POINT-BY-POINT RESPONSE TO REVIEWER 2

****

Independent Review Report, Reviewer 2

EVALUATION

This study was well designed and has a good research process, but there are still some research problems to be solved. I was concerned about the reliability of the results until these problems are solved and explained.

Response: We greatly appreciate the Reviewer's positive comments. We have carefully revised the manuscript according to the Reviewer's insightful comments and provided point-by-point responses as follows.

****

Methods

1. Why exclude the TCGA database in this study, it also contains a large number of bladder cancer data?

Response: We appreciate the Reviewer's careful review. This article mainly focuses on transcriptomics analysis of non-muscle-invasive bladder cancer (NMIBC), but the TCGA database consists of 1 Ta NMIBC, 3 T1 NMIBC and 407 T2-4 MIBC samples. We exclude this dataset because it contains too few NMIBC samples.

2. What is the full name of CIS?

Response: The full name of CIS is carcinoma in situ. CIS is noted in its full name when it first appears in our manuscript on Line 56. In addition, we have added an Abbreviations section following the main body of this article to make it easier to find. Readers can also find this full name and abbreviation on Line 384.

3. In the current results of Table S4, I can not directly see whether these immune cells are related to the prognosis, and whether can be used to predict the prognosis, perhaps there is a P-value not shown in Table S4?

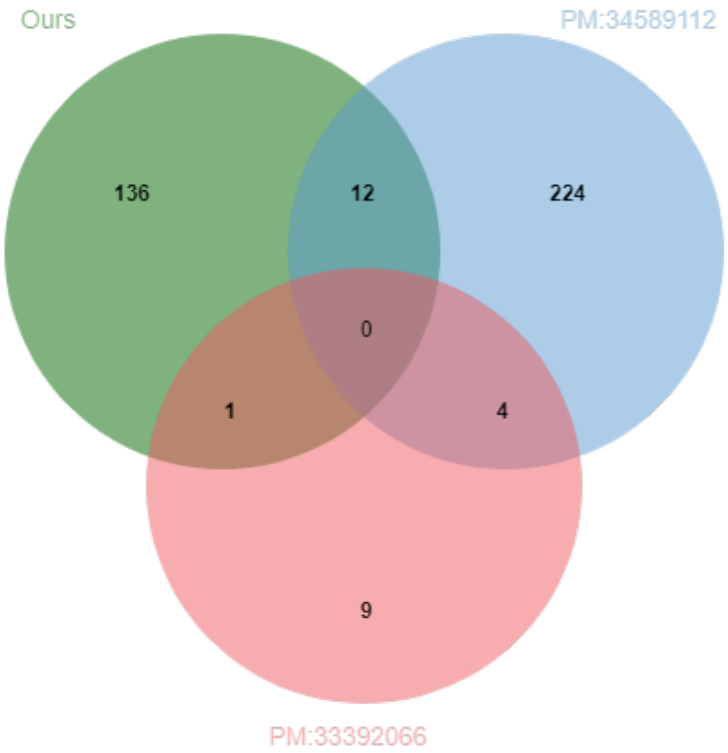
Response: Sorry for causing this ambiguity. Table S4 summarises the results of correlation analyses between selected clinicopathological risk factors and deconvoluted cell types. The original Table S3, which records the corresponding name of 64 cells in different deconvolution methods, gives little information. We change the contents of Table S3 to the significant correlation analysis results, providing P-values and False Discovery Rate (FDR) adjusted P-values. There are 332 records in Table S3, which just correspond to the 332 votes in Table S4. We hope this modification will make things more straightforward for our readers.

4. The identification of immune cell biomarkers seems to ignore the published and confirmed biomarkers in other studies or databases. This study only focused on differentially expressed genes (DEGs) associated with immune cell infiltration. If the proven immune cell biomarkers can be intersected with the current DEGs, it may be more representative and

reliable for subsequent ssGSEA analysis.

Response: In the third paragraph of the Introduction section (Line 64-90), we reviewed four previous studies developing prognostic/predictive biomarkers for NMIBC patients, including UROMOL molecular subtyping framework, immune prognostic signature (IPS) score, stromal tumor-infiltrating lymphocyte (sTIL) levels and luminal markers. Our work was built on these researches. The UROMOL study guided the direction of our research due to its confirmation that higher immune cell infiltration was strongly correlated with a lower recurrence rate. The remaining three studies helped us realize that the molecular subtype markers representing different immune cells might contribute most. So we designed our analysis framework emphasizing the precision evaluation of infiltrating immune cell levels. Identifying representative DEGs naturally became the focus of our study.

We understand the Reviewer's concern that our results may not be reproducible in other studies, so we checked the consistency between our DEGs and previous studies. Two (Zheng et al., 2020; He et al., 2021) out of the four studies we cited provided DEGs relating to the survival of NMIBC patients. We draw a Veen diagram visualizing the intersecting DEGs (Appendix Figure 5). The intersected genes are ANXA10, BTBD16, CRTAC1, CTSE, DEGS1, ENTPD3, FABP6, FAM174B, FAM3B, FGFR3, FN1, SPINK1, VSIG2. Furthermore, as one of the main feature DEGs identified in our study, the collagen family genes have been proven to have a biological function in tumor-infiltrating immune cells during tumor proliferation. The corresponding discussion (Line 327-334) would be another proof of the reliability of our results.



Appendix Figure 5. Veen diagram of infiltrating immune cell related DEGs with potential prognostic value identified by our study and two previous researches.

5. The appropriateness of this standard method "quantile normalization" for dealing with data from different sources needs to be explained. What about the combat method of the SVA package?

Response: Although the re-normalization procedure has been changed entirely according to the Reviewer 1's comments, we are still happy to explain the original considerations of our self-defined re-normalization procedure. At first, we tried to use the "removeBatchEffect" function of the "limma" package to correct the batch effect derived from different technology platforms and data sources. However, we got an expression matrix containing negative values, which is unacceptable. After that, we stopped trying similar batch effect correction methods, such as "Combat" of the "sva" package, because we thought they would cause the same over-adjusted problem. Then we shifted our attempts to batch effect correction methods for other omics data. For example, our experience dealing with metabolomics data led us to our original self-defined re-normalization procedure. The KNN imputing, normalization by housekeeping genes as internal standard substances, and quantile standardization are all commonly used methods in metabolomics data normalization tasks. In retrospect, our lack of experience and self-righteousness lead to this defect of our study. Thanks to both Reviewers' insightful comments, we now better understand the unbalanced transcripts data normalization methods and re-do the re-normalization step. Both the "RUVg" function of the "RUVSeq" package suggested by Reviewer 1 and the "sva" function of the "SVA" package suggested by Reviewer 2 are good options for the job. We choose to use the RUVg because it delivers better subsequential results than SVA. We highlighted all the related modifications in the manuscript and renewed Figures 2-5, Table 2 and Table S5-S8 accordingly.

Results

1. Figure 3 is misquoted in the article. For example, Figure3B is not KEGG enrichment analysis. The contents of C and D in Figure 3 are not explained in this paper.

Response: Sorry for this mistake. We have revised the quotation and description of Figure 3 in Line 247-249. The discussion about KEGG enrichment analysis results (Figure 3C) can be found in Line 249-256 and 333-334. Figure 3D-3F provided visual exhibitions of Gene Ontology (GO) enrichment results. Since the results were consistent with the conclusion suggested by the KEGG pathway enrichment results, we did not explain them separately in the manuscript.

2. Please provide sample information as suppl tables for training group and testing group respectively.

Response: As we introduced in the Method section (Line 190-191), we performed the randomly re-sampling, model building, and model evaluating process 5000 times for each survival type (including PFS, DFS, and OS) to find the optimal model strategy. So it is space-consuming and not necessary to provide all the re-sampling sets in supplementary tables. In the meantime, we could not add a new table because we have arrived at the maximum limit of permitted uploaded figures and tables. However, we can provide our final model's training and test group information by adding columns in the original Supplementary Table S2. Our final model was trained by predicting the PFS of NMIBC patients under the random

seed "6110". The sample size of the training group, test group with all the left available samples, and test group with balanced positive and negative samples were 384, 385, and 64, respectively. The Reviewer can find the corresponding record in the last three columns in Supplementary Table S2, titled with "Final_Model_Trianing", "Final_Model_Test_Left", and" Final_Model_Test_Balance".
In addition, the code to reproduce our modeling procedure, including the re-sampling results and model details, can be found in GitHub (https://github.com/XiaomengSun315/NMIBC_immuno-prognostic) and Gitee repository (https://gitee.com/xiaomengsun/nmibc_immuno-prognostic). Re-running the code entitled "10_Models_v4.2_cell_penalize_score.R" is another way to acquire the same information.

3. In Figure 2B, the blue circle size represents the number of DEGs, the thickness of lines indicates the negative log2 of mean p value of differential expression testing. But there are no legends to measure specific size data.
Response: Sorry for this omission. We have added the appropriate legends in Figure 2B.

4. What do Grade73 and Grade98 denote in Figure 4A? Maybe it is the Grade standard developed in different years, please specify them.
Response: Yes. Grade73 and Grade98 refer to the World Health Organization (WHO) 1973 and WHO 2004/2016 Classification Systems for Urothelial Carcinoma, respectively. The former is a three-tier (G1 vs G2 vs G3) grading system, first published in 1973 (MOSTOFI, 1973). The latter is a two-tier (Low-grade vs High-grade) grading system, first published in 1998 (Epstein et al., 1998). European Association of Urology (EAU) Prognostic Factor Risk Groups of NMIBC advocates the simultaneous use of both grading systems (Sylvester et al., 2021), so we include them both in our study. Their meanings are specified in the manuscript on Line 55. Moreover, we have also added corresponding explanations in the figure notes of Figure 4 on Line 592-593.

5. Please give the specific model formula you got, which immune cell types are included (whether there are any excluded cell types), as well as the coefficients in the model (if these analysis results are available).
Response: In our hypothesis, the combination of nine candidate immune cells plays an essential role in the survival of NMIBC patients. So we chose the modeling strategy that must include all input features in the final model. Please find the formula of our final model in Line 286-290.

Discussion

1. The authors emphasize that the re-normalization of data sets may become the innovation of this study, but as far as I know, re-normalization is a necessary step to integrate and analyze data sets from different sources. Many published studies have standardized data from different sources. You should find out these articles and discuss them to show the superiority of this re-normalization method.
Response: As we mentioned earlier, after evaluating our re-normalized data matrix according to Reviewer 1's suggestion, we have discovered limitations in our self-defined re-

normalization procedure. We have changed to RUVg, a well-established re-normalization method discovered in 2016, as our new re-normalization method. Comparisons between our original method and RUVg were made. Please kindly refer to the corresponding response to Reviewer 1's question "Major 1". In the meantime, we have modified our manuscript, only emphasizing that our analysis procedures were well-designed but eliminating the part of developing our self-defined re-normalization methods.

Reference

Epstein, J. I., Amin, M. B., Reuter, V. R., and Mostofi, F. K. (1998). The World Health Organization/International Society of Urological Pathology consensus classification of urothelial (transitional cell) neoplasms of the urinary bladder. Bladder Consensus Conference Committee. *Am J Surg Pathol* 22, 1435–1448. doi:10.1097/00000478-199812000-00001.

He, Y., Wu, Y., Liu, Z., Li, B., Jiang, N., Xu, P., et al. (2021). Identification of Signature Genes Associated With Invasiveness and the Construction of a Prognostic Model That Predicts the Overall Survival of Bladder Cancer. *Front Genet* 12, 694777. doi:10.3389/fgene.2021.694777.

MOSTOFI, F. (1973). Histological typing of urinary bladder tumors. *International histological classification of tumors* 17. Available at: https://ci.nii.ac.jp/naid/10011134792/ [Accessed March 10, 2022].

Sylvester, R. J., Rodríguez, O., Hernández, V., Turturica, D., Bauerová, L., Bruins, H. M., et al. (2021). European Association of Urology (EAU) Prognostic Factor Risk Groups for Non-muscle-invasive Bladder Cancer (NMIBC) Incorporating the WHO 2004/2016 and WHO 1973 Classification Systems for Grade: An Update from the EAU NMIBC Guidelines Panel. *Eur Urol* 79, 480–488. doi:10.1016/j.eururo.2020.12.033.

Zheng, Z., Mao, S., Zhang, W., Liu, J., Li, C., Wang, R., et al. (2020). Dysregulation of the Immune Microenvironment Contributes to Malignant Progression and Has Prognostic Value in Bladder Cancer. *Front Oncol* 10, 542492. doi:10.3389/fonc.2020.542492.