# Peer Review File

Reviewer A

- To understand the actual need for the algorithms used by the authors, it is necessary to make a direct comparison with prognostic systems currently used in clinical practice, such as UISS, Leibovich and GRANT (e.g. comparing c-index of the new algorithms with the "old" models).
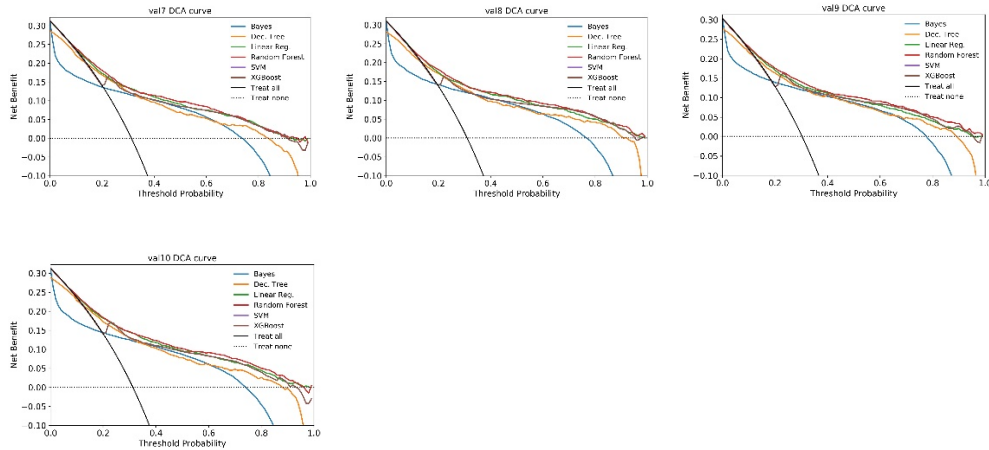
A: We are appreciated for your kind comment and indeed this is a really good problem worth to be explored. This manuscript was one of our whole project that included comparison with UISS as well as GRANT, and we are also collecting our own RCC patients and compared the difference between these prognostic systems. Thus, these results were next stage of our research and will be published in our following papers. Still, we do appreciate for these constructive comments that could really improve our research.

- C-index and Decision Curve Analysis (DCA), which are fundamental to assess the accuracy of prognostic models, are missing in the results.

A: Thanks for your question. The survival prediction in this paper is a two-classification task. The following reference shows that the C-index indicator could evaluate two- or multi-classification models and is equivalent to the area-under-curve (AUC) value of receiver-operating-characteristic-curve (ROC) only for two-classification task. That's to say, C-index is an extension of AUC and AUC is a special case of C-index. Besides, the DCA figures of all ML models are supplemented as Figure ***.

1. Harrell Jr F E, Lee K L, Mark D B. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors[J]. Statistics in medicine, 1996, 15(4): 361-387.
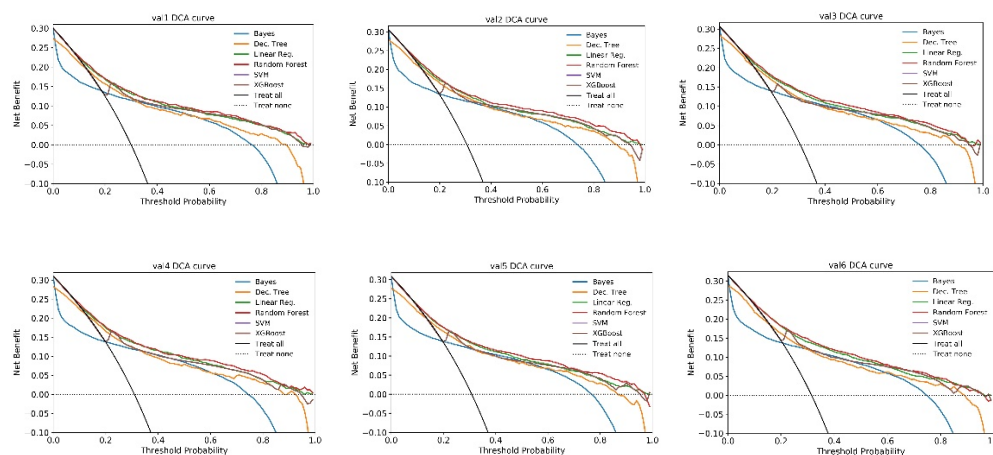
Minor comment:
- From line 250 to 271: this part should be in Methods, not in results

A: We are appreciated for your kind comment and have adjusted these parts in our revised manuscript.

First, the title needs to indicate the development and validation of a prognosis prediction model by using ML.

A: thank you for your kind comment and we have adjusted our title.

Second, the abstract is not adequate and needs further revisions. The background did not describe the limitations of prior prediction models using non-ML algorithms and why ML-based models are better. The methods need to describe the data collection of potential predictors, the measurement of prognosis outcomes, the generation of training and validation samples, and statistical indicators for assessing the predictive accuracy. The results need to summarize the baseline clinical characteristics and survival rates of the patient sample, as well as the predictive accuracy indicators in the validation samples. The conclusion needs more details comments for the clinical implications of the findings.

A: we are really appreciated for the kind comment and have revised our abstract. It really means a lot to improve the quality of our manuscript.

Third, in the introduction, the sentence "One of the most interesting and challenging aspects is accurate survival time prediction" is misleading since the focus is to predict OS, not survival time. The authors need to review all available prediction models in

RCC including predictors, algorithms, and predictive accuracy, have comments on their limitations, and analyze the strengths of ML, in particular why ML can more accurately predict prognosis.

A: Thank you for the valuable comment and we have added relevant information to the introduction.

Fourth, in the methodology of the main text, the authors need to describe the clinical research design, clinical sample and clinical variables and prognosis outcomes in the SEER dataset, and the generation of training and validation samples. My major concern for this part is the authors deleted cases with missing data, however, it would be more helpful and practical to develop the model with missing data by using ML. Such models would be more generalizable. In statistics, please report the threshold values of these accuracy indicators for a good predictive model. In discussion, the authors need to compare the accuracy of ML-based models with prior models to support the superiority of ML models.

A: We are appreciated for the kind and constructive comment and indeed this is a really nice perspective to improve our manuscript. As we explained to the another reviewer, this manuscript was one of our whole project that included comparison with other prognostic models, and we are also collecting our own RCC patients and compared the difference between these prognostic systems. Thus, these results were next stage of our research and will be published in our following papers. Still, we do appreciate for such constructive comments that could really improve our research.

Finally, some interesting related papers need to be reviewed and cited in this paper:
1. Wang C, Wu B, Shen D, Zhang B, Wang L, Xiao J. Establishment and validation of a prognostic nomogram for patients with renal cell carcinoma based on SEER and TCGA database. Transl Cancer Res 2023;12(6):1411-1421. doi: 10.21037/tcr-22-2692.
2. Yang T, Wu Y, Zuo Y, Fu S, Xu Z, Yu N. Development and validation of prognostic nomograms and a web-based survival rate calculator for sarcomatoid renal cell carcinoma in pre- and post-treatment patients. Transl Androl Urol 2021;10(2):754-764. doi: 10.21037/tau-20-1192.
3. Wang F, Gao SG, Xue Q, Tan FW, Gao YS, Wang DL, Mao YS, Zhao J, Li Y, Yu XY, Cheng H, Zhao CG, Yang D, Mu JW. Nomogram for predicting the overall survival of the patients with oesophageal signet ring cell carcinoma. J Thorac Dis 2021;13(3):1315-1326. doi: 10.21037/jtd-20-3084.

A: Thank you for the interesting papers and we have cited these papers to our manuscript.