

# Construction and validation of a clinicopathologic signature for predicting the prognosis of stage II and III colorectal cancer

**Weiqiang You**

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

**Zerong Cai**

Sun Yat- sen University

**Nengquan Sheng**

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

**Li Yan**

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

**Huihui Wan**

Shanghai Academy of Science and Technology

**Yongkun Wang**

Shanghai Academy of Science and Technology

**Jian Ouyang**

Shanghai Academy of Science and Technology

**Lu Xie**

Shanghai Academy of Science and Technology

**Xiaojian Wu**

Sun Yat- sen University

**Zhigang Wang** (✉ [surlab@hotmail.com](mailto:surlab@hotmail.com))

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

---

## Research Article

**Keywords:** colorectal cancer, prognosis, risk score, clinicopathologic signature

**Posted Date:** May 14th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-497202/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

# Abstract

**Background:** Various clinical and pathological indicators affect the prognosis of stage II and III colorectal cancer (CRC). Nevertheless, few studies have systematically integrated these indicators to construct a signature for assessing the prognostic risk.

**Methods:** Patients with stage II-III CRC underwent R0 radical resection from 2009 to 2016 were included in this study. All clinical and follow-up data were collected. Our study mainly incorporated an internal training cohort and an external validation cohort from two centers. Data processing and data analysis were performed in the R language. Cox proportional hazard regression was used for univariate and multivariable analyses. The log-rank test was performed to compare prognosis among groups.

**Results:** A total of 1200 eligible patients were included in our study. 8 variables, including T/N stage, lymphatic/vascular infiltration, preoperative CEA, CA125 and CA199, were included and the nomogram was established in the signature for predicting OS. The concordance index of the signature was 0.72. The 3-year and 5-year calibration curves of CRC based on the nomogram showed perfect correlation between predicted and observed outcomes. Obvious differences were observed in the survival of different risk groups ( $p < 0.001$ ). Patients with low risk score signature had a 5-year OS rate of 77%, whereas patients with high risk score signature had the worst 5-year OS rate (only 8%). Furthermore, our signature also achieved similar performance in external cohort validation.

**Conclusions:** The signature based on clinical and pathologic factors had favorable accuracy for predicting the prognosis of CRC. Therefore, our signature may provide significant suggestions for clinical work, particularly in identifying high-risk stage II and III CRC.

## Introduction

Colorectal cancer (CRC) is commonly diagnosed malignancies worldwide, which has the third highest incidence and the second highest mortality rate in cancers, causing more than 50,000 deaths(1). Despite continuous advances in treatment methods, such as improvements in surgery, and improvements in targeted drugs and the clinical application of immunotherapy, the prognosis of CRC is still not optimistic. Even after resection, the 5-year survival rate of CRC was 65%(2). Postoperative recurrence and metastasis are the main factors causing death and affecting the prognosis of patients.

How can we evaluate patient prognosis after surgery? It is a difficult clinical problem. Currently, the tumor-node-metastasis (TNM) staging system according to the 8th AJCC criterion is the main reference for adjuvant treatment and prognosis(3-5). Generally, patients with a higher stage have a worse prognosis. However, our previous study showed that the prognosis of stage II colon cancer may be worse than that of stage III colon cancer(6). This shows that there may be certain shortcomings in assessing the prognosis based on TNM staging alone. In addition, the TNM system cannot provide all the information about a patient's status after surgery. It is undeniable that the TNM system misses some indicators that are meaningful for the prognosis of CRC, such as high-risk factors, pathological variables

and serum tumor markers. Patients with lymphatic invasion in CRC liver metastasis demonstrated adverse survival, especially combined with vascular invasion(7). In addition, serum tumor markers, including CEA and CA19-9, are classic indicators closely related to the prognosis of CRC. Patients with elevated periprocedural CEA have a worse prognosis and a higher risk of recurrence than patients with normalized CEA(8, 9). The level of CA19-9 is usually normal in CRC. However, adding it to postoperative surveillance is recommended because of its significant value for prognosis, especially in patients with *BRAF* mutations(10).

Therefore, accurate evaluation of the prognosis of CRC is a complex and multidimensional problem. The TNM system cannot cover all the indicators, and the significance of a single serum tumor marker for gastrointestinal tumors is limited. How to systematically integrate these indicators, analyze the weight of these markers, and further develop a mathematical signature to accurately evaluate the prognosis are the primary questions of this study.

## Methods

### 1. Data Collection

This study was a double-center retrospective clinical study registered in the Chinese Clinical Trial Registry (Approval No. ChiCTR1800016906). Our study was approved by the ethics committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital (Approval No. 2018-KY-031K). The internal training cohort from Shanghai Jiao Tong University Affiliated Sixth People's Hospital was used to develop the signature, while the external cohort from the Sixth Affiliated Hospital of Sun Yat-sen University was used to validate our signature. All patients received radical resection (R0) and were pathologically diagnosed with CRC (stage II-III) from January 2009 to December 2016. The informed consent was waived by the ethics committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital because of the retrospective design. The criteria for exclusion were as follows: (1) stage I or IV CRC; (2) unsuitable pathological type, including neuroendocrine tumors and malignant melanomas; (3) no available preoperative serum tumor markers.

Patient clinical data, including sex, age, tumor markers and pathological reports, were mainly provided by medical histories and the electronic medical record department. Pathological staging was based on the 8th AJCC criterion for CRC. Preoperative tumor marker values were examined within one week before surgery. All patients were followed up according to current NCCN guidelines, including analysis of serum tumor markers, colonoscopy, chest X-ray and CT (or MRI). Patient follow-up data were updated by telephone, email and medical histories. Overall survival (OS) was defined as the time from surgery to death.

### 2. Statistical analysis

Data processing and data analysis were performed in the R language (version 3.4.4, <https://www.r-bloggers.com/r-3-4-4-released/>). Several missing data were estimated by multiple imputations in the

“mice” package.

For prognostic signature, we performed a univariate analysis of all variables by Cox proportional hazards regression. Then, the significant covariates (score test  $p < 0.05$ ) were retained in the multivariate Cox proportional hazards regression. We performed stepwise and backward selection processes to achieve the final signature using the Akaike information criterion (AIC). The discrimination of the signature was evaluated by the concordance index (C-index) and corrected 1,000 times by bootstrapping. Furthermore, patients were classified into three groups (high, intermediate, and low risk) based on risk score, which was calculated with the prognostic signature. The Kaplan-Meier (K-M) method was applied to compare survival curves for these three groups. Hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated, and the log-rank test was used to determine the significant differences. A nomogram was employed to predict the 3-year and 5-year patient survival rates. Additionally, calibration curves were drawn for 3-year and 5-year survival using the nomogram to evaluate our signature from observations and predictions. P values were two-sided, with statistically significant differences at  $p < 0.05$ .

## Results

### 1. Study design and patient data

Our study included two major databases, with an internal training cohort used to build the signature and an external cohort used to validate the signature. A total of 707 patients from Shanghai Jiao Tong University Affiliated Sixth People's Hospital were screened for internal cohort to construct our signature. An external cohort from the Sixth Affiliated Hospital of Sun Yat-sen University with 493 cases, validated our signature. In all, the two cohorts used in our signature included 1200 patients. (Fig. 1).

In our study, we collected clinical information and follow-up data from the internal training cohort patients. The characteristics of the training cohort were shown in Table 1. 431 patients (61.0%) were male, and others were female. The median age of these patients was 65. The Median follow-up time was 47 months. Majority of patients were diagnosed with T4 stage (74.8%), while patients with T1 and T2 accounted for only 2.7%. Most patients had N0 stage (59.3%), and 288 patients had lymph node metastasis, including 170 cases with T1 and 118 cases with T2. The mean values of CEA, CA125 and CA19-9 are 13.97 ng/ml, 21.62 U/ml and 23.8 U/ml, respectively. 495 patients were still alive by the time of follow-up, while 212 patients died.

Table 1  
 Patient demographics and clinicopathologic data of internal training cohort (NA, not available; CSA, cross section area)

<b>Variables</b>	<b>CRC patients (N = 707)</b>
Gender	
Male	431 (61.0%)
Female	276 (39.0%)
Median follow-up months	47.0
Median age (mean $\pm$ SD)	65 (64.40 $\pm$ 11.78)
T stage	
T1 + T2	19 (2.7%)
T3	159 (22.5%)
T4	529 (74.8%)
N stage	
N0	419 (59.3%)
N1	170 (24.0%)
N2	118 (16.7%)
Tumor location	
Right colon	224 (31.7%)
Left colon	216 (30.6%)
Rectum	267 (37.8%)
Tumor CSA (mean $\pm$ SD)	21.91 $\pm$ 29.03 cm <sup>2</sup>
Pathological type	
Adenocarcinoma	664 (93.9%)
Mucinous Adenocarcinoma and Signet-ring cell carcinoma	43 (6.1%)
Degree of histologic differentiation	
Moderate and well	420 (59.4%)
Poor	287 (40.6%)
Lymphatic infiltration	

<b>Variables</b>	<b>CRC patients (N = 707)</b>
Present	283 (40.0%)
Absent	424 (60.0%)
<b>Vascular infiltration</b>	
Present	78 (11.0%)
Absent	629 (89.0%)
<b>Nerve infiltration</b>	
Present	675 (95.5%)
Absent	32 (4.5%)
Median Ki67 (Range)	60% (5%-90%)
NLR (mean ± SD)	3.68 ± 5.01
PLR (mean ± SD)	179.30 ± 134.68
CEA value (mean ± SD, ng/ml)	13.97 ± 39.05
CA125 value (mean ± SD, U/ml)	21.62 ± 31.11
CA19-9 value (mean ± SD, U/ml)	23.98 ± 40.10
<b>Metastasis or recurrence</b>	
Yes	201 (28.4%)
No	506 (71.6%)
<b>Survival status</b>	
Alive	495 (70.0%)
Dead	212 (30.0%)

## 2. Univariate and multivariate analyses of clinical indicators for prognostic prediction in the internal training cohort

To select the variables that were suitable for inclusion in our signature, we performed univariate analyses. Our results suggested that 9 variables were prognostic factors for OS, including T stage, N stage, pathological type, histologic differentiation, lymphatic infiltration, vascular infiltration, CEA, CA125 and CA19-9 (Table 2;  $p < 0.05$ ).

Table 2  
Univariate and multivariate Cox regression analyses for OS (training cohort)

Variables	OS	
	HR (95% CI)	p-value
Univariate analysis		
Gender (male vs. female)	1.097 (0.826–1.456)	0.522
Age ( $\geq$ 65 vs. $\leq$ 65)	0.994 (0.756–1.307)	0.966
T stage (T4 vs. T3)	1.656 (1.155–2.373)	0.006
N stage		
N0	1 (Referent)	
N1	1.858 (1.324–2.606)	0.000
N2	3.514 (2.541–4.860)	0.000
Tumor location		
Right colon	1 (Referent)	
Left colon	1.050 (0.741–1.487)	0.784
Rectum	1.116 (0.802–1.552)	0.516
Tumor CSA	0.997 (0.990–1.004)	0.433
Pathological type (adenocarcinoma vs. other)	0.529 (0.340–0.823)	0.005
Histologic differentiation (poor vs. moderate and well)	1.583 (1.204–2.081)	0.001
Lymphatic infiltration (present vs. absent)	1.560 (1.159–2.099)	0.003
Vascular infiltration (present vs. absent)	2.154 (1.509–3.077)	0.000
Nerve infiltration (present vs. absent)	2.222 (0.826–5.978)	0.114
Ki67	0.997 (0.990–1.005)	0.494
NLR	1.017 (0.994–1.040)	0.159
PLR	1.001 (1.000-1.002)	0.148
CEA	1.005 (1.003–1.007)	0.000
CA125	1.004 (1.000-1.007)	0.024
CA19-9	1.006 (1.003–1.008)	0.000
Multivariate analysis		

Variables	OS	
T stage (T4 vs. T3)	1.468 (1.019–2.116)	0.039
N stage		
N0	1 (Referent)	
N1	1.891 (1.246–2.869)	0.003
N2	3.782 (2.471–5.787)	0.000
Pathological type (adenocarcinoma vs. other)	0.567 (0.357–0.902)	0.017
Histologic differentiation (poor vs. moderate and well)	1.332 (0.995–1.782)	0.054
Lymphatic infiltration (present vs. absent)	0.631 (0.420–0.948)	0.026
Vascular infiltration (present vs. absent)	1.597 (1.098–2.324)	0.014
CEA	1.003 (1.001–1.005)	0.012
CA125	1.005 (1.002–1.009)	0.002
CA19-9	1.003 (1.001–1.006)	0.005

Then, we used these variables as potential prognostic factors to construct prognostic signature with the Cox proportional hazards model. AIC, stepwise, and backward analyses were performed as variable selection methods. Finally, 8 indicators were selected to construct a prognostic signature for OS in CRC, including T stage, N stage, histologic differentiation, lymphatic infiltration, vascular infiltration, CEA, CA125 and CA19-9. Our multivariate analysis results suggested that these 8 variables were all independent prognostic factors except tumor histologic differentiation (table 3;  $p < 0.05$ ). The C-index was 0.72, corrected with 1,000 permutations (supplementary table 1).

### 3. Evaluation and determination of the accuracy and predictive power of the signature

To evaluate the prognosis of patients more intuitively, we developed nomogram with Cox regression model. All the variables in the nomogram had a weighted score, and we could predict the 3-year or 5-year survival outcome by the sum of the scores (Fig. 2). To further examine the importance of these variables and calculate the risk score, we developed a nonparametric approach in our signature using random survival forest. Logarithmic transformation was performed for CEA and CA125. Our results suggested that N stage had the largest influence on OS with a VIMP of positive value 0.7217, followed by vascular infiltration, tumor histologic differentiation, CA125, T stage, CEA and CA19-9 (supplementary Fig. 1). The predictive accuracy of the OS signature using time-dependent ROC analysis was relatively high. The AUC of our signature based on the risk score was 0.761 at 3 years and was 0.741 at 5 years (Fig. 3a and 3b). The calibration curves for CRC based on our signature showed excellent correlation between predicted and observed outcomes for OS prediction at 3 years and 5 years (Fig. 3c and 3d). All these results indicated that our signature had good accuracy and prediction ability.

#### **4. Prognosis Among Groups With Different Risk Scores**

We assessed patients according to the risk score achieved from the signature using these different variables. Then, patients were classified into three groups (high, intermediate, and low risk) by the cutoff of the risk score. Kaplan-Meier curves were applied to compare survival differences. Compared with the low-risk group, the intermediate- and high-risk groups had hazard ratios of 3.28 (95% CI, 2.37–4.52) and 8.67 (95% CI, 5.86–12.80), respectively (Fig. 4,  $p < 0.0001$ ). The 5-year OS rate of low risk group was 77%, much higher than intermediate group (46%) and high risk group (8%).

#### **5. Validation Of Our Signature In An External Cohort**

Based on the previous results, our signature showed good accuracy. To verify whether our signature was suitable for other hospitals or centers, we collected data from an external cohort for validation. Data were collected from a total of 493 CRC patients. Univariate analysis results showed that the variables included in our predictive OS signature of CRC were all prognostic factors (supplementary table 2,  $p < 0.05$ ) except CA125. The calibration curves suggested perfect correlation between predicted and observed outcomes for OS prediction at 2 years and 3 years (Fig. 5a and 5b). In addition, according to the scoring criteria of our signature, we divided the patients from the external cohort into three groups (high, intermediate, and low risk), and we obtained similar results. The survival among the groups was obviously different. The OS between the three groups were significantly different (supplementary Fig. 2,  $p < 0.001$ ).

#### **6. Website for predicting the prognosis of stage II and III CRC patients**

Based on our signature of 1200 patients from two cohorts, we developed a website for predicting the prognosis of stage II-III CRC (<http://115.28.66.83/liuyuan/coad.php>).

## **Discussion**

Recurrence and metastasis are characteristics of malignant gastrointestinal tumors and result in poor prognosis of patients. It is well known that the current guidelines for postoperative treatment and patients' follow-up are based on TNM stage, pathological factors, CEA and CA19-9. However, the guidelines do not provide individualized risk assessment for patients. Considering this, we conducted this study, which complements the current guideline evaluation system. How to evaluate the risk of postoperative survival has always been an area of interest in clinical practice. In our study, 1200 patients from two cohorts were included. Then, we systematically analyzed the effects of different variables on the prognosis of CRC and constructed signature that predicted postoperative OS. The C-index was higher than 0.7, indicating good discrimination ability. The calibration curves of the signature were quite good. These results indicated that our signature had promising accuracy and predictive ability. Furthermore, our validation results demonstrated that our signature was widely applicable and had prominent significance in clinical applications.

Multiple studies have focused on how to detect recurrence and metastasis and evaluate the survival of gastrointestinal tumors. Similar to our study, Martin R. Weiser et al. generated a signature for predicting

local recurrence from 1,320 CRCs by comprehensively assessing patient clinical characteristics. The nomogram scale found that patients in the high-risk group had a lower probability of relapse than those in the low-risk group(11). A risk classification system based on the expression of Apc, Fhit, and Her2 and five pathological parameters accurately divided the patients with gastric cancer into three classes(12). As reported by a previous study, four DNA methylation signatures could serve as an independent prognostic factors and predicted the prognosis of gastric cancer patients(13). Zhou Z et al. constructed a mathematical model with autophagy-related genes and assessed the risk of recurrence and metastasis in CRC patients according to the expression level of 5 genes. The patients were divided into a high-risk group and a low-risk group, and the results showed a significant difference in survival between the two groups(14). A signature for risk estimation taking into account genetic susceptibility also exists(15). Another study analyzed 735 specimens of stage II CRC. They found that comprehensive expression of six miRNAs could evaluate the risk of recurrence in patients, and patients in the low-risk group had longer survival than those in the high-risk group(16). Similarly, there was also evidence that microRNA polymorphisms were closely associated with the prognosis of gastric cancer(17). The immunoscore is another powerful factor affecting the prognosis of gastrointestinal tumors. Many studies have shown that the density of CD3<sup>+</sup> and CD8<sup>+</sup> lymphocyte populations in the stroma and invasion margins of the tumor microenvironment could effectively predict the prognosis of patients with gastrointestinal tumors(18–21).

These signatures have different characteristics and they need to detect additional metrics. This is not only technically difficult, but also increases the medical burden on patients. Therefore, the clinical application of these signatures is still not promising at present. Our signature is more intuitive, integrating the TNM staging system, pathological parameters and tumor markers without the need for additional testing. These characteristics make our signature more clinically convenient and easy to use. According to current guidelines, patients undergoing radical surgery for gastrointestinal tumors are usually followed up every 3–6 months(3, 4, 22, 23). In addition, the treatment of gastrointestinal tumors after radical surgery still has some problems, especially in CRC(24, 25). There was no significant disadvantage between 3-month adjuvant chemotherapy and 6-month chemotherapy in low-risk stage III CRC patients(26). However, these guidelines do not have individual follow-up and adjuvant therapy recommendations. According to the results predicted by our signature, patients with the same stage had distinct differences in the risk of recurrence and survival outcomes. Therefore, clinical staff can use our signature to evaluate patients when advising patients on postoperative adjuvant treatment and follow-up. For patients with a high risk of poor prognosis, adequate adjuvant therapy and intensive follow-up recommendations are necessary. In other words, patients at low risk can receive short doses of chemotherapy and long intervals of follow-up. It is worth noting that our signature also has some shortcomings. Our signature does not include the latest clinical indicators, such as microsatellite status, Ras status, and Braf status, because only a small population of patients (after 2015) had records on these indicators, whereas the patients included in our study were from 2009 to 2016. By the way, we are now building the second signature that will incorporate these new indicators. Additionally, the patients

included in our study were all Chinese, and there was a lack of international data cooperation. Finally, prospective clinical trials should be performed to validate our signature further.

## Conclusions

Our signature integrated various classical indicators, including the TNM staging system, pathological factors and preoperative serum tumor markers. Our signature is meaningful for predicting the risk of survival after surgery for CRC and have broad clinical application prospects.

## List Of Abbreviations

CRC: colorectal cancer; OS: overall survival; TNM: tumor–node–metastasis; CEA: carcinoembryonic antigen; CIs: confidence intervals; AIC: Akaike information criterion; HRs: hazard ratios; CSA: tumor cross-sectional area; NLR: neutrophil to lymphocyte ratio; PLR: platelet to lymphocyte ratio.

## Declarations

*Ethics approval and consent to participate.* This study was approved by the ethics committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital (Approval No. 2018-KY-031K). The requirement of informed consent from the patients was waived by the ethics committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital because of the retrospective design, and patients' information was protected. All procedures performed in the study involving human participants were in accordance with the Declaration of Helsinki.

*Consent for publication.* Not applicable.

*Availability of data and materials.* The datasets in our study are available from the corresponding author on reasonable request. No administrative permission was required to access the raw data from our databases. The clinical data used in our study is not publicly available because it is being used in other unpublished studies.

*Competing interests.* The authors have declared that no conflict of interest exists.

*Funding.* This work was supported by National Key Clinical Discipline, Shanghai Municipal Education Commission-Gaofeng Clinical Medicine Grant Support (no.20172023), Shanghai Science and Technology Commission Medical Project (no.16411953200), Shanghai Pujiang Program (no.16PJ1408200), Natural Science Foundation of Shanghai (no.16ZR1449600), National Natural Science Foundation of China (no.81602689).

*Author Contributions.* Z.W. conceived the project. W.Y., L.Y. and N. S. collected the clinical and follow-up data from Shanghai Jiao Tong University Affiliated Sixth People's Hospital. X.W. and Z.C. collected the

clinical and follow-up data from the Sixth Affiliated Hospital of Sun Yat-sen University. L.X., H.W., J.O. and Y.W. analyzed all data using the R language. W.Y. wrote the manuscript.

*Acknowledgments.* We sincerely thank the ethics committee and the electronic patients' record department. The authors would like to appreciate the language editing service provided by American Journal Experts.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7-30.
2. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin.* 2017;67(3):177-93.
3. Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen YJ, Ciombor KK, et al. NCCN Guidelines Insights: Colon Cancer, Version 2.2018. *J Natl Compr Canc Netw.* 2018;16(4):359-69.
4. Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen YJ, Ciombor KK, et al. Rectal Cancer, Version 2.2018, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2018;16(7):874-901.
5. Benson AB, Venook AP, Al-Hawary MM, Arain MA, Chen YJ, Ciombor KK, et al. NCCN Guidelines Insights: Rectal Cancer, Version 6.2020. *J Natl Compr Canc Netw.* 2020;18(7):806-15.
6. You W, Sheng N, Yan L, Chen H, Gong J, He Z, et al. The difference in prognosis of stage II and III colorectal cancer based on preoperative serum tumor markers. *J Cancer.* 2019;10(16):3757-66.
7. de Ridder JA, Knijn N, Wiering B, de Wilt JH, Nagtegaal ID. Lymphatic Invasion is an Independent Adverse Prognostic Factor in Patients with Colorectal Liver Metastasis. *Ann Surg Oncol.* 2015;22 Suppl 3:S638-45.
8. Konishi T, Shimada Y, Hsu M, Tufts L, Jimenez-Rodriguez R, Cercek A, et al. Association of Preoperative and Postoperative Serum Carcinoembryonic Antigen and Colon Cancer Outcome. *JAMA Oncol.* 2018;4(3):309-15.
9. Nakamura Y, Shida D, Tanabe T, Takamizawa Y, Imaizumi J, Ahiko Y, et al. Prognostic impact of preoperatively elevated and postoperatively normalized carcinoembryonic antigen levels following curative resection of stage I-III rectal cancer. *Cancer Med.* 2020;9(2):653-62.
10. Thomsen M, Skovlund E, Sorbye H, Bolstad N, Nustad KJ, Glimelius B, et al. Prognostic role of carcinoembryonic antigen and carbohydrate antigen 19-9 in metastatic colorectal cancer: a BRAF-mutant subset with high CA 19-9 level and poor outcome. *Br J Cancer.* 2018;118(12):1609-16.
11. Weiser MR, Landmann RG, Kattan MW, Gonen M, Shia J, Chou J, et al. Individualized prediction of colon cancer recurrence using a nomogram. *J Clin Oncol.* 2008;26(3):380-5.
12. Bria E, De Manzoni G, Beghelli S, Tomezzoli A, Barbi S, Di Gregorio C, et al. A clinical-biological risk stratification model for resected gastric cancer: prognostic impact of Her2, Fhit, and APC expression status. *Ann Oncol.* 2013;24(3):693-701.

13. Li C, Zheng Y, Pu K, Zhao D, Wang Y, Guan Q, et al. A four-DNA methylation signature as a novel prognostic biomarker for survival of patients with gastric cancer. *Cancer Cell Int.* 2020;20:88.
14. Zhou Z, Mo S, Dai W, Ying Z, Zhang L, Xiang W, et al. Development and Validation of an Autophagy Score Signature for the Prediction of Post-operative Survival in Colorectal Cancer. *Front Oncol.* 2019;9:878.
15. Hsu L, Jeon J, Brenner H, Gruber SB, Schoen RE, Berndt SI, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. *Gastroenterology.* 2015;148(7):1330-9 e14.
16. Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol.* 2013;14(13):1295-306.
17. Stenholm L, Stoehmacher-Williams J, Al-Batran SE, Heussen N, Akin S, Pauligk C, et al. Prognostic role of microRNA polymorphisms in advanced gastric cancer: a translational study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *Annals of Oncology.* 2013;24(10):2581-8.
18. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer.* 2012;12(4):298-306.
19. Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L, et al. ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. *Ann Surg.* 2018;267(3):504-13.
20. Galon J, Fridman WH, Pages F. The adaptive immunologic microenvironment in colorectal cancer: a novel perspective. *Cancer Res.* 2007;67(5):1883-6.
21. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou F-S, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *The Lancet.* 2018;391(10135):2128-39.
22. Ajani JA, D'Amico TA, Almhanna K, Bentrem DJ, Chao J, Das P, et al. Gastric Cancer, Version 3.2016, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2016;14(10):1286-312.
23. Glynne-Jones R, Wyrwicz L, Tiret E, Brown G, Rodel C, Cervantes A, et al. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2017;28(suppl\_4):iv22-iv40.
24. Lonardi S, Sobrero A, Rosati G, Di Bartolomeo M, Ronzoni M, Aprile G, et al. Phase III trial comparing 3-6 months of adjuvant FOLFOX4/XELOX in stage II-III colon cancer: safety and compliance in the TOSCA trial. *Ann Oncol.* 2016;27(11):2074-81.
25. Andre T, Vernerey D, Mineur L, Bennouna J, Desrame J, Faroux R, et al. Three Versus 6 Months of Oxaliplatin-Based Adjuvant Chemotherapy for Patients With Stage III Colon Cancer: Disease-Free Survival Results From a Randomized, Open-Label, International Duration Evaluation of Adjuvant (IDEA) France, Phase III Trial. *Journal of Clinical Oncology.* 2018;36(15):1469+.
26. Schilsky RL. A New IDEA in Adjuvant Chemotherapy for Colon Cancer. *N Engl J Med.* 2018;378(13):1242-4.

## Figures

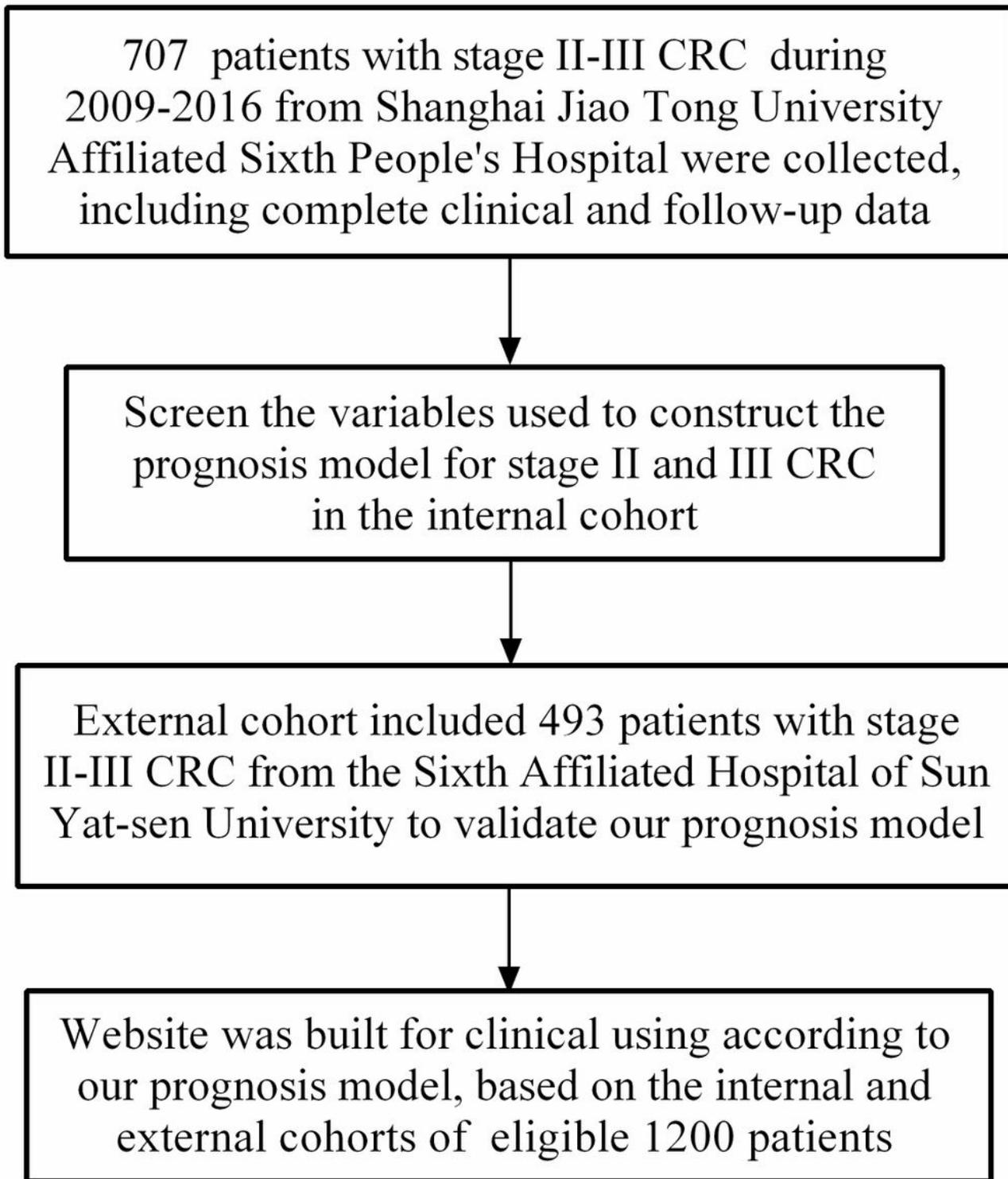
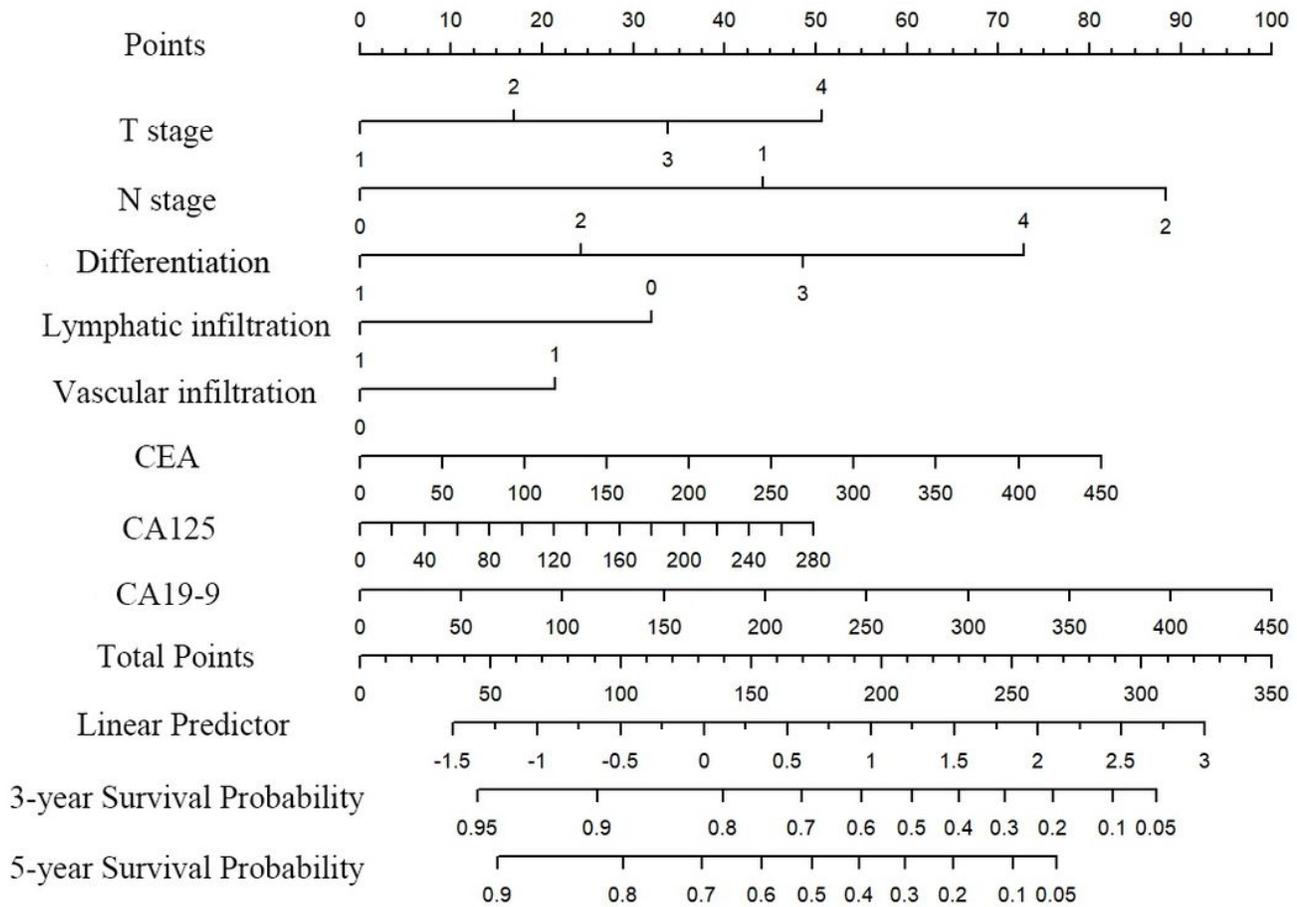


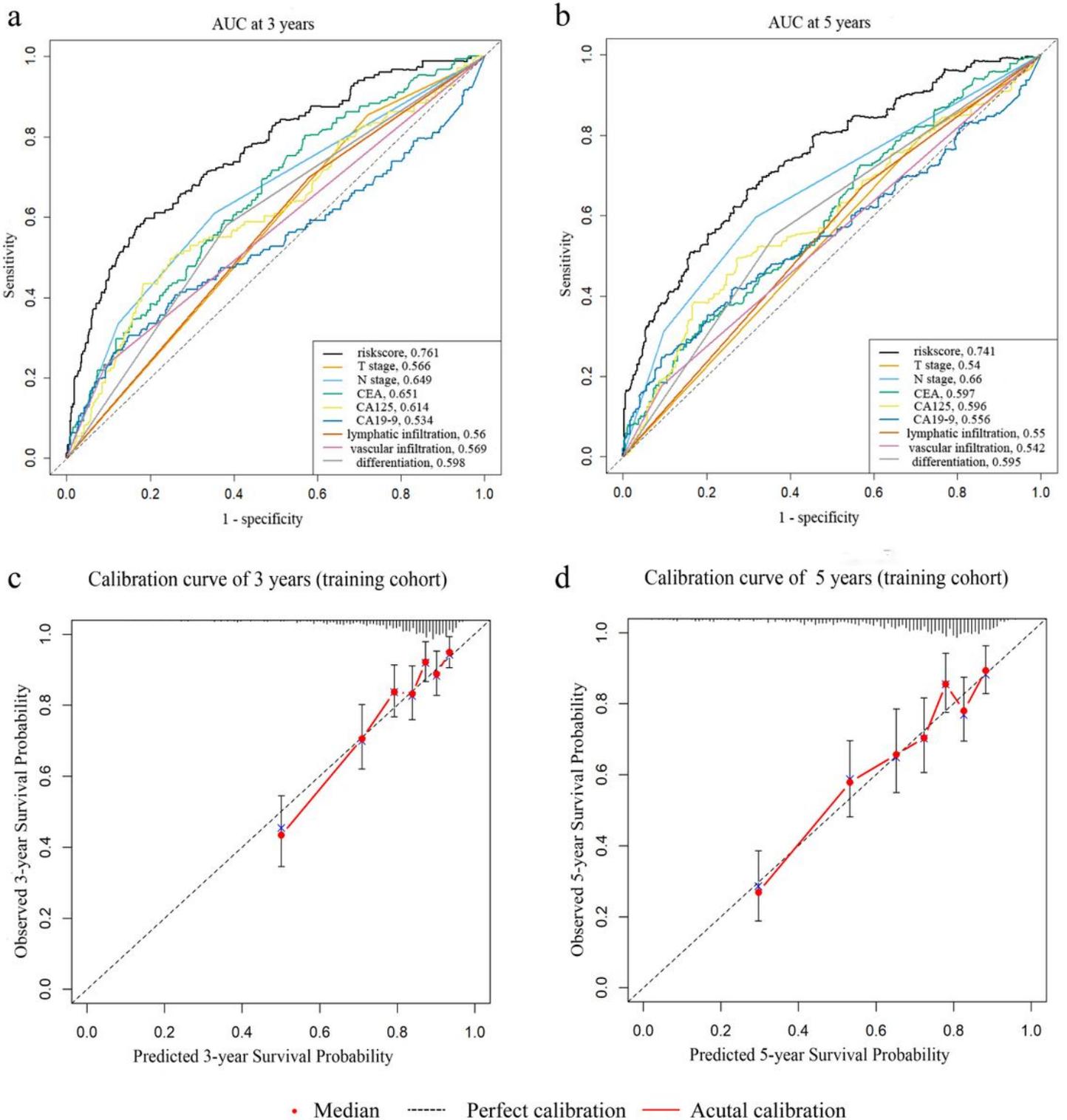
Figure 1

Study design.



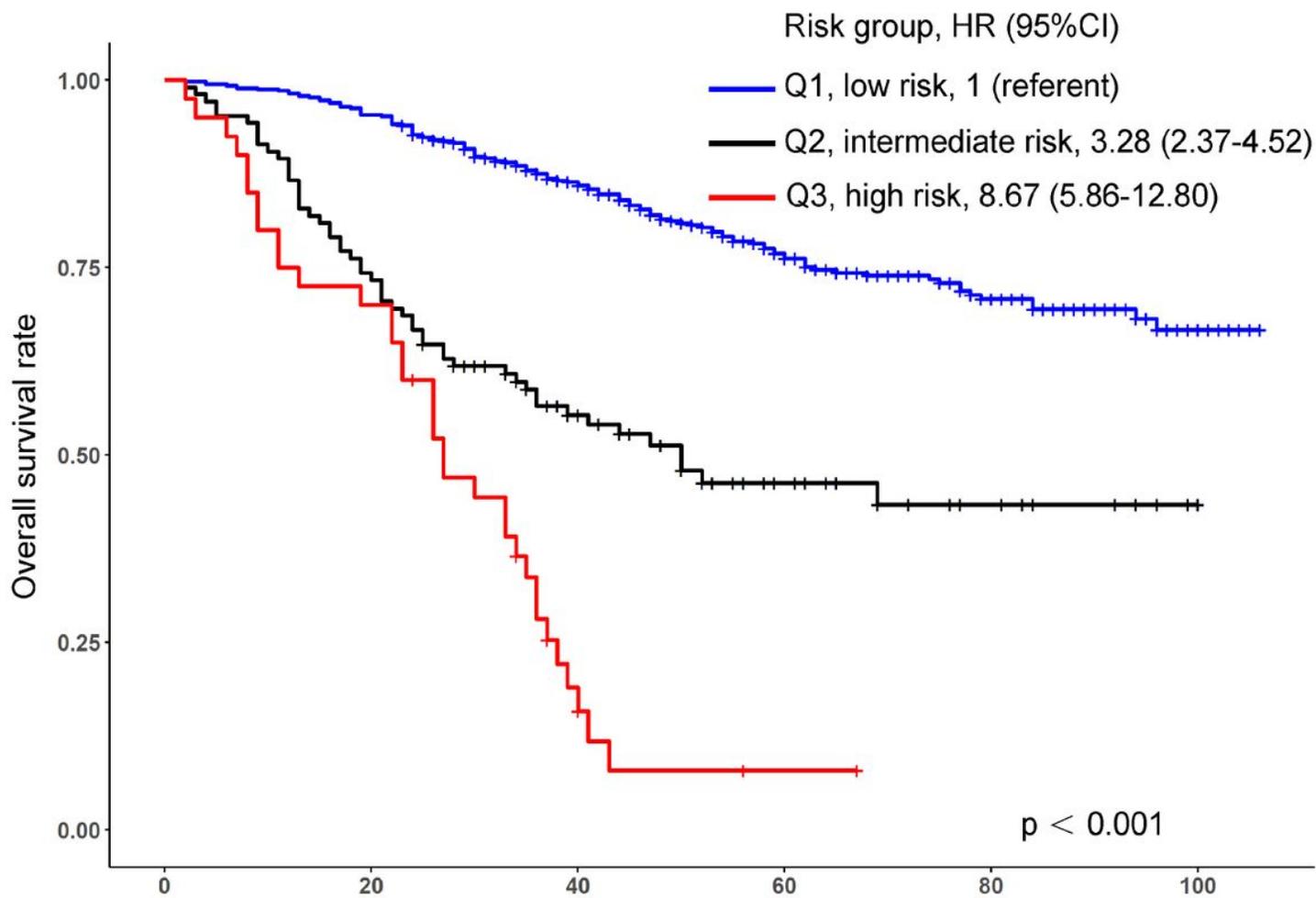
**Figure 2**

Nomograms for variables from the internal training cohort included in the signature. Lymphatic infiltration, vascular infiltration, nerve infiltration: 0-absent, 1-present. Histologic differentiation: 1-high, 2-moderate, 3-low, 4-unhistologic differentiation.



**Figure 3**

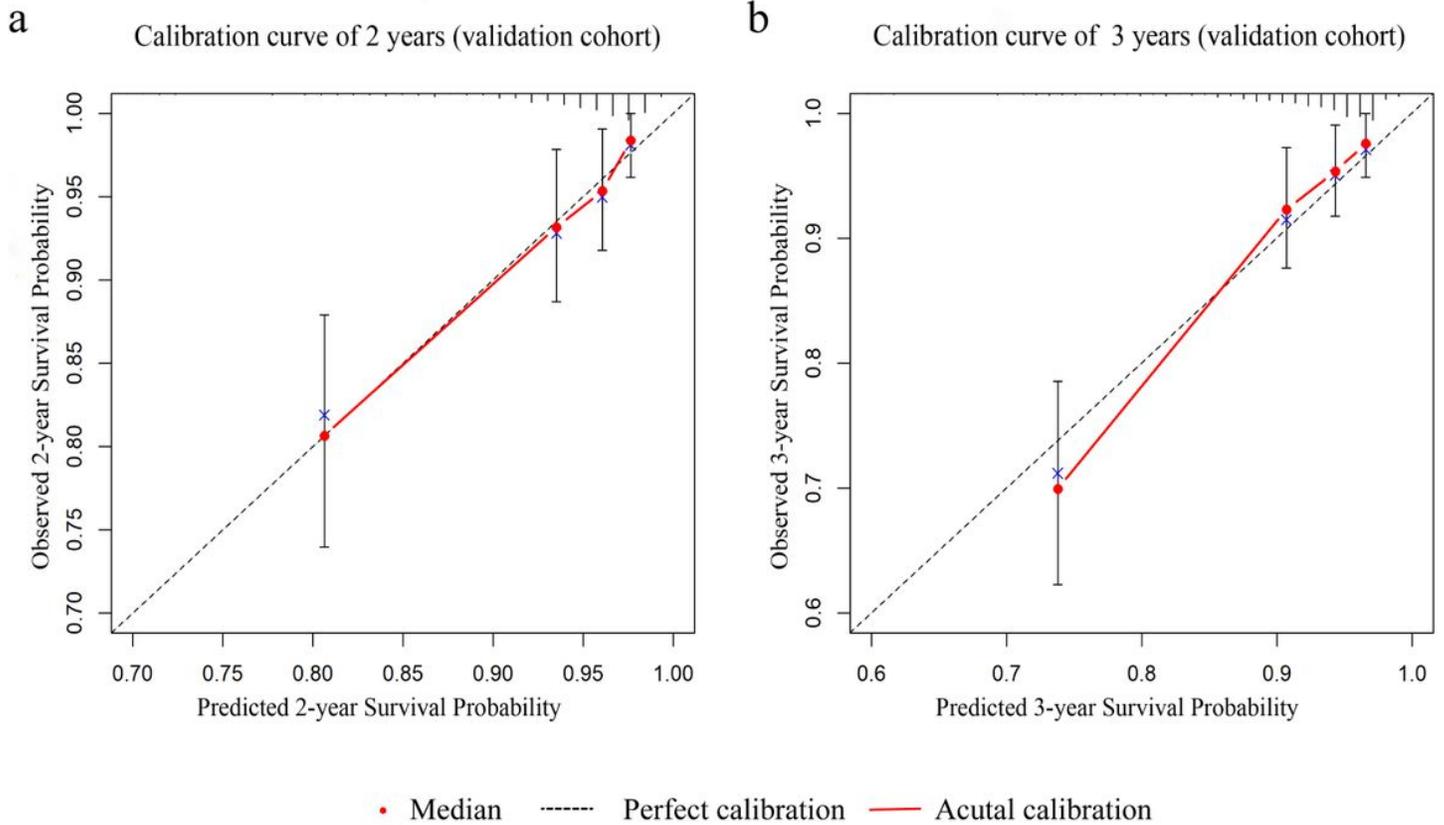
Evaluation the accuracy and predictive power of the OS signature. (a, b) Time-dependent receiver operating characteristic curves showing the predictive accuracy of the OS signature at 3 years (a) and 5 years (b). (c, d) The calibration curves of CRC based on the internal cohort nomograms between predicted and observed DFS and OS outcomes at 3 years (c) and 5 years (d).



	Months					
No. at risk	0	20	40	60	80	100
Q1	562	536	377	222	123	23
Q2	105	78	45	20	9	1
Q3	40	28	6	1	0	0

**Figure 4**

K-M curves for differences in patient survival in the high-, intermediate-, and low-risk groups from the internal cohort.



**Figure 5**

The calibration curves for external cohort validation between predicted and observed DFS and OS outcomes at 2 years (a) and 3 years (b).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfiguresandtables.docx](#)