to explore tissue microarrays and calculate the tumor-stroma ratio. X-tile was adopted to find the optimal cut-off value. Univariate and multivariate Cox analyses were used to screen out remarkable characteristics for constructing the nomogram in the total populations. A novel prognostic nomogram with clinical and pathological characteristics was constructed on the basis of the training cohort (n=1,144). What's more performance was validated in the validation cohort (n=490). Clinical-pathological nomogram were assessed by concordance index, time-dependent receiver operating characteristic, calibration curve and decision curve analysis.

Results: The patients can divide into two groups with cut-off value of 69.78 for the tumor-stroma ratio. It is noteworthy that the survival difference was noticeable (*P*<0.001). A clinical-pathological nomogram was constructed by combining clinical and pathological characteristics to predict the overall survival. In comparison with TNM stage, the concordance index and time-dependent receiver operating characteristic of the clinical-pathological nomogram showed better predictive value (*P*<0.001). High quality of calibration plots in overall survival

was noticed. As demonstrated by the decision curve analysis, the nomogram has better value than the TNM stage.

Conclusions: As evidently revealed by the research findings, tumor-stroma ratio is an independent prognostic factor in patients with esophageal squamous cell carcinoma. The clinical-pathological nomogram has an incremental value compared TNM stage in predicting overall survival.

KEYWORDS

esophageal squamous cell carcinoma, tumor-stroma ratio, nomogram, prognosis, overall survival

1 Introduction

Esophageal cancer (EC) is a highly malignant tumor with a poor prognosis. Apart from that, the 5-year survival rate of the EC patients with advanced stage is nothing more than about 10% (1). The morbidity and mortality of EC are considerably high in China. Worse still, about 200,000 people die of EC every year (2). The principal pathological types of EC are two types of esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC). ESCC is the predominant pathological type with 90% of the EC in China (3). So far, the determination of the prognosis of patients with ESCC primarily rests with the TNM stage, but the current research shows that the TNM staging method cannot predict the prognosis of patients well (4). What's more, the accuracy of the prognosis of patients can be heightened by adding clinical and pathological factors (5, 6). As a result, new indicators need to be explored to evaluate the prognosis of patients.

As persuasively demonstrated by a multitude studies (7–9), the interstitial components of tumor tissue affect the prognosis of tumor patients. Nevertheless, owing to the lack of accurate and reliable evaluation methods for tumor stroma, the association between tumor stroma and patient prognosis is still unclear. This retrospective study explored the correlation between the tumor stroma and the overall survival (OS) as well as clinical characteristics of patients with ESCC by detecting the tumor-stromal ratio (the ratio of the area of stromal to the cancerous, TSR), to establish a classification criteria for TSR and to propose a brand new prognostic indicator. Aside from that, this study combined the TSR with the clinical and pathological characteristics of patients to construct a clinical-pathological nomogram, aiming to provide a better prediction model for the survival of ESCC.

2 Materials and methods

2.1 Study populations and follow-up

This study included 1,634 patients with pathologically diagnosed ESCC who were treated with radical surgery. The information of these patients comes from the tissue bank, clinical diagnosis, treatment, pathology, follow-up information database of about 500,000 cases of esophageal and gastric cardia carcinoma (1973-2021), established by the State Key Laboratory of Esophageal Cancer Prevention & Treatment and Henan Key Laboratory for Esophageal Cancer Research of The First Affiliated Hospital, Zhengzhou University. The clinical information of patients collected from the database, including gender, age at diagnosis, smoking history, drinking history, family history, high and low incidence area [the mortality, and mortality of esophageal cancer greater than 50/100,000 is defined as high incidence area (10)], and pathological information, including the degree of differentiation, T stage, N stage, M stage, as well as the patient's H&E tissue microarrays. All patients were staged in accordance with to the 2002 American Joint Committee on Cancer (AJCC) esophageal cancer staging system. The documentation of informed consent was waived owing to anonymity of the participants. This study was approved by the Ethical Committee of The First Affiliated Hospital of Zhengzhou University. All procedures involving human participants were performed in accordance with the ethical committee's standards and the Helsinki declaration.

Inclusion criteria: 1. Detailed and completed clinical data; 2. Received radical surgery for ESCC; 3. Postoperative pathological diagnosis was ESCC; Exclusion criteria: 1. Received radiotherapy or (and) chemotherapy before surgery; 2. Patients with distant metastasis; 3. Patients with other malignant tumors; 4. Patients who died within 30 days after surgery.

The follow-up of this study is principally through letters, telephone calls, home visits, and direct contact between village doctors and patients/their families, or through the new cooperative medical database, the medical security bureau database, and citizen death information registration management systems. It is carried out in an annual follow-up to record the time of death and the predominant cause of death to January 27, 2022. The primary clinical outcome was overall survival (OS).

2.2 Detection of tumor-stroma ratio

Tissue microarrays (TMAs) is a brand new method for high-throughput analysis of diagnostic or predictive markers in cancer specimens (11). What's more, it has excellent applications in immunohistochemical analysis and observation of pathological features (12, 13). A total of 1,634 samples were histologically

screened by H&E staining, and the sections were observed under a microscope. The 40 magnification was adopted to find the most infiltrating area of tumor tissue, and the representative area was marked in the paraffin block. Two cores (a cylinder with a diameter of 1.5 mm and a height of 3-5 mm) were taken out from the paraffin block to prepare a tissue chip. Each TMAs contains 150 cores (TMAs were constructed by Wuhan Servicebio Technology Co., Wuhan, China).

AIPATHWELL software (developed by Wuhan Servicebio Technology Co.) was utilized to scan and probe deep into the tissue microarray. The software distinguishes cancerous and stroma in accordance with tissue characteristics such as cell morphology and nucleus morphology. Apart from that, the output results are data such as cancerous area, stroma area. The tumor-stroma ratio (TSR) was expressed as the area of stroma/total area (stroma area + cancerous area). To ensure the accuracy of the results, we selected two cores for each patient and placed them on the tissue microarray. Subsequently, we took the mean as the final TSR value. The X-tile software was adopted to determine the cut-off value for the TSR. The X-tile software provides a convenient and comprehensive method to stratify the population by finding the optimal cut-off value based upon the patient's overall survival time, when the difference in survival between these two groups is most striking (14).

2.3 Construction of the nomogram

The clinical and pathological information of patients were first considered as candidate variables for the prognostic model. It is noteworthy that all variables are categorical variables. The factors which affecting OS in univariate analysis ($P < 0.05$) were then included in Cox analyses. Subsequently independent prognostic risk factors ($P < 0.05$) were included as the final variables to construct the nomogram model. In the training set data, these characteristics were used to construct nomogram to predict the probability of patient survival time up to 1-year, 3-year, and 5-year by employing the rms package.

2.4 Calibration, validation and risk stratification of nomogram

Predictive performance was quantified with reference to accuracy (calibration), discrimination, and clinical application. The agreement between the predicted survival probability calculated by the nomogram and the patient's actual survival probability was assessed by plotting 1-, 3-, and 5-year calibration curves for a bootstrapped with 1,000 sample. The discrimination of the nomogram model was assessed by calculating the concordance index (C-index), NRI and time-dependent receiver operating characteristic (ROC) (15). The clinical application and net benefit at diverse threshold probabilities of the nomogram model was assessed by decision curve analysis (DCA) (16). Patients in the study were divided into high-risk or low-risk groups in accordance with the optimal cutoff point for the total prognostic score (TPS) decided by the X-Tile software. Survival curves were drawn by adopting Kaplan-Meier analysis. The

potential correlations between risk scores and OS among dissimilar subgroups were investigated in the overall patient population (14).

2.5 Statistical analysis

All statistical analyses were performed by employing the R software (version 4.2.0). T test was adopted to probe into continuous variables, categorical variables by adopting χ2 test and Fisher's exact two-sided test. Survival analysis was drawn by employing Cox analysis and the Kaplan-Meier method. The nomogram and calibration plot were constructed using the "rms" package; the time-dependent receiver operating characteristic (ROC) was established by adopting the "timeROC" package; the DCA was established by employing the "rmda" and "ggDCA" package. $P < 0.05$ (two-sided) was considered statistically significant.

3 Results

3.1 Patient characteristics

The characteristics information of the 1,634 patients is shown in Table 1. The male to female ratio is 1.9:1; the mean age is 61 years, and the age range is 33-85. The median survival time of patients with ESCC was 2.0 years, and the 1-, 3- and 5-years OS rates were 77.2%, 35.1% and 23.7% respectively.

Patients were randomly divided into training and validation sets in a ratio of 7:3. The male to female ratio in the training set is 1.9:1; the mean age is 62 years, and the age range is 33-85. The median survival time of patients with ESCC was 2.0 years, and the 1-, 3- and 5-years OS rates were 76.5%, 35.1% and 24.1% respectively. The ratio of male to female in the validation set is 1.9:1; the average age is 61 years, and the age range is 38-82. The median survival time of patients with ESCC was 2.0 years, and the 1-, 3- and 5-years OS rates were 77.6%, 34.9% and 22.9% respectively.

3.2 Association analysis of TSR and clinical characteristics

All patients successfully collected the detection value of TSR. The classification criteria for TSR were explored by comparing the relationship between TSR and survival for all patients. And the optimal cut-off value will be found when the survival difference was greatest. The optimal cut-off value was found by X-tile software. The result showed that patients were divided into two groups when the TSR cutoff value was 69.78, at which time the survival difference between the two groups was the greatest (P<0.001), and 69.78 was the optimal cutoff value for TSR. According to the optimal cut-off value, patients with TSR more than 69.78 were divided into stroma-rich group and patients with TSR less than 69.78 were divided into stroma-poor group (Figure 1). And a total of 1324 patients were divided into stroma-poor group and 310 patients were divided into stroma-rich group. The results showed that the abundance of the stroma in tumor tissue significantly affects the prognosis of patients,

TABLE 1   Clinical characteristics of patients with esophageal squamous cell carcinoma.

Characteristics	Case (%)
Gender	
Male	1061 (64.93)
Female	573 (35.07)
Age at diagnosis in year	
Mean ± SD	61 ± 8
Smoking history	
Yes	781 (47.80)
No	853 (52.20)
Drinking history	
Yes	527 (32.25)
No	1107 (67.75)
Family history	
Yes	898 (54.96)
No	736 (45.04)
Area	
High incidence area	1138 (69.65)
Low incidence area	496 (30.35)
Tumor location	
Upper	271 (16.59)
Middle	1110 (67.93)
Lower	253 (15.48)
Differentiation	
Well	134 (8.20)
Moderate	975 (59.67)
Poor	525 (32.13)
T stage	
T1	55 (3.37)
T2	375 (22.95)
T3	1177 (72.03)
T4	27 (1.65)
N stage	
N1	832 (50.92)
N0	802 (49.08)

and patients survived significantly better when the stroma was abundant, and the TSR is an independent prognostic factor in patients with ESCC. The survival curves of the two groups are shown in Figure 2G. We applied the χ2 test to assess associations between TSR and clinical characteristics. We found that the TSR was significantly associated with the differentiation ($P = 0.004$) and depth of invasion ($P = 0.002$). In the stroma-rich group, the

proportion of well-differentiated patients was relatively high, while, patients in the stroma-rich group had a relatively higher depth of infiltration. The relationship between TSR and differentiation and T stage may be the reason why it affects the prognosis of patients. The relationship between the clinical characteristics and the TSR is shown in Table 2.

3.3 Prognostic roles of clinical and pathological characteristics

To evaluate the potential association of clinical factors with overall survival, we performed univariate and multivariate Cox analysis in the total population, the prognostic role of each characteristic in OS was tested in the overall populations (Figure 2). And the univariate analysis results showed that age, drinking history, family history, area, TSR, differentiation, T stage and N stage were significantly associated with OS ($P < 0.05$) (Figure 2A), and multivariate Cox analysis results showed that age, drinking history, family history, area, TSR, T stage and N stage were independent predictors of OS ($P < 0.05$) (Figure 2B). Data are also represented using Kaplan-Meier curves (Figures 2C–J).

3.4 Construction of nomogram

To address the patient's prognosis, we constructed a clinical-pathological nomogram (Figure 3A), and age, drinking history, family history, area, TSR, differentiation, T stage, and N stage were finally included into the model. Age was the age of patients at the time of diagnosis of the disease and was divided into two age groups, <60 and ≥60; drinking history was divided into positive and negative groups according to whether patients drank alcohol; family history was divided into positive and negative groups according to whether patients had tumor patients in their immediate family; area was divided into two groups, high and low incidence areas, according to the mortality rate of patients with ESCC in the area where patients lived (10); TSR was divided into stroma rich and stroma poor groups according to the results of this study; according to the AJCC stage, differentiation was divided into well differentiated, moderately differentiated, and poor differentiated, and T-stage was divided into T1, T2, T3 and T4; and N-stage was divided into N0 and N1. Among them, the differentiation was not an independent predictor for prognosis in the multivariate analysis, because it was included in the latest version of AJCC stage, differentiation was also included in the final nomogram. Since there was no difference in survival between moderate differentiation and poor differentiation patients in this study, thus combining them into one category. To investigate whether the clinical-pathological nomogram has incremental prognostic value for individualized OS prediction, a TNM stage nomogram model was also constructed, including only differentiation, T stage, N stage (Figure 3B). In the nomogram, each characteristic has its corresponding points, and total points is obtained by summing the scores corresponding to each characteristic. The scale at the bottom of the total points indicates the survival rates at 1, 3 and 5 years. The probability of survival at
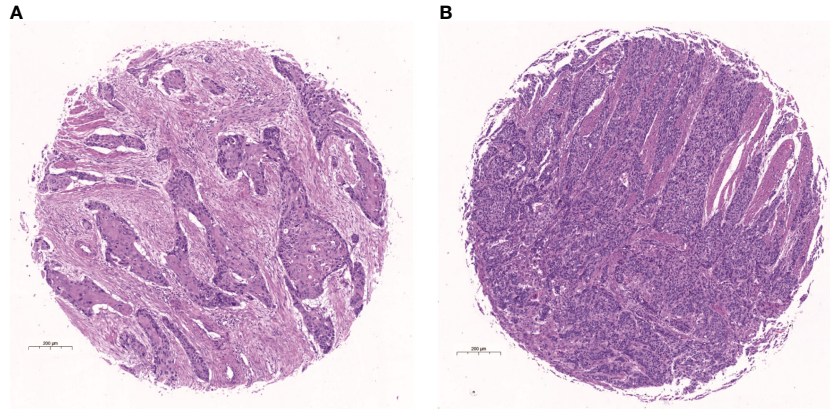
FIGURE 1
Hematoxylin-eosin (H&E) stained picture of the core of tissue microarray. **(A)** Example of stroma-rich group. **(B)** Example of stroma-poor group.

3.5 Calibration of clinical-pathological nomogram models

Figure 4 shows the calibration curve for the clinical-pathological nomogram model. The results show that the 1-, 3- and 5-year OS predictions of the clinical-pathological nomogram model well matched the actual outcomes in the training and validation cohort.

3.6 Validation of clinical-pathological nomogram model

The C-index of the clinical-pathological nomogram model and the TNM stage nomogram model in the training cohort were 0.643 (0.642-0.644) and 0.600 (0.599-0.601), respectively, and the clinical-pathological nomogram was better than the TNM stage in OS prediction. The 1-, 3- and 5-year NRI values of the clinical-pathological nomogram compared with TNM stage in the training cohort were 0.219 (0.039-0.433), 0.555 (0.396-0.677) and 0.491 (0.358-0.638), respectively, and the results showed that the
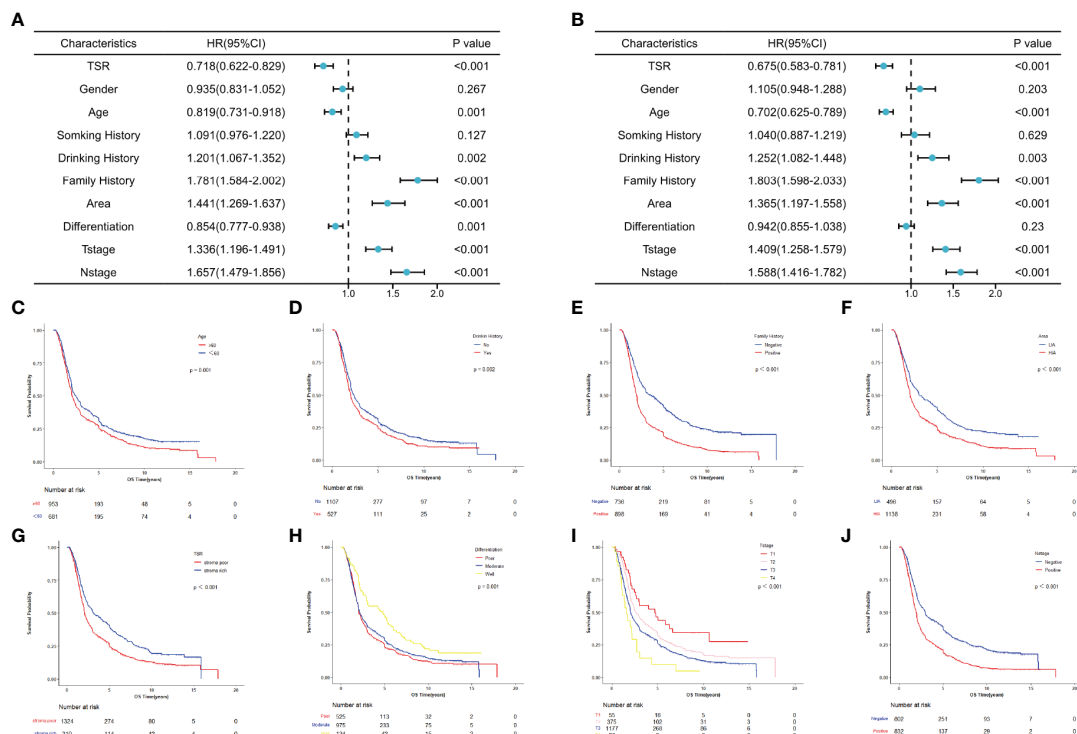
FIGURE 2
Survival analysis for patients with esophageal squamous cell carcinoma in overall populations. **(A)** Univariate analysis. **(B)** Multivariate analysis. **(C–J)** Survival curves for patients with age, drinking history, family history, area, TSR, differentiation, T stage and N stage, separately.

TABLE 2 Relationship between TSR and clinical characteristics, n (%).

Characteristics	No. of The Patients Examined	Stroma Rich	Stroma Poor	P	OR	CI (95%)
Gender				0.375	0.888	0.683-1.154
Male	1061 (64.93)	208 (67.09)	853 (64.43)			
Female	573 (35.07)	102 (32.91)	471 (35.57)			
Age				0.98	0.997	0.886-1.281
≥60	953 (58.32)	181 (58.39)	772 (58.31)			
<60	681 (41.68)	129 (41.61)	552 (41.69)			
Area				0.501	0.913	0.699-1.191
High incidence area	1138 (69.65)	211 (68.06)	927 (70.01)			
Low incidence area	496 (30.35)	99 (31.94)	397 (29.98)			
Smoking History				0.983	0.997	0.779-1.277
Yes	781 (47.80)	148 (47.74)	633 (47.81)			
No	853 (52.20)	162 (52.26)	691 (52.19)			
Drinking History				0.684	1.056	0.812-1.374
Yes	527 (32.25)	105 (33.23)	424 (32.02)			
No	1107 (67.75)	207 (66.77)	900 (67.98)			
Family History				0.963	0.994	0.775-1.275
Positive	898 (54.96)	170 (54.84)	728 (54.98)			
Negative	736 (45.04)	140 (45.16)	596 (45.02)			
Location				0.184	–	–
Upper	271 (16.59)	50 (16.13)	221 (16.69)			
Middle	1110 (67.93)	222 (71.61)	888 (67.07)			
Lower	253 (15.48)	38 (12.26)	215 (16.24)			
Differentiation				0.004	–	–
Poor	525 (32.13)	81 (26.13)	444 (33.53)			
Moderate	975 (59.67)	192 (61.93)	783 (59.14)			
Well	134 (8.20)	37 (11.94)	97 (7.33)			
T stage				0.002	–	–
T1	55 (3.37)	10 (3.23)	45 (3.40)			
T2	375 (22.95)	46 (14.84)	329 (24.85)			
T3	1177 (72.03)	248 (80.00)	929 (70.16)			
T4	27 (1.65)	6 (1.93)	21 (1.59)			
N stage				0.984	1.002	0.783-1.284
N0	802 (49.08)	152 (49.03)	650 (49.09)			
N1	832 (50.92)	158 (50.97)	674 (50.91)			

prediction of the clinical-pathological nomogram was better than that of the traditional TNM stage in 1-, 3-, and 5-year OS prediction.

Figure 5 shows the time-dependent ROC curves of the clinical-pathological nomogram and the TNM stage in the training and validation cohort. The results showed that in the training cohort, the area of under curve (AUC) of the 3 years and 5 years of the clinical-pathological nomogram model was significantly larger than that of the TNM stage nomogram model, while the AUC of the 1 year of the two models was similar, thus the clinical-pathological nomogram was better than the TNM stage at 3 years and 5 years ($P < 0.001$), and the two models were similar at 1 year ($P = 0.414$); in the validation cohort the results were similar to the training cohort, the clinical-pathological nomogram model was better than the TNM stage nomogram model at 3 years and 5 years ($P < 0.001$), and the two models were similar at 1 year ($P = 0.306$).
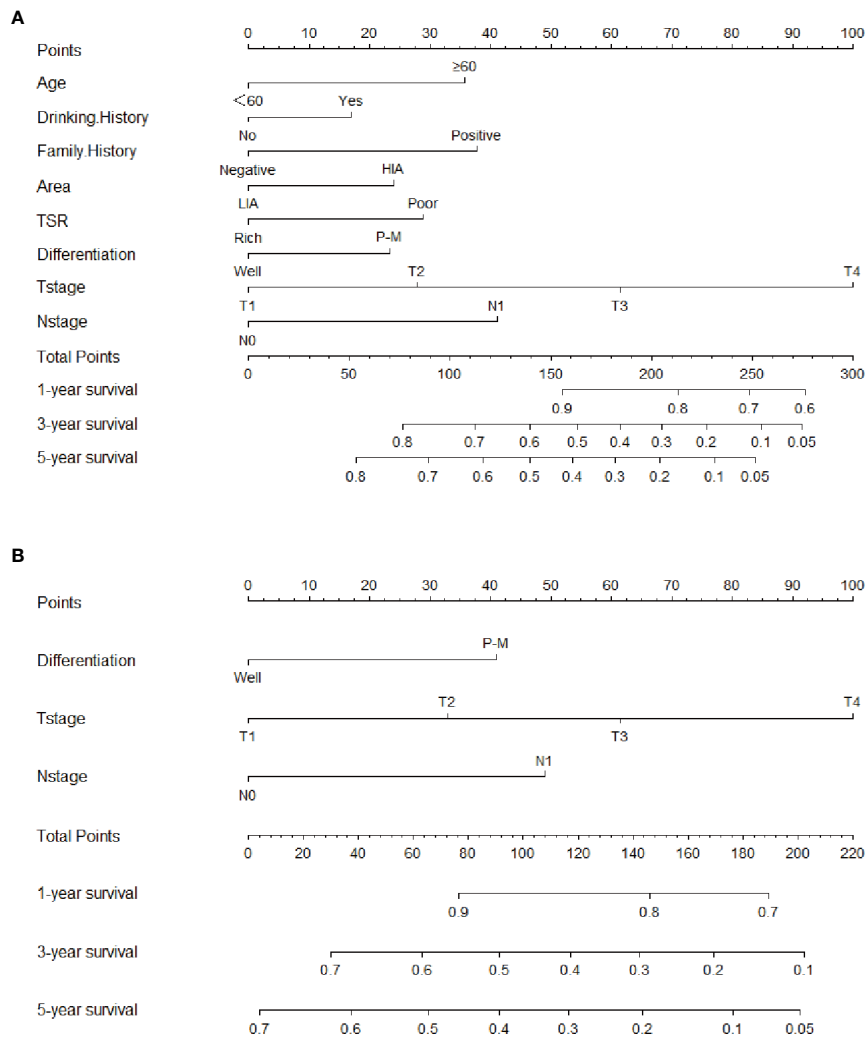
FIGURE 3
Clinical-pathological nomogram and TNM stage nomogram for 1-, 3-, and 5-year overall survival predicting in patients with esophageal squamous cell carcinoma. **(A)** Clinical-pathological nomogram. **(B)** TNM stage nomogram. HIA, High incidence area; LIA, Low incidence area; P-M, Poor or moderate differentiation.

Clinical benefits of clinical-pathological nomogram model

The clinical benefit of the clinical-pathological nomogram model was compared with the TNM stage nomogram model. The Decision curve analysis (DCA) curve showed that when the threshold probability was > 0.5, the area under the decision curve of the clinical-pathological nomogram model was larger than that of the TNM stage nomogram model in the 3-year and 5-year, thus the clinical-pathological nomogram model could better predict the 3-year and 5-year OS of patients. And the clinical-pathological nomogram model yielded a larger net benefit than TNM stage nomogram model when the threshold probability was > 0.5 (Figure 6).
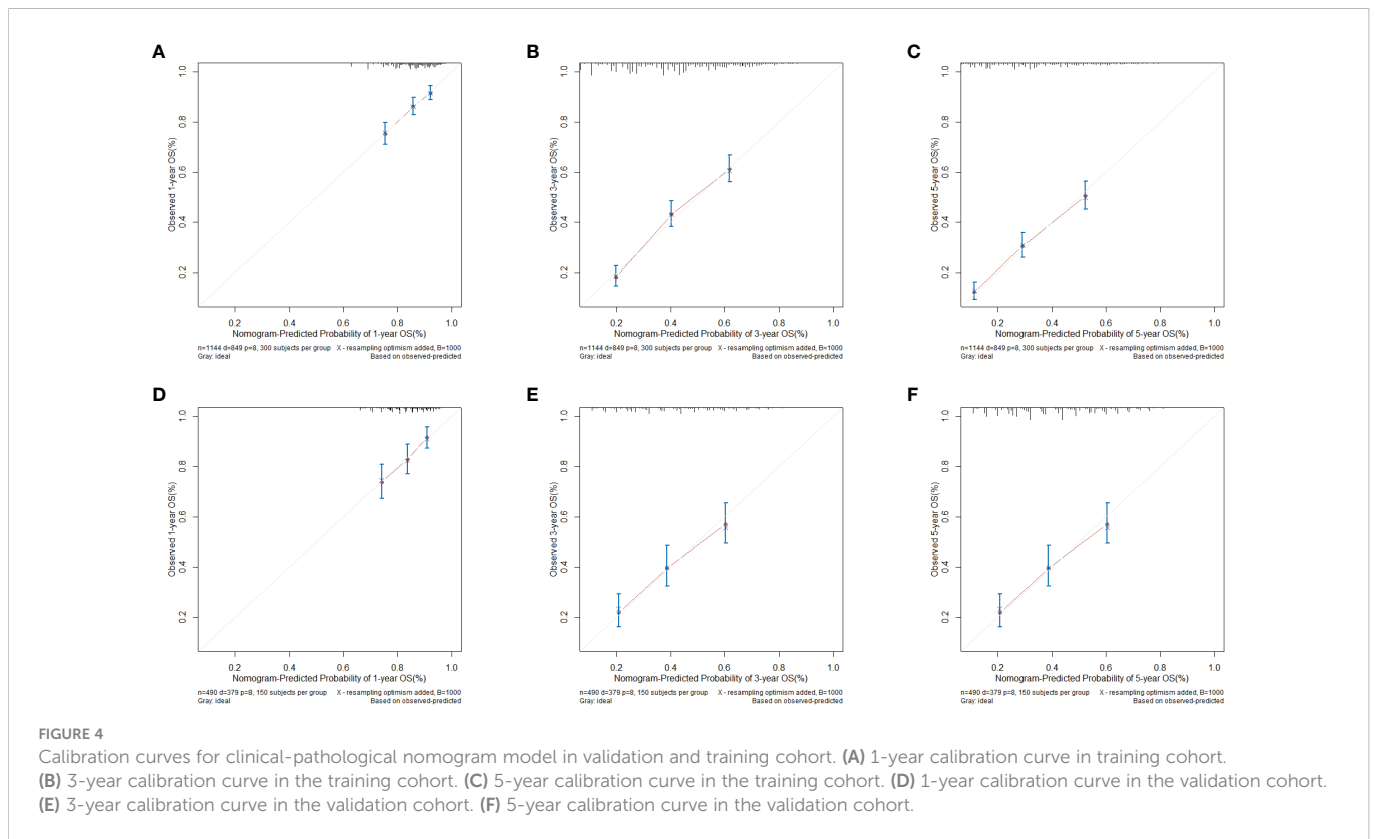
3.7 Performance of prognostic nomogram model

For further revealing the prognostic prediction value of the clinical-pathological nomogram model, the X-tile software was used to categorized into high-risk and low-risk groups according to the optimal cut-off value of the total predictive scores in the training cohort. The relationships among risk scores and survival status of patients in training cohort and validation cohort are shown in Figures 7A, B. Kaplan-Meier survival curves revealed that patients with higher risk scores had significantly poorer OS (P<0.001, Figures 7C, D). Subgroup analyses were performed in the overall populations, and the clinical-pathological model also maintained a good and stable prediction performance (Figure 8).

4 Discussion

Notwithstanding the fact that great progress has been made in the prevention and treatment of ESCC in recent decades, the prognosis of ESCC is still poor with the high mortality rate (17, 18). The TNM stage system is adopted to predict the prognosis of numerous cancers, but TNM stage system is difficult to effectively predict the prognosis of individual patients due to multitudinous factors affecting the prognosis. For instance, patients with the same pathological stage,

FIGURE 4
Calibration curves for clinical-pathological nomogram model in validation and training cohort. **(A)** 1-year calibration curve in training cohort.
(B) 3-year calibration curve in the training cohort. **(C)** 5-year calibration curve in the training cohort. **(D)** 1-year calibration curve in the validation cohort.
(E) 3-year calibration curve in the validation cohort. **(F)** 5-year calibration curve in the validation cohort.

and treated by the same treatment plan, have completely dissimilar prognosis, indicating that there are biological differences between distinct patients. Furthermore, it is not enough to judge the prognosis on the basis of the TNM stage alone. As a consequence, there is a need to add new factors that affect survival and establish more reliable prognostic tools that can be used for risk stratification and further clinical decision-making, thereby which is advantageous for patients to ameliorate their prognosis.

Tumor is composed of cancerous cells and their surrounding complex stroma, stroma play a paramount role in tumor genesis and tumor metastasis (19). The occurrence and development of tumor is a complex process in which cancerous cells and stroma continuously interact and influence each other (20). There are complex stroma surrounding the tumor, including blood and lymphatic vessels, fibroblasts, endothelial cells, immune cells, cytokines, extra-cellular vesicles, and extra-cellular matrix (21). The extra-cellular matrix is the primary component of stroma. Apart from that, its function is not only limited to connect, support and protect cells, but also play a crucial role in cell proliferation and differentiation, cell migration, cell adhesion, angiogenesis, tissue development and repair (22–25). Aside from that, the accumulation and quantity of extra-cellular matrix components can result in tumor tissue fibrosis, which is required for wound healing and tissue repair to fight tissue infection, inflammation, autoimmune diseases, degenerative diseases, and tumors, resulting in prolonged survival (26). The stroma also contains a multitude of immune cells, such as NK cells, CD8+ cytotoxic T cells and antigen presenting cells (APCs), which can inhibit tumor growth and the immune response generated by the immune cells has anti-tumor properties (27). The stroma affects tumor growth and biological behavior. What's more, the value of

stroma in tumor prognosis prediction has been paid an heightening amount of attention (28). Tumor-stromal ratio (TSR) is a common stroma statistical method by calculation the ratio of area occupation of the stroma to tumor tissue. For the time being, a myriad of studies (29–31) have used the TSR in the prediction of postoperative prognosis of tumor patients. Furthermore, the results show that the TSR is independent predictor of prognosis in cancer patients. In the practice of Chinese doctors, those tumor with rich stroma are called hard cancers, and with poor stroma are called soft cancers. Aside from that, they found that the prognosis of hard cancers is better than that of soft cancers. In this investigation, we found a novel classification criteria by comparing the correlation between the richness of the stroma (tumor-stromal ratio) and survival in patients with ESCC. This classification criteria can be adopted to group patients by adopting TSR, which provides a basis for subsequent studies on the role of stroma in tumor development. In line with the classification criteria in this research, the stroma-rich group survived remarkably better than stroma-poor group. Apart from that, we also found that in the stromal rich group, the differentiation of patients was better than that in the stromal poor group, nevertheless, the depth of invasion was higher than that in the stromal poor group. With this classification criteria, the survival and clinical characteristics of patients in the two groups differed strikingly, which demonstrated that the classification criteria had great value. As a consequence, the TSR was incorporated into the construction of the prognostic model.

Esophageal cancer in China has the characteristics of conspicuous familial aggregation phenomenon and marked regional distribution differences (forming apparent high and low incidence areas) (1, 32). As already demonstrated by the epidemiological studies, individuals with a family history of esophageal cancer have a higher probability of
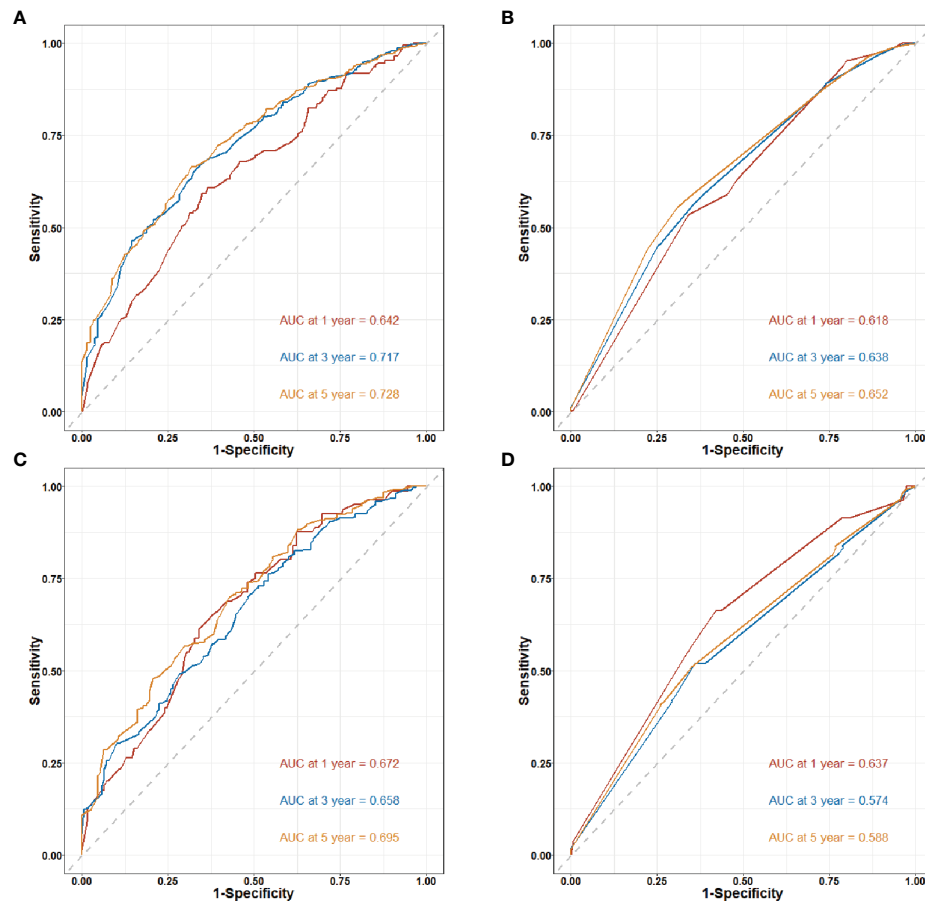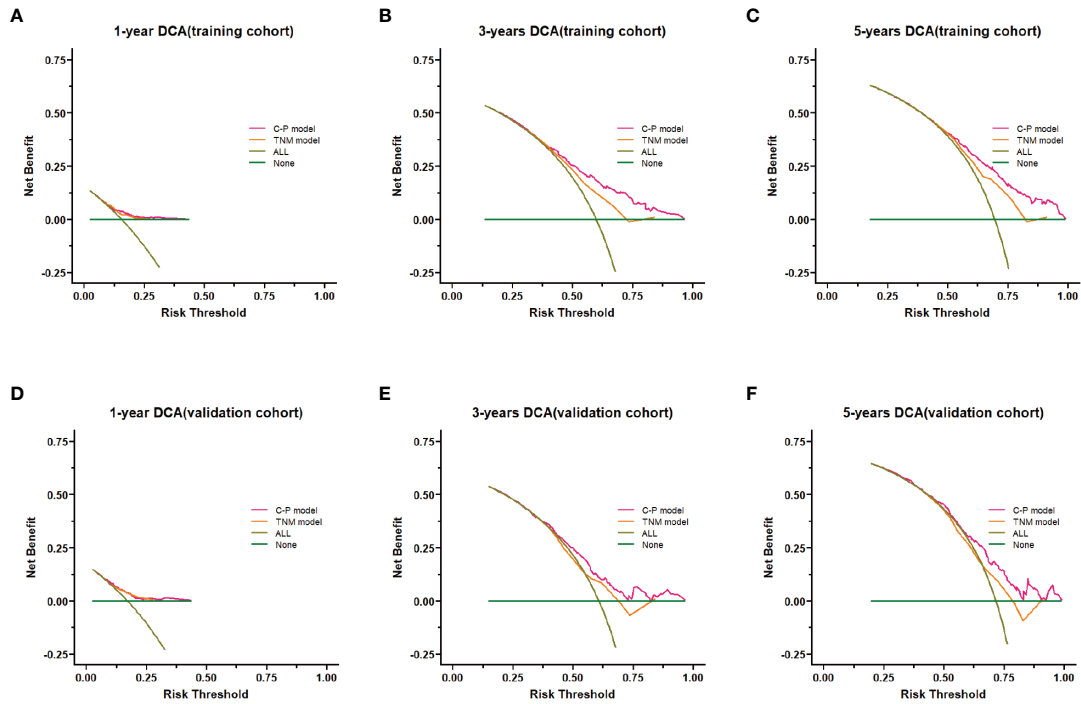
FIGURE 5
Time-dependent receiver operating characteristic (ROC) curves for nomogram in training and validation cohort. AUC. Area under the ROC curve; CI: confidence interval. **(A)** 1-year, 3-year and 5-year ROC curves of the clinical-pathological nomogram model in the training cohort. **(B)** 1-year, 3-year and 5-year ROC curves of the TNM stage nomogram model in the training cohort. **(C)** 1-year, 3-year and 5-year ROC curves of the clinical-pathological nomogram model in the validation cohort. **(D)** 1-year, 3-year and 5-year ROC curves of the TNM stage nomogram model in the validation cohort.

developing esophageal cancer and a worse prognosis than individuals without a family history (33, 34). The reasons of higher risk of individuals with a positive family history may be on account of genetic inheritance. Apart from that, patients with a family history of esophageal cancer are more likely to have a genetic predisposition than those without a family history of cancer (34). Since family history is a reflection of an individual genetic risk factor, and it is believed that family history can be used as an indicator to be included in the predictive model for identification of high-risk groups (35). What's more, the environment is also a pivotal factor affecting the occurrence of esophageal cancer. In China, there are striking geographical differences in the incidence of esophageal cancer, with pronounced high-incidence areas and low-incidence areas, the natural environment and living customs in diverse regions are crucial factors affecting the incidence of esophageal cancer (1, 32). As persuasively demonstrated by relevant studies (34, 36), environmental and genetic risk factors may exert collective influence on individual morbidity risk and prognosis survival. Shared genetic susceptibility and environmental exposure, or the possibility of their interaction, may be paramount factors in the development and progression of esophageal cancer. In this exploration, we found that patients with negative family history and patients in low-incidence areas had markedly better survival than

patients with positive family history and patients in high-incidence areas ($P < 0.05$), so high and low incidence areas were incorporated into the model construction. Furthermore, age, drinking history, differentiation, T stage, and N stage are evidently prognostic factors that affect the survival of patients with esophageal cancer (37, 38). What's more, they were also included in the construction of the prognostic model in this research.

In this investigation, we constructed a nomogram model to predict the prognosis of patients with ESCC. The nomogram has been extensively employed as a predictive method in oncology (39–41), it meets the requirements of an ensemble model and plays a role in driving personalized medicine for clinicians to use for prognosis prediction (42). Traditionally, tumor staging on the basis of the AJCC/UICC criteria has been the initial choice for predicting the prognosis of ESCC. Nevertheless, diverse prognoses were observed in patients at the same stage. This prognostic heterogeneity can be explained by age, family history, high and low incidence regions, and other factors that are not considered in tumor staging based upon AJCC/UICC criteria. This study included the largest sample size of patients with ESCC and the longest follow-up time to explored the association between TSR and clinical characteristics and prognosis of patients with ESCC, clarified that TSR is an crucial factor affecting the prognosis of patients with ESCC, and developed a novel criteria for

FIGURE 6

Decision curve analysis (DCA) of the nomogram model. **(A–C)** DCA curves of 1-,3- and 5-years survival prediction for the clinical-pathological nomogram model and the TNM stage nomogram model in the training cohort. **(D–F)** DCA curves of 1-,3- and 5-years survival predictions for the clinical-pathological nomogram model and the TNM stage nomogram model in the validation cohort. C-P model: clinical-pathological nomogram model.

the classification of TSR and include TSR in the nomogram model for patients with ESCC for the first time. A clinical-pathological nomogram model was constructed by combining the clinical information (age, drinking history, family history, and high-low risk area) and pathological information (TSR, differentiation, T stage and N stage) of patients. The C-index and NRI of the internally validated clinical-pathological model constructed in this

exploration were remarkably higher than those of the TNM staging. In the training set and validation set, the 3-year and 5-year predictive ability of the clinical-pathological nomogram model was found to be strikingly better than that of the TNM staging (P < 0.001), indicating that the predictive performance of nomogram model constructed in this investigation was superior to that of TNM analysis. The results of decision curve analysis (DCA) indicated that the clinical-pathological
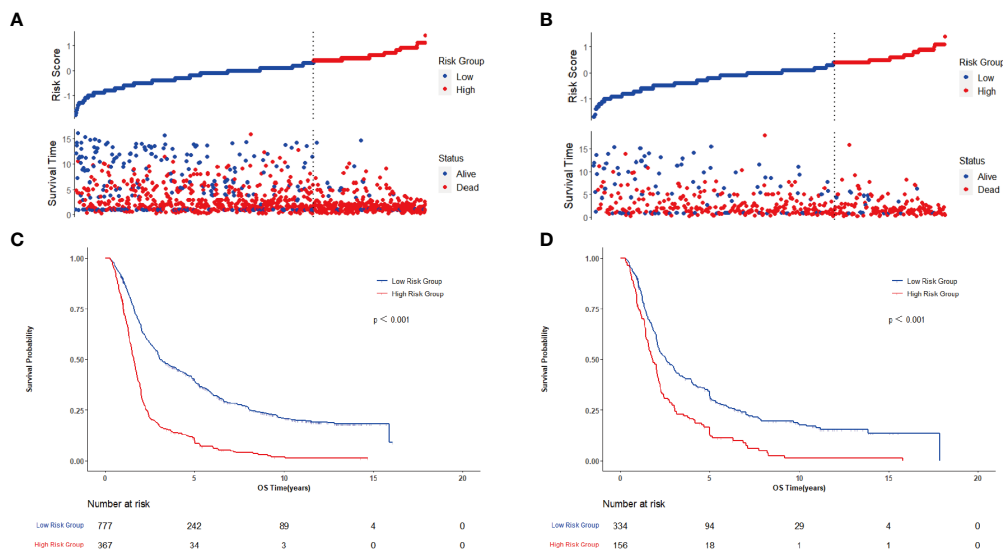
FIGURE 7

Evaluates of clinical-pathological nomogram model performance in training cohort and confirms on the basis of validation cohort. **(A)** Distribution of risk scores and survival status in training cohort; **(B)** Distribution of risk scores and survival status in validation cohort; **(C)** Kaplan-Meier survival curve in training cohort; **(D)** Kaplan-Meier survival curve in validation cohort.

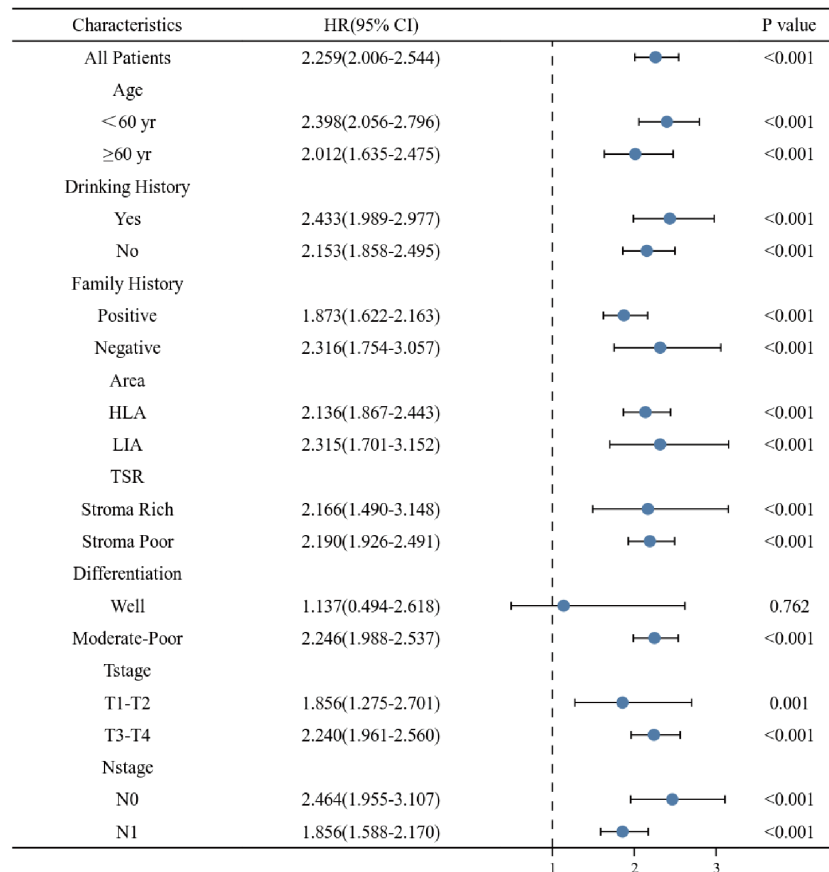| Characteristics	HR(95% CI)		P value
All Patients	2.259(2.006-2.544)		<0.001
Age			
<60 yr	2.398(2.056-2.796)		<0.001
≥60 yr	2.012(1.635-2.475)		<0.001
Drinking History			
Yes	2.433(1.989-2.977)		<0.001
No	2.153(1.858-2.495)		<0.001
Family History			
Positive	1.873(1.622-2.163)		<0.001
Negative	2.316(1.754-3.057)		<0.001
Area			
HLA	2.136(1.867-2.443)		<0.001
LIA	2.315(1.701-3.152)		<0.001
TSR			
Stroma Rich	2.166(1.490-3.148)		<0.001
Stroma Poor	2.190(1.926-2.491)		<0.001
Differentiation			
Well	1.137(0.494-2.618)		0.762
Moderate-Poor	2.246(1.988-2.537)		<0.001
Tstage			
T1-T2	1.856(1.275-2.701)		0.001
T3-T4	2.240(1.961-2.560)		<0.001
Nstage			
N0	2.464(1.955-3.107)		<0.001
N1	1.856(1.588-2.170)		<0.001

FIGURE 8

Stratified analysis of the clinical-pathological nomogram model in different subgroups.

model had a better net benefit than TNM staging in both the training and validation sets for 3-year and 5-year OS prediction. Based upon the risk scores of patients predicted by the nomogram model, patients were divided into two groups (high-risk and low-risk). Aside from that, the survival of patients in the high-risk group was conspicuously worse than that in the low-risk group (P < 0.001); in the subgroup analysis, the survival of patients in the high-risk group was noticeably worse than that in the low-risk group (P < 0.001). As evidently displayed by all these the research findings, the prognostic prediction model constructed in this research has better predictive and clinical application value than TNM staging.

There were some limitations to our research. First and foremost, our analysis may have been affected by bias owing to the retrospective of the study. Aside from that, in this exploration, whether the patients received chemoradiotherapy before surgery was considered, but may receive dissimilar treatments during the long postoperative course, but this factor was not addressed detailedly in our present study. Furthermore, this study used TMAs to observe TSR, which has the limitation of under-representation compared to whole sections.

5 Conclusion

This study confirmed that the TSR was an independent prognostic factor in patients with ESCC. Combined with tumor-

stromal ratio and clinical and pathological factors, a prognostic nomogram is constructed, which has higher accuracy and better clinical value than traditional TNM stage. Furthermore, it provides a method for clinicians to predict ESCC prognosis and to adjust the treatment strategy.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Committee of The First Affiliated Hospital of Zhengzhou University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

LDW, JFH and XS conceptualized and designed the study. FYZ, JLL, XZW, XML, RW, ZMF, XNH, BL, HJY, FBC, JLR and SLZ

provided tumor samples collection and assembly. XKZ, RHX, LLL, YC, LYL JJJ, YZY and MMY provided patients' information collection and follow-up. JFH, XS and KZ provided analyzed and interpreted the data. JFH, PPW and MXW drafted the manuscript. All authors reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal cancer. *Lancet* (2017) 390(10110):2383–96. doi: 10.1016/S0140-6736(17)31462-9

2. Zhao XK, Xing P, Song X, Zhao M, Zhao L, Dang Y, et al. Focal amplifications are associated with chromothripsis events and diverse prognoses in gastric cardia adenocarcinoma. *Nat Commun* (2021) 12(1):6489. doi: 10.1038/s41467-021-26745-3

3. Chen T, Cheng H, Chen X, Yuan Z, Yang X, Zhuang M, et al. Family history of esophageal cancer increases the risk of esophageal squamous cell carcinoma. *Sci Rep* (2015) 5:16038. doi: 10.1038/srep16038

4. Wang S, Yang L, Ci B, Maclean M, Gerber DE, Xiao G, et al. Development and validation of a nomogram prognostic model for SCLC patients. *J Thorac Oncol* (2018) 13 (9):1338–48. doi: 10.1016/j.jtho.2018.05.037

5. Adam MA, Thomas S, Hyslop T, Scheri RP, Roman SA, Sosa JA. Exploring the relationship between patient age and cancer-specific survival in papillary thyroid cancer: Rethinking current staging systems. *J Clin Oncol* (2016) 34(36):4415–20. doi: 10.1200/JCO.2016.68.9372

6. Pan JJ, Ng WT, Zong JF, Lee SW, Choi HC, Chan LL, et al. Prognostic nomogram for refining the prognostication of the proposed 8th edition of the AJCC/UICC staging system for nasopharyngeal cancer in the era of intensity-modulated radiotherapy. *Cancer* (2016) 122(21):3307–15. doi: 10.1002/cncr.30198

7. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013

8. Lv Z, Cai X, Weng X, Xiao H, Du C, Cheng J, et al. Tumor-stroma ratio is a prognostic factor for survival in hepatocellular carcinoma patients after liver resection or transplantation. *Surgery* (2015) 158(1):142–50. doi: 10.1016/j.surg.2015.02.013

9. Chen Y, Zhang L, Liu W, Liu X. Prognostic significance of the tumor-stroma ratio in epithelial ovarian cancer. *BioMed Res Int* (2015) 2015:589301. doi: 10.1155/2015/589301

10. Guanrei Y, Songliang Q. Endoscopic surveys in high-risk and low-risk populations for esophageal cancer in China with special reference to precursors of esophageal cancer. *Endoscopy* (1987) 19(3):91–5. doi: 10.1055/s-2007-1018250

11. Liu CL, Prapong W, Natkunam Y, Alizadeh A, Montgomery K, Gilks CB, et al. Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. *Am J Pathol* (2002) 161(5):1557–65. doi: 10.1016/S0002-9440(10)64434-3

12. Zhang D, Salto-Tellez M, Putti TC, Do E, Koay ES. Reliability of tissue microarrays in detecting protein expression and gene amplification in breast cancer. *Mod Pathol* (2003) 16(1):79–84. doi: 10.1097/01.MP.0000047307.96344.93

13. Cooper T, Biron VL, Adam B, Klimowicz AC, Puttagunta L, Seikaly H. Association of keratinization with 5-year disease-specific survival in oropharyngeal squamous cell carcinoma. *JAMA Otolaryngol Head Neck Surg* (2015) 141(3):250–6. doi: 10.1001/jamaoto.2014.3335

14. Camp RL, Dolled-Filhart M, Rimm DL. X-Tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* (2004) 10(21):7252–9. doi: 10.1158/1078-0432.CCR-04-0713

15. Li L, Greene T, Hu B. A simple method to estimate the time-dependent receiver operating characteristic curve and the area under the curve with right censored data. *Stat Methods Med Res* (2018) 27(8):2264–78. doi: 10.1177/0962280216680239

16. Talluri R, Shete S. Using the weighted area under the net benefit curve for decision curve analysis. *BMC Med Inform Decis Mak* (2016) 16:94. doi: 10.1186/s12911-016-0336-x

17. Abnet CC, Arnold M, Wei WQ. Epidemiology of esophageal squamous cell carcinoma. *Gastroenterology* (2018) 154(2):360–73. doi: 10.1053/j.gastro.2017.08.023

18. Thrumurthy SG, Chaudry MA, Thrumurthy SSD, Mughal M. Oesophageal cancer: risks, prevention, and diagnosis. *BMJ* (2019) 366:l4373. doi: 10.1136/bmj.l4373

19. Casazza A, Di Conza G, Wenes M, Finisguerra V, Deschoemaeker S, Mazzone M. Tumor stroma: A complexity dictated by the hypoxic tumor microenvironment. *Oncogene* (2014) 33(14):1743–54. doi: 10.1038/onc.2013.121

20. Salvatore V, Teti G, Focaroli S, Mazzotti MC, Mazzotti A, Falconi M. The tumor microenvironment promotes cancer progression and cell migration. *Oncotarget* (2017) 8 (6):9608–16. doi: 10.18632/oncotarget.14155

21. Wei R, Liu S, Zhang S, Min L, Zhu S. Cellular and extracellular components in tumor microenvironment and their application in early diagnosis of cancers. *Anal Cell Pathol (Amst)* (2020) 2020:6283796. doi: 10.1155/2020/6283796

22. Khawar IA, Kim JH, Kuh HJ. Improving drug delivery to solid tumors: Priming the tumor microenvironment. *J Control Release* (2015) 201:78–89. doi: 10.1016/j.jconrel.2014.12.018

23. Nissen NI, Karsdal M, Willumsen N. Collagens and cancer associated fibroblasts in the reactive stroma and its relation to cancer biology. *J Exp Clin Cancer Res* (2019) 38 (1):115. doi: 10.1186/s13046-019-1110-6

24. Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. *Adv Drug Delivery Rev* (2016) 97:4–27. doi: 10.1016/j.addr.2015.11.001

25. Karamanos NK, Theocharis AD, Piperigkou Z, Manou D, Passi A, Skandalis SS, et al. A guide to the composition and functions of the extracellular matrix. *FEBS J* (2021) 288(24):6850–912. doi: 10.1111/febs.15776

26. Zhang M, Zhang S. T Cells in fibrosis and fibrotic diseases. *Front Immunol* (2020) 11:1142. doi: 10.3389/fimmu.2020.01142

27. Helfen A, Roth J, Ng T, Eisenblaetter M. *In vivo* imaging of pro- and antitumoral cellular components of the tumor microenvironment. *J Nucl Med* (2018) 59(2):183–8. doi: 10.2967/jnumed.117.198952

28. Hussain S, Peng B, Cherian M, Song JW, Ahirwar DK, Ganju RK. The roles of stroma-derived chemokine in different stages of cancer metastases. *Front Immunol* (2020) 11:598532. doi: 10.3389/fimmu.2020.598532

29. Yan D, Ju X, Luo B, Guan F, He H, Yan H, et al. Tumour stroma ratio is a potential predictor for 5-year disease-free survival in breast cancer. *BMC Cancer* (2022) 22(1):1082. doi: 10.1186/s12885-022-10183-5

30. Wang K, Ma W, Wang J, Yu L, Zhang X, Wang Z, et al. Tumor-stroma ratio is an independent predictor for survival in esophageal squamous cell carcinoma. *J Thorac Oncol* (2012) 7(9):1457–61. doi: 10.1097/JTO.0b013e318260dfe8

31. Vangangelt KMH, van Pelt GW, Engels CC, Putter H, Liefers GJ, Smit VTHBM, et al. Prognostic value of tumor-stroma ratio combined with the immune status of tumors in invasive breast carcinoma. *Breast Cancer Res Treat* (2018) 168(3):601–12. doi: 10.1007/s10549-017-4617-6

32. Lin Y, Totsuka Y, Shan B, Wang C, Wei W, Qiao Y, et al. Esophageal cancer in high-risk areas of China: Research progress and challenges. *Ann Epidemiol* (2017) 27 (3):215–21. doi: 10.1016/j.annepidem.2016.11.004

33. Wang F, Fan L, Zhao Q, Liu Y, Zhang Z, Wang D, et al. Family history of malignant tumor is a predictor of gastric cancer prognosis: Incorporation into a nomogram. *Med (Baltimore)* (2022) 101(35):e30141. doi: 10.1097/MD.0000000000030141

34. Su Z, Zou GR, Mao YP, OuYang PY, Cao XL, Xie FY, et al. Prognostic impact of family history of cancer in southern Chinese patients with esophageal squamous cell cancer. *J Cancer* (2019) 10(6):1349–57. doi: 10.7150/jca.26511

35. Chen W, Li H, Ren J, Zheng R, Shi J, Li J, et al. Selection of high-risk individuals for esophageal cancer screening: A prediction model of esophageal squamous cell carcinoma based on a multicenter screening cohort in rural China. *Int J Cancer* (2021) 148(2):329–39. doi: 10.1002/ijc.33208

36. Jia N, Wen X, Zhang N, Yang Y, Zhang L, Wang X, et al. Younger age of onset and multiple primary lesions associated with esophageal squamous cell carcinoma cases with a positive family history of the cancer suggests genetic predisposition. *Chin Med J (Engl)* (2014) 127(15):2779–83.

37. Hu SJ, Zhao XK, Song X, Lei LL, Han WL, Xu RH, et al. Preoperative maximal voluntary ventilation, hemoglobin, albumin, lymphocytes and platelets predict postoperative survival in esophageal squamous cell carcinoma. *World J Gastroenterol* (2021) 27(4):321–35. doi: 10.3748/wjg.v27.i4.321

38. Chen P, Zhao X, Zhou F, Song X, Hu S, Jin Y, et al. Characterization of 500 Chinese patients with cervical esophageal cancer by clinicopathological and treatment outcomes. *Cancer Biol Med* (2020) 17(1):219–26. doi: 10.20892/j.issn.2095-3941.2019.0268

39. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: More than meets the eye. *Lancet Oncol* (2015) 16(4):e173–80. doi: 10.1016/S1470-2045 (14)71116-7

40. Yang J, Pan Z, Zhou Q, Liu Q, Zhao F, Feng X, et al. Nomogram for predicting the survival of patients with malignant melanoma: A population analysis. *Oncol Lett* (2019) 18(4):3591–8. doi: 10.3892/ol.2019.10720

41. Kong J, Zheng J, Cai J, Wu S, Diao X, Xie W, et al. A nomogram for individualized estimation of survival among adult patients with adrenocortical carcinoma after surgery: A retrospective analysis and multicenter validation study. *Cancer Commun (Lond)* (2019) 39(1):80. doi: 10.1186/s40880-019-0426-0

42. Wu J, Zhang H, Li L, Hu M, Chen L, Xu B, et al. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: A population-based analysis. *Cancer Commun (Lond)* (2020) 40(7):301–12. doi: 10.1002/cac2.12067