Peer Revie File

Article Information: https://dx.doi.org/10.21037/jgo-22-758

Response to Reviewer A:

Comment 1: The study does not include any validation of the results. The best solution would be to add a new cohort. If not possible, can the performance of the nomogram be tested using bootstrap validation or cross-validation?

Reply 1: We understand your concerns. The data are based on a single cancer center, and the lack of external validation is one of the limitations in our study. Therefore, to evaluate the predictive accuracy of the nomograms, a calibration curve with bootstrapping, which measured the agreement between the nomogram-predicted outcome probability and the average actual probability, was plotted. In Fig 3b, the solid line is bias-corrected by bootstrapping (1000 repetitions), indicating observed nomogram performance. In the future, patients will continue to be collected and followed up to validate this prognostic model.

We have modified the description of the calibration curve (see Page 9, line 192-193).

Changes in the text (Modified sections are marked in red):

To evaluate the performance of the prediction model, a calibration curve with bootstrapping was constructed by measuring the agreement between the predicted and actual 3-year RFS.

Comment 2: The lack of external validation should be reflected in the conclusion. Though mentioned as a limitation, the need for validation in an external cohort should also be clearly stated in the study and abstract conclusion. The authors should also avoid recommendations of DR or the nomogram before being externally validated.

Reply 2: We thank reviewer comments. We have added this part in the conclusion (see Page 4, line 65-68).

Changes in the text (Modified sections are marked in red):

The novel nomogram combining these factors may be used as a reliable and effective tool for the prediction of RFS in stage II CRC, thus helping optimize therapeutic regimens under cooperation of oncologists and surgeons, further multicentric studies are required for validation of this novel, simple and cost-effective prognostic model.

Comment 3: Did any patients receive adjuvant therapy? If patients received adjuvant therapy, this should be considered as a potential confounder?

Reply 3: Thanks for the suggestion. We have added the history of postoperative adjunctive chemotherapy as a clinicopathological factors (see Page 8,11 line 157, 222-225; Table 1-5).

Changes in the text (Modified sections are marked in red):

Page 8 line 157: In total, 207 patients with stage II CRC were included. Data on clinicopathological factors including age, gender, hypertension history, diabetes history, smoking history, family history of cancer, preoperative CEA, CA724, CA199, and CA242 levels, histological type, differentiation grade, T stage, presence or absence of lymphovascular invasion (LVI), perineural invasion (PNI), internal obstruction or perforation

(IOP) status, number of lymph nodes examined (⩾12 vs. <12), mismatch repair status, Ki-67 expression level and history of postoperative adjunctive chemotherapy (present vs. absent) were obtained from medical records and pathology reports.
Page 11 line 222-225: DR was significantly associated with CEA level (P = 0.035), PNI (P = 0.024), IOP status (P = 0.030), TBd (P < 0.001), TSR (P < 0.001) and adjuvant therapy (P = 0.015). TBd was significantly associated with histological grades (P = 0.046), hypertension (P = 0.004), DR (P < 0.001), TSR (P < 0.001) and adjuvant therapy (P < 0.001);

Comment 4: There seems to be an error in the reported CI95 and p-value from multivariate cox regression in both abstract and main text. CI95% includes 1, but both have p-value below 0.05. DR (HR 1.378; 95% CI 0.890-2.134; P =0.048), LVI (HR 49 1.978; 95% CI 0.056-3.708; P =0.033))". Are the values correct?

Reply 4: Sorry for this mistake. With the addition of the history of postoperative adjunctive chemotherapy, we recalculate all the data and double-checked the results (see Page 3, line 58-60; Page11- 12, line 237-244; 247-252; Table 5).

Changes in the text (Modified sections are marked in red):

Page 3, line 58-60: In multivariable analysis, DR (HR 2.111; 95% CI 1.184-3.766; P = 0.011), LVI (HR 1.919; 95% CI 1.004-3.669; P = 0.049) and PNI (HR 2.724; 95% CI 1.362-5.448; P = 0.005) were prognostic factors to RFS.

Page11- 12, line 237-244: Patients with higher TB (BD3) and low TSR (stroma-high tumors) tended to have a worse RFS; however, the difference was not statistically significant (P = 0.13 and 0.15, respectively). In addition, it is worth noting that the findings from Kaplan–Meier curves showed similar survival curves for the DR-intermediate and DR-mature groups, but only the DR-immature group had a significantly worse prognosis compared to the other two groups. According to this, in the univariate and multivariate analyses that followed, DR was categorized into 2 groups: immature DR and other DR types, as previously reported (15, 26).

Page 12, line 247-252: The univariate and multivariate Cox proportional hazards regression analyses for RFS in CRC are detailed in Table 5. In the univariate analyses of RFS, pT stage (P =0.040), LVI (P < 0.001), PNI (P < 0.001), IOP status (P = 0.001), number of nodes examined (P < 0.001), mismatch repair status (P = 0.019), CEA (P = 0.001), CA242 (P = 0.024), DR (P < 0.001) and adjuvant chemotherapy (P = 0.001) were associated with RFS. When subjected to multivariate analysis, DR remained prognostic for RFS (HR 2.111; 95% CI 1.184-3.766; P = 0.011). LVI (HR 1.919; 95% CI 1.004-3.669; P = 0.049) and PNI (HR 2.724; 95% CI 1.362-5.448; P = 0.005) were the only other parameters significantly associated with RFS.

Comment 5: It is not directly assessed whether location of primary tumor (colon or rectum) influence the prognostic use of the markers. Perhaps location should be added in the multivariate model even though it is not significant in the univariate model?

Reply 5: Thanks for the suggestion. We divided all patients into colon group and rectum group according to their location and performed univariate and multivariate analyses. The result of the univariate analyses showed that the location of primary tumor (colon or rectum) was not associated with RFS (HR 0.598; 95% CI 0.342-1.045; P = 0.071). In this study, we only attempted to propose a way to assist clinicians in patient stratification preliminarily and may not include all meaningful factors. Future multicenter research will validate our prediction model and may further supplement the prognostic factors of this study.

Changes in the text (Modified sections are marked in red): No changes in the manuscript.

Methods

Comment 1: Please define RFS in the statistics section.

Reply 1: Thanks for the suggestion. We added this part in the manuscript (see Page 8, line 157-161).

Changes in the text (Modified sections are marked in red):

All patients were followed-up for at least 3 years from the date of surgical resection. The RFS as the primary endpoint was defined as the date from surgery to confirmed clinical recurrence, death or last available contact.

Comment 2: Nomogram: It is not clear how exactly the nomogram is generated. Please elaborate on how you developed the nomogram and how coefficients from the multivariate cox regression are incorporated into the model.

Reply 2: Thanks for the suggestion. We added this part in the manuscript (see Page 10, line 203-206).

Changes in the text (Modified sections are marked in red):

R packages of "rms," "survival," "calibrate," "pROC," and "OptimalCutpoints" were used to construct the prognostic nomogram, plot calibration curves, ROC curves and cutoff value. The optimal cutoff point was estimated by Youden's J index (25).

Comment 3: ROC and AUC – please state what groups the model is used to discriminate between (patients with recurrence before 3 years vs. no recurrence within 3 years?). How do you select the optimal cut-off point? Please avoid calling it "the diagnostic performance" of the model when you are testing the prognostic performance.

Reply 3: We understand the reviewers concern. Yes, discrimination is the ability to differentiate between patients with recurrence before 3 years and patients without recurrence within 3 years, and it is measured by AUC. The optimal cutoff point was estimated by Youden's J index (see Page 10, line 203-206). We changed "diagnostic" to "prediction" (see Page 10, line 195).

Changes in the text (Modified sections are marked in red): The optimal cutoff point was estimated by Youden's J index (24).

Furthermore, a receiver operating characteristic (ROC) curve was produced to further verify the prediction performance of the model by calculating the area under the curve (AUC) as

well as accuracy, sensitivity and specificity.

Results:

Comment 1: Please specify the number of patients having recurrence within 3 years.

Reply 1: Thanks for the suggestion. We modified this part in the manuscript (see Page 10, line 213-215).

Changes in the text (Modified sections are marked in red):

50 (24.15%) patients of the entire population had disease recurrence, and all of them experienced recurrence within 3 years.

Comment 2: When presenting results from cox regression, please specify, at least in table 5 and preferably also in the text, what levels are compared (categorical variable) or the scale that is used (continuous variables).

Reply 2: Thanks for the suggestion. We modified this part in the Table 5 (see table 5).

Changes in the text (Modified sections are marked in red): Please see table 5.

Comment 3: Please provide the optimal cut-off point from where you derive sensitivity and specificity of the model using ROC analyses

Reply 3: Thanks for the suggestion. We added this part in the manuscript (see Page 12-13, line 258-263).

Changes in the text (Modified sections are marked in red):

Moreover, an ROC curve was plotted for the evaluation of the model performance (Fig. 4). An optimal cut-off value that maximized the sum of sensitivity and specificity in the ROC curve was identified. The best cut-off point was 0.235. The prediction model yielded an AUC of 0.826 (95% CI 0.758-0.895; accuracy 0.749; sensitivity 0.780; specificity 0.739) in our cohort.

Comment 4: Figure 3: Please describe how the nomogram is interpreted in the figure captions.

Reply 3: Thanks for the suggestion. We added this description in the figure captions (see Page 25, line 532-536).

Changes in the text (Modified sections are marked in red):

Instructions for use of the nomogram: First, assign the points of each characteristic of the patient by drawing a vertical line from that variable to the points scale. Then, sum all the points and draw a vertical line from the total points scale to disease recurrence axis to obtain the probability.

Discussion

Comment 1: Line 308-309: Authors write: “Notably, DR was the only independent factor for

RFS". Should state that PNI and LVI were also significant in the multivariate model.

Reply 1: Thanks for the suggestion. We modified this part in the manuscript (see Page 15, line 319).

Changes in the text (Modified sections are marked in red):

Notably, compared to TSR and TBd, DR was the only independent factor for RFS, as demonstrated by multivariate analysis in the current study.

Comment 2: Line 320: "we constructed a diagnostic nomogram" – the nomogram is prognostic – please rephrase.

Reply 2: We understand the concern and thanks for the suggestion. We have revised this sentence in the discussion section to follow your suggestions (see Page 16, line 330).

Changes in the text (Modified sections are marked in red):

In the present study, we constructed a predictive nomogram that can generate an individual probability of having disease recurrence by integrating both the significant preoperative clinicopathologic variables and DR, which is easily evaluated using hematoxylin-eosin-stained specimens.

Comment 3: The author concludes "such a model with DR had the potential in predicting patients who are more likely to benefit from postoperative adjuvant chemotherapy". This is purely hypothetical and should be addressed as such. The study does not include any information about adjuvant therapy, and whether the model can predict this is not assessed.

Reply 3: We understand the concern and thanks for the suggestion. We modified this part in the discussion section (see Page 16-17, line 347-352).

Changes in the text (Modified sections are marked in red):

Furthermore, the significant correlation of DR with several conventional high-risk clinicopathological factors suggests that such a model with DR may have the potential to substratify the high-risk patients and predict patients who are more likely to benefit from postoperative adjuvant chemotherapy, although future multicentric and prospective studies are needed before widespread implementation in clinical practice.

Minor comments:

Background

Comment 1: A short description of the desmoplastic reaction in the introduction would make it easier for readers unfamiliar with the concept to understand the manuscript.

Reply 1: Thanks for the suggestion. We add the description of the desmoplastic reaction in the introduction and references section (see Page 6, line 122-123).

Changes in the text (Modified sections are marked in red):

Desmoplastic reaction (DR) is the proliferation of myofibroblasts in the stroma of invasive

cancer (13,14).

13.Angeli F, Koumakis G, Chen MC, et al. Role of stromal fibroblasts in cancer: Promoting or impeding? Tumour Biol 2009;30,109-120.

14.Ohno K, Fujimori T, Okamoto Y, et al. Diagnosis of desmoplastic reaction by immunohistochemical analysis, in biopsy specimens of early colorectal carcinomas, is efficacious in estimating the depth of invasion. Int J Mol Sci 2013;14,13129-13136.

Comment 2: Figure 2: Please state time unit and definition of DR 1 – 3 and TSR 0 and 1, in captions or figure legend.

Reply 2: Thanks for the suggestion. We add this part in the figure legend (see Page 25, line 527-530).

Changes in the text (Modified sections are marked in red):

FIGURE 2: Kaplan-Meier survival curves for RFS (Time unit: days). (a) DR; (b) TBd; (c) TSR. DR1, DR-mature; DR2, DR-intermediate; DR3, DR-immature; TBd1, BD1; TBd2, BD2; TBd3, BD3; TSR0, high TSR (stroma-low); TSR1, low TSR (stroma-high).

Comment 3: Tables: Please describe all abbreviations.

Reply 2: Thanks for the suggestion. We checked and added this part in the figure legend (see Table 1-5).

Response to Reviewer B

Fan and colleagues have submitted a manuscript evaluating the prognostic significance of desmoplastic stromal reaction, tumor budding, and tumor‐stroma ratio in a cohort of 200 Stage II CRC patients. The manuscript is well-written and the study question is valuable. The prognostic benefit of histologic markers can assist in treatment decisions and counseling patients. The only concern of the manuscript is the generalizability given the small cohort and may not have broad application or change clinical practice. Acceptance is deferred to the Editorial team to determine if this is appropriate for this publication.

Reply: We thank reviewer comments. The data are based on a single cancer center, and the lack of external validation is one of the limitations in our study. Our study highlighted the prognostic value of DR in stage II CRC, meanwhile, we attempted to propose a way to assist clinicians in patient stratification preliminarily. Future multicentric and prospective studies are needed before widespread implementation in clinical practice.