Peer Review File
Article information: https://dx.doi.org/10.21037/jgo-23-427

Reviewer A

1) First, the title is unclear and inaccurate since the identification of clinical features and prognostic factors cannot be equal to the development of the nomogram. I suggest the authors to directly indicate the focus of this this study since the identification of potential predictors is only a step towards to the development of the nomogram.

Response: We have modified our title as advised. (see Page1, line2-3)

2) Second, the abstract is problematic and needs further revisions. The background did not indicate the limitations and knowledge gaps of precious studies and why there is a clinical need for the novel nomogram. The methods need to describe the inclusion of subjects, the generation of training and validation samples, the assessment of potential predictors, follow up, and measurements of prognosis outcomes. The indicators for the predictive accuracy of the nomogram should be described. The results need to describe the basic clinical characteristics of the study sample, the predictors identified in the predictive model, and the sensitivity and specificity of the nomogram. The AUC of TNM system needs to be reported. The conclusion needs to have more detailed comments for the clinical implications of the findings.

Response: We have modified our abstract as advised. (see Page2, line5-30)

3) Third, the introduction of the main text is poor. Since the focus is the "novel" nomogram, the authors need to review "old" studies on the prognostic factors and prognosis prediction models including known predictors, analyze their limitations and knowledge gaps including predictive accuracy, and indicate the clinical needs for the new predictive models. Please clearly indicate the novel aspects of the current predictive model.

Response: We have modified our introduction as advised (see page3, line17-26)

4) Fourth, the methodology of the main text is not adequate. The authors need to describe the research design, inclusion criteria of eligible subjects, the generation of training and validation samples, baseline factors in the SEER dataset, follow up procedures, and measurements of prognosis outcomes. The authors need to explain why the outcomes were limited to 2- and 5-year CSS only. In statistics, the authors need to provide the threshold values of AUC for a good nomogram. I also suggest the authors to calculate the sensitivity and specificity of the nomogram, since AUC is only an indicator of overall predictive accuracy, which cannot ensure a good sensitivity or specificity.

Response: We have described the research design, eligible subjects the generation of training and validation samples, baseline factors and prognosis outcomes in the Methods part (see page4, line4-21) .

For choosing the 2- and 5-year CSS as the outcomes, 5-year survival is a commonly and generally

accepted approach for assessing the prognosis of cancer and effectiveness of treatment strategies, including HCC. Because in this study we predicted the prognostic probability incorporating the therapies patients received, treatments like systemic chemotherapy and radiotherapy need months or more than one year to be operated and evaluated survival benefit, we applied 2-year CSS as another indicator for short-term prognostic assessment.

The goal of an ROC analysis is to assess the discrimination of the prognostic model. We focused here on using Cox model methods to generate a model score to evaluate the rate of survival, here sensitivity measured the expected fraction of subjects with a rate more than cut-off among subpopulation of individuals who survive beyond time *t*, while specificity measures the fraction of subjects with a rate less than cut-off among those who die at time *t*. When no priori cut-off value is defined the full spectrum of sensitivities and specificities can be characterized using ROC curve and AUC value that plots sensitivity versus (1-specifity) for all cut-offs(1). The greater the area under the ROC curve, the better the prediction model. A generally accepted approach suggests that AUC less than 0.60 reflects poor discrimination; 0.60-0.75, possibly helpful discrimination; and more than 0.75, clearly useful discrimination(2).

Reviewer B

(1) The novel nomogram for HCC was constructed in the study. The "novel" included what? Please state in the introduction.

Response: We have modified our introduction as advised (see page3, line17-26).

(2) It was better to add related reference (DOI: 10.21037/atm.2019.09.01) about HCC in SEER database in the introduction.

Response: This is a good work (DOI: 10.21037/atm.2019.09.01) about non-small cell lung cancer based on SEER database, we focused in this study about the prognosis of HCC, which was not very relevant.

(3) The nomogram was the crucial topic in the study. How about the progress of independent prognostic factors of HCC? Please supplement in the introduction.

Response: We have modified our introduction as advised (see page3, line17-26)

(4) How to determine the independent prognostic factors of HCC in the study? Please state in the methods.

Response: We have modified our methods as advised (see page4, line29-33)

(5) The figure 1 was not clear enough. Please replace it with a new.

Response: The figure 1 has been replaced by a new one and also provided in a separate jpg file.

(6) It was more convincing to validate the constructed nomogram by the data in your hospital.

Response: We agree with your opinion about external validation of the nomogram, unfortunately, due to the limited time and funding, we could not validate the nomogram by our data from hospital

right now. Therefore, we performed the internal validation by split-sample validation and bootstrap approach to account for model overfitting or uncertainty, which diminished when the sample size was large. Furthermore, we are collecting the data in our hospital right now and it will take a long time to complete this work and we will publish the results in the future.

(7) The targeting therapy and immunotherapy were more applied for HCC. Whether the targeting therapy and immunotherapy could be an independent prognostic factor of HCC? Please state in the discussion.

Response: TKIs and ICIs have significantly improved the survival of patients from clinical trial data since 2017, now they are routinely considered for patients with advanced hepatocellular carcinoma. Due to the limited time of widely application in clinical practice, we could not assess the prognostic ability based on SEER data. However, we will collect information about TKIs and ICIs in our hospital and evaluate their prognostic strength in the future study.

(8) Compared to other constructed nomogram for HCC, what were the advantages of new constructed nomogram in the study? Please state in the discussion.

Response: In the discussion part about the strengths and limitations. We have already talked about that our study focused on the application of the model in the general population of all HCC stages, rather than the subpopulation previous studies interested. Additionally, compared with nomograms constructed before, we incorporated the AFP level and bone metastasis to the nomogram based on the evidences of prior researches and our preselection analysis from SEER data. (see page9, line4-8, 14-17)

References

1. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. Biometrics 2005;61:92-105.

2. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. Jama 2017;318:1377-84.