Research Article



Development and validation of nomogram combining serum biomarker for predicting survival in patients with resected rectal cancer

💿 Shaonan Fan, Ting Li, Ping Zhou, Qiliang Peng and 💿 Yaqun Zhu

Department of Radiation Oncology, Second Affiliated Hospital of Soochow University, Suzhou 215004, China

Correspondence: Yaqun Zhu (szzhuyaqun@sina.com)



Purpose: Nomogram is a widely used tool that precisely predicts individualized cancer prognoses. We aimed to develop and validate a reliable nomogram including serum tumor biomarkers to predict individual overall survival (OS) for patients with resected rectal cancer (RC) and compare the predictive value with the American Joint Committee on Cancer (AJCC) stages.

Patients and methods: We analyzed 520 patients who were diagnosed with non-metastatic rectal cancer as training cohort. External validation was performed in a cohort of 11851 patients from the Surveillance, Epidemiology, and End Results (SEER) database. Independent prognostic factors were identified and integrated to build a nomogram using the Cox proportional hazard regression model. The nomogram was evaluated by Harrell's concordance index (C-index) and calibration plots in both training and validation cohort.

Results: The calibration curves for probability of 1-, 3-, and 5-year OS in both cohorts showed favorable accordance between the nomogram prediction and the actual observation. The C-indices of the nomograms to predict OS were 0.71 in training cohort and 0.69 in the SEER cohort, which were higher than that of the seventh edition American Joint Committee on Cancer TNM staging system for predicting OS (training cohort, 0.71 vs. 0.58, respectively; *P*-value < 0.001; validation cohort, 0.69 vs. 0.57, respectively; *P*-value < 0.001). Conclusion: We developed and validated a novel nomogram based on CEA and other factors for predicting OS in patients with resected RC, which could assist clinical decision making and improvement of prognosis prediction for individual RC patients after surgery.

Introduction

Colorectal cancer (CRC) is the fourth most common cancer and ranks second in terms of mortality in the United States [1]. In 2018, it is estimated that more than 1.8 million new cases of colorectal cancer and 881,000 deaths occurred in the world, accounting for approximately 10% cancer cases and deaths [2]. According to the latest data released by the National Cancer Center, there were 370,000 new cases of colorectal cancer in China in 2015, accounting for 26% of the world. The incidence and mortality were 28 and 13.61 per 100,000, ranking third and fifth among all malignant tumors, respectively [3]. Favorable long-term survival and a better quality of life after a curative resection are usually achieved in patients with non-metastatic rectal cancer (RC) [4]; however, distant metastases and local recurrences that result in poor survival still remain the primary concerns in this entity [5].

The tumor, node, metastasis (TNM) staging system is a widely used cancer staging system for predicting the prognosis and determining the appropriate treatment strategy. In the seventh edition of the American

Received: 05 August 2019 Revised: 26 September 2019 Accepted: 21 October 2019

Accepted Manuscript online: 22 October 2019 Version of Record published: 13 November 2019



Joint Committee on Cancer TNM classification, patients with non-metastatic rectal cancer (RC) are stratified according to the depth of invasion and extent of lymph node involvement [6]. Although the TNM stage is important and useful, it is well known that the survival time varies widely, even in patients with the same stage of disease [7]. What convinced is that there are other independent prognostic factors such as sex, age, histology, and treatment-related factors could be significantly devoted to individualized survival prediction [6]. Recently, serum biomarkers have been widely used in clinical diagnosis, post-operative monitoring and prognosis of CRC patients [8]. Common biomarkers for RC including carcinoembryonic antigen (CEA) and tumor-associated antigens, such as cancer antigen 19-9 (CA19-9), cancer antigen 50 (CA50), and cancer antigen 72-4 (CA72-4), whose predictive significance of pre-operative levels in different tumor populations has been extensively documented in numerous studies [9–12].

Nomogram is an intuitive graph of a statistical predictive model that generates a numerical probability of a clinical event, such as overall survival (OS) by incorporating important factors for oncologic prognosis, and it has been widely practiced in clinical application [13,14]. Comparing with the traditional TNM staging systems or other staging systems, the nomogram has been proved with the precise predictive value in other types of tumors [14–16]. However, a reliable nomogram based on common serum tumor biomarkers and other clinicopathologic index for predicting long-term survival outcome of patients with resected RC was scarce.

Therefore, the aim of this study was to develop a nomogram for individual post-operative prediction in patients with RC by combining known clinicopathologic variables (including tumor biomarkers) based on the data from a single-institutional registry in China. In addition, we used a separate cohort from the Surveillance, Epidemiology, and End Results (SEER) database to externally validate it [17].

Methods and materials Patient selection

For the SEER cohort, by using the SEER*Stat, version 8.3.5 software, we selected RC patients diagnosed between 2004 and 2015 [18]. The inclusion criteria for selected patients were as follows: (1) Patients aged 18 years or older; (2) patients with RC diagnosed as the first primary cancer; (3) patients with RC underwent surgical resection without neoadjuvant therapy; (4) the International Classification of Diseases for Oncology third edition (ICD-O-3) codes (8010-8231 and 8255-8576); and (5) patients who were followed up for at least 3 months after surgery were included in the study. In the training cohort, patients with RC received surgery between January 2008 and December 2015 in the Second Affiliated Hospital of Soochow University were included and following inclusion criteria: Patients aged 18 years or older; underwent surgical resection; patients who were followed up for at least 3 months after surgery were included in the study; no synchronous distant metastases or other malignancies; no neoadjuvant therapy; serum biomarker levels within 6 months of surgery were available. The same variables used were age at diagnosis, gender, T stage, N stage, histological subtypes, histological grade, surgical margin, tumor deposits, and the levels of preoperative CEA.

Statistical analysis

Continuous variables were compared using the Student's *t*-test (normally distributed data) or the Mann–Whitney *U*-test (non-normal distribution data), as appropriate. Categorical variables were compared using the Chi-square test. The primary endpoint of this study was overall survival (OS), which was defined as the interval between the initial diagnosis of RC and the last follow-up. The Kaplan–Meier method was used to estimate the survival probabilities and the log-rank test was used to assess the survival differences of OS. Univariate and multivariate Cox regression hazards models were employed to identify prognostic factors. The development of the final nomogram was based on the prognostic factors extracted from the training cohort and validation of the nomogram was carried out using the data from the SEER cohort. The validation of the nomogram was performed using the concordance index (C-index), calibration curves. Finally, a risk classification system was established according to the total scores of each RC patient in the training cohort by using the nomogram to separate patients into three prognostic groups, i.e. the low-, intermediate-, and high-risk groups. The X-tile (Yale University, 3.6.1) was implemented to calculate cut-off points for prognostic groups. The risk classification system was further validated in the SEER cohort. All statistical analyses were completed using R (version 3.5.2, R Foundation for Statistical Computing, Vienna, Austria). All statistical tests were two-sided and statistical significance was set at *P*-value < 0.05.

Results Clinical characteristics

The baseline characteristics of the training and validation cohort were shown in Table 1. All patients received surgery.



Table 1 Baseline characteristics of RC patients in the training and SEER cohort

<table-cell></table-cell>	Patient characteristics		Training cohort (<i>n</i> =520)	SEER cohort (n=11,851)	P-value																																																																																																
Gender (%)Fenale200 (38.5)4682 (39.5)0.666Male320 (61.5)7168 (0.5)7Surgical type (%)Dixon330 (61.5)-T stage (%)T1 nu910 (35.5)-T stage (%)T1 nu224 (4.6)2223 (19.7)<.0.011	Age (mean (SD))		62.86 (11.17)	60.91 (12.65)	0.001																																																																																																
MaleMale320 (3)7169 (6)Surgitype (%)Dkon30 (3)-Others30 (3)T stage (%)T124 (4)162 (1)T stage (%)T250 (2)323 (1)T stage (%)T372 (2)323 (1)N stage (%)T372 (2)30 (2)N stage (%)Na72 (2)30 (2)30 (2)N stage (%)Na74 (1)34 (6)30 (2)N stage (%)Na34 (6)480 (3)40 (3)N stage (%)Na34 (6)102 (1)N stage (%)Aleocarcinoma36 (2)102 (1)N stage (%)Na36 (2)102 (1)102 (1)N stage (%)Na36 (2)102 (1)102 (1)102 (1)N stage (%)Na36 (2)<	Gender (%)	Female	200 (38.5)	4682 (39.5)	0.666																																																																																																
		Male	320 (61.5)	7169 (60.5)																																																																																																	
InterpretationInterpretationInterpretationInterpretationTatage %142462620.01Tata15082319.0700260.01Tata150700260.0170027002Natage %7424904700260.01Natage %140846.0390460.037002Natage %610.021002100270027002Natage %610.021002100270027002Natage %610.021002100270027002Natage %610.02610.02100270027002Natage %610.02610.02610.0270027002Natage %610.02610.02610.0270027002Natage %610.02610.02700270027002Natage %610.02610.0270170027002Natage %610.0270170170027002Natage %610.027017017017012Natage %7017017017017012Natage %70170170170127012Natage %70170170170127012Natage %701701701270127012Natage %701701701270127012Natage %7017012701270127012Natage %701	Surgical type (%)	Dixon	330 (63.5)	_																																																																																																	
$egin{tabular}{lllllllllllllllllllllllllllllllllll$		Others	190 (36.5)	_																																																																																																	
I2150 (28.8)2323 (19.6)T3272 (52.3)7002 (50.1)Natage (%)T4 (14.2)904 (7.6)Natage (%)N06466 (54.6)<0.001	T stage (%)	T1	24 (4.6)	1622 (13.7)	<0.001																																																																																																
Image by the series of the s		T2	150 (28.8)	2323 (19.6)																																																																																																	
Index14 (14.2)904 (7.6)N stage (%)NG348 (66.9)6469 (54.6)<.0.01		ТЗ	272 (52.3)	7002 (59.1)																																																																																																	
N stage (%)N0348 (66.9)6469 (54.6)<.0.01N1116 (22.3)4180 (55.3)N156 (10.8)1202 (10.1)Histopathology (%)Adenocarcinoma488 (93.8)11194 (94.5)0.62Grade (%)Cthers32 (6.2)657 (5.5)Grade (%)G1/G2476 (91.5)10369 (87.5)<.0.01		T4	74 (14.2)	904 (7.6)																																																																																																	
N116 (2.3)4180 (35.3)N256 (10.8)1202 (1.1)Histopathology (%)Adenocarcinoma48 (9.8)1194 (94.5)0.62Othera32 (6.2)67 (5.5)0.01Grade (%)G1/G2476 (9.5)10369 (87.5)<0.01	N stage (%)	NO	348 (66.9)	6469 (54.6)	<0.001																																																																																																
N256 (0.8)1202 (0.1)Histopathology (%)Adenocarcinoma488 (0.3.)11194 (0.4.5.)0.62Others0.0 (0.1.1		N1	116 (22.3)	4180 (35.3)																																																																																																	
Histopathology (%)Adenocarcinoma488 (93.8)11194 (94.5)0.62Others32 (6.2)657(5.5)Grade (%)G1/G2476 (91.5)0308 (87.5)<0.001		N2	56 (10.8)	1202 (10.1)																																																																																																	
Others32 (62)657(5.5)Grade (%)G1/G2476 (91.5)10309 (87.5)<0.001	Histopathology (%)	Adenocarcinoma	488 (93.8)	11194 (94.5)	0.62																																																																																																
Grade (%)G1/G2476 (91.5)10369 (87.5)<0.001G3/G426 (5.0)1482 (12.5)NR18 (3.5)-Vascular invasion (%)No426 (81.9)-Yes94 (18.1)Perineural involvement (%)No452 (86.9)-Yasqui (%)Yes68 (13.1)-Margin (%)No518 (99.6)7071 (59.7)<0.001		Others	32 (6.2)	657(5.5)																																																																																																	
G3/G426 (5.0)1482 (12.5)NR18 (3.5)-Vascular invasion (%)No426 (81.9)-Yes94 (18.1)Perineural involvement (%)No452 (86.9)-Yes68 (13.1)Margin (%)No518 (99.6)7071 (59.7)<0.001	Grade (%)	G1/G2	476 (91.5)	10369 (87.5)	<0.001	NR18 (3.5)-Vascular invasion (%)No426 (81.9)-Yes94 (18.1)Perineural involvement (%)No452 (86.9)-Yes68 (13.1)Margin (%)No518 (99.6)7071 (59.7)<0.001		G3/G4	26 (5.0)	1482 (12.5)		Vascular invasion (%)No426 (81.9)-Yes94 (18.1)-Perineural involvement (%)No452 (86.9)-Yes68 (13.1)-Margin (%)No518 (99.6)7071 (59.7)Yes2 (0.4)532 (4.5)Yumor Deposits (%)NR-NR-4248 (35.8)Yes60 (11.5)3781 (31.9)Preoperative CEA (%)Negative362 (69.6)7534 (63.6)Preoperative CA199 (%)Negative482 (92.7)-Negative38 (7.3)Negative38 (7.3)Negative-<		NR	18 (3.5)	-		Yes 94 (18.1) - Perineural involvement (%) No 452 (86.9) - Yes 68 (13.1) - Margin (%) No 518 (99.6) 7071 (59.7) <0.001	Vascular invasion (%)	No	426 (81.9)	-		Perineural involvement (%) No 452 (86.9) - Yes 68 (13.1) - Margin (%) No 518 (99.6) 7071 (59.7) <0.01		Yes	94 (18.1)	-		Yes 68 (13.1) - Margin (%) No 518 (99.6) 707 (59.7) <0.01	Perineural involvement (%)	No	452 (86.9)	-		Margin (%) No 518 (99.6) 7071 (59.7) <0.001 Yes 2 (0.4) 532 (4.5) </td <td></td> <td>Yes</td> <td>68 (13.1)</td> <td>-</td> <td></td>		Yes	68 (13.1)	-		Yes 2 (0.4) 532 (4.5) NR - 4248 (35.8) Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001	Margin (%)	No	518 (99.6)	7071 (59.7)	<0.001	NR - 424 8 (35.8) Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001		Yes	2 (0.4)	532 (4.5)		Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001 NR 10 (1.9) 3781 (31.9)		NR	-	4248 (35.8)		NR 10 (1.9) 3781 (31.9) Yes 60 (11.5) 619 (5.2) Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Preoperative CA199 (%) Negative 482 (92.7) - - Preoperative CA199 (%) Negative 38 (7.3) - -	Tumor Deposits (%)	No	450 (86.5)	7451 (62.9)	<0.001	Yes 60 (11.5) 619 (5.2) Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Preoperative CA199 (%) Negative 158 (30.4) 4317 (36.4) - Preoperative CA199 (%) Negative 387.3) - -		NR	10 (1.9)	3781 (31.9)		Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Postive 158 (30.4) 4317 (36.4) -		Yes	60 (11.5)	619 (5.2)		Preoperative CA199 (%) Postive 158 (30.4) 4317 (36.4) Preoperative CA199 (%) Negative 482 (92.7) – Postive 38 (7.3) –	Preoperative CEA (%)	Negative	362 (69.6)	7534 (63.6)	0.006	Preoperative CA199 (%) Negative 482 (92.7) - Postive 38 (7.3) -		Postive	158 (30.4)	4317 (36.4)		Postive 38 (7.3) –	Preoperative CA199 (%)	Negative	482 (92.7)	-				Postive	38 (7.3)	-	
Grade (%)	G1/G2	476 (91.5)	10369 (87.5)	<0.001																																																																																																	
NR18 (3.5)-Vascular invasion (%)No426 (81.9)-Yes94 (18.1)Perineural involvement (%)No452 (86.9)-Yes68 (13.1)Margin (%)No518 (99.6)7071 (59.7)<0.001		G3/G4	26 (5.0)	1482 (12.5)																																																																																																	
Vascular invasion (%)No426 (81.9)-Yes94 (18.1)-Perineural involvement (%)No452 (86.9)-Yes68 (13.1)-Margin (%)No518 (99.6)7071 (59.7)Yes2 (0.4)532 (4.5)Yumor Deposits (%)NR-NR-4248 (35.8)Yes60 (11.5)3781 (31.9)Preoperative CEA (%)Negative362 (69.6)7534 (63.6)Preoperative CA199 (%)Negative482 (92.7)-Negative38 (7.3)Negative38 (7.3)Negative-<		NR	18 (3.5)	-																																																																																																	
Yes 94 (18.1) - Perineural involvement (%) No 452 (86.9) - Yes 68 (13.1) - Margin (%) No 518 (99.6) 7071 (59.7) <0.001	Vascular invasion (%)	No	426 (81.9)	-																																																																																																	
Perineural involvement (%) No 452 (86.9) - Yes 68 (13.1) - Margin (%) No 518 (99.6) 7071 (59.7) <0.01		Yes	94 (18.1)	-																																																																																																	
Yes 68 (13.1) - Margin (%) No 518 (99.6) 707 (59.7) <0.01	Perineural involvement (%)	No	452 (86.9)	-																																																																																																	
Margin (%) No 518 (99.6) 7071 (59.7) <0.001 Yes 2 (0.4) 532 (4.5) </td <td></td> <td>Yes</td> <td>68 (13.1)</td> <td>-</td> <td></td>		Yes	68 (13.1)	-																																																																																																	
Yes 2 (0.4) 532 (4.5) NR - 4248 (35.8) Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001	Margin (%)	No	518 (99.6)	7071 (59.7)	<0.001																																																																																																
NR - 424 8 (35.8) Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001		Yes	2 (0.4)	532 (4.5)																																																																																																	
Tumor Deposits (%) No 450 (86.5) 7451 (62.9) <0.001 NR 10 (1.9) 3781 (31.9)		NR	-	4248 (35.8)																																																																																																	
NR 10 (1.9) 3781 (31.9) Yes 60 (11.5) 619 (5.2) Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Preoperative CA199 (%) Negative 482 (92.7) - - Preoperative CA199 (%) Negative 38 (7.3) - -	Tumor Deposits (%)	No	450 (86.5)	7451 (62.9)	<0.001																																																																																																
Yes 60 (11.5) 619 (5.2) Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Preoperative CA199 (%) Negative 158 (30.4) 4317 (36.4) - Preoperative CA199 (%) Negative 387.3) - -		NR	10 (1.9)	3781 (31.9)																																																																																																	
Preoperative CEA (%) Negative 362 (69.6) 7534 (63.6) 0.006 Postive 158 (30.4) 4317 (36.4) -		Yes	60 (11.5)	619 (5.2)																																																																																																	
Preoperative CA199 (%) Postive 158 (30.4) 4317 (36.4) Preoperative CA199 (%) Negative 482 (92.7) – Postive 38 (7.3) –	Preoperative CEA (%)	Negative	362 (69.6)	7534 (63.6)	0.006																																																																																																
Preoperative CA199 (%) Negative 482 (92.7) - Postive 38 (7.3) -		Postive	158 (30.4)	4317 (36.4)																																																																																																	
Postive 38 (7.3) –	Preoperative CA199 (%)	Negative	482 (92.7)	-																																																																																																	
		Postive	38 (7.3)	-																																																																																																	

Abbreviations: RC, rectal cancer; SD, Standard Deviation; SEER, Surveillance, Epidemiology and End Results; NR, not recorded.

In training cohort, a total of 158 (30.4%) patients had positive preoperative CEA (\geq 6.5 ng/ml) and 38 (7.3%) patients had positive preoperative CA19-9 (\geq 37.0 U/ml). Post-operative pathological tumor staging suggested that 24 patients were in the T1 stage, 150 patients in the T2 stage, 272 in the T3 stage, and 74 were in the T4 stage, while 348 patients were in the N0 stage, 116 in the N1 stage, and 56 in the N2 stage. The positive rates of vascular invasion and perineural involvement were 18.1% (94/520) and 13.1% (68/520), respectively. For both training and validation cohorts, the majority of patients were with T3 and N0 stages and the most common histological subtype was adeno-carcinoma. Ignoring the absence of certain parameters (e.g. surgical type, vascular invasion, perineural involvement, and preoperative CA19-9 level) in the validation cohort, which justified their values as training and validation cohorts.

Univariate and multivariate Cox regression of the training cohort

For the training cohort, univariate analyses were performed to identify clinical variables that were significantly associated with OS. As shown in Figure 1A, female sex and younger age were associated with better prognosis (*P*-value < 0.05). With respect to factors associated with surgery, we found that pathologic T and N stages, tumor differentiation, histological subtype, vascular invasion, perineural involvement (PNI), and tumor deposits (TDs) status were greatly associated with OS (*P*-value < 0.05). In addition, the negative correlation of preoperative CEA (HR: 1.67 95% CI: 1.3–2.15) and CA19-9 (HR: 1.58; 95% CI: 1.05–2.37) levels with OS were significant (*P*-value < 0.05). After univariable analysis in the training cohort, the variables that achieved significance (*P*-value < 0.05) were factored into multivariable analysis in the Cox proportional hazard regression model. Finally, only eight factors (age, sex, grade, PNI,







TDs, preoperative CEA, T stage and N stage) remained independent prognostic factors for overall survival (Figure 1B).

Prognostic nomogram for OS

Nomogram for predicting 1-, 3-, and 5-year OS was constructed by incorporating the above independent prognostic factors (Figure 2). The nomogram illustrated age and N stage as sharing the largest contribution to prognosis, followed by the T category and TDs status. Each subtype within these variables was assigned a score on the point scale. A straight line can be easily drawn down by adding the scores of all the selected variables to determine the survival probabilities of the individual patients.

Validation of the nomograms

The C-indices of the nomogram to predict OS were 0.71 (95%CI:0.64–0.79) in training cohort and 0.69 (95%CI:0.61–0.78) in validation cohort. The calibration plots presented an acceptable agreement in the training cohort and an excellent agreement in SEER validation cohort between the nomogram prediction and actual observation for 1-, 3-, and 5-year OS (Figure 3). We further compared the OS predictive ability between the nomogram and TNM staging system. In the training cohort, the Harrell's C-index for the established nomogram to predict OS (0.71; 95%CI: 0.64–0.79) was significantly higher than that of the TNM staging system (0.58; 95%CI: 0.54–0.62; P-value < 0.001). In the SEER cohort, the difference is also statistically significant (0.69 vs. 0.57; P-value < 0.001).

Performance of the nomogram in stratifying risk of patients

Based on the quartiles of the nomogram score from the training cohort, we determined the cut-off values by grouping the patients in both cohorts evenly into three proposed risk groups; each group represented a distinct prognosis. As

Figure



Points	0 	10		20	30		40	50	60		0 80	90	100
Age	20	25	30	35	40	45	50	55	60	65	70 75	80 85	90
Gender	Femal	Male e											
T.stage	T1	T2 T3			T4								
N.stage	NO	N1					N2						
Grade	G1/G2	200		NR G	⊐ 3/G4								
Perineural.involvement	N	1	,										
Tumor deposits	No		N	R	1	Yes							
Preoperative CEA	negativ	posti e	ve										
Total Points	۲. 0	2	0	40		60	80	10	00	120	140	160	180
1-year survival							0.99			0.9	5 0.5	9 0.8	
3-year survival					0.9:	5	0.9	0.8	3 0.	7 0.6	0.5 0.4 0.1	3 0.2 0.1	
5-year survival 2. Nomogram integratin	g tumo	r mark	0. ers a	9 Ind cli	0.8	0.7 athol	0.6 0 ogical	.5 0.4 ().3 0.2 s	2 0.1			



Figure 3. Evaluation of nomogram using calibration curves

1-, 3-, and 5-year nomogram calibration curves in the training cohort (**A**,**C**,**E**) and SEER cohort (**B**,**D**,**F**). The dashed line represents an ideal evaluation, whereas the red line represents the performance of the nomogram.

shown in Figure 4, the subgroups experienced distinct overall survival in the training cohort and the SEER cohort (*P*-value < 0.001).

Comparisons of blood biomarkers in pre- and post-operative

Our results indicated the profile of CEA and CA19-9 undergone profound changes before and after surgery in RC patients. The mean values of the pre-operative CEA and CA19-9 level were 23.09 ± 5.85 ng/ml and 52.34 ± 7.19 U/ml,





Figure 4. Kaplan–Meier curves of OS for patients in the low-, intermediate-, and high-risk groups in the training cohort (A) and SEER cohort (B)

respectively. However, the changes in pre-operative and post-operative CA50 levels (pre-post-CA50) were not statistically significant. Due to the limitation of the number of cases, the changes in the levels of pre- and post-operative CA72-4 remain ambiguous (Figure 5).

Discussion

Due to the heterogeneity of RC in the survival of individual patients, the TNM staging system for predicting survival is inaccurate and insufficient to satisfy the medical demand [4]. Based on multivariate analysis to identify proper risk factors that are significantly associated with decreased survival, the nomogram could help clinicians to select patients with high risk of recurrence and metastasis, obtain better personalized prognostic assessments in clinical practice, which may be better than traditional TNM staging system in several cancers [19]. In current research, our multivariate analysis revealed that several factors are independent negative predictive determinants for OS (Figure 1B). Thus, we strive to develop a postoperative nomogram to predict long-term survival of patients with resected RC.

Tumor biomarkers, substances expressed and released during tumorigenesis and progression, give information about tumor biology in an indirect way and may reveal the presence of a new growth that indicates the relative tumor burden and biological aggressiveness [20]. For example, disruption of normal tissue structure and loss of polarity of tumor cells leads to CEA expression on the entire cell surface, which is then secreted into the bloodstream, ultimately increasing serum CEA levels. Because of its convenience and cost-effectiveness, tumor biomarkers are expected to guide therapy schedules and monitor prognosis [21]. Although CEA is not a disease-specific marker and National Comprehensive Cancer Network guidelines also did not include elevated serum CEA levels as a high-risk factor, recent studies, as well as our findings support its effectiveness for monitoring or evaluating treatment in RC patients. Moreover, changes between the preoperative and postoperative CEA levels are also suggested as indicators of higher recurrence rate.

Undoubtedly, the suggested cut-off values of preoperative CEA values were different, ranging from 2.5 to 50 ng/ml [22–24]. Generally, the cut-off level of 5 ng/ml was used to define CEA elevation. However, in our study, in order to make the nomogram better used in clinical practice of our institution, we defined CEA \geq 6.5 ng/ml as positive in the training cohort, which is based on the reference value provided by the clinical testing center of our institution.

In the present study, we performed a retrospective study in patients with resected RC who had four tumor biomarkers (CEA, CA19-9, CA50 and CA724) examined preoperatively. However, the number of cases of CA724 and CA50 in the training cohort was relatively insufficient, especially preoperative CA724, which was only effectively recorded in 27 cases (Figure 5), less than 10%. Thus, in the subsequent univariate analysis, we did not include it in the scope of the analysis. In addition, information of these two serum markers is not available in the open-accessed SEER database. Our results indicated the profile of CEA and CA19-9 undergone profound changes before and after surgery in RC





Figure 5. Distribution of serum biomarkers, including Carcinoembryonic antigen (CEA) (A), Cancer antigen 724 (CA724) (B), Cancer antigen 199 (CA199) (C) and Cancer antigen 50 (CA50) (D) in rectal cancer patients pre- and postoperative Abbreviations: NS, no significance; ***, *P*-value < 0.001

patients. Through univariable analysis and subsequent multivariable analysis, we identified age, sex, grade, perineural involvement, tumor deposits, preoperative CEA, T stage and N stage as independent prognostic factors (Figure 1B). These findings were in high concordance with previous reports on risk factors for RC [25,26].

It is essential to verify the nomogram in order to avoid over-fitting of the matrix and determine its universality [27]. We externally validated the nomogram using an independent cohort from the SEER database. In the present study, the calibration plot showed a first-rank consistency between the predicted and actual observation, supporting the reproducibility and reliability of the created nomogram (Figure 3). The C-index of the nomogram achieved 0.71 (95%CI:0.64–0.79) in the training cohort and 0.69 (95%CI:0.61–0.78) in the SEER cohort, both of which were comparable to those of other nomograms for rectal cancer [4,13]. Compared with the TNM staging system in the training group, the C-index of the nomogram is significantly higher, thus revealing discrimination and the discriminative ability was only slightly reduced in validation cohort. In addition, Kaplan–Meier estimates of the event rates over time for the validation cohort showed statistically different outcomes for the three proposed risk groups, which displayed a favorable discrimination and generalizability. More importantly, the model also fit the SEER cohort, which supports multi-country use of this nomogram, regardless of race and health care disparities.

Previous studies have established nomograms containing serum markers such as CEA to predict the prognosis for CRC patients. Zhang et al. [28] selected eligible CRC patients from the SEER database and created a nomogram



containing CEA expressions and other clinic-pathological factors, which showed accuracy and risk assessment ability than the AJCC staging system. Although this study is based on a large number of real-world populations, the SEER database records information on malignant tumors in the United States, and prognosis of colorectal cancer varies among different populations. Therefore, whether it is universal or not requires further verification. Another eastern-population based retrospective study have comprehensively considered the expression of CEA, CA50, and CA724 to assess their clinical significances in predicting prognosis and built nomograms to predict the overall survival (OS) and disease-free survival (DFS) in CRC patients after radical resection, indicating the above tumor biomarkers have an important role in monitoring recurrence and metastasis [4]. However, as a retrospective study, the limitations caused by the sample size are inevitable. Therefore, their nomograms may represent a better preferential ability to improve prognosis prediction for CRC patients, If could verify by using a large population-based approach with longer durations of follow-up. Compared to previous studies, our study has established a nomogram based on the real-world data. This nomogram contains not only preoperative CEA but also other clinical pathologic variables (age, sex, grade, PNI, TDs, T, and N Stage). In addition, we conducted external validation by using the SEER database, which was mainly derived from the western population and increased the reliability of the nomogram.

We acknowledged several limitations in our study. First, our study was depended exclusively on a single institutional retrospective analysis, although eligibility criteria have been developed to minimize the selective bias, and we failed to incorporate some important molecular factors (e.g. KRAS, NRAS, BRAF mutation, MSI, or MMR). Second, the information of PNI was unavailable in the SEER cohort, we excluded this parameter during verification, which may limit the power of external validation. In addition, detailed information on recurrence, performance status and postoperative treatment, especially the pharmacological treatments (neoadjuvant/ adjuvant/ others) received by the patients, which are not available in the open-accessed SEER database, was not evaluated in our study.

At present, neoadjuvant therapy for rectal cancer has become the standard treatment strategy for locally advanced rectal cancer (cT3-4, or N+) patients. The complete pathological response after neoadjuvant therapy may be the most reliable prognostic factor, and building a nomogram containing the T downstaging after neo-adjuvant treatment may better predict the survival of patients with rectal cancer. However, limited by the sample size, the patients with neoadjuvant therapy were excluded in our study.

Despite the above-mentioned limitations, we established and validated a novel nomogram for predicting survival of patients with resected RC. To the best of our knowledge, this is the first nomogram for predicting survival of patients with resected RC that is based on a single institutional database with long-term follow-up and validated by SEER database. This model could help physicians consider preoperation serum levels of CEA and other factors to more precisely estimate the survival of individual patients after surgery, and identify those high-risk patients and support decision-making in clinical practice and follow-up strategies. Further studies are warranted on prospective, multi-institutional, large sample size data collection and individualized follow-up, broader geographic recruitment, and incorporation of some other factors (e.g. important molecular factors, the pharmacological treatments, the complete pathological response and T downstaging after neo-adjuvant treatment, etc.) to validate and improve this model in the treatment decision-making field of rectal cancer.

Acknowledgements

We thank International Science Editing (http://www.internationalscienceediting.com) and Dr Junmiao Wen from the Department of Radiation Oncology, Fudan University Shanghai Cancer Center for editing this manuscript.

Ethical approval

All studies involving human participants are based on the ethical standards of institutions and national research committees, as well as the Helsinki Declaration of 1964, its later amendments, or comparable ethical standards. The retrospective analysis was approved by the Ethical Committee and Institutional Review Board of the Second Affiliated Hospital of Soochow University. Since the study does not include interaction with humans or the use of personally identifiable information, the informed consent from patients were waived by our institutional review board. The SEER research data was based on the publicly available SEER database and was permitted by the National Cancer Institute to access the database for research purposes (Private SEER ID 10425-Nov2018).

Competing Interests

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



Author Contribution

Shaonan Fan and Yaqun Zhu conceived the study concept and design, was involved with patient care, and drafted the manuscript. Shaonan Fan, Ting Li, Ping Zhou, Qiliang Peng, Yaqun Zhu were involved with formation of the study concept and design, patient care, and drafting of the manuscript. All authors have read and approved the final version of the manuscript.

Funding

This study was supported by the Jiangsu Commission of Health medical research project [grant number H2018115]; the Second Affiliated Hospital of Soochow University Preponderant Clinic Discipline Group Project funding [grant number XKQ2015005] and Sthe uzhou Introduction Project of Clinical Medical Expert Group [grant number SZYJTD201804].

References

- 1 Siegel, R.L., Miller, K.D. and Jemal, A. (2018) Cancer statistics, 2018. CA Cancer J. Clin. 68, 7–30, https://doi.org/10.3322/caac.21442
- 2 Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424, https://doi.org/10.3322/caac.21492
- 3 Chen, W., Zheng, R., Baade, P.D., Zhang, S., Zeng, H., Bray, F. et al. (2016) Cancer statistics in China, 2015. CA Cancer J. Clin. 66, 115–132, https://doi.org/10.3322/caac.21338
- 4 Li, Q., Dai, W., Li, Y., Xu, Y., Li, X. and Cai, S. (2017) Nomograms for predicting the prognostic value of serological tumor biomarkers in colorectal cancer patients after radical resection. *Sci. Rep.* **7**, 46345, https://doi.org/10.1038/srep46345
- 5 Sauer, R., Liersch, T., Merkel, S., Fietkau, R., Hohenberger, W., Hess, C. et al. (2012) Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. J. Clin. Oncol. 30, 1926–1933, https://doi.org/10.1200/JCO.2011.40.1836
- 6 Trotti, A., Edge, S.B., Byrd, D.R., Compton, C.C., Fritz, A.G. and Greene, F.I. (2010) Colon and Rectum. AJCC Cancer Staging Manual, 7th ed., pp. 143–164
- 7 Maeda, K., Shibutani, M., Otani, H., Nagahara, H., Ikeya, T., Iseki, Y. et al. (2015) Inflammation-based factors and prognosis in patients with colorectal cancer. *World J. Gastrointest Oncol.* **7**, 111–117, https://doi.org/10.4251/wjgo.v7.i8.111
- 8 Zhang, F., Zhang, Y., Zhao, W., Deng, K., Wang, Z., Yang, C. et al. (2017) Metabolomics for biomarker discovery in the diagnosis, prognosis, survival and recurrence of colorectal cancer: a systematic review. *Oncotarget* **8**, 35460–35472
- 9 Malati, T. (2007) Tumour markers: An overview. Indian J. Clin. Biochem. 22, 17–31, https://doi.org/10.1007/BF02913308
- 10 Hotta, T., Takifuji, K., Yokoyama, S., Matsuda, K., Oku, Y., Nasu, T. et al. (2014) Impact of the post/preoperative serum CEA ratio on the survival of patients with rectal cancer. *Surg. Today* 44, 2106–2115, https://doi.org/10.1007/s00595-014-0852-1
- 11 Luo, G., Liu, C., Guo, M., Cheng, H., Lu, Y., Jin, K. et al. (2017) Potential biomarkers in lewis negative patients with pancreatic cancer. *Ann. Surg.* 265, 800–805, https://doi.org/10.1097/SLA.00000000001741
- 12 Sun, Z.-Q., Ma, S., Zhou, Q.-B., Yang, S.-X., Chang, Y., Zeng, X.-Y. et al. (2017) Prognostic value of lymph node metastasis in patients with T1-stage colorectal cancer from multiple centers in China. *World J. Gastroenterol.* **23**, 8582–8590, https://doi.org/10.3748/wjg.v23.i48.8582
- 13 Valentini, V., van Stiphout, R.G.P.M., Lammering, G., Gambacorta, M.A., Barba, M.C., Bebenek, M. et al. (2011) Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. *J. Clin. Oncol.* **29**, 3163–3172, https://doi.org/10.1200/JC0.2010.33.1595
- 14 Wen, J., Liu, D., Xu, X., Chen, D., Chen, Y., Sun, L. et al. (2018) Nomograms for predicting survival outcomes in patients with primary tracheal tumors: a large population-based analysis. *Cancer Manag. Res.* **10**, 6843–6856, https://doi.org/10.2147/CMAR.S186546
- 15 Liang, W., Zhang, L., Jiang, G., Wang, Q., Liu, L., Liu, D. et al. (2015) Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J. Clin. Oncol. 33, 861–869, https://doi.org/10.1200/JC0.2014.56.6661
- 16 Wang, Y., Li, J., Xia, Y., Gong, R., Wang, K., Yan, Z. et al. (2013) Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J. Clin. Oncol.* **31**, 1188–1195, https://doi.org/10.1200/JC0.2012.41.5984
- 17 National Cancer Institute, The surveillance, epidemiology, and end results (SEER) program. https://seer.cancer.gov/ (accessed 31 October 2018)
- 18 Surveillance, Epidemiology and End Results Program, About the SEERProgram, 2018. http://seer.cancer.gov/ about, 2018 (Accessed 05 May 2018)
- 19 Dai, D., Chen, B., Tang, H., Wang, B., Zhao, Z., Xie, X. et al. (2016) Nomograms for Predicting the Prognostic Value of Pre-Therapeutic CA15-3 and CEA Serum Levels in TNBC Patients. *PLoS One* **11**, e0161902, https://doi.org/10.1371/journal.pone.0161902
- 20 Kanakis George., Kaltsas Gregory (2012) Biochemical markers for gastroenteropancreatic neuroendocrine tumours (GEP-NETs). Best Pract. Res. Clin. Gastroenterol. 26, 791–802, https://doi.org/10.1016/j.bpg.2012.12.006
- 21 Buijsen, J., van Stiphout, R.G., Menheere, P.P.C.A., Lammering, G. and Lambin, P. (2014) Blood biomarkers are helpful in the prediction of response to chemoradiation in rectal cancer: a prospective, hypothesis driven study on patients with locally advanced rectal cancer. *Radiother. Oncol.* **111**, 237–242, https://doi.org/10.1016/j.radonc.2014.03.006
- 22 Moreno, G.V., Cejas, P., Blanco, C.M., Feliu, B.J., de Castro, C.J., Belda-Iniesta, C. et al. (2009) Prognostic value of carcinoembryonic antigen level in rectal cancer treated with neoadjuvant chemoradiotherapy. *Int. J. Colorectal Dis.* **24**, 741–748, https://doi.org/10.1007/s00384-009-0682-6
- 23 Kim, C.W., Yoon, Y.S., Park, I.J., Lim, S.-B., Yu, C.S. and Kim, J.C. (2013) Elevation of preoperative s-CEA concentration in stage IIA colorectal cancer can also be a high risk factor for stage II patients. Ann. Surg. Oncol. 20, 2914–2920, https://doi.org/10.1245/s10434-013-2919-4



- 24 Takagawa, R., Fujii, S., Ohta, M., Nagano, Y., Kunisaki, C., Yamagishi, S. et al. (2008) Preoperative serum carcinoembryonic antigen level as a predictive factor of recurrence after curative resection of colorectal cancer. *Ann. Surg. Oncol.* **15**, 3433–3439, https://doi.org/10.1245/s10434-008-0168-8
- 25 Compton, C., Fenoglio-Preiser, C.M., Pettigrew, N. and Fielding, L.P. (2000) American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer* 88, 1739–1757,

https://doi.org/10.1002/(SICI)1097-0142(20000401)88:7%3c1739::AID-CNCR30%3e3.0.C0;2-T

- 26 Patel, S.A., Chen, Y.-H., Hornick, J.L., Catalano, P., Nowak, J.A., Zukerberg, L.R. et al. (2014) Early-stage rectal cancer: clinical and pathologic prognostic markers of time to local recurrence and overall survival after resection. *Dis. Colon Rectum* 57, 449–459, <u>https://doi.org/10.1097/DCR.0b013e3182a70709</u>
- 27 Iasonos, A., Schrag, D., Raj, G.V. and Panageas, K.S. (2008) How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26, 1364–1370, https://doi.org/10.1200/JC0.2007.12.9791
- 28 Zhang, Z.-Y., Gao, W., Luo, Q.-F., Yin, X.-W., Basnet, S., Dai, Z.-L. et al. (2016) A nomogram improves AJCC stages for colorectal cancers by introducing CEA, modified lymph node ratio and negative lymph node count. *Sci. Rep.* **6**, 39028, https://doi.org/10.1038/srep39028