Peer Review File

Article Information: https://dx.doi.org/10.21037/tlcr-21-751

Reviewer A:

The authors tested the association of somatic mutation profile, copy number alterations, immune infiltrate composition and location, TMB and PD-L1 expression in stage I NSCLC with survival and integrated these parameters to develop a prognostic model. The study is interesting and add relevant information to the field. However, some clarification and corrections are necessary.

Major comments

Comment 1: The authors should make clear that they, in fact, purposefully compared two cohorts with opposite outcomes (early x late/non recurring cancer), and not that they included these populations in the cohort.

Reply 1: We thank the reviewer for this comment. We have made clear that we compared early recurrence group and non-recurrence group before this resubmission.

Change in text: We have modified our text as advised in title, abstract, and whole text.

Comment 2: What was define as tumor recurrence? Metastasis? Second primaries? Local relapse?

Reply 2: We thank the reviewer for this comment. Tumor recurrence was defined as local relapse or distant metastasis in our study.

Change in text: We have modified our text as advised (see Page 6, line 128-129)

Comment 3: Was the test of association between recurrence and gene mutations performed for each gene separately? All of them? One by one? Was any correction for type I error performed? Maybe, it would be interesting to look if there was some clusterization go genes capable of segregating the two cohorts instead.

Reply 3: We thank the reviewer for this comment. Chi-square test or fisher exact test

were used to analyzed the association between recurrence and gene mutations for genes with mutation number >2 (n=97). And P<0.005 was considered statistically significant for correction of type I error. As a result, MUC4 mutation significantly correlated recurrence (*P*=0.0008). We performed GO enrichment analysis for mutant genes and found several clusterization go genes were enriched only and significantly enriched in early-recurrence or non-recurrence groups (Figure S1).

Change in text: We have modified our text as advised (see Page 12, line 243-245 and Page 16, 332-338)

Comment 4: In how many samples TMB was calculated from WES data and how many from targeted sequencing panels? These platforms may give different TMB estimates. How was TMB normalized and calibrated between them? I suggest to use the public tool to adjust TMB read-outs (Vega et al. Ann Oncol. 2021).

Reply 4: We thank the reviewer for this comment. We calculated TMB in 74 samples from WES data and 14 samples form targeted sequencing panels. Considering the data from WES might be enough for analysis and the heterogeneity between different platform might lead to more statistical errors even after adjustment, we did not use the TMB data from targeted sequencing panels.

Change in text: We have modified our text as advised (see Page 15, line 316-320)

Comment 5: Immune infiltrate was analyzed in whole section slides or in TMAs? Was any type of normalization of performed to improve comparability between samples, for example, divide the number of CD3 or CD8 positive cells by tumor area analyzed or by the number of CD45 positive cells?

Reply 5: We thank the reviewer for this comment. To normalize density of CD3+ and CD8+ TILs, we performed immunohistochemistry in whole section slides, selected all areas under 200x microscope, and calculated the number of CD3 or CD8 positive cells per square millimeter.

Change in text: We have modified our text as advised (see Page 9, line 188-191)

Comment 6: Why not investigate PD-L1 expression as a continuous variable? How were the cutoffs determined (TPS>1% and CPS>5%)?

Reply 6: We thank the reviewer for this comment. We agree that it is better to investigate PD-L1 expression as a continuous variable in our research. Consistent with previous results, we found no association between PD-L1 expression and early recurrence (P=0.8727 for TPS and P=0.3578 for CPS).

Change in text: We have modified our text as advised (see Page 17, line 362-364)

Comment 7: The validation cohort for IS was obtained from the TIMER database. Nevertheless, the TIMER databank employed gene expression to estimate immune cell abundance in their samples while the authors used immunehistochemistry to assess immune cell abundance and distribution. How might this affect comparability?

Reply 7: We thank the reviewer for this comment. The data form the TIMER databank could not be used to for validation of IS. We were hoping to validate the prognostic value of CD8+T infiltration through TIMER database. However, we agree that the method to estimate immune cell abundance of CD8+ TILs were different between TIMER databank and our cohort, and results from TIMER might not be able to support our conclusion. Thus, we decided to delete the results of CD8+ T cells from TIMER database before this submission. In future study, we will validate the prognostic value of IS testing with same method.

Change in text: We deleted the results of CD8+ T cells from TIMER database.

Comment 8: The tumor staging system has been modified along the 9 years during which patients were accrued for this study. What edition of TNM was employed? Was TNM staging reviewed?

Reply 8: We thank the reviewer for this comment. We reviewed the imaging system and surgical report for all patients, and restaged according to the eight edition of AJCC TNM staging system.

Change in text: We have modified our text as advised (see Page 6, line 126-128)

Comment 9: What was the minimum follow-up time and the rate of follow-up loss in each cohort? This may skew median survival in each cohort and should be clarified.

Reply 9: We thank the reviewer for this comment. The median follow-up time was 13.5 months for recurrence group and 77 months for non-recurrence group. Patients who had missing value for follow-up loss were not enrolled in our cohort.

Change in text: We have modified our text as advised (see Page 12, line 250-252)

Comment 10: Was there any further selection besides timer recurrence? The prognostic variables evaluated are too perfectly matched for a retrospective cohort. It would be interesting to compare the distribution of variables in each cohort in relation to that observed in all patients diagnosed with stage I NSCLC in the same period. For instance, there are more patients with adenocarcinoma and pleural invasion in the early recurrence group.

Reply 10: We thank the reviewer for this comment. Besides timer recurrence, several criteria were used for selection of patients. Patients that met the following criteria were included in this study: (1) pathological diagnosis of stage I NSCLC; (2) surgical margins were negative; (3) complete clinical data and pathological tissue specimen were available. Patients with the following conditions were excluded from the study: (1) had any other malignancies; (3) received neoadjuvant chemotherapy; (4) could not obtain qualified tumor tissue; (5) refused to participate in this study. We selected 97 stage I NSCLC patient who had postoperative recurrence within 3 years (early recurrence group) and had no recurrence more than 5 years (non-recurrence group) from TCGA cohort. Early recurrence group had more patients with adenocarcinoma, while age, gender, smoking history, and TNM stage were similar between two groups in TCGA cohort, which was consistent with our cohort (Table S1).

Change in text: We have modified our text as advised (see Page 6-7, line 131-137 and Page 12, 261-264)

Comment 11: It would be desirable to evaluate split T (T1a x T1b x T1c) stage separately.

Reply 11: We thank the reviewer for this comment. We evaluated split T (T1a x T1b x T1c) stage separately and found no association between T stage and recurrence (*P*=0.261; Table 1).

Change in text: We have modified our text as advised (see Table 1)

Comment 12: Authors should clarify how variables were selected for multivariate logistic regression. Since staging and grade are well defined prognostic variables, they should have been included in the multivariate model, although they were not statistically significant in the univariate model.

Reply 12: We thank the reviewer for this comment. We selected variables that were statistically significant in the univariate analysis, including TMB, histologic type, IS, and TMS for multivariate. We enrolled staging and grade in multivariate logistic regression before this submission and found that TMS and IS were still independent prognostic factors for recurrence, and TMB, histologic type, staging, and grade were not significantly associated with early recurrence (Figure 8).

Change in text: We have modified our text as advised (see Page 18, line 375-378)

Minor comments

Comment 13: There are some errors regarding syntax error and use of prepositions that must be revised in the text.

Reply 13: We thank the reviewer for this comment. We have carefully edited the entire manuscript and had the revised manuscript corrected again by professional editors before this re-submission.

Change in text: We have modified our text as advised.

Comment 14: Frequencies in Table 1 should be calculated in relation to number of all informative cases for that specific parameter and and not inside each category.

Reply 14: We thank the reviewer for this comment. We have modified Table 1.

Change in text: We have modified our text as advised (see Table 1)

Comment 15: Include a bar above the graph in figure 1 indicating which samples come from cases who had recurrence or not.

Reply 15: We thank the reviewer for this comment. We added a bar above the graph in figure 1 to distinguish different groups.

Change in text: We have modified our text as advised (see Figure 1)

Comment 16: I think the formula to calculate the TMS should be moved to the supplementary material section.

Reply 16: We thank the reviewer for this comment. We moved the formula for calculating the TMS to the supplementary material section.

Change in text: We moved the formula for calculating the TMS from results to the supplementary material section.

Comment 17: The discussion is quite long. It should be more focused on the added benefit of using the model the authors developed, instead of exploring each prognostic parameter tested.

Reply 17: We thank the reviewer for this comment. We revised the discussion section before this submission.

Change in text: We have modified our text as advised (see Page 20-25, line 421-549)

Reviewer B:

In this study, a recurrence risk prediction model in resected stage I NSCLC was generated by integrated genomic alterations and immune signatures of tumor immune microenvironment. One hundred thirty patients were enrolled. Whole exome sequencing was performed to evaluate gene mutation, copy number variation, and tumor mutation burden (TMB). Immunohistochemistry was carried out to assess expression of PD-L1 and level of CD3+ and CD8+ tumor-infiltrating lymphocytes

(TILs). Tumor mutation score (TMS) was constructed with the recurrence-associated genes that identified by Lasso regression. This study included relatively large number of stage I NSCLC patients for generation of a recurrence risk prediction model. However, several critical questions need to be answered.

Comment 1. In stage IB NSCLC patients, the adjuvant chemotherapy after surgical resection is usually recommended. How many stage IB NSCLC patients received postoperative adjuvant chemotherapy in this study? Have you analysed the impact of postoperative adjuvant chemotherapy in prognosis?

Reply 1: We thank the reviewer for this comment. There were two stage IB NSCLC patients received postoperative adjuvant chemotherapy in local cohort and adjuvant chemotherapy had no correlation with early recurrence (P=0.498; Table 1).

Change in text: We have modified our text as advised (see table 1)

Comment 2. For those who were not qualified for WES, targeted gene panel sequencing consisted of 425 genes was performed to detect the somatic mutation. How many patients were analysed by targeted panel sequencing?

Reply 2: We thank the reviewer for this comment. Fourteen samples were analyzed by targeted panel sequencing.

Change in text: We have modified our text as advised (see Page 8, line 162-164)

Comment 3. TMB was defined as the number of missense mutations per megabase of coding regions of the genome sequenced in this study. What was the definition of high TMB? Every studies adopted different definition and the cut-off level of high TMB can be different between LUAD and LUSC.

Reply 3 : We thank the reviewer for this comment. In uni-variate analysis, we investigated TMB as a continuous variable. In multivariate analysis, we divided TMB into high and low groups by cut off >50% for whole NSCLC cohort.

Change in text: We have modified our text as advised (see Figure 8A and Page 18, line 397)

Comment 4. TMB can be defined as the number of missense mutations per megabase of coding regions of the genome sequenced in WES. How did you define the TMB in patients with targeted panel sequencing?

Reply 4: We thank the reviewer for this comment. Panel TMB was counted by summing all base substitutions and indels in the coding region of targeted genes, including synonymous alterations to reduce sampling noise and excluding known driver mutations as they are over-represented in the panel, as previously described [1].

Change in text: We put the definition of TMB in panel in supplementary methods.

Reference:

[1] Fang, W., et al. (2019). "Comprehensive Genomic Profiling Identifies Novel Genetic Predictors of Response to Anti-PD-(L)1 Therapies in Non-Small Cell Lung Cancer." Clin Cancer Res 25(16): 5015-5026.

Comment 5. The authors compared genes with top 40 mutation frequency in study cohort and TCGA cohort and found that only two genes (TP53 and TTN) were common in both cohorts. Only three common genes (TP53, TTN, and AKNAK2) were identified in study cohort and matched the TCGA cohort. How do you explain the reason? Is it ethnic difference?

Reply 5: We thank the reviewer for this comment. We think the ethnic difference contributed to difference of genomic background between local cohort and TCGA cohort. Previous studies reported that gene mutation were different between western and eastern NSCLC patients[1-2].

Change in text: We have modified our text as advised (see Page 21, line 457-461)

Reference:

[1] Chen, J., et al. (2020). "Genomic landscape of lung adenocarcinoma in East Asians." Nature Genetics 52(2): 177-186.

[2] Jiang, T., et al. (2019). "Genomic landscape and its correlations with tumor mutational burden, PD-L1 expression, and immune cells infiltration in Chinese lung squamous cell carcinoma." J Hematol Oncol 12(1): 75.

Reviewer C:

The report by Hu et al. investigates biomarkers for the risk of recurrence of stage I NSCLC in terms of genetic abnormalities and immunological environmental factors using mathematical methods.

I hope that each of the following items will be clarified before publication of the paper.

Comment 1. What was the percentage of patients in this study who received postoperative adjuvant therapy (excluding the TCGA cohort)? If the cases who underwent postoperative adjuvant therapy are to be included in the prognostic analysis, I think a survival analysis (landmark analysis, multi-state model analysis) that takes into account the cases who underwent postoperative adjuvant therapy is necessary.

Reply 1: We thank the reviewer for this comment. There were 2 (1.5%) patients received postoperative adjuvant chemotherapy in local cohort and adjuvant chemotherapy had no correlation with early recurrence (P=0.498; Table 1)

Change in text: We have modified our text as advised (see table 1)

Comment 2. What is the hypothesis and primary endpoint of this study?

Reply 2: We thank the reviewer for this comment. The hypothesis of this study was that gene mutation and immune microenvironment are closely correlated with early recurrence, and prediction model integrated both is able to provide guide for individualized treatment in patients with stage I NSCLC. The primary endpoint of this study was early recurrence after surgical resection.

Change in text: We have modified our text as advised (see Page 7, line 138-139)

Comment 3. Is the WES evaluation done in CLIA (Clinical Laboratory Improvement

Amendments) lab?

Reply 3: We thank the reviewer for this comment. The WES was evaluated in CLIA lab.

Change in text: We have modified our text as advised (see Page 8, line 161)

Comment 4. Describe the reason why you chose Lasso as the method for selecting TMS genes.

Reply 4: We thank the reviewer for this comment. Lasso regression can select variables and built a integrated model by forcing the sum of the absolute value of the regression coefficients to be less than a fixed value, which forces certain coefficients to zero, excluding them from impacting prediction. We chose Lasso regression because it is one of the most popular variable selection and risk prediction algorithms in dealing with high dimensional data [1].

Reference:

[1] Cheung-Lee, W. L. and A. J. Link (2019). "Genome mining for lasso peptides: past, present, and future." J Ind Microbiol Biotechnol 46(9-10): 1371-1379.

Change in text: No change in text.

Comment 5. Please explain why you chose the tumor center and invasive margin as the sites for TIL assessment, and provide evidence that assessment of these sites indicates tumor immune activity in lung cancer.

Reply 5: We thank the reviewer for this comment. We chose the tumor center and invasive margin because recent studies reported that TILs in these two region were might closely correlated with cancer metastasis [1]. Immune score integrated CD3+ and CD8+ TILs from tumor center and invasive margin showed great prognostic value in colon cancer [2].

Change in text: We have modified our text as advised (see Page 22, line 468-472)

Reference:

[1] Pagès, F., et al. (2018). "International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study." The Lancet

391(10135): 2128-2139.

[2] Pagès, F., et al. (2018). "International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study." The Lancet 391(10135): 2128-2139.

Comment 6. Please provide evidence that this study's assessment method (lines 180-183) is suitable for assessing tumor immune activity and IS in lung cancer.

Reply 6: We thank the reviewer for this comment. The cut off value is controversial for PD-L1 TPS and CPS, thus we decided to investigate PD-L1 expression as a continuous variable in this version. Consistent with previous results, we found no association between PD-L1 expression and early recurrence (P=0.8727 for TPS and P=0.3578 for CPS).

Change in text: We have modified our text as advised (see Page 17, line 362-364)

Comment 7. Regarding the assessment of the density of immune microenvironment factors, the density of CD8, CD3 in the whole slide should be assessed on whole slide imaging, as in a previous study (Lancet 2018;391:2128).

Reply 7: We thank the reviewer for this comment. We described the method in a wrong way. In previous manuscript, we described that we quantitatively detected cell density of CD3+ and CD8+ TILs. However, the description was not accurate. In fact, we made a semi-quantitative analysis of the cell density for each sample. Although whole slide imaging is more accurate, semi-quantitation is more widely used in clinical application, which will make our model more applicable.

Change in text: We have modified our text as advised (see Page 9, line 187)

Comment 8. When you built the IS, how did you set the density cutoff values for CD3 and CD8?

Reply 8: We thank the reviewer for this comment. We calculated the density cutoff values of CD3 and CD8 by performing ROC curve analysis.

Change in text: We have modified our text as advised (see Page 9, line 192-193)

Comment 9. What is the definition of combined proportion score (CPS)?

Reply 9: We thank the reviewer for this comment. CPS was defined as the number of positive tumor cells, lymphocytes and macrophages, divided by the total number of viable tumor cells multiplied by100.

Change in text: We have modified our text as advised (see Page 10, line 202-205)

Comment 10. line 308; High TMB→Higher TMB?

Reply 10: We thank the reviewer for this comment. We have modified our text as advised.

Change in text: We have modified our text as advised (see Page 15, line 321-327)

Comment 11. Are there any biostatisticians involved in the analysis of this study? Are there any co-authors?

Reply 11: We thank the reviewer for this comment. There is no biostatistician involved in the co-authors. However, a biostatisticians helped us when we revised this manuscript in this version. We will include him as a co-authors if the editor agree our application.

Change in text: No change in text.

Comment 12: Please describe the biological and clinical significance of the mathematically selected TMS 11 gene abnormality in Lasso.

Reply 12: We thank the reviewer for this comment. We described the biological and clinical significance of MUC4 and other TMS genes in discussion before this submission. Besides MUC4, MEOX2 was also closely associated with tumor progression and metastasis [1-2]. While few studies reported the role of the other TMS 11 genes in cancer.

Change in text: We have modified our text as advised (see Page 20, line 433-435)

[1] Wang Z, Yang H, Zhang R, et al. MEOX2 serves as a novel biomarker associated with macrophage infiltration in oesophageal squamous cell carcinoma and other

digestive system carcinomas. Autoimmunity. 2021;54(6):373-383.
[2] Liu Y, Cheng L, Li C, Zhang C, Wang L, Zhang J. Identification of tumor microenvironment-related prognostic genes in colorectal cancer based on bioinformatic methods. Sci Rep. 2021;11(1):15040.

Comment 13: Please show whether the TMS 11 genes mathematically selected by the Lasso method can be reproducibly selected by other analysis methods.

Reply 13: We thank the reviewer for this comment. Except Lasso regression, we also performed random forest for variable selection and found that the result form random forest was consistent with Lasso (Figure 1). In addition, several methods can also be considered for variables selection in our research, including multivariate regression model, classification and regression tree (CART), bagged trees, bootstrap. However, Lasso regression was one of the most popular algorithms for variables selection and model construction when data werelinearly dependent and high-dimensional [1]. Thus, we only reported the results form Lasso.

Change in text: No change in text.

[1] Cheung-Lee, W. L. and A. J. Link (2019). "Genome mining for lasso peptides: past, present, and future." J Ind Microbiol Biotechnol 46(9-10): 1371-1379.

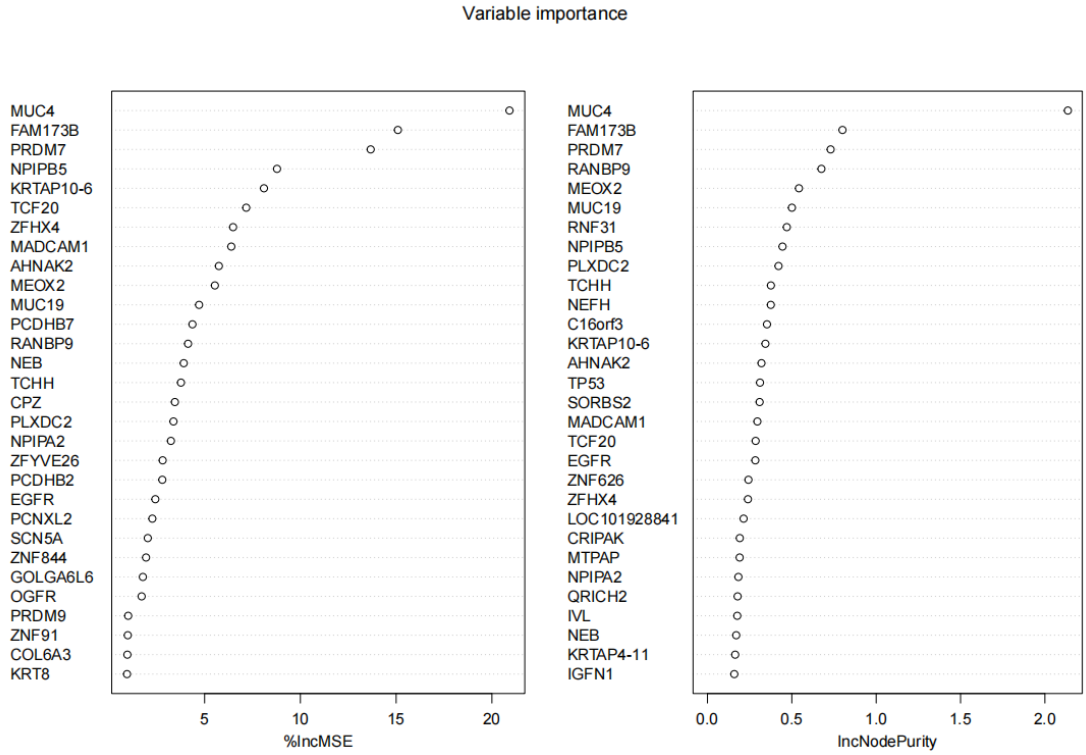

Figure 1. Random Forest for selecting early recurrence associated gene mutations.

Comment 14. Please describe the biological and immunological significance of the 14 IS-associated TMS genes mathematically selected by the Lasso method. Please indicate whether the selection can be reproducibly made by other methods.

Reply 14: We thank the reviewer for this comment. We described the relationship between the 14 IS-associated gene mutations before this submission.

Change in text: We have modified our text as advised (see Page 25, line 530-540)

Comment 15. What does TIME (line 441, 445, etc.) stand for?

Reply 15: We thank the reviewer for this comment. The TIME meant tumor immune microenvironment.

Change in text: We have modified our text at the Page 5, line 103, where TIME showed first time in text.

Comment 16. Both Ly and v factors, which is already a prognostic factor for stage I NSCLC, should also be considered.

Reply 16: We thank the reviewer for this comment. We agree that other clinical factors like staging and grade are well defined prognostic variables. Although they were not statistically significant in the univariate analysis, we enrolled them in the multivariate model. The new multivariate analysis showed that TMS and IS were still independent predictors (Figure 8A).

Change in text: We have modified our text as advised (Figure 8A and Page 18, lines 376-379)

Comment 17. How can the results of TMS and IS presented in this study be used in clinical practice?

Reply 17: We thank the reviewer for this comment. We created nomogram to visualize the integrated model, which could be used in clinical practice. Doctors can calculate the early recurrence risk through the nomogram model.

Change in text: We have modified our text as advised (see Page 23, line 502-505)

Reviewer D:

The authors painstakingly performed WES of stage 1 NSCLC and muc4 mutation was revealed to predict the prognosis, which I think is relatively new except recently published paper BIOENGINEERED 2021, VOL. 12, NO. 1, 791–802.

The lung cancer cohort in China and other east Asia may be unique in that many have EGFR mutation and occur in never smoker, and in the female. In the analysis of the authors, the readers would like to know the concordance or correlation of MUC4 mutation, EGFR mutation and DETAILED histological subtype (WHO; lepidic growth or non-mutinous BAC type depending on the versions).

Reply : We thank the reviewer for this comment. In our study, all EGFR mutation were identified in patients with LUAD, though the difference was not statistically significant might due to the sample size (P=0.0590). The MUC4 mutation had no correlation with histologic type (P>0.999). Compared with wild-type patients, MUC4 mutation was higher in patients with EGFR mutation (42.8% vs 20%, P=0.0727).

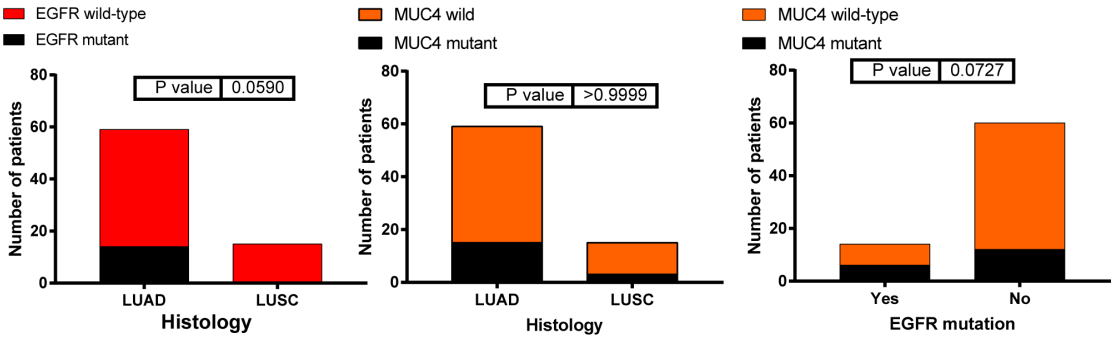

Figure 1. The relationship between EGFR mutation, MUC4 mutation and histology.

Change in text: We have modified our text as advised (see Page 14, line 293-297)

Reviewer E:

In this study, the authors performed whole exome sequencing (for some tumors, targeted sequence was performed instead) and immune marker staining (CD3, CD8, PD-L1) to identify prognostic markers in surgically resected stage I NSCLC patients. The authors suggested a recurrence-model by integrating TMS (tumor mutation score) and immune score. TCGA data was also employed to validate their results (some results were inconsistent between the authors' data and TCGA data). In general, the reviewer thinks the study was well performed.

Major comments

Comment 1. Some patients received target sequencing but not whole exome sequencing. However, it is not clear if these analyses (TMB analysis, TMS analysis, etc) were based on data from WES only or based on the combined data of WES and target sequencing.

Reply 1: We thank the reviewer for this comment. Since target sequencing did not include the data of MUC4 and many other genes, EGFR and TP53 mutation were analyzed based on combined data of WES and target sequencing, while TMB, TMS were analyzed based on WES data only.

Change in text: We have modified our text as advised (see Page 16, line 346-347; Page 15, line 317)

Comment 2. The reviewer thinks surgical procedure (lobectomy or limited resection) is also important in Table 1 and later analyses.

Reply 2: We thank the reviewer for this comment. There were two patients received sublobar resection and 128 patients received lobectomy resection. The surgical procedure had no correlation with early recurrence in our research (Table 1).

Change in text: We have modified our text as advised (Table 1)

Comment 3. Some tumors were resected in almost 10 years ago (study period: 2011 - 2020 May). It is generally believed that older samples (usually over 3 years) may cause false positive or false negative results at both of NGS and immunohistochemistry. How the authors overcame this issue?

Reply 2: We thank the reviewer for this comment. As a retrospective analysis, we are not able to avoid the influence of time on tumor tissues. Thus we performed WES in more 100 sample, while only 74 samples had qualified WES data. However, most of tissues were qualified for immunohistochemistry as the tissues were well protected in our department of pathology.

Change in text: No changes in text.

Minor comments

Comment 1. Please describe the mutation rate of MUC4 in TCGA.

Reply 1: We thank the reviewer for this comment. The mutation rate of MUC4 in TCGA is 5%.

Change in text: We have modified our text as advised (see Page 13, line 284-285)

Comment 2. Please describe the background of patients who had MUC4 mutation. Smokers? Adenocarcinomas?

Reply 2: We thank the reviewer for this comment. MUC4 mutation had no association with smoking history and histologic subtype, while it was significantly higher in stage IB patients.

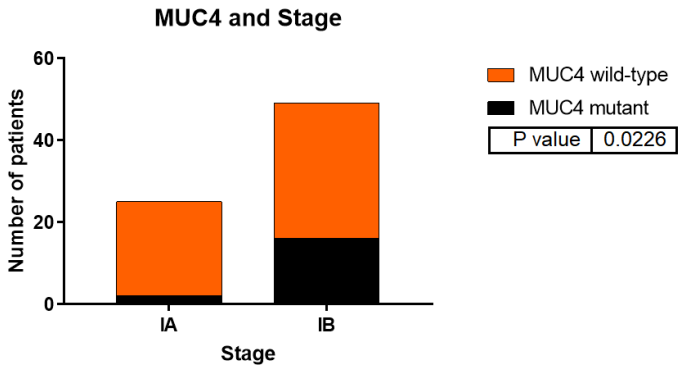

Figure. MUC4 mutation was significantly higher in patients with stage IB (*P*=0.026)

Change in text: No change in text.