A prognostic nomogram and web-based real-time calculator for predicting CSS and OS of primary rectal adenocarcinoma patients

Zhenshuang Wang¹, Yi Yang⁴, Jingjing Jing¹,²,³, Yuan Yuan¹,²,³*, Chenzhong Xing¹,³*

* Correspondence:
Dr. Chenzhong Xing, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, the First Hospital of China Medical University, No.155 NanjingBei Street, Heping District, Shenyang, Liaoning Province, P.R. China 110001. Email: xcz1966@126.com

Dr. YuanYuan, Tumor Etiology and Screening Department of Cancer Institute and General Surgery, the First Hospital of China Medical University, No.155 NanjingBei Street, Heping District, Shenyang, Liaoning Province, P.R. China 110001. Email: yuanyuan@cmu.edu.cn

Abstract

Background:
The purpose of this study was to construct a clinical prognostic model of primary RAC to assist in clinical diagnosis and treatment.

Methods:
Primary RAC patients from 2010 to 2015 were selected from the surveillance, epidemiology and end results (SEER). The relevant significant variables were used to develop the prognosis model. A score was determined for each prognostic factor in the
model. The Kaplan–Meier method and the log-rank test were used to establish and distinguish the survival curves. The accuracy of predictive model was assessed by receiver operating characteristics (ROCs) curve, concordance index (C-index), and decision curve analysis (DCA).

**Results:**

A total of 8069 primary RAC were retrieved in this study. The overall survival (OS) model was established based on 14 variables. The CSS nomogram were constructed using 12 variables. The C-indexes for the training set of OS and CSS were 0.769 (95% confidence interval (CI) 0.761–0.777) and 0.793 (95% CI 0.784–0.802) respectively. The C-indexes for the validation set of OS and CSS were 0.776 (95% CI 0.768–0.784) and 0.794 (95% CI 0.785–0.803) respectively. High-quality calibration plots were observed and the model displayed a favorable outcome compared with (TNM) stage and SEER stage of primary RAC based on DCA curve. We then divided the patients into low-risk, medium-risk, and high-risk groups, showing that patients with primary RAC in the high-risk group had a poor prognosis. A primary RAC prognosis prediction model has been shown by a predicted nomogram for 3 or 5 years and a real-time web-based calculator.

**Conclusions:**

In conclusion, we established a clinical prognosis model for primary RAC for the first time based on a variety of risk factors including individual differences, tumor-related factors, and diagnostic and therapeutic factors, which is more effective than TNM stage in predicting the prognosis of patients. The clinical prediction model
visualized by the nomogram for 3 and 5 years and the web-based real-time calculator
are helpful to optimize clinical work, facilitate patient consultation and clinical
individualized treatment.

**Keywords:**
- Nomogram, rectal adenocarcinoma (RAC), SEER, overall survival (OS),
cancer-specific survival (CSS), web-based calculator, Prognosis

**Background**

Rectal adenocarcinoma (RAC) is a common malignant tumor of digestive tract with
high morbidity and mortality(Zhou et al., 2019a). So far, as we know, the prognosis of
RAC is connected with several factors, such as the depth of bowel-wall
invasion(Harrison et al., 1994), the number of lymph nodes with tumor invasion(Tang
et al., 1995), the infiltration of external intestinal veins(Talbot et al., 1980), mesenteric
fascia(Adam et al., 1994), and peritoneal invasion(Shepherd et al., 1995). There are
also other factors strongly influencing RAC prognosis: For example, individual
patient differences, including age, sex, marital status, and insurance; treatment
measures including surgery, radiation, chemotherapy and specific biological markers;
and other factors associated with the tumor, including tumor size and perineural
invasion(Soto-Ferrari et al., 2017; Liu et al., 2018; Fan et al., 2019). It is significant to
predict the prognosis of the patient with primary RAC by combining all
above-mentioned factors, thus enabling to improve the quality of individualized
survival and the strategy of treatment, according to the risk such as health disparities,
therapy method.
For the moment, it is most common to guide treatment strategy and predict the prognosis of primary RAC with the AJCC TNM staging system, which is for a widely applied tool of tumor assessment (Cong et al., 2018; Fujima et al., 2019). Though the TNM stage is vital and useful, it may be inadequate that it takes into account only some of the factors associated with the tumor itself, such as the depth of tumor invasion, numbers of lymph nodes and whether it has distant metastasis. As is known to us all, when patients are at the same TNM stage, their prognosis usually varies widely (Zhao and Chen, 2015; Gao et al., 2018). The nomogram or a web-based calculator is a multivariable model that considers multiple prognostic factors simultaneously, which is more suitable for the study of multivariable effects in large data sets to accurately predict the prognosis of cancer patients (Ullman et al., 2015; Jorissen et al., 2019; Hu et al., 2020; Jehi et al., 2020; Kim et al., 2020).

Approximately 90% of colorectal cancer is pathologically diagnosed as adenocarcinoma (Evdokimova et al., 2018). However, the nomogram or a web-based calculator for evaluating the prognosis of patients with primary RAC has not been constructed.

In this study, we used the SEER database combining tumor-related variables, patient-specific variational factors, therapeutic measures, and tumor biomarkers to analyze a series of factors that may affect primary RAC prognosis aiming to establish a primary RAC prognosis prediction model shown by a predicted nomogram for 3 or 5 years and a real-time web-based calculator.

Methods
Case collection from SEER database

The patients diagnosed with the primary RAC between 2010 and 2015 were retrieved from the newest SEER database (version 8.3.5; https://seer.cancer.gov/). Since the database is publicly supported for use, a local ethics approval or a statement was not required. Tumor location and histological type code references international classification of oncology 3rd edition (icd-o-3). The inclusion criteria included: (I) Histologic type was RAC; (II) The site of the tumor was the rectosigmoid junction and rectum (C19.9/C20.9); (III) Age > 18 years old; (IV) Information of all the patients was complete and available. The extraction criteria included: (I) Patients were diagnosed with more than one primary cancer; (II) The information needed by the analysis was incomplete. Finally, data of 8069 patients diagnosed with primary RAC were retrieved from the SEER database in this retrospective study. (Figure 1) On the basis, the data were further randomly divided into two groups, the training set containing 3999 patients and the verification set containing 4070 patients.

Variable data extraction

Variables analyzed in this study were extracted including age, sex, race, marital status, insurance, tumor size, Grade, T stage, N stage, M stage, SEER stage, tumor deposits, regional LN examined, perineural invasion, Mets at dx-bone, Mets at dx-brain, Mets at dx-liver, Mets at dx-lung, surgery, chemotherapy, radiation sequence with surgery, tumor biomarker CEA, survival related information, etc. Overall survival (OS) was considered as the primary endpoint and cancer-specific survival (CSS) was considered as the secondary endpoint. The optimal cut-off value of age,
tumor size and regional LN examined was obtained by X-tile software which can quickly find the optimal truncation value when searching for truncation value of survival statistics (Guan et al., 2016) and can transform these three continuous variables into component type variables.

The construction of prognostic model for the OS and CSS of the primary RAC

The Kaplan-Meier method and log-rank test were applied to conduct survival analysis. Univariate Cox analysis was applied to select significant variables (p<0.05), which were further submitted to multivariate COX analysis. Independent prognostic factors filtered by multivariable Cox proportional hazards models were used to construct a prognosis prediction model of primary RAC based on the training set. In this model, a score was defined for each risk factor. Accordingly, a new risk classification system was established in the training set. All the patients in the training set were divided into the low-risk group, the middle-risk group and the high-risk group respectively by the X-tile software, according to the patient total risk score in the model. To display the prognostic model better, we visualize it by generating a nomogram for 3 or 5 years and a real-time web-based calculator.

Validation and calibration of the model

Harrell's concordance index (C-index) was assessed to discriminate difference between training set and validation set. The calibration chart showed the difference visually using a 45-degree line as a selected model. Receiver operating characteristics (ROCs) curve was applied to assess the precision of the 3- and 5-year survival of the nomogram. To verify the accuracy and effectiveness of the new risk classification
system based on this model, we calculated the total score in the validation set and aggregate data for each patient and the risk score cutoff value is derived from the risk score cutoff value of the training set. The new risk classification system was validated in the validation set and aggregate set.

**Practicability analysis of the prognosis prediction model**

Decision curve analysis (DCA) is a new method to evaluate different prognostic strategies. Compared with traditional ROC curve which only pursues accuracy, DCA curve can find a method with the greatest net gain to avoid false positive and false negative simultaneously. To find the best way to assess prognosis strategies for primary RAC, we simultaneously mapped DCA decision analysis curves of our constructed model, TNM stage, SEER stage, and Grade.

**Statistical analyses**

SPSS version 22.0 was used to process the data. R software version 4.0.2 was used to draw pictures. Adobe Illustrator CS6 software was applied to further process the pictures. P<0.05 was used as the criteria for screening single factor and multivariate analysis variables.

**Results**

**Basic characteristics of study cases**

Data from a total of 8069 patients with primary RAC diagnosed through January 2010 to December 2015 were retrospectively gained from the SEER database, of which 3999 were randomly allocated to the training set while 4070 cases were into validation set. The specific screening process of the demographic features and
clinicopathological characteristics was described in Table 1.

**Independent prognostic factors in primary RAC patients**

We used the X-tile software to find the optimal cutoff values for age, tumor size, and number of lymph nodes examined, which were at 66 and 77 years, 33 mm and 61 mm, nine and fourteen, respectively (Figure 2). Univariate and multivariate analyses of overall survival in the training set were described in Table 2. According to the results, fourteen variables including age at diagnosis, sex, marital status, insurance, tumor size, T stage, N stage, CEA, regional LN examined, perineural invasion, Mets at dx-liver, Mets at dx-lung, surgery and radiation sequence with surgery were associated with OS of the primary RAC. Results of univariate and multivariate analyses of CSS in the training set were presented in Table 3. To more vividly depict the relationship between independent prognostic risk factors and survival time, the survival curve of 3999 patients with primary RAC was plotted by the Kaplan–Meier method. According to the results, 12 variables including age at diagnosis, marital status, insurance, tumor size, Grade, T stage, N stage, CEA, regional LN examined, perineural invasion, Mets at dx-liver and Mets at dx-lung were related to CSS of the primary RAC.

**Construction of prognostic model for the OS and CSS of the primary RAC**

According to the consequences of the multivariate Cox regression analysis, the prognosis model for CSS and OS was constructed on the basis of all significant variables. In this model, a score was defined for each risk factor. (Table4) Then, according to the total score of each patient in the training set, the cutoff value was
selected using X-tile software, and a risk classification system was established. The patients were divided into three groups: low-risk group, medium-risk group and high-risk group. (Figure 3) Moreover, the Kaplan–Meier method and the log-rank test were used to establish and distinguish the survival curves of all risk groups in the training set. In the training set, patients in the low-risk group (<124 points) had higher OS than the medium-risk group (124-141 points) and the high-risk group (>141 points). Patients in the low-risk group (<116 points) had higher CSS than the medium-risk group (116-132 points) and the high-risk group (>132 points). There were statistically significant differences in OS and CSS survival curves among the three groups in the training set, (p<0.001), and patients with primary RAC in the high-risk group had a poor prognosis. (Figure 4) When the cutoff scores are the same, the same trends were observed in the validation set cohort and in the entire cohort study, and we can see statistically remarkable differences in OS and CSS survival curves among the three groups in the validation set and entire cohort (p<0.001), suggesting the accuracy of the overall points-scoring system. (Figure 4)

Visualization of a prognostic model of primary RAC

To visually predict patient OS and CSS based on the model at 3 and 5 years. We developed the nomogram for 3- and 5-year OS and CSS shown respectively in Figure 5 based on the prognostic model. A straight line come out easily by summing the scores of all variables. Then the survival probabilities of the individual patients could be determined.

To predict overall survival for each patient at a different point in time, we
developed a web-based real-time calculator for patients with primary RAC, which can be used by clinicians and others easily. The "https://yiyang188.shinyapps.io/3999NomogramOS/" site predicts overall survival at different time points in patients with primary RAC. The "https://yiyang188.shinyapps.io/3999NomogramCSS/" site predicts tumor-specific survival at different time points in patients with primary RAC. We can select the specific variable value of a patient in the variable column, and then click "Predicted Survival at this Follow Up" to select the Predicted time point. Then go to the "Predict" option. In the right display column, click "survival plot" to get the survival curve of the patient. Click "Predict Survival" to see the patient's predicted survival probability and 95% confidence interval at the set time point. Click on the option "Numerical Summary" to see the specific forecast information of each patient. In the "Model Summary" option, we can see the details of how the Model is constructed. (Figure 6)

**Validation and calibration of the model**

In the training set, the C-indexes of OS and CSS were 0.769 (95% CI 0.761–0.777) and 0.793 (95% CI 0.784–0.802) respectively. In the validation set, the C-indexes for the validation set of OS and CSS were 0.776 (95% CI 0.768–0.784) and 0.794 (95% CI 0.785–0.803) respectively. In addition, calibration curves for the nomogram did not have deviated from the reference line, which can better predict credibility. (Figure 7) In order to confirm that the nomogram could predict the prognosis of primary RAC well, the time-dependent ROC curve of the OS and CSS in the training cohort and validation set were shown. In the training set, the AUC values of the 3- and 5-year OS
rates were 0.798 and 0.784, and the AUC values of the 3- and 5-year CSS rates were 0.829 and 0.815. (Figure 8). In the validation set, the AUC values of the 3- and 5-year OS rates were 0.797 and 0.785, and the AUC values of the 3- and 5-year CSS rates were 0.811 and 0.802. (Figure 8)

**The predictive effectiveness analysis of the prognosis models**

To test whether the model has better clinical applicability, the DCA curve was plotted. (Figure 9) From the diagram, we can see that the nomogram has higher net benefits than the TNM staging system, SEER stage and Grade at the same threshold, indicating that this nomogram has better clinical predictive and practical value than TNM stage, SEER stage, and Grade.

**Discussion**

There are many factors affecting the prognosis of cancer. Previous studies have paid too much attention to the autogenous variables from cancer itself, often ignoring the individual differences from cancer patients and external factors applied to patients. In this study, we conducted a set of integration analyses covering tumor-related variables, patient-specific variational factors, therapeutic measures, and tumor biomarkers. Based on these variables, we established a primary RAC prognosis prediction mode, which was shown by nomograms for 3 or 5 years and a web-based real-time calculator.

Firstly, we used SEER database to analyze independent variables that affect primary RAC prognosis including age at diagnosis, marital status, insurance, tumor size, regional nodes examined, surgery, chemotherapy, radiotherapy and CEA.
expression were independent prognostic risk factors for primary RAC, which is consistent with previous studies. Besides, we also had some novel findings. First of all, we found that large tumors (>61 mm) may lead to poor prognosis. Secondly, the number of lymph node samples greater than 14 was related to poor prognosis. This result may lead clinicians to obtain more accurate lymph node sampling compared to 12 lymph node sampling as the cutoff value in the NCCN guidelines (Zhou et al., 2019b). In addition, CEA has been confirmed to be a reasonable biomarker for the prognosis of patients with rectum adenocarcinoma cancer in this study. Until now, the treatment for patients with RAC is still unclear and remains a bone of contention. In our study, local resection and total rectal resection can improve the survival rate to some extent. At the same time, we can also see that partial resection can better improve the prognosis of patients compared with total resection, indicating that it is unnecessary for RAC patients to process a broader range surgery. Although total proctectomy is the most important treatment for rectal cancer, it can also leads to a large number of related complications such as permanent urinary and sexual dysfunctions and so on (Garcia-Aguilar et al., 2000). Finally, chemotherapy and radiation therapy are controversial for patients with colorectal cancer. It has been reported that preoperative chemoradiotherapy can reduce the clinicopathological stage of rectal cancer, which is not only beneficial to surgery, but also can improve the prognosis (Morohashi et al., 2019). Our study found that both radiotherapy and chemotherapy are independent prognostic risk factors for patients with primary RAC. Importantly, we found that postoperative radiotherapy was obviously effective in
improving prognosis. Although preoperative radiotherapy can appropriately improve the tumor-specific survival rate, it can reduce the overall survival rate of patients and increase their economic burden, so it is not recommended. At the same time, we found that radiotherapy before and after surgery increased the risk of death in patients with primary RAC. Maybe the radiation damaged the patient's body (Joel et al., 2018; Marill et al., 2019; Wei et al., 2019).

Based on the independent assessment of prognostic factors, we conducted visual integration analysis by using nomogram that's believed as more effective patterning than traditional predictive models and staging systems, such as TNM staging, in predicting the prognosis of patients (Nam et al., 2017). Our prognostic model includes more prognostic factors than TNM staging system, such as, the patient's individual difference factors, including age, marital status, and insurance, treatment, including surgery, chemotherapy and radiotherapy, tumor markers CEA and some other tumor related factors, including tumor size and perineural invasion, etc. In this model, each predictive variable was defined by a score. The greater the score, the more significant the influence of this variable on the prognosis. We can preliminarily predict the patient's prognosis by determining which prognostic risk group the patient belongs to by calculating the patient's total score. The higher the individual's overall score, the greater the risk of death. In the OS scoring system, <124 is classified as low-risk group, 124-141 as medium-risk group, and >141 as high-risk group. In the CSS scoring system, <116 is classified as low-risk group, 116-132 as medium-risk group, and >132 as high-risk group. In this study, we found that if the patient's overall OS
score was $>141$ or CSS score was $>132$, it should attract the attention of doctors and patients, and reasonable intervention measures should be taken timely to improve the prognosis of the patient. The overall points-scoring system was verified sufficiently in the validation set and entire cohort. We can also directly read the probability of OS or CSS in 3 or 5 years by the nomogram. Furthermore, we developed a web-based real-time calculator for gaining the patient's likelihood of survival at different points in time, which provides a user-friendly graphical interface instead of a computational approximation. It is easy to use, and suitable for the electronic health record system to provide real-time feedback to doctors.

To explore the accuracy of the nomogram model in predicting primary RAC outcomes, this study compared it in parallel with TNM and SEER stage. From the DCA carve, we can see that the predicted model has higher net benefits than the TNM staging system and seer stage at the same threshold, indicating that this nomogram has better clinical predictive and practical value than TNM stage and seer stage. In the OS risk score, we added up the maximum score of all risk factors except T/N stages, and the score was 151. We added up the maximum score of the T and N stages, and the score is 113. Similarly, in the CSS risk score, except for T/N stages, the total score was 151. The value of T stage plus the value of N stage was 116. This indicates that the TNM staging system does play a great role in predicting the prognosis of patients, but we found that the total score of other factors was higher than that of the TNM staging system, although the score of each factor was not large. Therefore, in order to better evaluate the prognosis of patients, our model has obvious advantages over the
traditional TNM staging system. The establishment of this model may be more conducive to the realization of personalized medicine, which has objective guiding significance for both doctors and patients.

There are some limitations in this paper. First of all, because all data we used were obtained from retrospective information in the SEER database, possible selection bias were inevitable. Secondly, the SEER database was lack of certain vital clinicopathological indicators associated with prognosis, such as vascular invasion, surgical margin status and molecular information. Last but not least, though we have used an independent cohort from SEER dataset to verify our nomogram, it is essential to give an external validation. Finally, it is encouraged to enroll more prognostic indicators for more accurate decision-making.

**Conclusion**

In conclusion, for the first time, we established a clinical prognosis model for primary RAC and a web-based real-time calculator based on a variety of risk factors including individual differences, tumor-related factors, and diagnostic and therapeutic factors, which is more effective than TNM stage in predicting the prognosis of patients. Doctors and patients can calculate a total score based on each risk factor score to assess whether a patient is in the low-risk, medium-risk or high-risk group, and can visually see the 3- and 5-year survival probability of the patient by the nomogram for 3- and 5-years. They can use the newly developed web-based calculator to accurately predict the probability of survival at different time points in patients with primary RAC. The clinical prediction model was helpful to optimize
clinical work, facilitate patient consultation and clinical individualized treatment. However, it is necessary to explore more clinical prognostic risk factors to optimize this model and the web-based calculator, and more external validation is needed.

Abbreviations

OS: Overall survival
CSS: Cancer-specific survival
HR: Hazards ratio
CI: Confidence interval
ROC: Receiver-operating characteristic
AUC: Area under the curve
C-index: Concordance index
DCA: Decision Curve Analysis
SEER: Surveillance, Epidemiology, and End Results.

Acknowledgments

We thank the Surveillance, Epidemiology, and End Results (SEER) program for providing access to its research data file for free.

Authors’ contributions

Chenzhong Xing and Yuan Yuan designed and conducted this research. Zhenshuang Wang collected all data, wrote the manuscript, and analyzed them with Yi Yang; Yuan Yuan and Jing Jing revised the manuscript and processed some data.

Funding

The study was supported by the National Key R&D Program of China (Grant #2018YFC1311600).
Availability of data and materials

In our study, all data were selected from the Surveillance, Epidemiology, and End Results (SEER) database at https://seer.cancer.gov/. Relevant data can be accessed through proper request, from the first author.

Ethics approval and consent to participate

We obtained the permission to access the SEER database with the ID number 13594-Nov2019 via Internet access method. Ethical approval was waived by the local ethics committee, as SEER data is publicly available, deidentified.

Consent for publication

Not applicable.

Competing interests

The authors report no conflicts of interest.

Author details

1Tumor Etiology and Screening Department of Cancer Institute and General Surgery, The First Hospital of China Medical University, Shenyang 110001 Liaoning Province, China

2Key Laboratory of Cancer Etiology and Prevention (China Medical University), Liaoning Province Education Department, Shenyang 110001, China

3Key Laboratory of Gastrointestinal Cancer Etiology and Screening, Shenyang 110001 Liaoning Province, China

4Department of Neurosurgery of the First Hospital of China Medical University, Shenyang 110001, China
Adam, I., Mohamdee, M., Martin, I., Scott, N., Finan, P., Johnston, D., et al. (1994). Role of circumferential margin involvement in the local recurrence of rectal cancer. *Lancet (London, England)* 344(8924), 707-711. doi: 10.1016/s0140-6736(94)92206-3.

Cong, L., Liu, Q., Zhang, R., Cui, M., Zhang, X., Gao, X., et al. (2018). Tumor size classification of the 8(th) edition of TNM staging system is superior to that of the 7(th) edition in predicting the survival outcome of pancreatic cancer patients after radical resection and adjuvant chemotherapy. *Sci Rep* 8(1), 10383. doi: 10.1038/s41598-018-28193-4.

Evdokimova, A., Laurent, S., and Dorthu, L. (2018). Recurrent Acute Pancreatitis Caused by Mucin-producing Liver Metastases. *J Belg Soc Radiol* 102(1), 19. doi: 10.5334/jbsr.1440.

Fan, S., Li, T., Zhou, P., Peng, Q., and Zhu, Y. (2019). Development and validation of nomogram combining serum biomarker for predicting survival in patients with resected rectal cancer. *Biosci Rep* 39(11). doi: 10.1042/BSR20192636.

Fujima, N., Shimizu, Y., Yoshida, D., Kano, S., Mizumachi, T., Homma, A., et al. (2019). Machine-Learning-Based Prediction of Treatment Outcomes Using MR Imaging-Derived Quantitative Tumor Information in Patients with Sinonasal Squamous Cell Carcinomas: A Preliminary Study. *Cancers (Basel)* 11(6). doi: 10.3390/cancers11060800.

Gao, W.L., Ye, G.C., Liu, L.W., and Wei, L. (2018). The downregulation of Rap1 GTPase-activating protein is associated with a poor prognosis in colorectal cancer and may impact on tumor progression. *Oncol Lett* 15(5), 7661-7668. doi: 10.3892/ol.2018.8305.

Garcia-Aguilar, J., Mellgren, A., Sirivongs, P., Buie, D., Madoff, R.D., and Rothenberger, D.A. (2000). Local excision of rectal cancer without adjuvant therapy: a word of caution. *Ann Surg* 231(3), 345-351. doi: 10.1097/00000658-200003000-00007.

Guan, X., Chen, W., Jiang, Z., Liu, Z., Miao, D., Hu, H., et al. (2016). Exploration of the Optimal Minimum Lymph Node Count after Colon Cancer Resection for Patients Aged 80 Years and Older. *Scientific reports* 6, 38901. doi: 10.1038/srep38901.

Harrison, J., Dean, P., el-Zeky, F., and Vander Zwaag, R. (1994). From Dukes through Jass: pathological prognostic indicators in rectal cancer. *Human pathology* 25(5), 498-505. doi: 10.1016/0046-8177(94)90122-8.

Hu, P., Bai, J., Liu, M., Xue, J., Chen, T., Li, R., et al. (2020). Trends of incidence and prognosis of gastric neuroendocrine neoplasms: a study based on SEER and our multicenter research. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* 23(4), 591-599. doi: 10.1007/s10120-020-01046-8.

Jehi, L., Ji, X., Milinovich, A., Erzurum, S., Rubin, B., Gordon, S., et al. (2020). Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing: Results from 11,672 Patients. *Chest* doi: 10.1016/j.chest.2020.05.580.

Joel, E., Maxwell, O., Adewoyin, O., Ehi-Eromosele, C., and Saeed, M. (2018). Comparative Analysis of Natural Radioactivity Content in Tiles made in Nigeria and Imported Tiles from China. *Scientific reports* 8(1), 1842. doi: 10.1038/s41598-018-20309-0.

Jorissen, R., Croxford, M., Jones, I., Ward, R., Hawkins, N., Gibbs, P., et al. (2019). Evaluation of the transferability of survival calculators for stage II/III colon cancer across healthcare systems. *International journal of cancer* 145(1), 132-142. doi: 10.1002/ijc.32100.

Kim, B., Kim, S., Kim, Y., Chang, J., Hwang, K., Kim, S., et al. (2020). Development of an
Individualized Prediction Calculator for the Benefit of Postoperative Radiotherapy in Patients with Surgically Resected De Novo Stage IV Breast Cancer. *Cancers* 12(8). doi: 10.3390/cancers12082103.

Liu, Q., Luo, D., Cai, S., Li, Q., and Li, X. (2018). Real-World Implications of Nonbiological Factors with Staging, Prognosis and Clinical Management in Colon Cancer. *Cancers (Basel)* 10(8). doi: 10.3390/cancers10080263.

Marill, J., Mohamed Anesary, N., and Paris, S. (2019). DNA damage enhancement by radiotherapy-activated hafnium oxide nanoparticles improves cGAS-STING pathway activation in human colorectal cancer cells. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 141, 262-266. doi: 10.1016/j.radonc.2019.07.029.

Morohashi, S., Morohashi, H., Seino, H., Yoshizawa, T., Haga, T., Goto, S., et al. (2019). Investigating the association between radiological images and the pathology of rectal cancer treated with neoadjuvant chemotherapy. *Mol Clin Oncol* 11(6), 573-579. doi: 10.3892/mco.2019.1931.

Nam, K., Hwang, D., Shim, J., Song, T., Lee, S., Seo, D., et al. (2017). Novel Preoperative Nomogram for Prediction of Futile Resection in Patients Undergoing Exploration for Potentially Resectable Intrahepatic Cholangiocarcinoma. *Scientific reports* 7, 42954. doi: 10.1038/srep42954.

Shepherd, N., Baxter, K., and Love, S. (1995). Influence of local peritoneal involvement on pelvic recurrence and prognosis in rectal cancer. *Journal of clinical pathology* 48(9), 849-855. doi: 10.1136/jcp.48.9.849.

Soto-Ferrari, M., Prieto, D., and Munene, G. (2017). A Bayesian network and heuristic approach for systematic characterization of radiotherapy receipt after breast-conservation surgery. *BMC Med Inform Decis Mak* 17(1), 93. doi: 10.1186/s12911-017-0479-4.

Talbot, I., Ritchie, S., Leighton, M., Hughes, A., Bussey, H., and Morson, B. (1980). The clinical significance of invasion of veins by rectal cancer. *The British journal of surgery* 67(6), 439-442. doi: 10.1002/bjs.1800670619.

Tang, R., Wang, J., Chen, J., Chang-Chien, C., Tang, S., Lin, S., et al. (1995). Survival impact of lymph node metastasis in TNM stage III carcinoma of the colon and rectum. *Journal of the American College of Surgeons* 180(6), 705-712.

Ullman, H., Spencer-Smith, M., Thompson, D., Doyle, L., Inder, T., Anderson, P., et al. (2015). Neonatal MRI is associated with future cognition and academic achievement in preterm children. *Brain : a journal of neurology* 138, 3251-3262. doi: 10.1093/brain/awv244.

Wei, J., Wang, B., Wang, H., Meng, L., Zhao, Q., Li, X., et al. (2019). Radiation-Induced Normal Tissue Damage: Oxidative Stress and Epigenetic Mechanisms. *Oxidative medicine and cellular longevity* 2019, 3010342. doi: 10.1155/2019/3010342.

Zhao, R., and Chen, G. (2015). Role of GSTP1 Ile105Val and XRCC1 Arg194Trp, Arg280His and Arg399Gln gene polymorphisms in the clinical outcome of advanced non-small cell lung cancer. *Int J Clin Exp Pathol* 8(11), 14909-14916.

Zhou, S., Tian, Y., Wang, X., Zhao, C., Ma, S., Jiang, J., et al. (2019a). Application of indocyanine green-enhanced near-infrared fluorescence-guided imaging in laparoscopic lateral pelvic lymph node dissection for middle-low rectal cancer. *World journal of gastroenterology* 25(31), 4502-4511. doi: 10.3748/wjg.v25.i31.4502.
Zhou, Z., Mo, S., Dai, W., Xiang, W., Han, L., Li, Q., et al. (2019b). Prognostic nomograms for predicting cause-specific survival and overall survival of stage I-III colon cancer patients: a large population-based study. *Cancer cell international* 19, 355. doi: 10.1186/s12935-019-1079-4.

**Figure legends**

**Fig.1** | A flowchart of the selection process for patients with primary rectal adenocarcinoma (RAC)

**Fig.2** | The best cutoff values for age, tumor size, and number of lymph node tests were found by the X-tile software based on prognosis. The age optimal cutoff value tagged by the black circle in the (A) is presented by a histogram of the whole cohort (B), which was 67 and 78, respectively. The tumor size optimal cutoff value tagged by the black circle in the (B) is presented by a histogram of the whole cohort (D), which was 33mm and 61mm, respectively. The number of lymph node optimal cutoff value tagged by the black circle in the (E) is presented by a histogram of the whole cohort (F), which was 9 and 14, respectively.

**Fig.3** | The best cutoff values for total risk score of OS and CSS tests were found by the X-tile software based on prognosis. The total risk score of OS optimal cutoff value tagged by the black circle in the (A) is presented by a histogram of the whole cohort (B), which was 124 and 141, respectively. The total risk score of CSS optimal cutoff value tagged by the black circle in the (C) is presented by a histogram of the whole cohort (D), which was 116 and 132, respectively.

**Fig.4** | Survival analysis curve for total risk score of OS and CSS in the training set, the validation set and the whole queue. Survival analysis curve for total risk score of OS in the training set(A), the validation set(C) and the whole queue(E). Survival analysis curve for total risk score of OS in the training set(B), the validation set(D) and the whole queue(F).

**Fig.5** | The nomograms of the prediction of 3- and 5-year OS and CSS in primary RAC based on the training set.

**Fig.6** | A web-based survival prediction calculator demonstrating the real-time probability of OS of primary RAC, which is available at “https://yiyang188.shinyapps.io/3999NomogramOS/”. A web-based survival prediction calculator demonstrating the real-time probability of CSS of primary RAC, which is available at “https://yiyang188.shinyapps.io/3999NomogramCSS/”.

**Fig.7** | Calibration plots of the predicted nomogram for 3- and 5-year OS prediction in the training set (A) and the validation set (C). Calibration plots of the predicted nomogram for 3- and 5-year OS prediction in the training set (B) and the validation set (D).

**Fig.8** | Discriminatory accuracy for predicting OS evaluated by receiver operator characteristics (ROC) analysis calculating area under the curve (AUC). The 3- and 5-year OS and 3- and 5-year CSS in the training set (A). The 3- and 5-year OS and 3- and 5-year CSS in the validation set (B).

**Fig.9** | Decision curve analysis for the nomograms in the training set and the
validation set. Decision curve analysis of the predicted nomogram for 3- (A) and 5-(C) year OS prediction in the training set. Decision curve analysis of the predicted nomogram for 3- (B) and 5-(D) year CSS prediction in the training set. Decision curve analysis of the predicted nomogram for 3- (E) and 5-(F) year OS prediction in the validation set. Decision curve analysis of the predicted nomogram for 3- (G) and 5-(H) year CSS prediction in the validation set.

Tables

Table 1 | Baseline characteristics of primary RAC patients from SEER database

| Patient characteristics | All patients | Training set | Validation set |
|-------------------------|--------------|--------------|----------------|
|                         | N=8069       | N=3999       | N=4070         |
|                         | No. of patients (%) | No. of patients (%) | No. of patients (%) |
| Age (years)             |              |              |                 |
| 19-66                   | 5398(66.9)   | 2672(66.8)   | 2726(67.0)     |
| 67-78                   | 1823(22.6)   | 895(22.4)    | 928(22.8)      |
| 79-101                  | 848(10.5)    | 432(10.8)    | 416(10.2)      |
| Sex                     |              |              |                 |
| Male                    | 4795(59.4)   | 2397(59.9)   | 2398(58.9)     |
| Female                  | 3274(40.6)   | 1602(40.1)   | 1672(41.1)     |
| Race                    |              |              |                 |
| White                   | 6454(80.0)   | 3199(80.0)   | 3255(80.0)     |
| Black                   | 628(7.8)     | 320(8.0)     | 308(7.6)       |
| Other                   | 987(12.2)    | 480(12.0)    | 507(12.5)      |
| Marital status          |              |              |                 |
| Married                 | 4847(60.1)   | 2397(59.9)   | 2450(60.2)     |
| Single                  | 1418(17.6)   | 707(17.7)    | 711(17.5)      |
| Separated/Divorced     | 1006(12.5)   | 506(12.7)    | 500(12.3)      |
| Widowed                 | 798(9.9)     | 389(9.7)     | 409(10.0)      |
| Insurance               |              |              |                 |
| Insured                 | 6730(83.4)   | 3318(83.0)   | 3412(83.8)     |
| Any Medicaid            | 1053(13.0)   | 536(13.4)    | 517(12.7)      |
| Uninsured              | 286(3.5)     | 145(3.6)     | 141(3.5)       |
| Tumor Size(mm)          |              |              |                 |
| <33                     | 2591(32.1)   | 1266(31.7)   | 1325(32.6)     |
| 33-61                   | 4079(50.6)   | 2018(50.5)   | 2061(50.6)     |
| Grade  | 1399(17.3) | 715(17.9) | 684(16.8) |
|--------|------------|------------|------------|
| Grade I| 516(6.4)   | 243(6.1)   | 273(6.7)   |
| Grade II| 6578(81.5) | 3274(81.9) | 3304(81.2) |
| Grade III| 846(10.5)  | 423(10.6)  | 423(10.4)  |
| Grade IV| 129(1.6)   | 59(1.5)    | 70(1.7)    |

| T stage | 1469(18.2) | 741(18.5) | 728(17.9) |
|---------|------------|------------|------------|
| Tis     | 23(0.3)    | 11(0.3)    | 12(0.3)    |
| T1      | 504(6.2)   | 248(6.2)   | 256(6.3)   |
| T2      | 1469(18.2) | 741(18.5)  | 728(17.9)  |
| T3      | 5240(64.9) | 2581(64.5) | 2659(65.3) |
| T4a     | 419(5.2)   | 213(5.3)   | 206(5.1)   |
| T4b     | 414(5.1)   | 205(5.1)   | 209(5.1)   |

| N stage | 4387(54.4) | 2183(54.6) | 2204(54.2) |
|---------|------------|------------|------------|
| N0      | 2479(30.7) | 1235(30.9) | 1244(30.6) |
| N1      | 8069(14.9) | 581(14.5)  | 622(15.3)  |
| N2      | 7295(90.4) | 3603(90.1) | 3692(90.7) |
| M stage | 774(9.6)   | 396(9.9)   | 378(9.3)   |

| M0      | 4732(58.6) | 2337(58.4) | 2395(58.8) |
|---------|------------|------------|------------|
| M1      | 3337(41.4) | 1662(41.6) | 1675(41.2) |
| SEER stage | 2908(36.0) | 1462(36.6) | 1446(35.5) |
| Localized only | 4284(53.1) | 2090(52.3) | 2194(53.9) |
| Regional | 877(10.9)  | 447(11.2)  | 430(10.6)  |

| CEA     | 7265(90.0) | 3625(90.6) | 3640(89.4) |
|---------|------------|------------|------------|
| Positive| 804(10.0)  | 374(9.4)   | 430(10.6)  |

| Tumor Deposits | 1037(12.9) | 548(13.7) | 489(12.0) |
| Negative      | 2383(29.5) | 1194(29.9) | 1189(29.2) |
| Positive      | 4649(57.6) | 2257(56.4) | 2392(58.8) |

| Regional LN examined | 6991(86.6) | 3449(86.2) | 3542(87.0) |
| <9                  | 1078(13.4) | 550(13.8) | 528(13.0) |
| 9-14                | 8051(99.8) | 3990(99.8) | 4061(99.8) |
| Yes                 | 18(0.2)    | 9(0.2)    | 9(0.2)     |
| Variables          | Univariate analysis | Multivariate analysis |
|--------------------|---------------------|-----------------------|
|                    | P       | HR     | 95%CI     | P  | HR     | 95%CI     |
| Age (years)        | <0.00  | 1.667  | 1.535-1.810 | <0.00 | 1 | Reference | 1.374-1.91 |
|                    | 19-66   |        |            | 67-78 |        |            |
|                    |        |        |            | <0.00 | 1.621  | 3 |
|                    | 79-101  |        |            | <0.00 | 2.789  | 3 |
| Sex                | 0.507  | 0.956  | 0.837-1.092 | 0.1 | 2.789  | 3 |
|                    | Male    |        |            | 0.1 | 2.789  | 3 |

Table 2 | Univariate and multivariate analyses of overall survival in the training set

aIncluding American Indian/Alaskan Native and Asian/Pacific Islander.
| Variable             | p-value | Odds Ratio | 95% CI          | Reference Category |
|----------------------|---------|------------|-----------------|--------------------|
| Female               | 0.649   | 0.978      | 0.888-1.077     |                    |
| Race                 |         |            |                 |                    |
| White                | <0.00   | 1.295      | 1.223-1.372     | 0.003              |
| Black                | 1.295   | 1.223      | 1.188-1.262     | 0.649              |
| Other                | 1.295   | 1.223      | 1.188-1.262     | 0.649              |
| Marital status       |         |            |                 |                    |
| Married              | <0.00   | 1.285      | 1.073-1.53      | 1.073-1.53         |
| Single               | 0.006   | 1.285      | 1.073-1.53      | 0.003              |
| Separated/Divorced   | 0.030   | 1.253      | 1.022-1.53      | 0.030              |
| Widowed              | 0.005   | 1.344      | 1.092-1.65      | 0.005              |
| Insurance            | <0.00   | 1.399      | 1.248-1.567     | 1.073-1.53         |
| Insured              | 0.006   | 1.285      | 1.073-1.53      | 0.003              |
| Any Medicaid         | <0.00   | 1.526      | 1.277-1.82      | 1.073-1.53         |
| Uninsured            | 0.016   | 1.488      | 1.075-2.05      | 1.073-1.53         |
| Tumor Size(mm)       | <0.00   | 1.488      | 1.356-1.633     | 1.073-1.53         |
| <33                  | 0.022   | 1.268      | 1.035-1.55      | 1.073-1.53         |
| 33-61                | 0.772   | 1.025      | 0.865-1.21      | 1.073-1.53         |
| >61                  |         |            |                 |                    |
| Grade                | <0.00   | 1.336      | 1.172-1.523     | 1.073-1.53         |
| Grade I              | 1.336   | 1.172      | 1.073-1.523     | 0.009              |
| Grade II             | 0.551   | 0.911      | 0.672-1.23      | 0.009              |
| Grade III            | 0.198   | 1.252      | 0.889-1.76      | 0.198              |
| Grade IV             | 0.572   | 1.178      | 0.667-2.08      | 0.198              |
| T stage              | <0.00   | 1.715      | 1.591-1.848     | 1.073-1.53         |
|                      | 1.715   | 1.591      | 1.530-1.653     | 0.009              |
|        | Tis | Reference | T1 | 0.721 | 272.005 | 0-6.52E+1 |
|--------|-----|-----------|----|-------|---------|------------|
|        | T2  | 0.711     | 339.186 | 0-8.12E+1 |
|        | T3  | 0.697     | 457.329 | 0-1.09E+1 |
|        | T4a | 0.670     | 809.049 | 0-1.94E+1 |
|        | T4b | 0.678     | 679.801 | 0-1.63E+1 |
| N stage|     |           | <0.00 | 1.583 | 1.458-1.720 | <0.00 |
| N0    |     | Reference |       | 1     |          |           |
| N1    |     |           | 0.010 | 1.292 | 1.064-1.57 | 0       |
| N2    |     |           | <0.00 | 1.828 | 1.470-2.27 | 3       |
| M stage|     |           | <0.00 | 4.139 | 3.550-4.826 | 0.182 |
| M0    |     | Reference |       | 1     |          |           |
| M1    |     |           |       | 1.471 | 0.835-2.59 |         |
| SEER stage |   |           | <0.00 | 2.403 | 2.169-2.662 | 0.028 |
| Localized only | | Reference |       | 1     |          |           |
| Regional |   |           | 0.019 | 1.319 | 1.047-1.66 | 2       |
| Distant |    |           | 0.039 | 1.747 | 1.030-2.96 | 5       |
| CEA    |     |           | <0.00 | 2.123 | 1.863-2.421 | <0.00 |
| Negative |   | Reference |       | 1     |          |           |
| Positive |  |           | 1.478 | 1.285-1.69 | 9       |
| Tumor Deposits |   |           | <0.00 | 2.195 | 1.821-2.646 | 0.233 |
| Negative |   | Reference |       | 1     |          |           |
| Positive |  |           | 1.135 | 0.921-1.39 | 9       |
| Regional LN | |           | <0.00 | 0.836 | 0.767-0.910 | <0.00 |
| <9     |     | Reference |       | 1     |          |           |
| 9-14   |     |           | <0.00 | 0.674 | 0.552-0.82 | 3       |
| >14    |     |           | <0.00 | 0.553 | 0.456-0.67 | 1       |

25
| Variable                   | OR   | 95% CI       | P-value |
|---------------------------|------|--------------|---------|
| Perineural invasion       | <0.00| 2.198        | <0.00   |
| No                        | 1    | 1.878-2.571  |         |
| Yes                       | 1    | <0.00        |         |
| Mets at DX-bone           | <0.00| 9.184        | <0.00   |
| No                        | 1    | 4.571-18.45  |         |
| Yes                       | 1    | 0.061        |         |
| Mets at DX-brain          | <0.00| 10.94        | <0.00   |
| No                        | 1    | 3.512-34.10  |         |
| Yes                       | 1    | 0.223        |         |
| Mets at DX-liver          | <0.00| 4.120        | <0.00   |
| No                        | 1    | 3.478-4.880  |         |
| Yes                       | 1    | 0.025        |         |
| Mets at DX-lung           | <0.00| 4.122        | <0.00   |
| No                        | 1    | 3.184-5.338  |         |
| Yes                       | 1    | 0.021        |         |
| Surgery                   | 0.015| 1.220        | 0.002   |
| No surgery                | 1    | 1.039-1.432  |         |
| Local excision/partial    | 0.001| 0.368        | 0.006   |
| Total proctectomy         | 0.006| 0.424        | 0.001   |
| Chemotherapy              | 0.001| 0.692-0.905  | <0.00   |
| No/unknown                | 1    | 0.411-0.61   |         |
| Yes                       | 1    | 0.503        |         |
| Variables                        | Univariate analysis | Multivariate analysis |
|----------------------------------|---------------------|-----------------------|
|                                  | P value             | HR        | 95%CI       | P value | HR        | 95%CI       |
| Age (years)                      | <0.00               | 1.300     | 1.172-1.443 | <0.00   | 1.517     | 1.251-1.840 |
| 19-66                            | 1                   |           |             |         |           |             |
| 67-78                            | <0.00               | 1.973     | 1.833-2.116 | 1       | 2.338     | 1.811-3.019 |
| 79-101                           | <0.00               | 1.029     | 0.923-1.148 | 1       |           |             |
| Sex                              | 0.730               | 0.973     | 0.833-1.136 |         |           |             |
| Male                             |                     |           |             |         |           |             |
| Female                           |                     |           |             |         |           |             |
| Race                             | 0.606               | 1.029     | 0.923-1.148 |         |           |             |
| White                            |                     |           |             |         |           |             |
| Black                            |                     |           |             |         |           |             |
| Other                            |                     |           |             |         |           |             |

Table 3 | Univariate and multivariate analyses of CSS in the training set
| Marital       | Reference | 1.209 | 1.128-1.295 | 0.004  |
|--------------|-----------|-------|-------------|--------|
| Married      | Single    | 0.001 | 1.404       | 1.148-1.718 |
| Single       | Separated/Divorced | 0.052 | 1.265       | 0.998-1.603 |
| Separated/Divorced | Widowed | 0.058 | 1.294       | 0.992-1.688 |
| Insurance    | Reference | 1.476 | 1.297-1.680 | <0.00  |
| Insured      | Any Medicaid | <0.00 | 1.454       | 1.179-1.792 |
| Any Medicaid | Uninsured | 0.007 | 1.611       | 1.137-2.283 |
| Tumor Size(mm) | Reference | 1.572 | 1.409-1.753 | 0.012  |
| <33          | 0.787     | 0.973 | 0.796-1.18  |
| 33-61        | 0.031     | 1.293 | 1.024-1.633 |
| >61          | 0.035     | 1.590 | 1.032-2.450 |
| Grade        | Reference | 1.561 | 1.350-1.805 | <0.00  |
| Grade I      | Grade II  | 0.893 | 1.028       | 0.690-1.531 |
| Grade II     | Grade III | 0.035 | 1.590       | 1.032-2.450 |
| Grade III    | Grade IV  | 0.199 | 1.545       | 0.796-2.999 |
| Grade IV     | T         | Reference | 2.001 | 1.839-2.178 | <0.00  |
| Tis          | T1        | 0.799 | 242.8588    | 0-6.158E+2 |
|     |   |     |     |     |     |
|-----|---|-----|-----|-----|-----|
| T2  | 0.800 | 238.202 | 0-6.027E+2 0 |
| T3  | 0.777 | 455.992 | 0-1.153E+2 1 |
| T4a | 0.754 | 871.893 | 0-2.207E+2 1 |
| T4b | 0.762 | 693.062 | 0-1.755E+2 1 |
| N   | <0.00 | 1.999 | 1.818-2.199 | <0.00 |
| N0  | 1.428 | 1.140-1.789 |
| N1  | 0.002 | 1.428 | 1.140-1.789 |
| N2  | <0.001 | 2.097 | 1.638-2.684 |
| M   | <0.00 | 5.696 | 4.819-6.732 | 0.438 |
| M0  | 1.274 | 0.691-2.349 |
| M1  |        |     |     |     |
| SEER stage | <0.00 | 3.522 | 3.120-3.977 | 0.013 |
| Localized only |     |     |     |     |
| Regional | 0.010 | 1.475 | 1.098-1.983 |
| Distant  | 0.014 | 2.119 | 1.165-3.855 |
| CEA | <0.00 | 2.331 | 1.997-2.721 | <0.00 |
| Negative |     |     |     |     |
| Positive | 1.435 | 1.216-1.693 |
| Tumor Deposits | <0.00 | 2.803 | 2.293-3.426 | 0.056 |
| Negative |     |     |     |     |
| Positive | 1.242 | 0.995-1.551 |
| Regional nodes examined | <0.00 | 0.818 | 0.740-0.903 | <0.00 |
| 1.000 | 0.740-0.903 | 0.056 |
| 1.242 | 0.995-1.551 | 0.056 |
|                          | <9             | 9-14           | >14           |                     |
|--------------------------|----------------|----------------|---------------|---------------------|
|                          |                |                |               | Referenc e          |
| Perineural invasion      | <0.00          | 1              | 2.700         | 2.268-3.214        |
|                          |                |                | <0.00         | 1                   |
|                          | Yes            |                |               | Referenc e          |
| Mets at DX-bone          | <0.00          | 1              | 12.50         | 6                   |
|                          |                |                | 6.215-25.16   |                     |
|                          | No             |                |               | Referenc e          |
|                          | Yes            |                |               | Referenc e          |
| Mets at DX-brain         | 0.001          | 3              | 10.17         |                     |
|                          |                |                | 2.530-40.90   |                     |
|                          | No             |                |               | Referenc e          |
|                          | Yes            |                |               | Referenc e          |
| Mets at DX-liver         | <0.00          | 1              | 5.564         |                     |
|                          |                |                | 4.640-6.673   |                     |
|                          | No             |                |               | Referenc e          |
|                          | Yes            |                |               | Referenc e          |
| Surgery                  | 0.028          | 1.233          | 1.023-1.487   |                     |
|                          | No surgery     |                |               | Referenc e          |
|                          | Local excision/partial | 0.052 | 0.496 | 0.245-1.007 |
|                          | Total proctectomy | 0.098 | 0.541 | 0.262-1.120 |
| Variable | Risk score of OS | Risk score of CSS |
|----------|-----------------|------------------|
| **Chemotherapy** | 0.539 1.053 0.893-1.242 | No/unknown |
| Yes | | |
| **Radiation sequence with** | 0.031 0.892 0.803-0.989 0.005 | Reference |
| No radiation prior to surgery | 0.074 0.837 0.688-1.018 | |
| after surgery | 0.002 0.669 0.516-0.867 | |
| before and after surgery | 0.240 1.423 0.790-2.565 | |

Table 4 | Prognostic risk score from nomograms

| Variable | Risk score of OS | Risk score of CSS |
|----------|-----------------|------------------|
| **Age(years)** | | |
| <67 | 0 | 0 |
| 67-78 | 10 | 9 |
| >78 | 20 | 18 |
| **Marital** | | |
| Married | 0 | 0 |
| Single | 5 | 7 |
| Separated/Divorced | 5 | 5 |
| Widowed | 6 | 5 |
| **Tumor Size(mm)** | | |
| <33 | 0 | 0 |
| 33-61 | 0 | 0 |
| >61 | 5 | 6 |
| **Insurance** | | |
| Insured/No specifics | 0 | 0 |
| Any Medicaid | 8 | 8 |
| Uninsured | 8 | 10 |
| **T stage** | | |
| Tis | 0 | 0 |
| T1 | 77 | 73 |
| T2 | 82 | 72 |
| T3 | 88 | 86 |
|              | Count 1 | Count 2 |
|--------------|---------|---------|
| T4a          | 100     | 100     |
| T4b          | 95      | 94      |
| **N stage**  |         |         |
| N0           | 0       | 0       |
| N1           | 6       | 8       |
| N2           | 13      | 16      |
| **SEER stage** |       |         |
| Localized only | 0     | 0       |
| Regional     | 5       | 8       |
| Distant      | 16      | 19      |
| **CEA**      |         |         |
| Negative     | 0       | 0       |
| Positive     | 8       | 8       |
| **Regional LN examined** |   |         |
| <9           | 12      | 17      |
| 9-14         | 4       | 7       |
| >14          | 0       | 0       |
| **Perineural invasion** |   |         |
| No           | 0       | 0       |
| Yes          | 8       | 8       |
| **Mets at DX-bone** |       |         |
| No           |         | 0       |
| Yes          |         | 20      |
| **Mets at DX-liver** |       |         |
| No           | 0       | 0       |
| Yes          | 11      | 11      |
| **Mets at DX-lung** |       |         |
| No           | 0       | 0       |
| Yes          | 9       | 12      |
| **Surgery**  |         |         |
| No surgery   | 22      |         |
| Local excision/partial proctectomy | 0 |         |
| Total proctectomy | 3  |         |
| **Chemotherapy** |     |         |
| No/unknown   | 14      |         |
| Yes          | 0       |         |
| **Radiation sequence with surgery** |   |         |
| No radiation | 3       | 8       |
| prior to surgery | 6  | 5       |
| after surgery | 0     | 0       |
before and after surgery | 12 | 15 |

3-year survival probability

| 0.95 | 101 | 107 |
| 0.9  | 115 | 122 |
| 0.6  | 147 | 154 |
| 0.3  | 164 | 172 |
| 0.1  | 177 | 185 |

5-year survival probability

| 0.95 | 87  | 94  |
| 0.9  | 102 | 108 |
| 0.6  | 133 | 141 |
| 0.3  | 150 | 158 |
| 0.1  | 163 | 172 |

Abbreviations: OS, overall survival; CSS, cancer-specific survival.