Original Article

Clinicopathological significance of cancer stem cell marker CD44/SOX2 in esophageal squamous cell carcinoma (ESCC) patients and construction of a nomogram to predict overall survival

Siyue Tian[1]^, Ruibin Ma[2], Yingmin Liu[1], Fei Chen[1], Xiaotong Huang[1], Qianqian Yang[1], Wei Nian[1], Zhiqin Fan[1,3]

[1]Department of Daily Surgery, Affiliated Tumor Hospital, Xinjiang Medical University, Urumqi, China; [2]Department of Characteristic Specialty One Group, Xinjiang Municipal Corps Hospital of the Chinese People's Armed Police Force, Urumqi, China; [3]State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia, Xinjiang Medical University, Urumqi, China

Contributions: (I) Conception and design: Z Fan, S Tian; (II) Administrative support: Z Fan; (III) Provision of study materials or patients: R Ma, X Huang; (IV) Collection and assembly of data: S Tian, Q Yang; (V) Data analysis and interpretation: Y Liu, F Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Zhiqin Fan, PhD. Candidate, Department of Daily Surgery, Affiliated Tumor Hospital, Xinjiang Medical University, 789 Suzhou East Street, Xinshi District, Urumqi 830000, China; State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia, Xinjiang Medical University, Urumqi, China. Email: 491962188@qq.com.

Background: Esophageal squamous cell carcinoma (ESCC), a prevalent malignancy within the upper gastrointestinal system, is characterized by its unfavorable prognosis and the absence of specific indicators for outcome prediction and high-risk case identification. In our research, we examined the expression levels of cancer stem cells (CSCs), markers CD44/SOX2 in ESCC, scrutinized their association with clinicopathological parameters, and developed a predictive nomogram model. This model, which incorporates CD44/SOX2, aims to forecast the overall survival (OS) of patients afflicted with ESCC.

Methods: Immunohistochemistry was utilized to detect the expression levels of CD44 and SOX2 in both cancerous and paracancerous tissues of 68 patients with ESCC. The correlation between CD44/SOX2 expression and clinicopathological parameters was subsequently analyzed. Factors impacting the prognosis of ESCC patients were assessed through univariate and multivariate Cox regression analyses. Leveraging the results of these multivariate regression analyses, a nomogram prognostic model was established to provide individualized predictions of ESCC patient survival outcomes. The predictive accuracy of the nomogram prognostic model was evaluated using the consistency index (C-index) and calibration curves.

Results: The expression levels of CD44 were markedly elevated in the tumor tissues of ESCC patients. Similarly, SOX2 was significantly overexpressed in the tumor tissues of ESCC patients. The positive expression of SOX2 in ESCC demonstrated a strong correlation with both the pathological T-stage and the presence of carcinoembryonic antigen. CD44 and SOX2 co-positive expression was significantly associated with the pathological T-stage and tumor node metastasis (TNM) stage. Furthermore, ESCC patients exhibiting CD44-positive expression in their tumor tissue generally had a more adverse prognosis. The co-expression of CD44 and SOX2 resulted in a grimmer prognosis compared to patients with other combinations. Multivariate Cox regression analysis identified the co-expression of CD44 and SOX2, the pathological T-stage, and lymph node metastasis as independent prognostic indicators for ESCC patients. The three identified variables were subsequently incorporated into a nomogram for predicting OS. The C-index of the measurement model and the area under the curve of the subjects' work characteristics showed good individual prediction. This prognostic model stratified patients into low- and high-risk categories. Analysis revealed that the 5-year OS rate was significantly higher in the low-risk group compared to the high-risk group.

^ ORCID: 0009-0006-3346-724X.

Conclusions: Elevated CD44 levels, indicative of CSC presence, are intimately linked with the oncogenesis of ESCC and are strongly predictive of unfavorable patient outcomes. Concurrently, the SOX2 gene exhibits a heightened expression in ESCC, markedly accelerating tumor progression and fostering more extensive disease infiltration. The co-expression of CD44 and SOX2 correlates significantly with ESCC patient prognosis, serving as a reliable, independent prognostic marker. Our constructed nomogram, incorporating CD44/SOX2 expression, enhances the prediction of OS and facilitates risk stratification in ESCC patients.

Keywords: Esophageal squamous cell carcinoma (ESCC); CD44; SOX2; cancer stem cells (CSCs); nomogram model

Submitted Dec 16, 2023. Accepted for publication Apr 24, 2024. Published online May 30, 2024.
doi: 10.21037/tcr-23-2313
View this article at: https://dx.doi.org/10.21037/tcr-23-2313

Introduction

Background

Esophageal carcinoma (EC) ranks as one of the predominant malignancies within the upper gastrointestinal system worldwide. Globally, esophageal cancer is recognized as the seventh most prevalent cancer and the sixth leading cause of cancer-related mortality (1). Pathologically, esophageal cancer is bifurcated into two primary subtypes: esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC). ESCC is characterized by unique pathogenetic features, exhibiting insidious as well as aggressive early symptoms. At the time of initial consultation, approximately 70% of patients present with local infiltration or distant metastasis. Despite substantial advancements in diagnostic and therapeutic strategies for ESCC, the prognosis remains dismal. The 5-year survival rate for patients with ESCC is less than 15% (2). Given this context, there is an escalating interest in investigating ESCC-specific biomarkers. The objectives of such studies are twofold: to prognosticate patient outcomes and to identify high-risk cases warranting immediate attention.

Highlight box

Key findings
- Pathological features based on CD44/SOX2 can contribute to prognostic evaluation and prediction in patients with oesophageal squamous carcinoma.

What is known and what is new?
- CD44 and SOX2, recognized as cancer stem cell markers, are integral to the processes of tumor self-renewal and oncogenesis. Their involvement extends to a multitude of biological procedures within the realm of cancer.
- A prognostic nomogram model combining CD44/SOX2 co-expression and clinicopathological factors based on the clinicopathological significance of the tumor stem cell marker CD44/SOX2 in esophageal squamous cell carcinoma (ESCC) was constructed. This provides risk stratification for the prognosis of patients with esophageal squamous carcinoma, provides a basis for the development of personalized follow-up strategies after surgery, and provides theoretical guidance for precise targeted therapy for patients.

What is the implication, and what should change now?
- Identification of CD44/SOX2 co-positive expression predicted survival and identified high-risk cases in ESCC patients.
- The CD44/SOX2 co-positive effect should be validated in additional datasets or prospective studies for further applications.

Rationale and knowledge gap

The tumor stem cell hypothesis posits that cancerous tissues harbor cancer stem cells (CSCs), which orchestrate tumor proliferation. These CSCs, constituting less than 5% of the total tumor cell population, are characterized by their self-renewal capacity, unrestricted proliferation, and differentiation potential (3). They play a pivotal role in disease progression, metastasis, and recurrence in patients who have undergone successful initial chemotherapy or radiation therapy (4,5). CSCs are propelled by factors such as cellular plasticity, senescence and quiescence, augmented DNA repair capability, deregulation of anti-apoptotic proteins, and modifications in cell cycle dynamics and the tumor microenvironment. These factors enable CSCs to sustain their self-renewal capacity and resist the stress of the tumor microenvironment and therapy (6,7). In esophageal squamous carcinoma, CSCs can be discerned using specific

cell surface and intracellular markers. For instance, cell surface markers like CD44, CD90, and CD133 are employed to identify esophageal squamous carcinoma CSCs (8).

CD44, a set of transmembrane glycoproteins located on the cell surface, undergoes post-translational modifications and selective splicing that influence its binding affinity to ligands, thereby regulating the activity of CSCs (9). It is predominantly expressed in tumor cells as an extracellular matrix adhesion protein, playing a crucial role in their adhesion and migration (10). As the most prevalent marker for CSCs, CD44's expression is directly linked with the self-renewal and tumorigenic capacity of CSCs, contributing significantly to tumorigenesis (11). Numerous studies have indicated that heightened CD44 expression can escalate the tumorigenic potential in various cancers such as colorectal (12), breast (13) and gastric (14) cancers. Moreover, CD44 is extensively overexpressed in other cancer types like ovarian cancer (15) and oral squamous carcinoma (16), in which it is associated with aggressive biological behavior leading to a poor prognosis.

SOX2, a member of the HMG protein family, plays a pivotal role in maintaining the self-renewal and multidirectional differentiation of stem cells (17). It is identified as one of the CSCs markers, including Nanog and SALL4, of which its overexpression enhances the properties of tumor stem cells (18). In various cancers such as bladder cancer, enhanced expression of SOX2 has been found to be associated with an increase in tumor cell stemness and the maintenance of their self-renewal capacity (19). The long non-coding RNA (lncRNA) SOX2-OT, localized in the cytoplasm, has been found to promote SOX2 expression, thereby augmenting bladder tumor cell stemness and progression (20). Moreover, studies have demonstrated that overexpression of SOX2 significantly amplifies ESCC cell migration, invasion, and resistance to cisplatin. This correlation with poor patient prognosis underlines its potential value in therapeutic interventions for cancer (21). Research indicates that both SOX2 amplification and the acquisition of chromosome 3 are phenomena observed in ESCC. These occurrences serve as potential biomarkers, linked with tumor progression and risk stratification within ESCC cases (22,23).

Objective

This research employed immunohistochemistry (IHC), survival analysis, and regression analysis to explore the association between CD44/SOX2 protein expression and the clinicopathological traits and prognosis of patients. We developed a CD44-centric nomogram to forecast individual overall survival (OS) rates, thereby offering a potential stem cell marker compilation for evaluating patient clinical outcomes. We present this article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2313/rc).

Methods

Patients and clinical samples

A total of 68 ESCC specimens were randomly selected from patients who underwent surgical treatment at the Department of Clinical Pathology in Xinjiang Medical University Affiliated Tumor Hospital (Xinjiang Uygur Autonomous Region) between January 1st, 2016 and December 31st, 2016. Each patient had comprehensive clinical, pathological, and follow-up records. The follow-up period extended until June 2021 or until the patient's death, with an average duration of 34 months and survival times ranging from 4 to 68 months. Every selected case was diagnosed as ESCC by two pathologists. The esophageal squamous carcinoma stages were determined based on histological classification, degree of differentiation, and detailed postoperative pathological staging in accordance with American Joint Committee on Cancer (AJCC)/International Union against Cancer Tumor Node Metastasis (TNM) classification system. The associated clinicopathological characteristics are enumerated in *Table 1*. The study's inclusion criteria were as follows: (I) first-time oesophageal cancer patients treated at our hospital; (II) preoperative confirmation of squamous oesophageal cancer via imaging, endoscopy, and histological biopsy; (III) absence of other interventions such as neoadjuvant radiotherapy or chemotherapy prior to surgery; and (IV) availability of complete clinicopathological data. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of Xinjiang Medical University Affiliated Tumor Hospital (No. K-2019053) and informed consent was obtained from all individual participants.

Immunohistochemical analysis

IHC was employed to ascertain the expression levels of SOX2 and CD44 proteins in ESCC and adjacent tissue specimens. The respective antibodies used for CD44 and

Table 1 Baseline characteristics of patients

Patient characteristic	Frequency (n)	Percentage (%)
Age (years)		
≤62	33	48.53
>62	35	51.47
Sex		
Male	46	67.65
Female	22	32.35
CEA		
Negative	64	94.12
Positive	4	5.88
SCC		
Negative	61	89.71
Positive	7	10.29
Tumor size (cm)		
≤4	36	52.94
>4	32	47.06
Histological grade		
Grade 1	3	4.41
Grade 2	43	63.24
Grade 3	22	32.35
Nerve invasion		
Negative	49	72.06
Positive	19	27.94
pT stage		
pT 1	8	11.75
pT 2	8	11.75
pT 3	51	75.00
pT 4	1	1.47
N categories		
N0	31	45.59
N1	19	27.94
N2	15	22.06
N3	3	4.41
TNM stage		
Stage I	6	8.82
Stage II	27	39.71
Stage III	35	51.47
CD44		
Negative	19	27.94
Positive	49	72.06
SOX2		
Negative	24	35.29
Positive	44	64.71
$CD44^+/SOX2^+$		
Negative	36	52.94
Positive	32	47.06

Age is based on a median age of 62 years. CEA, carcinoembryonic antigen; SCC, squamous cell carcinoma.

SOX2 were procured from Zhongshan Jinqiao, China (ZM-0052) and Proteintech, USA (11064-1-AP). A cell membrane staining pattern signified positive CD44 expression in ESCC and normal oesophageal mucosal cells, while a nuclear staining pattern indicated positive SOX2 expression. In each section, ten stained areas were randomly selected under a high magnification field of view (×400), and the positive staining was evaluated using an immunoreactivity score, which was calculated as the product of staining intensity and the percentage of positively stained cells. The intensity of staining was rated on a scale of 0 to 2: weak staining [0]; moderate staining [1]; strong staining [2]. The percentage of positively stained cells was assessed and categorized as follows: 0% [0], 1–10% [1], 11–50% [2], 51–75% [3], and 76–100% [4]. Scores were interpreted by two senior pathologists in a double-blinded manner and averaged. A score of 3 or higher was deemed positive, while a score between 0 and 2 was considered negative (24).

Statistical analyses

The experiment's results were examined using SPSS 26.0 statistical software. Measurements are expressed as mean ± standard deviation; conforming to normal distribution, differences between the two groups were tested by independent samples *t*-test. The Chi-squared test was used to compare the count data with the following rules (N is the total number of cases, and E is the expected value of each cell): when N<40 or E <1, the analysis was performed using the Fisher exact test; when N≥40 and 1≤ E ≤5, the

analysis was performed using the successive-corrected chi-square test; and when N≥40 and E≥5, the Pearson chi-square analysis was performed. The correlation between the expression levels of CD44/SOX2 and OS in patients with ESCC was evaluated utilizing Kaplan-Meier survival analysis, complemented by a log-rank test. Hazard ratios, accompanied by 95% confidence intervals, were utilized for analysis. A P value of less than 0.05 was deemed statistically significant.

Construction of nomogram

The analysis of experimental results was conducted using R software (version 4.2.1). The initial stage involved testing the proportional risk hypothesis through the survival package (3.3.1). Subsequently, the Cox proportional risk regression model was employed for a univariate analysis of potential risk factors associated with OS in patients diagnosed with squamous esophageal cancer. Variables that yielded P values less than 0.1 in the univariate analysis, as well as gender and age, were selected for further examination through multivariate Cox regression analysis. The rms package (6.3-0) was then utilized to construct a nomogram prediction model, based on the multivariate Cox regression analysis results, for forecasting OS in patients suffering from esophageal squamous carcinoma. The predictive capability of the nomogram was evaluated using the concordance index (C-index). To assess the consistency between observed and predicted outcomes, a calibration plot was created using 800 bootstrap resamples.

Results

The association between the expression levels of CD44 and SOX2 with clinicopathological parameters.

IHC was utilized to examine the expression of CD44 and SOX2 in both cancerous and paracancerous tissue samples obtained from ESCC patients. CD44 was discernible in both the cell membrane and cytoplasm, with the definition of positive expression being a cell membrane staining pattern. Out of 68 cases, positive CD44 staining was identified in 49 tumor tissue samples (72.06%) and 39 corresponding paracancerous tissue samples (57.35%), a difference which was statistically significant (P=0.004) (*Figure 1A*). Furthermore, SOX2 was observable in the nucleus. Positive SOX2 staining was found in 44 tumor tissue samples (64.71%) and only ten corresponding paracarcinomatous tissue samples (14.71%), again, a statistically significant difference (P<0.001) (*Figure 1B*). No significant association was discovered between the expression level of CD44 in ESCC and factors such as patient's gender, age, tumor size, pathological grade, depth of tumor infiltration, TNM stage, and lymph node metastasis. However, the expression level of SOX2 in ESCC demonstrated a correlation with pathological T-staging and carcinoembryonic antigen (CEA) expression level (P<0.05), but not with age, gender, or TNM staging. A significant correlation was established between CD44/SOX2 co-positive expression and both pathological T stage and TNM stage. Comprehensive statistical results are presented in *Table 2*.

Survival analysis

The study was conducted over a median follow-up period of 34 months, ranging from 4 to 68 months. The OS rate for all patients after 5 years was observed to be 25%. A significant relationship was identified between the expression of CD44 and the clinical prognosis in patients with ESCC. Kaplan-Meier analysis highlighted a strong association between CD44 expression in ESCC and the OS rate. Notably, patients with positive CD44 expression had a significantly reduced survival time compared to those with negative CD44 expression (log-rank =4.19, P=0.041, *Figure 2A*). Cases of ESCC exhibiting SOX2-positivity did not demonstrate a significant correlation with patient survival rates (log-rank =0.20, P=0.65, *Figure 2B*). However, the study revealed that patients with a positive co-expression of CD44 and SOX2 markers experienced a poorer prognosis compared to those expressing other marker combinations (log-rank =4.77, P=0.029, *Figure 2C*).

Construction of nomogram for OS in ESCC patients

Table 3 demonstrates that univariate Cox regression analysis identified squamous cell carcinoma (SCC), maximum diameter of tumor, pathological T-stage, lymphatic metastasis, TNM stage, and CD44/SOX2 co-expression as factors influencing OS (P<0.05). Multivariate analysis, however, pinpointed CD44/SOX2 co-expression, pathological T-stage, and lymphatic metastasis as independent prognostic indicators for OS in ESCC patients.

Utilizing these three independent prognostic indicators, age and gender, a nomogram predictive model was constructed to estimate OS rates, as illustrated in *Figure 3A*. The procedure for utilizing the nomogram plots to forecast

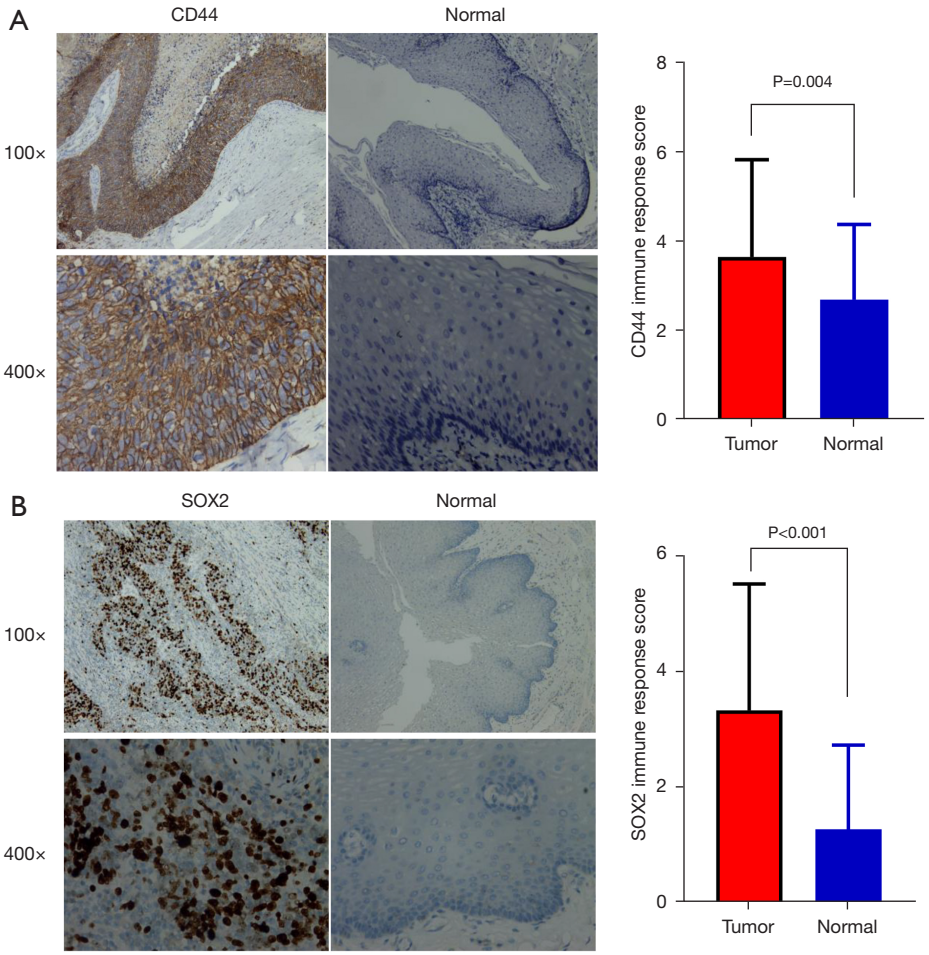

Figure 1 Immunohistochemical staining of CD44 and SOX2 in ESCC. (A) Staining of CD44 in ESCC cancer tissues and paracancerous tissues (left), immunohistochemical scoring of CD44 (right). (B) Staining of SOX2 in ESCC cancer tissues and paracancerous tissues (left), immunohistochemical scoring of SOX2 (right). ESCC, esophageal squamous cell carcinoma.

OS in patients is as follows: firstly, the values for CD44/SOX2 co-expression, pathological T-stage, lymphatic metastasis, age and gender are identified on their respective variable axes. A vertical line drawn upwards from the variable axis to the Points axis will provide the score for each variable. The total score is obtained by summing these five-factor scores and locating this sum on the total points axis. Lastly, a line drawn vertically downwards from the total points axis to the survival axis provides the predicted 1, 3, and 5-year OS rates. The nomogram model was used as follows: a patient with ESCC who had a pathological T-stage of T3/T4, then the tumor invaded or infiltrated the fibrous outer lining of the esophagus, which corresponded to a score of 100, and if it was a T1/T2 stage then the tumor growth was confined to the intrinsic muscular layer of the esophagus, which was predicted to be a score of zero. If an ESCC patient's N stage is N0+, then regional lymph node metastasis is seen, and the corresponding score is 37.5; if the patient's N stage is N0, then no regional lymph node metastasis is seen, and the corresponding score for this patient is 0. The corresponding score for gender is 32.5 for females and 0 for males. If a tissue section from 1 ESCC patient was immunohistochemically stained for $CD44^+$/$SOX2^+$ co-positive expression, then the predicted score was 40. If both CD44/SOX2 staining is negative or one negative and one positive, then the corresponding score is 0.

C-index and AUC of the nomogram model

The findings demonstrated that the nomogram prognostic

Table 2 Relationship between CD44/SOX2 and clinicopathological variables in ESCC

Variable	Cases (n=68)	CD44				SOX2				$CD44^+/SOX2^+$			
		Negative (n=19)	Positive (n=49)	χ^2 value	P value	Negative (n=24)	Positive (n=44)	χ^2 value	P value	Negative (n=36)	Positive (n=32)	χ^2 value	P value
Sex				0.24	0.62			3.08	0.08			3.03	0.08
Male	46	12	34			13	33			21	25		
Female	22	7	15			11	11			15	7		
Age (years)				0.01	0.91			0.70	0.40			1.44	0.23
≤62	33	9	24			10	23			15	18		
>62	35	10	25			14	21			21	14		
CEA				0.02	0.89			7.79	0.005			3.78	0.05
Negative	64	18	46			20	44			32	32		
Positive	4	1	3			4	0			4	0		
SCC				0.72	0.40			0.20	0.66			0.32	0.57
Negative	61	18	43			21	40			33	28		
Positive	7	1	6			3	4			3	4		
Maximum tumor diameter (cm)				0.00	0.98			1.36	0.24			3.68	0.06
≤4	36	10	26			15	21			23	13		
>4	32	9	23			9	23			13	19		
Nerve invasion				0.17	0.68			2.34	0.13			0.33	0.57
Negative	49	13	36			20	29			27	22		
Positive	19	6	13			4	15			9	10		
Histological grade				3.47	0.18			0.02	0.99			1.76	0.41
G1	3	0	3			1	2			1	2		
G2	43	10	33			15	28			21	22		
G3	22	9	13			8	14			14	8		
T stage				2.65	0.45			8.75	0.03			10.26	0.02
T1	8	2	6			6	2			8	0		
T2	8	2	6			3	5			5	3		
T3	51	14	37			14	37			22	29		
T4	1	1	0			1	0			1	0		
Lymph node metastasis				3.79	0.29			6.44	0.09			4.24	0.24
N0	31	10	21			10	21			18	13		
N1	19	7	12			5	14			9	10		
N2	15	2	13			9	6			9	6		
N3	3	0	3			0	3			0	3		
Stage				0.93	0.63			2.98	0.23			6.08	<0.05
Stage I	6	2	4			4	2			6	0		
Stage II	27	9	18			8	19			14	13		
Stage III	35	8	27			12	23			16	19		

ESCC, esophageal squamous cell carcinoma; CEA, carcinoembryonic antigen; SCC, squamous cell carcinoma.

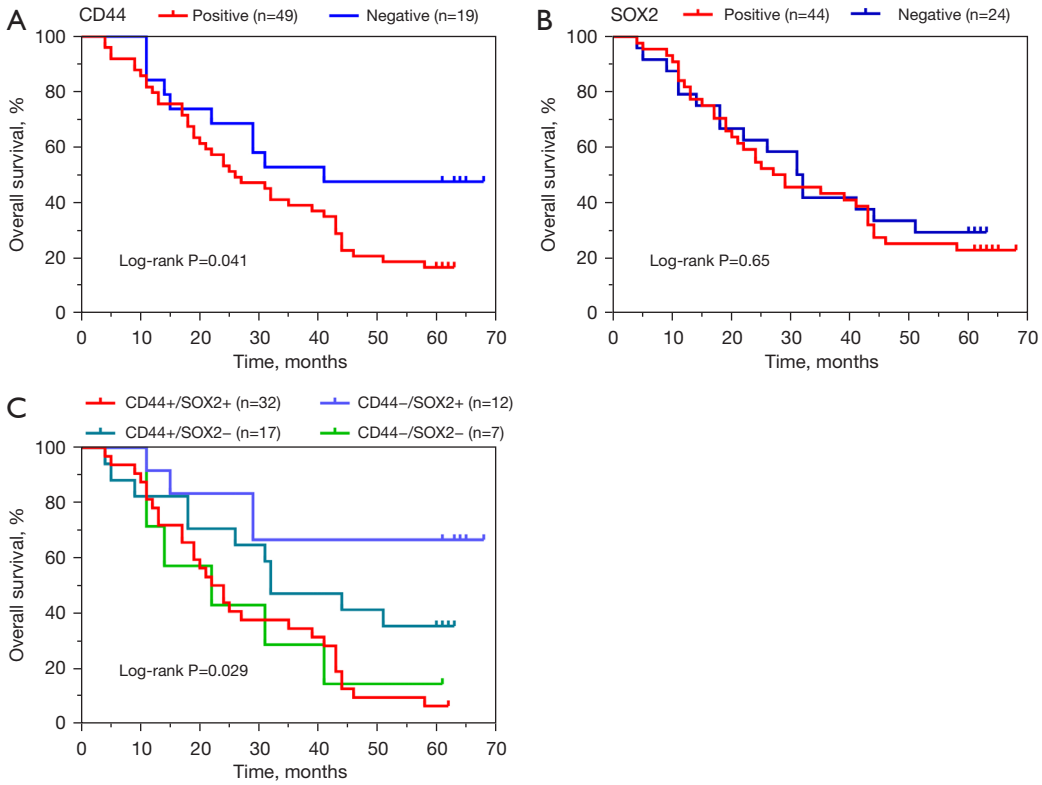

Figure 2 Kaplan-Meier analysis of survival in ESCC patients. (A) Overall patient survival correlated with CD44 expression (log-rank =4.19, P=0.041). (B) There was no significant correlation between patients' overall survival and SOX2 expression (log-rank =0.20, P=0.65). (C) Overall survival of patients was associated with co-expression of CD44 and SOX2 (log-rank =4.77, P=0.029). ESCC, esophageal squamous cell carcinoma.

Table 3 Analysis of risk factors for disease overall survival in ESCC patients

Variables	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
Age (≤62/>62 years)	0.99 (0.57–1.72)	0.98	1.11 (0.61–2.03)	0.73
Sex (male/female)	0.91 (0.50–1.64)	0.75	1.44 (0.72–3.85)	0.30
CEA (negative/positive)	1.17 (0.36–3.75)	0.80		
SCC (negative/positive)	3.03 (1.33–6.87)	0.008	1.63 (0.67–3.97)	0.29
Maximum tumor diameter (≤4/>4 cm)	1.87 (1.07–3.25)	0.03	1.28 (0.70–2.24)	0.43
Histological grade (G1/G2/G3)	1.06 (0.62–1.82)	0.84		
Nerve invasion (negative/positive)	1.38 (0.76–2.53)	0.29		
T stage (T1 + T2/T3 + T4)	2.67 (1.50–4.74)	<0.001	3.34 (1.33–8.39)	0.002
Lymph node metastasis (negative/positive)	1.84 (1.05–3.24)	0.03	7.43 (1.52–36.41)	0.04
Stage (I/II/III)	1.66 (1.07–2.59)	0.02	0.22 (0.05–1.08)	0.14
$CD44^+/SOX2^+$ (negative/positive)	2.32 (1.32–4.09)	<0.001	1.86 (1.01–3.43)	0.04

ESCC, esophageal squamous cell carcinoma; CI, confidence interval; CEA, carcinoembryonic antigen; SCC, squamous cell carcinoma.

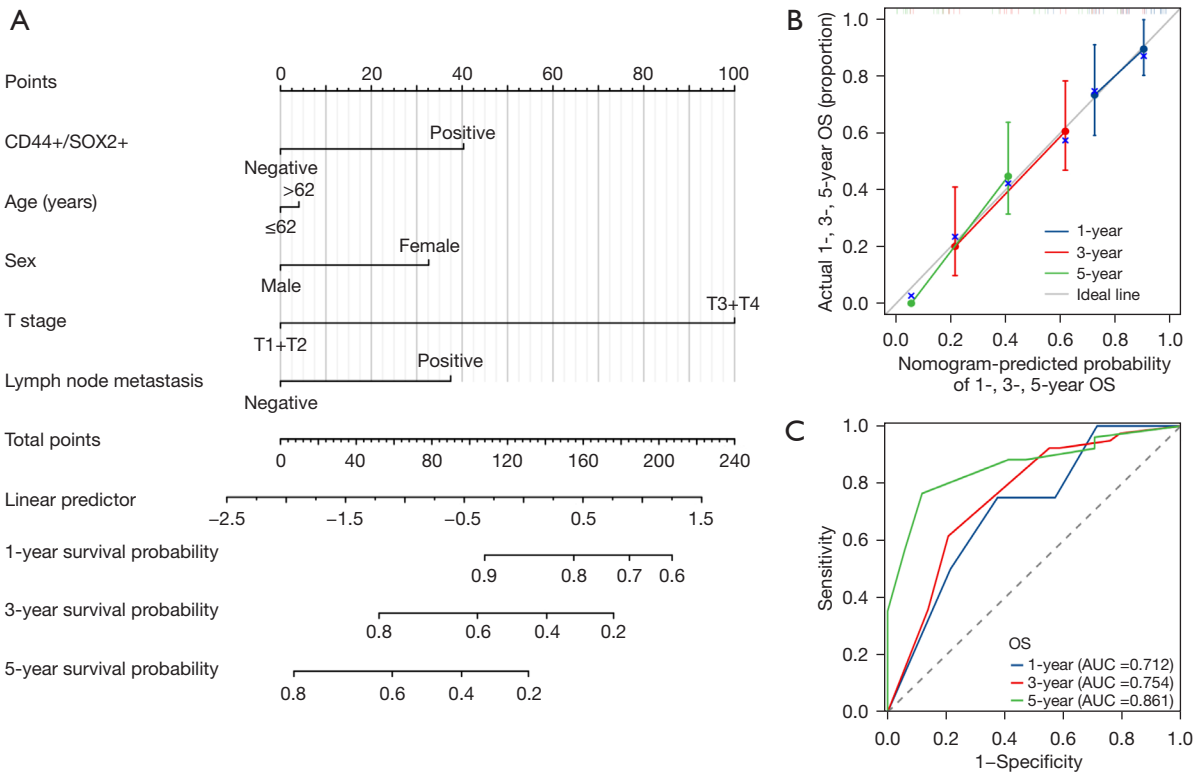

Figure 3 Construction of nomogram for OS in ESCC patients. (A) Nomogram prediction model of OS rate in ESCC patients. (B) Calibration curves of 1-, 3-, 5-year OS rate of nomogram prediction model of OS in ESCC patients. (C) Nomogram prediction model 1-year (blue line), 3-year (red line), and 5-year (green line) OS profiles of subjects' work characteristics. OS, overall survival; ESCC, esophageal squamous cell carcinoma; AUC, area under curve.

model for OS exhibited a C-index of 0.691 [95% confidence interval (CI): 0.66–0.73]. Additionally, the calibration plots confirmed substantial congruence between OS rates, as depicted in *Figure 3B*. The receiver operating characteristic curve (ROC) curves for the nomogram predictive model, complemented by the area under curve (AUC) values, are presented in *Figure 3C*.

Differentiation of high-risk and low-risk patients based on nomogram scores

The TNM staging system exhibited limited differentiation capability between OS in stage II and stage III ESCC patients (*Figure 4A*). Utilizing the nomogram prediction model, each patient's cumulative OS score was computed. The optimal cut-off value for the total OS score, determined by the maximum Youden index of the ROC curve, was set at 100. This facilitated stratification of patients into high-risk and low-risk groups based on their total scores. In terms of OS, a total of 48 patients were classified into the high-risk category, while 20 patients were allocated to the low-risk category. A significant statistical difference was observed in the 5-year OS rate between the low-risk and high-risk groups (35% *vs.* 8.3%, P<0.005; HR: 4.107, 95% CI: 2.303–7.327) as depicted in *Figure 4B*. The nomogram, based on the OS survival analysis results, demonstrated a robust predictive capability when patients were stratified into high-risk and low-risk categories according to their cumulative scores.

Discussion

The CSCs represent a primary cellular source in tumor recurrence and play a crucial role in the pathogenesis, progression, recurrence, and resistance to treatment of various cancer types. Their direct involvement regulates the initiation, progression, metastasis, and resistance to treatment of cancer, thereby leading to recurrence (25,26).

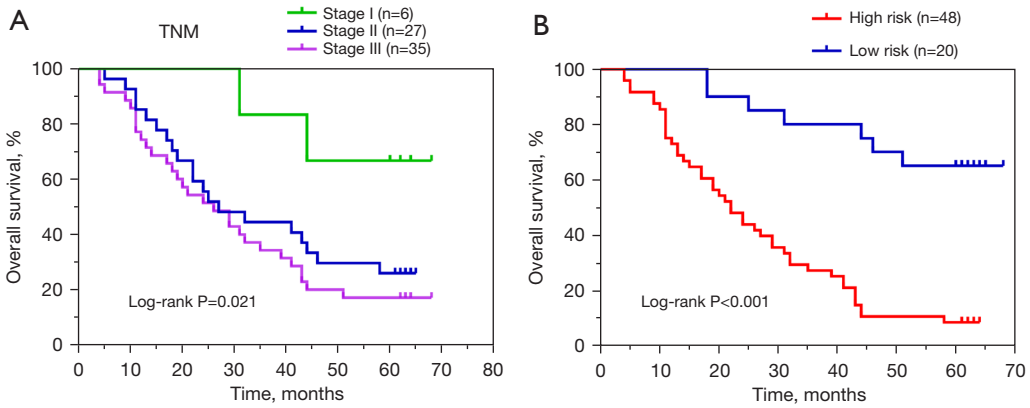

Figure 4 Kaplan-Meier survival curves in different subgroups of ESCC patients. (A) Kaplan-Meier survival curves for ESCC patients with the TNM staging system. (B) Kaplan-Meier survival curves for the nomogram prediction model. ESCC, esophageal squamous cell carcinoma; TNM, tumor node metastasis.

Specific markers of esophageal squamous cell carcinoma stem cells (ECSCs) aid in the extraction of this tumor cell subset. These markers have significant clinical implications, not just for the diagnosis, metastasis, and recurrence of ESCC, but also for its targeted therapy. They are instrumental in enhancing patient prognosis and managing tumor advancement, thereby facilitating a more efficient treatment approach for esophageal cancer (27). Hence, the examination of ECSCs markers is imperative for the prognosis of patient conditions and the detection of high-risk instances.

CD44, a prevalent CSC marker, exhibits elevated expression in ESCC tissues compared to adjacent non-cancerous tissues. An increasing body of research indicates a correlation between CD44, the CSC marker, and malignant tumorigenesis. Parul Gupta *et al.*'s study (28) reinforces this notion, revealing a substantial increase in CD44 expression in ESCC compared to atypical hyperplasia, thereby suggesting its involvement in esophageal carcinogenesis. In the research conducted by Awasthi *et al.*'s study (29), they demonstrated that concurrent alterations in CD44 and IL-1β within tumor stem cells facilitate the early identification of oral submucous fibrosis and oral squamous carcinoma. Dhar *et al.* (30) found that CD44 expression in hepatocytes was dramatically upregulated when hepatocytes were exposed to carcinogens, and damaged hepatocytes maintained this induced mutation by activating AKT via CD44 to promote MDM2 phosphorylation and nuclear translocation to escape from p53-mediated death, and passed on from the parental cells to the daughter cells, which further became hepatocellular carcinoma stem cells, leading to tumorigenesis. These findings are consistent with the results of the present study. Gong *et al.*'s study (31) proposed a correlation between positive CD44 expression, T-staging, and tumor pathology lymph node metastasis in ESCC, a conclusion that contradicts the findings of our research. Our analysis did not reveal a significant link between positive CD44 expression and the clinicopathological data of ESCC patients. We found that CD44 expression was up-regulated with deeper tumor infiltration and later clinical stage in 68 ESCC tissue samples, but the difference was not statistically significant. We analyzed that it was due to the inconsistency of the results with the literature caused by the small sample size analyzed in this study. We hypothesized that the differences in CD44 expression between patients' clinical profiles might be statistically different after expanding the clinical tissue sample size. We speculate that CD44, an identified marker of CSCs, contributes to tumourigenesis in ESCC. Therefore, we will also conduct further studies on the issue of differential expression of CD44 between low- and high-grade atypical hyperplasia of esophageal squamous epithelium and carcinoma *in situ* with the aim of discovering the mechanistic role of CD44 in esophageal squamous carcinogenesis. It is important to highlight that the patients with high CD44 expression in their tumors exhibited a reduced overall OS and a more severe prognosis compared to those with low CD44 expression. This suggests that CD44 could serve as a valuable prognostic tool for CSCs in predicting the future clinical course of ESCC patients.

SOX2, a transcription factor distinctly expressed in CSCs, plays a significant role in tumourigenesis,

progression, and the preservation of cellular stem cells (32). Our research has revealed a significant correlation between the heightened expression of SOX2 in ESCC cancer tissues and the depth of tumor infiltration, as well as the positive expression of CEA. This suggests a potential role of SOX2 in the progression of ESCC tumor, possibly facilitating their deeper tissue infiltration. Chai *et al.*'s cellular study (33) discovered a substantial enhancement in migration, invasion, and cisplatin chemoresistance of ESCC cells as a result of SOX2 overexpression. This is in alignment with the observations made in this study. There was no discernible correlation between SOX2 positive expression and OS in ESCC patients noted in this study. However, we found that among the patients whose 5-year follow-up outcome was death, there were 34 cases of positive SOX2 expression (34/51, 66.7%) and 17 cases of negative SOX2 expression (17/51, 33.3%), suggesting that the number of SOX2-positive expression cases was higher among the death cases. This suggests a stage-dependent effect of protein expression of SOX2 on the prognosis of patients with esophageal squamous carcinoma; stage IV patients were not included in this study, and it is hypothesized that the difference in the effect of SOX2 expression on 5-year survival may be statistically significant with the inclusion of a sample size of clinical tissues in the advanced stages of progression. The prognostic value of SOX2 remains a topic of contention. A separate study involving 113 ESCC patients revealed that low SOX2 expression is an independent factor associated with poor prognosis in ESCC patients (34). Conversely, it has been proposed that high SOX2 expression fosters ESCC migration and invasion, and is an independent risk factor for OS in ESCC patients (21). We will also further investigate the issue of differential expression of SOX2 between *in situ* and invasive carcinomas in ESCC, aiming to discover the mechanistic role of SOX2 in promoting tumor infiltration in ESCC, leading to tumor progression.

The objective of this investigation was to ascertain the ideal threshold value for CD44/SOX2 and to create and corroborate a CD44/SOX2-based nomogram prognostic model tailored for individual ESCC patients. Typically, the predictive efficacy of nomogram models is evaluated using discrimination and calibration methods (35-37). The evaluation of nomograms was conducted through calibration curves, illustrating the congruence between actual and anticipated clinical results. Patients with ESCC expressing both CD44 and SOX2 indicators were found to have a poorer prognosis. Multifactorial Cox analysis results identified the co-expression of CD44/SOX2 as an independent risk determinant for OS prognosis in ESCC patients, suggesting the potential use of this CSCs marker combination as a prognostic indicator. Further empirical validation is necessary to establish the clinical utility of CD44/SOX2 co-expression as a prognostic factor, which would require a larger sample size.

According to multifactorial Cox analysis results, C-index for the construction of a nomogram prediction model that includes CD44/SOX2, pathological T-stage, and lymph node metastasis in this study was 0.691 (95% CI: 0.656–0.727). Both the C-index and ROC curves demonstrated the nomogram's robust individual predictive capacity for OS in ESCC patients. A total nomogram score exceeding 100 could be categorized as a high-risk group with a 5-year OS rate of 7.1%.

Our findings confirm the efficacy of the predictive nomogram model as a swift, intuitive, and precise tool for predicting individual OS rates. The nomogram-based risk stratification approach exhibited strong discriminatory effects. A risk stratification system was developed based on the nomogram model to distinguish between patients exhibiting varying mortality risk levels. When juxtaposed with the AJCC staging system, the 5-year OS curves for the low/high-risk groups using the nomogram prediction model demonstrated significant divergence. However, this divergence was not observed among clinical stages. As a result, the ESCC prediction model outperformed the AJCC staging system in predicting OS, with the latter providing more precise treatment guidance. Individuals with an elevated risk of ESCC appear to be inadequately served by existing radical therapies. There is an urgent need to develop and implement more intensive treatment plans to enhance patient prognosis.

Our study, while comprehensive, is not without its limitations. The first limitation stems from the fact that we utilized retrospective analysis to conduct our research, inevitably introducing treatment selection bias. To confirm the prognostic accuracy of ESCC prediction models, a future study should be prospective and randomized. The second limitation pertains to the narrow scope of our study, which focused solely on ESCC patients who underwent surgical procedures. This resulted in a relatively small sample size, excluding patients with distant metastases. To address this, future studies conducted by our team will incorporate a larger sample size and encompass a broader range of patient conditions. This will aid us in constructing a more comprehensive column-line diagram, showcasing

the composition of prognostic factors in greater detail.

Conclusions

CSCs markers, CD44 and SOX2, are significant contributors to the tumorigenesis and progression of ESCC, respectively. CD44 is instrumental in early tumor development, making it a valuable early warning indicator, while SOX2 facilitates ESCC invasion. Furthermore, the co-expression of CD44 and SOX2 is strongly correlated with patient outcomes.

In this research, we have successfully developed a nomogram that integrates CD44 and SOX2 for the accurate prediction of OS in ESCC patients. This innovative tool not only offers a reliable clinical reference for precise OS prediction in ESCC patients, but also facilitates the stratification of high-risk individuals. Furthermore, it lays a solid foundation for directing clinical interventions and enhancing therapeutic strategies.

Acknowledgments

Funding: This study was supported by the Xinjiang Uygur Autonomous Region Postgraduate Research Innovation Project (No. XJ2023G182) and the "Tianshan Talents" Medical and Health High-level Talents Training Plan (No. TSYC202301B035).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2313/rc

Data Sharing Statement: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2313/dss

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2313/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-23-2313/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics committee of Xinjiang Medical University Affiliated Tumor Hospital (No. K-2019053) and informed consent was obtained from all individual participants.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
2. Ji Z, Bao S, Li L, et al. Hypoxia-inducible factor-2α promotes EMT in esophageal squamous cell carcinoma through the Notch pathway. Adv Clin Exp Med 2022;31:795-805.
3. Zeng Z, Fu M, Hu Y, et al. Regulation and signaling pathways in cancer stem cells: implications for targeted therapy for cancer. Mol Cancer 2023;22:172.
4. Shi ZD, Pang K, Wu ZX, et al. Tumor cell plasticity in targeted therapy-induced resistance: mechanisms and new strategies. Signal Transduct Target Ther 2023;8:113.
5. Kong W, Gao Y, Zhao S, et al. Cancer stem cells: advances in the glucose, lipid and amino acid metabolism. Mol Cell Biochem 2023. [Epub ahead of print]. doi: 10.1007/s11010-023-04861-6.
6. Paul R, Dorsey JF, Fan Y. Cell plasticity, senescence, and quiescence in cancer stem cells: Biological and therapeutic implications. Pharmacol Ther 2022;231:107985.
7. Izadpanah A, Mohammadkhani N, Masoudnia M, et al. Update on immune-based therapy strategies targeting cancer stem cells. Cancer Med 2023;12:18960-80.
8. Zhou C, Fan N, Liu F, et al. Linking Cancer Stem Cell Plasticity to Therapeutic Resistance-Mechanism and Novel Therapeutic Strategies in Esophageal Cancer. Cells 2020;9:1481.
9. Hassn Mesrati M, Syafruddin SE, Mohtar MA, et al.

CD44: A Multifunctional Mediator of Cancer Progression. Biomolecules 2021;11:1850.
10. Yaghobi Z, Movassaghpour A, Talebi M, et al. The role of CD44 in cancer chemoresistance: A concise review. Eur J Pharmacol 2021;903:174147.
11. Gomez KE, Wu F, Keysar SB, et al. Cancer Cell CD44 Mediates Macrophage/Monocyte-Driven Regulation of Head and Neck Cancer Stem Cells. Cancer Res 2020;80:4185-98.
12. Wang Z, Tang Y, Xie L, et al. The Prognostic and Clinical Value of CD44 in Colorectal Cancer: A Meta-Analysis. Front Oncol 2019;9:309.
13. Escudero Mendez L, Srinivasan M, Hamouda RK, et al. Evaluation of CD44+/CD24- and Aldehyde Dehydrogenase Enzyme Markers in Cancer Stem Cells as Prognostic Indicators for Triple-Negative Breast Cancer. Cureus 2022;14:e28056.
14. Gómez-Gallegos AA, Ramírez-Vidal L, Becerril-Rico J, et al. CD24+CD44+CD54+EpCAM+ gastric cancer stem cells predict tumor progression and metastasis: clinical and experimental evidence. Stem Cell Res Ther 2023;14:16.
15. Iżycka N, Zaborowski MP, Ciecierski Ł, et al. Cancer Stem Cell Markers-Clinical Relevance and Prognostic Value in High-Grade Serous Ovarian Cancer (HGSOC) Based on The Cancer Genome Atlas Analysis. Int J Mol Sci 2023;24:12746.
16. Mirhashemi M, Sadeghi M, Ghazi N, et al. Prognostic value of CD44 expression in oral squamous cell carcinoma: A meta-analysis. Ann Diagn Pathol 2023;67:152213.
17. Mirzaei S, Paskeh MDA, Entezari M, et al. SOX2 function in cancers: Association with growth, invasion, stemness and therapy response. Biomed Pharmacother 2022;156:113860.
18. Li Y, Wang M, Yang M, et al. Nicotine-Induced ILF2 Facilitates Nuclear mRNA Export of Pluripotency Factors to Promote Stemness and Chemoresistance in Human Esophageal Cancer. Cancer Res 2021;81:3525-38.
19. Chen G, Chen Y, Xu R, et al. Impact of SOX2 function and regulation on therapy resistance in bladder cancer. Front Oncol 2022;12:1020675.
20. Zhan Y, Chen Z, He S, et al. Long non-coding RNA SOX2OT promotes the stemness phenotype of bladder cancer cells by modulating SOX2. Mol Cancer 2020;19:25.
21. Li C, Ma YQ. Prognostic significance of sex determining region Y-box 2, E-cadherin, and vimentin in esophageal squamous cell carcinoma. World J Clin Cases 2022;10:9657-69.
22. Wang X, Ge X, Wang H, et al. SOX2 amplification and chromosome 3 gain significantly impact prognosis in esophageal squamous cell carcinoma. Ann Transl Med 2021;9:321.
23. Xiao S, Huang G, Zeng W, et al. Roles of oncogenes in esophageal squamous cell carcinoma and their therapeutic potentials. Clin Transl Oncol 2023;25:578-91.
24. Pedregal-Mallo D, Hermida-Prado F, Granda-Díaz R, et al. Prognostic Significance of the Pluripotency Factors NANOG, SOX2, and OCT4 in Head and Neck Squamous Cell Carcinomas. Cancers (Basel) 2020;12:1794.
25. Das PK, Pillai S, Rakib MA, et al. Plasticity of Cancer Stem Cell: Origin and Role in Disease Progression and Therapy Resistance. Stem Cell Rev Rep 2020;16:397-412.
26. Duan H, Liu Y, Gao Z, et al. Recent advances in drug delivery systems for targeting cancer stem cells. Acta Pharm Sin B 2021;11:55-70.
27. Chen L, Kong R, Wu C, et al. Circ-MALAT1 Functions as Both an mRNA Translation Brake and a microRNA Sponge to Promote Self-Renewal of Hepatocellular Cancer Stem Cells. Advanced Science 2019;7:- 1900949.
28. Gupta P, Rizvi SZ, Lal N, et al. Expression of CD44 and CD133 stem cell markers in squamous cell carcinoma of esophagus. Indian J Pathol Microbiol 2021;64:472-8.
29. Awasthi S, Ahmad S, Gupta R, et al. Differential expression of cancer stem cell markers and pro-inflammatory cytokine IL-1β in the oral squamous cell carcinoma and oral submucosal fibrosis. Int J Health Sci (Qassim) 2023;17:28-38.
30. Dhar D, Antonucci L, Nakagawa H, et al. Liver Cancer Initiation Requires p53 Inhibition by CD44-Enhanced Growth Factor Signaling. Cancer Cell 2018;33:1061-1077.e6.
31. Gong X, Wang A, Song W. Clinicopathological significances of PLOD2, epithelial-mesenchymal transition markers, and cancer stem cells in patients with esophageal squamous cell carcinoma. Medicine (Baltimore) 2022;101:e30112.
32. Zhang J, Wang Z, Zhao H, et al. The roles of the SOX2 protein in the development of esophagus and esophageal squamous cell carcinoma, and pharmacological target for therapy. Biomed Pharmacother 2023;163:114764.
33. Chai Y, Li Q, Zhao H, et al. SOX2 antagonizes WWC1 to drive YAP1 activation in esophageal squamous cell carcinoma. Cancer Med 2019;8:7055-64.
34. Maehara R, Fujikura K, Takeuchi K, et al. SOX2-silenced squamous cell carcinoma: a highly malignant form of esophageal cancer with SOX2 promoter hypermethylation. Mod Pathol 2018;31:83-92.
35. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing

the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.
36. Zhang Z, Wang D, Zhang J, et al. Comparison of the effectiveness of chemotherapy combined with immunotherapy and chemotherapy alone in advanced biliary tract cancer and construction of the nomogram for survival prediction based on the inflammatory index and controlling nutritional status score. Cancer Immunol Immunother 2023;72:3635-49.
37. Zhang W, Ji L, Wang X, et al. Nomogram Predicts Risk and Prognostic Factors for Bone Metastasis of Pancreatic Cancer: A Population-Based Analysis. Front Endocrinol (Lausanne) 2022;12:752176.

Cite this article as: Tian S, Ma R, Liu Y, Chen F, Huang X, Yang Q, Nian W, Fan Z. Clinicopathological significance of cancer stem cell marker CD44/SOX2 in esophageal squamous cell carcinoma (ESCC) patients and construction of a nomogram to predict overall survival. Transl Cancer Res 2024;13(6):2971-2984. doi: 10.21037/tcr-23-2313

