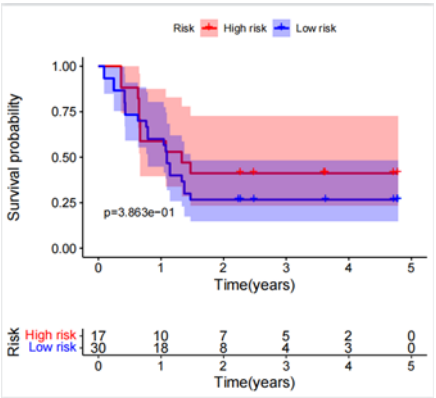
Peer Review File

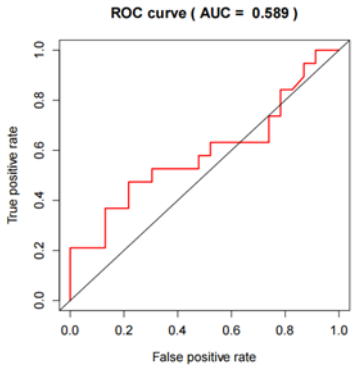
Article information: https://dx.doi.org/10.21037/tcr-23-128

Reviewer A

Wang et al. constructed a prognostic risk model of 15 genes based on the data of HCC from the TCGA database, which has certain predictive capacities. However, some issues should be concerned.

1. The authors used transcriptome data of tissue samples from 6 normal people and 160 patients with HCC for differential gene expression analysis. The number of normal samples was too small.
   In this study, we selected the Asian HCC population in TCGA database as the target population, and we screened the genetic sequencing and clinical information of 161 Asian HCC patients, the transcriptome sequencing information of these 161 patients included 6 normal samples and 160 tumor samples, that is, 6 patients had sequencing information of both normal and tumor tissues. The genes with p less than 0.05 were selected as statistically significant genes, and the subsequent series of analyses were also statistically validated.
2. Although the authors showed validation such as the K-M survival curve, the prognostic model constructed needs to be validated in an internal dataset or a separate dataset.
   After we constructed the model, we performed internal validation, and the K-M survival curve showed that the survival rate in the low-risk group was higher than that in the high-risk group, and the p-value was less than 0.05, which was statistically significant. In addition, we also plotted the ROC curve, and the AUC value was 0.901, which indicated that the model was more accurate in predicting prognosis.
   We tried to select external datasets for further validation, and we screened two sets of Asian HCC patients through GEO database, one set was GSE45114, but the survival curve was not ideal because the sample size was too small and the follow-up time was too short (see below), but by plotting the ROC curve (see below), we found that the AUC value was greater than 0.5, which still had some clinical significance. The other dataset, GSE27150, did not detect the expression of the target gene ACVR2B, so we could not further validate it.





3. The authors mentioned that the protein encoded by MICB is associated with NK cell activation, and why correlation analysis with NK cells was missing in the analysis of this model with immune cells?
   We obtained the immune cell content through the timer website and included only the six immune cells we analyzed in each sample, so we did not perform a correlation analysis for NK cells.
4. Clinical features showed that there was only 1 patient in T1 or N1, I question the reliability of the analysis results of COX univariate and multivariate regression in Figure 10
   The study screened a population with staging of T1 in 82 patients and not 1, while N staging had a p-value greater than 0.05 in both univariate and multivariate regression analyses, which were analyzed and learned to be not statistically significant.

Reviewer B

The paper titled “Construction and validation of a prognostic risk score model based on immune genes for Hepatocellular carcinoma in an Asian population” is interesting. The prognostic risk score model constructed by immune genes based on the Asian HCC population has certain predictive capacity. However, there are several minor issues that if addressed would significantly improve the manuscript.

1) The tumor heterogeneous make-up of immune cell infiltrates is a key factor for the therapy response and prognosis of HCC. How to comprehensively understand the tumor immune microenvironment at the gene and cell levels? Please answer according to the content of this study.
   We give the results of the correlation analysis between risk score and immune cells in lines 206 to 209 of the manuscript, which shows that the risk value is positively correlated with each immune cell content, indicating that the higher the immune cell content in the tumor microenvironment, the higher the risk value and the worse the prognosis.

2) What is the greatest advantage of the prognostic risk score model in this study? What is the biggest problem we are facing? It is suggested to add relevant content to the discussion.

We have added the corresponding discussion in lines 314-320 of the manuscript.

3) The description of some figure legends in this study is too simplistic, please describe in detail.

Some of the images that have been described in detail in similar images earlier have not been further elaborated to avoid unnecessary repetition. If you still have reservations, could you please provide details on which images need further description and we will be happy to revise them.

4) It may be more meaningful to add functional research on key genes.

Our main objective in conducting this study was to construct a value-at-risk score prognostic model and validate its accuracy, and in the next phase of our future work, we may conduct further basic research.

5) Figures 3-4, 9 and 11 are not clear enough. It is recommended to provide clearer figures again.

I have remade these images（lines528,566,613,640）.

6) This study is based on bioinformatics analysis. It is recommended to increase in vivo and in vitro experimental studies, which may be more meaningful.

Adding in vitro experiments does have greater significance, but at this stage our time and funding do not allow it, and if conditions are ripe at a later stage, we will also consider further improving in vitro experiments.

7) The introduction part of this paper is not comprehensive enough, and the similar papers have not been cited, such as “Characterization of diagnostic and prognostic significance of cell cycle-linked genes in hepatocellular carcinoma, Transl Cancer Res, PMID: 35116320”. It is recommended to quote this article.

We added a new paragraph in the introduction section and added citations (after the original 55 lines).

8) This study focuses on the Asian population, but what is the current situation of other populations? What are the specificity and limitations of the Asian population compared to other populations? It is recommended to add relevant content.

We explore the epidemiological characteristics of hcc in the first paragraph of the introduction (lines 36 to 43) and also describe our reasons for choosing the Asian population for the study in the third paragraph of the introduction (lines 56-62).