Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-23-2332

Reviewer A

This is an interesting attempt to detect biomarkers for predicting succesful immunotherapy of lung adenocarcinoma. However, the paper lacks clear descriptions of the work carried out, lacks proper referencing in some places and most importantly does not compare the data with related studie.

1. Description of the work: The authors used a number of algorithms, but the description is insufficient for otherrs to repeat this. It is not clear why the IMvigor cohort was used as only source for validation.

Answer: Dear reviewer, I would like to express my genuine appreciation for your important comments. We acknowledge the importance of providing a comprehensive description of the algorithms and methodologies employed in our study to facilitate reproducibility. In light of your comments, we have revised the manuscript to include a more detailed account of the algorithms used, as well as the specific processes involved in their application.

The detailed process of signature generation unfolded as follows:

(1) Initially, we conducted differential expression analysis between tumor and normal samples in the TCGA-LUAD dataset using the Wilcoxon rank-sum test. Genes were selected based on the following criteria: p-value < 0.05 and |log2FC| > 1. Additionally, in the TCGA-LUAD cohort, we carried out univariate Cox regression analysis and selected genes using the following criterion: p-value < 0.05, and TISCH2 served as a valuable resource for scRNA-seq data from both human and mouse tumors, enabling a comprehensive characterization of gene expression within the TME. In this study, we retrieved CD8T-related genes from TISCH2 using specific criteria: log2FC > log2 (1.5) and adjusted p-value < 0.05. Subsequently, an intersection of these three gene sets was performed, yielding a consolidated list of 33 CD8T cell-related prognostic genes (CD8TRPGs).

(2) Following this, the 101 combinations of algorithms were utilized to independently construct prognostic signatures based on the expression profiles of the 33 CD8TRPGs within the TCGA-LUAD training cohort.

(3) Based on the above results, we selected the combination of RSF and StepCox[forward], which achieved the highest average C-index (0.707). This combination identified a final model named riskScore consisting of 23 CD8TRPGs

(4) We calculated a riskScore for each patient using the expression of 23 CD8TRPGs weighted by their regression coefficients in a Cox model. RiskScores were computed for each validation dataset, namely GSE31210, GSE3141, and GSE72094, were calculated using the signature derived from the training cohort.

(5) Harrell's concordance index (C-index) was then computed for each model across all validation datasets, and the model with the highest average C-index was selected as the optimal one. We updated the content on lines 163-172, 182-205.

Regarding your second point on the use of the IMvigor210 cohort for validation, we

would like to clarify our rationale. The IMvigor210 cohort was selected as an independent validation set due to its large sample size, high-quality clinical and expression data, and the fact that it consists of patients treated with immunotherapy. By using this well-established and reputable cohort, we aimed to validate our findings in a robust and clinically relevant setting. However, we acknowledge that the inclusion of multiple independent validation cohorts would further strengthen the generalizability of our results. In the revised manuscript, to further strengthen the generalizability of our results, we have additionally validated our findings using the GSE135222 dataset, which consists of RNA-seq data from advanced non-small cell lung carcinoma patients treated with anti-PD-1/PD-L1 immunotherapy. Our analysis revealed that overall survival (OS) significantly differed between high-risk and low-risk subgroups in the GSE135222 cohort, supporting the prognostic value of our CD8T+ cell-related gene signature in an independent immunotherapy-treated lung cancer population. The inclusion of both the IMvigor210 and GSE135222 cohorts as validation sets demonstrates the robustness and reproducibility of our findings across different cancer types and immunotherapy settings. We updated the content on lines 67, 156, 396-398, 526-528, 873.

2. Proper referencing: Online 128 Gregory should be Jones, on line 387 the reference to Thorson is lacking, line 537 refers to various signatures, but no references are given and the link between riskScore and these stusie is not discussed.

Answer: Dear reviewer, I would like to express my genuine appreciation for your important comments. We have replaced Gregory with Jones on line 130. We also added the reference to Thorson on line 401. We acknowledge that the link between our riskScore and the other prognostic signatures mentioned in the manuscript could be better discussed. While we have previously described the biological functions of the overlapping genes on lines 542-565, we would like to further elaborate on the connection between our study and the existing literature. For instance, LDHA was also identified as one of the genes within a disulfidoptosis signature in LUAD. KRT18 was also identified as one of the genes within a programmed cell death signature in LUAD. DDIT4 was also identified as one of the genes within a TKI resistant-based prognostic immune related gene signature in LUAD. BTG2 was also identified as one of the genes within a novel mTOR-associated gene signature for predicting prognosis and evaluating tumor immune microenvironment in lung adenocarcinoma. The identification of common genes, such as LDHA, KRT18, DDIT4, and BTG2, across multiple prognostic models highlights their consistent prognostic value in LUAD. Despite the different methodologies and focus of these studies, the recurrent inclusion of these genes underscores their reliability as prognostic markers and supports the validity of our findings. This also highlighted the potential complementarity of these prognostic models, suggesting that integrating our riskScore with other models could provide a more comprehensive and accurate assessment of LUAD prognosis. Furthermore, our study extended the current understanding of the prognostic landscape in LUAD by introducing a novel CD8+ T cell-related gene signature. While building upon the existing knowledge, our riskScore offered a new angle by specifically investigating the impact of the tumor immune microenvironment on LUAD prognosis. In conclusion, we recognize the importance of discussing the link between our riskScore and other prognostic signatures. The identification of common genes across

multiple studies reinforces the prognostic significance of these markers in LUAD. Moreover, the unique focus of our riskScore on CD8+ T cell-related genes complements existing models and expands our understanding of LUAD prognosis. We believe that integrating our findings with those from other studies will facilitate a more comprehensive and precise prognostic assessment in LUAD. We hope this explanation addresses your concern and clarifies the connection between our riskScore and the broader context of LUAD prognostic research. We updated the content on lines 571-590.

3. Clinical relevance and comparison with other relevant studies: Similar recent studies are e.g. Immunoscore immune checkpoint using spatial quantitative analysis of CD8 and PD-L1 markers is predictive of the efficacy of anti- PD1/PD-L1 immunotherapy in non-small cell lung cancer By François Ghiringhelli et al. and Biomarkers for Immune Checkpoint Inhibitor Response in NSCLC: Current Developments and Applicability by Katiane Tostes et. al. I would have expected a discussion of the results obtained by the authors in the light of these and other relevant papers.

Answer: Dear reviewer, I would like to express my genuine appreciation for your important comments. Our study introduces a novel artificial intelligence-derived prognostic signature, riskScore, based on 23 consensus prognostic genes related to CD8 T cell marker genes in LUAD. This tool not only predicts overall survival, progression-free interval, and disease-specific survival but also correlates with key immunological features of the tumor microenvironment (TME), offering insights into immunotherapy responses.

Comparing our findings with the studies mentioned (Ghiringhelli et al., and Tostes et al.), our research aligns with and extends the current understanding of the predictive value of immune markers in NSCLC treatment outcomes. Ghiringhelli et al. highlighted the significance of a spatial quantitative analysis of CD8 and PD-L1 markers, which is predictive of the efficacy of anti-PD1/PD-L1 immunotherapy in NSCLC. Our riskScore model, while focusing on a broader set of CD8 T cell-related genes, underscores the complexity of immune interactions within the TME and their implications for LUAD patient prognosis and therapy response. This complements the spatial and density analysis of CD8 and PD-L1 markers, suggesting a multi-faceted approach to understanding and predicting immunotherapy outcomes.

Furthermore, Tostes et al. reviewed biomarkers for immune checkpoint inhibitor response, emphasizing the need for efficient tools in clinical decision-making. In addition to the research by Ghiringhelli et al. and the review by Tostes et al., several other studies have explored prognostic biomarkers and risk models for lung adenocarcinoma (LUAD) based on the tumor immune microenvironment. Rizvi et al. found that higher TMB was associated with improved survival in LUAD patients treated with immune checkpoint inhibitors (DOI: 10.1126/science.aaa1348). Wang et al. found that predictive power of tumor mutational burden in lung cancer immunotherapy response is influenced by patients' sex (DOI: 10.1002/ijc.32327). Our study directly contributes to this need by providing a prognostic tool that not only predicts survival outcomes but also identifies potential therapeutic targets within specific risk subgroups. The riskScore's correlation with tumor-promoting biological

functions, TMB, NEO, MSI, and immune cell infiltration offers a comprehensive understanding of the TME, facilitating tailored treatment strategies.

By establishing a link between our riskScore model and the well-documented predictive markers (CD8 and PD-L1), we can offer a more nuanced understanding of the TME and its impact on immunotherapy efficacy. This integration not only validates our model's significance but also opens avenues for refining patient selection and treatment personalization in LUAD.

In conclusion, our study's findings, in concert with the insights from Ghiringhelli et al. and Tostes et al., contribute to a growing consensus on the critical role of immune markers in predicting immunotherapy responses. The riskScore model enhances our capability to stratify LUAD patients more effectively, offering a promising tool for guiding clinical management and exploring new therapeutic avenues. We updated the content on lines 454-461.

Reviewer B

This interesting original article may support clinical decision-making in patients with lung adenocarcinoma. The authors demonstrate the utility of risk score as a robust tool for guiding clinical management and tailoring personalized treatments for lung adenocarcinoma patients, supported by its association with immunotherapy response and tumor characteristics. The manuscript is well-written, and the results are clearly presented. However, I have a few major comments, which are explained below.

1.In this study, data from TCGA LUAD show that the determined risk score is associated with CD8 and reflects the tumor immune microenvironment. Genetic variation has been reported to affect the immune microenvironment. Were there any genes associated with risk scores in TCGA LUAD genetic variant data?
Answer: Dear reviewer, I would like to express my genuine appreciation for your important comments. This insightful question has drawn our attention to the potential impact of specific gene mutations on patient prognosis in lung adenocarcinoma, which we had previously overlooked in our manuscript. To address this point, we have carefully analyzed the TCGA LUAD MAF mutations and found that mutations in TP53, PTEN, and SMARCA4 are indeed associated with significantly higher risk scores, which is consistent with our earlier conclusion that higher risk scores indicate poorer prognosis.

The TP53 gene, often referred to as P53, encodes a crucial tumor suppressor protein that plays a vital role in preventing cancer formation. TP53 is the most frequently mutated gene in lung adenocarcinoma, and its mutations have been associated with various clinical and molecular features. TP53 mutations predict poor response to immunotherapy in patients with metastatic solid tumors, including NSCLC (DOI: 10.1002/cam4.5953).

A study by Biton et al. showed that TP53 mutations in lung adenocarcinoma were major determinants of the tumor immune profile (TIP) and of the expression of PD-L1 by malignant cells, which may contribute to a more aggressive tumor phenotype and worse clinical outcomes [2] (DOI: 10.1158/1078-0432.CCR-18-0163). Mendoza et al. showed that TP53 mutation alone may have the capacity to act as an oncogenic driver in lung adenocarcinomas, and its presence in isolation may be sufficient to trigger poor prognostic features in lung tumors (DOI:

10.1002/cam4.6873). With multi-omic cellular and spatial tumor atlas of 23 treatment-naïve NSCLC human tumors, Zhao et al. found that highly-entropic TP53mut malignant cells lose alveolar identity and coincide with an increased abundance of exhausted T cells and immune checkpoint interactions with implications for response to checkpoint blockade (DOI: 10.1101/2023.06.28.546977).

PTEN functions to suppress cell survival, proliferation, and growth, thereby acting as a potent tumor suppressor. PTEN mutations have been reported in a subset of lung adenocarcinomas and are associated with poor prognosis. Gkountakos et al. demonstrated that PTEN expression is usually determined by immunohistochemistry and low protein levels have been associated with decreased survival in lung cancer. Moreover, available data involve PTEN mutations and loss of activity with resistance to targeted treatments and immunotherapy (DOI: 10.3390/cancers11081141). Parikh et al. reported the case of a patient with metastatic NSCLC harboring a PTEN mutation and PD-L1 positivity, refractory to a checkpoint inhibitor (DOI: 10.2147/LCTT.S161738). Chen et al. found that PTEN mutation was only found in non-responders in NSCLC (DOI: 10.1111/cas.14113).

The SMARCA4 gene encodes a member of the SWI/SNF family of proteins that are part of the ATP-dependent chromatin remodeling complexes. These complexes are crucial for regulating gene expression by altering chromatin structure to allow access to transcription factors. The loss of function of the SMARCA2/A4 catalytic subunit within the SWI/SNF complex due to the inactivation of the SMARCA4 gene will lead to an increase in DNA damage functionality (DOI: 10.3390/cells10081920). SMARCA4 gene mutation is reported to encompass 12% of non-oncogenic addicted lung adenocarcinomas and 5% are co-altered among oncogenic-driven lung adenocarcinomas, SMARCA4-deficient NSCLCs with co-alterations are known to be unfavorable and with almost 50% of these tumors harboring biallelic truncating SMARCA4 alterations, which causes highly negative effects in the patient, leading to a poor survival outcome (DOI: 10.1016/j.jtho.2020.01.002, DOI: https://doi.org/10.1101/2020.04.18.043927). These findings suggest that SMARCA4 mutations may contribute to an immune-evasive phenotype, leading to poorer prognosis in lung adenocarcinoma.

In summary, our analysis of the TCGA LUAD dataset has identified mutations in TP53, PTEN, and SMARCA4 as potential determinants of patient prognosis in lung adenocarcinoma, with higher risk scores observed in the mutated groups. These findings are supported by recent studies that have uncovered the complex interplay between these gene mutations and the tumor immune microenvironment, providing mechanistic insights into how these mutations may contribute to a more aggressive tumor phenotype and worse clinical outcomes. We believe that these results not only strengthen our original conclusions but also highlight the importance of considering specific gene mutations when assessing patient prognosis and developing personalized treatment strategies for lung adenocarcinoma. We updated the content on lines 438-448, 887-890, Figure S1.

2.In lung adenocarcinoma, EGFR mutations are clinically and fundamentally important. Did the risk scores differ between lung adenocarcinomas with EGFR mutations and wild-type EGFR?

Answer: Dear reviewer, I would like to express my genuine appreciation for your important comments. We have carefully analyzed the differences in risk scores between lung adenocarcinoma patients with EGFR mutations and those with wild-type EGFR, the results showed no statistically significant difference between the two groups (P=0.73) in TCGA LUAD samples. And combined with current research progress, we provide the following detailed response:

In the clinical diagnosis and treatment of lung adenocarcinoma, genetic testing has become routine to guide personalized treatment decisions. The gene panel of the internationally renowned genetic testing company Foundation One can simultaneously analyze multiple genes

related to the occurrence, progression, and treatment of non-small cell lung cancer (NSCLC), including EGFR, ALK, BRAF, MET, and ROS1 (https://www.foundationmedicine.com/test/foundationone-cdx). NSCLC can further subdivide into two most common subtypes, lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), representing 50–60% and 20–30% of total NSCLC cases, respectively [DOI: 10.1016/S0140-6736(21)00312-3, DOI: 10.1097/JTO.0000000000000630]. The mutation or gene-fusion information of these genes helps identify patients suitable for corresponding targeted drug treatments [DOI: 10.1038/nbt.2696, DOI: 10.3390/ijms241713390, DOI: 10.3389/fmed.2022.758464]. In addition, indicators such as PD-L1 expression level, tumor mutation burden (TMB), and microsatellite instability (MSI) are also included in the detection to predict the efficacy of immune checkpoint inhibitors [DOI: 10.1016/j.annonc.2020.07.014, DOI: 10.1200/JCO.2017.75.3384, DOI: 10.1016/j.jtho.2018.05.013, DOI: 10.1001/jamaoncol.2022.1981, DOI: 10.1038/s41588-023-01355-5]. This indicates that in the precision diagnosis and treatment of NSCLC, other molecular markers besides EGFR also play an important role.

In our study, Wilcoxon rank-sum test was used to compare the risk scores of lung adenocarcinoma patients with EGFR mutations and those with wild-type EGFR. The results showed no statistically significant difference between the two groups (P=0.73). This finding suggests that although EGFR mutations are of great clinical and basic research significance in the occurrence, progression, and treatment of lung adenocarcinoma [DOI: 10.1056/NEJMoa1913662, DOI: 10.1056/NEJMoa1713137], the EGFR mutation status may not be the main factor causing prognostic differences in our risk score model.

In our study, the previous analysis found that patients in the low-risk score group had higher levels of TMB, MSI, and neo-antigen load (NEO), and higher expression of immune checkpoint molecules PD-1 and PD-L1. This suggests that the low-risk score group may have better immunogenicity, leading to better prognosis. This result is consistent with many recent studies, which have shown that indicators such as TMB, MSI, and NEO can predict the efficacy of immune checkpoint inhibitors and patient prognosis [DOI: 10.1038/s41588-018-0312-8, DOI: 10.1016/j.annonc.2021.02.006, DOI: 10.3389/fimmu.2023.1227797, DOI: 10.1016/j.lungcan.2023.107255, DOI: 10.1016/j.ccell.2022.08.003, DOI: 10.1126/sciimmunol.abm6359].

Moreover, our literature found that although EGFR plays a key role in lung adenocarcinoma, patient prognosis and survival do not entirely depend on the EGFR mutation status. Multiple studies have shown that in addition to EGFR, other genomic changes such as ALK, ROS1, BRAF, and KRAS also play important roles in the occurrence and progression of lung adenocarcinoma [DOI: 10.1016/j.annonc.2022.07.005, DOI: 10.1038/s41568-019-0179-8, DOI: 10.1002/cpt.810]. Furthermore, factors such as tumor heterogeneity, immune microenvironment, and treatment plan selection can also significantly affect patient prognosis [DOI: 10.1056/NEJMoa1616288, DOI: 10.1016/j.lungcan.2019.06.020, DOI: 10.1056/NEJMra1703413]. This evidence supports our finding that the EGFR mutation status may not be the decisive factor influencing risk scores and prognosis.

In summary, our study results indicate that although there was no significant difference in risk scores between lung adenocarcinoma patients with EGFR mutations and those with wild-type EGFR, patients in the low-risk score group may have better prognosis due to higher levels of TMB, MSI, NEO, and PD-1/PD-L1 expression. This finding highlights the importance of immune-related biomarkers in predicting the prognosis of lung adenocarcinoma and reflects the complexity of lung adenocarcinoma biology and clinical management. Our study provides new ideas and evidence for further development and optimization of prognostic prediction models for lung adenocarcinoma.

We hope this detailed answer can address your concerns. We believe that with the development

of precision medicine, integrating multi-omics data and clinical information to establish more comprehensive and optimized prognostic prediction models will help guide personalized diagnosis and treatment decisions for NSCLC patients and ultimately improve patient survival benefits. We updated the content on lines 438-448, 887-890, Figure S1.