Peer Review File

Available at http://dx.doi.org/10.21037/tcr-21-37

Comment 1: While I have no major difficulties in understand, there are a few statements that are hard to understand, for instance the lines 105-104 and 213-215. English needs a carefully editing.
Reply 1: Thank you for your advice. I made a professional language editor of the full text.

Changes in the text: The paper has been modified by native speaker, and the polish proof is provided to the journal.

Comment 2: Numerous errors throughout the manuscript should be corrected. 1) using either "COX" or "Cox"; 2) "focus" adhesion (line 21); 3) define abbreviations at their first appearance (like NGS, line 47); 4) "… … the high-risk group and high-risk group, … …" line 149; 5) line 175, "four" should be five; and 5) line 202, "MUT, MUT".
Reply 2:
1) The text is changed to Cox;
Changes in the text: Full text.
2) Replace "focus adhesion" with "Focus adhesion";
Changes in the text: see Page 3, line 46.
3) All abbreviations in the article have been annotated;
Changes in the text: see Page 4, line 79.
4) "the high-risk group and high-risk group" changed to "the high-risk group and low-risk group"; "four" is changed to "five"
Changes in the text: see Page 11, line 220 and see Page 13, line 249.
5). "MUT" is changed to "PCCA"
Changes in the text: see Page 14, line 284.

Comment 3: For ccRCC general knowledge, please cite more recent articles. For instance, the ref 1 (line 34).

Reply 3: Thank you for your advice. I have added the latest report on ccRCC.

Changes in the text: see Page 4, line 60 and see Page 4, line 63-66.

Comment 4: Line 65 – "… new therapeutic targets." Please supported the claim with references.

Reply 4: MiRNA mimics and miRNA inhibitors have been used in clinic as new

therapeutic drugs, and relevant references have been added to this article.

Changes in the text: see Page 22, line 474-496.

Comment 5: Please provide the ratio used to separate the TCGA dataset into a Training and Testing group (line 89).
Reply 5: The differentially expressed miRNA was randomly divided into training group and experimental group according to 1:1.

Changes in the text: see Page 7, line 126-127.

Comment 6: For the Risk Score formula (line 89), please define “gene 1” and “exprgene1” and others; what were they representing for, expression or coefficient? For the same reasons, please give details on “mirR-13-3p*0.278 … …” lines 128-129. Was “0.278” coefficient? How was it derived?

Reply 6: "gene 1" and "exprgene 1" represent the screened miRNA and the corresponding correlation coefficient, respectively. Similarly, we obtained five kinds of prognosis-related miRNA and correlation coefficients through multivariate analysis. The details are shown in **Table 2.**

Changes in the text: see Page 7, line 133-135 and see Page 10, line 187-188.

Comment 7: Please indicate the R packages used (lines 98, 112, and others).
Reply 7: The R software package used in this article has been added.

Changes in the text: see Page 7, line 135-136 and see Page 7, line 143-144 and see Page 8, line155-156.

Comment 8: Please organize a Table for all 211 DE-miRNAs identified with fold changes, p and q values included.
Reply 8: See Annex S1 for details, including details of 5p and 3p.

Changes in the text: see Annex S1.

Comment 9: The meaning of lines 121-122 is not clear. Not clear how the 5 miRNAs were selected from the 11 miRNAs using multivariate Cox regression; please clarify it.
Reply 9: According to the median of miRNAs expression, patients were divided into experimental group and training group at 1:1, and circulation analysis was carried out, and the respective expression levels of the two groups were extracted at the same time. Univariate Cox regression analysis was carried out in the training group, and the

relationship between each miRNA, survival time and survival state was compared in turn, and 11 kinds of miRNA with prognostic value were selected ($p<0.05$). In order to further screen the miRNA, in the model to eliminate the inherent correlation between each miRNA, we then carried out a multi-factor Cox regression analysis, and finally identified five miRNA ($p<0.05$).

Changes in the text: see Page 9, line 171-179.

Comment 10: Fig 3 – Were the cutoff values (median expression) used for the Training, Testing, and Full cohort (A-C) all derived from the Training cohort or individual cohorts? The same applies for panels D-F, i.e. were they all based on the risk scores obtained from the Training group?
Reply 10: We are based on the risk model and calculation formula obtained by the training group. However, the training group and the verification group were still grouped according to 1:1, and each patient in the two groups obtained the corresponding risk value through the calculation formula, and was divided into high-risk group and low risk group according to the median risk value. No matter it is the training group, the verification group and the whole group, the risk value of the patients is calculated according to the model constructed by the training group. The main purpose of Fig3 is to verify whether there is a significant difference in the risk value obtained by the model among the high and low risk groups.

Changes in the text: see Page 10, line 198- Page 11, line 206.

Comment 11: Lines 159-160 – "The feasibility of the model was further verified from the side (Fig.4A)." Not sure how fig 4A can support this claim.
Reply 11: By analyzing the relationship between clinical factors and prognosis, we can see that there is a certain correlation between age, stage, grade and prognosis. By comparing the risk value calculated by the model with clinical factors, we find that the relationship between risk value and prognosis is more significant, and the ROC curve also shows that the AUC value of risk-value is 0.724, which is more significant than other factors. To a certain extent, this shows that the risk value obtained by the model can well predict survival, and also shows the accuracy of the model in this aspect.

Comment 12: Lines 177-179 – "If we use Perl language to find at least two target genes of miRNA recognized by the software, we think that the gene is the target gene of miRNA." The meaning of it is not clear; as well, prediction is always prediction regardless how many programs are used. Please rephrase it for a more precise statement.

Changes in the text: see Page 13, line 251-254.

Comment 13: Lines 200 – 202 – Please provide details regarding how these 12 genes were selected from these overlapped genes included in Fig 5. Please comment by which miRNA these 12 genes are regulated.
Reply 13: MRNA (expression quantity of all target genes) was combined with survival time and survival status data. According to the median of mRNA expression, they were divided into high-risk group and low risk group. The relationship between target gene expression and prognosis was compared. In order to make the relationship between target gene and prognosis most significant, we set $p< 0.00001$. Finally, we got 12 target genes most closely related to prognosis. The corresponding miRNA of these 12 target genes is shown in **Table 4**.
Changes in the text: see Page 14, line 278-286.

Comment 14: Please include the cutoff values in Figs 2, 3, and 8 legend.
Reply 14: In figure 2, we divided patients into high and low expression groups according to the median of five miRNA expressions in each patient. , We divided patients into high-risk group and low-risk group according to the median risk value of each type of miRNA in different patients. in figure 3. In figure 8, according to the median expression of each gene in these samples, all samples are divided into two equal groups, high expression group and low expression group. The red line represents the high expression group and the blue line represents the low expression group.

Changes in the text: see Figs 2, 3, and 8 legend.

Comment 15: Fig 4 – Indicate the panel B and C being derived from univariate and multivariate Cox analysis.

Changes in the text: see Fig 4 legend.

Comment 16: Fig 7 is not legible. Can it be enlarged? As well, please indicate the programs used in the enrichment analyses and graph construction. Please also provide the program identity used to construct the Fig 6 graph.
Reply 16: We have redrawn figure 7 and enlarged it to some extent. The procedures used for GO and KEGG rich analysis are attached to the document. Together with the program identification used to build the diagram in figure 6.
Changes in the text: see Fig 7 and Attachment (R program).