**Reviewer A**

At first, it was well-written and well-constructed. The figures are clear and provide a clear view on the most relevant outcomes of the study.

The study has obvious clinical relevance. Although the pneumonectomy is utilized in a small number of patients with NSCLC, it has additional risks for patients when compared to other anatomical resections. It is also relevant to provide patients with adequate prognostic information. The use of an online prediction tool is valuable for physicians and surgeons to properly inform their patients.

The study has adequate methodology, where a SEER database is used to obtain a cohort with an adequate number of patients to perform the study. The number of missing values is small. The evaluation and verification of the model were carried out with the data of the training and verification set. The AUC is adequate for short- and long-term survival.

The main limitation of the model is that it mainly includes tumor characteristics. As the author correctly state, cardiopulmonary comorbidities which might influence long term survival. On the other hand, without a doubt the presented nomogram is the best there is for pneumonectomy and adequate AUC are presented.

In the discussion, I miss a part about the future perspectives of the pneumonectomy in the treatment of NSCLC. Because of expected stage and age shifts due to screening programs for NSCLC, what opinion do the authors have regarding the role of the pneumonectomy? It seems it is decreasing.

To conclude, thanks for this work and I will use nomogram model calculation in the future once available.

Reply: Thanks so much for your comments. The authors much agree with your opinion that the proportion of pneumonectomy is decreasing in recent years which was also concluded by analyzing the data in the SEER database year by year (2010-2015). We believe that the reason was related not only to more and more NSCLC patients who were diagnosed at an early stage and age by screening programs but also to the widespread use of sleeve lobectomy and pulmonary artery reconstruction as an alternative treatment. But in the future, pneumonectomy will still be an inevitable surgery approach to treat NSCLC, even if it might be less performed. Considering that the prognosis and quality of life for patients undergoing pneumonectomy are significantly inferior to that of patients undergoing lobectomy or sleeve lobectomy, the specific study for this small but special group of patients is of great clinical significance.

Changes in the text: We have added some content about the future perspectives of the pneumonectomy as advised (Marked version without title page: Line 185-186).

**Reviewer B**

Comment 1: - Introduction, Line 33: the indication "of " pneumonectomy

Reply 1: Thanks so much for your advice. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without author title page: Line 35).

Comment 2: - Introduction, Line 34: ... and the "predicted" postoperative survival...

Reply 2: Thanks so much for your advice. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 37).

Comment 3: - Introduction, Line 35 - 36: You say the aggressive surgical approach makes it rare to study the prognostic factors in large scale. Please explain why.

Reply 3: Thanks so much for your correction and comments. The authors also believed that pneumonectomy is recently being less performed in daily clinical practice. We have reviewed the literature on pneumonectomy. So far, the number of prospective studies was very limited, and the sample size was also very small (1,2). Therefore, with the development of minimally invasive surgery technology and the wide application of sleeve lobectomy for centrally located cancer, some scholars hold the view that a prospective prognosis study with a larger sample size may be challenging (3).

Changes in the text: The previous expression might not be very clear, so we have modified our text as advised (Marked version without title page: Line 38-40).

Comment 4: - Discussion, Line 221: access "to " more

Reply 4: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 238).

Comment 5: As you were not able to use the 8th Edition of TNM staging criteria, Charlson comorbidity index and pulmonary function due to the lack of data in the SEER database, are you planning on recalculating the nomogram model with this data?

Reply 5: Thanks so much for your comments. We did take all the relevant variables from the SEER database into the comprehensive analyses and obtained the nomogram

predictive model by the significant independent prognosis indicators. It is true that the data of the 8th staging criteria, the Charlson complication index, and pulmonary function are not available under the current situation. Considering the limited sample size of most single-center studies at present, we are also making other research protocols, plan to organize multi-center cooperation, developing a prospectively designed database with more variables and closer follow-up, so as to further update this predictive model.

Changes in the text: We added a brief explanation on this concern. (Marked version without title page: Line 237-240). If needed, we will explain more in the next version of manuscript.


Comment 6: How do you think the use of the nomogram will help physicians to improve the outcome of their patients?

Reply 6: Thanks so much for your comments. The prognosis predictive model developed in this manuscript is mainly used to evaluate the long-term survival of patients undergoing pneumonectomy. Good results are obtained in both the training and verification cohort and can be used to more specifically evaluate patients after pneumonectomy. We reviewed a number of nomogram-related literature, most of which focused on the evaluation of the prognosis of specific diseases and did not study the measures to improve the prognosis (4,5). To a certain extent, the independent prognostic risk factors involved in this study may be used as a reminder in clinical practice, try to avoid or pay more attention to these risk factors, and give closer follow-up and supportive care to the high-risk patients, so as to improve the long-term survival.

Changes in the text: We added a brief explanation on this concern. (Marked version without title page: Line 42-46). If needed, we will explain more in the next version of manuscript.

**Reviewer C**

The authors presented an interesting paper. Doctors sometimes consider pneumonectomy as controversial operation, but it is still an effective method of treatment for some patients with locally advanced lung cancer.

The results of the author's univariate and multivariate analysis are not novel. Features such as gender, age, pathology, tumor size, N stage, chemotherapy, and radiotherapy are well-known predictors for post-pneumonectomy overall survival that have been sufficiently broadly described in the literature.

However, creating nomograms and risk classification system in the context of pneumonectomy is worth paying attention. Although there are already available papers describing this topic in the literature, the collected group of over 2000 pneumonectomies is impressive. Also, the good agreement between calibration plots and the predictive model is worth highlighting. Authors have proposed a method of determining overall survival, which is willingly used in modern oncology.

Please allow some comments meant to improve the quality of the paper.

The paper needs some language editing.

Comment 1: - Line 33- "The indication pneumonectomy should be cautiously weighed against the loss of additional lung function" - the meaning of the sentence seems understandable. However, I think it should be written more simply because it is confusing to the reader, for example, "The indications for pneumonectomy should be carefully considered because this resection is connected with the extensive removal of lung tissue".

Reply 1: Thanks so much for your comments. We have corrected this inappropriate expression.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 34-36).

Comment 2: -Line 34 - prediction of postoperative.

Reply 2: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 37).

Comment 3: -Line 36- "on a large scale" instead of "in large scale",

Reply 3: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 41).

Comment 4: -Line 40- of the disease,

Reply 4: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 46).

Comment 5: -Line 199- with the asymmetric position

Reply 5: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 218).

Comment 6: -Line 36-38- In recent years, nomograms are widely used tools for predicting prognosis which generate individual quantified probabilities of clinical events by integrating prognostic clinicopathological variables. - this sentence is too long. I suggest dividing it into 2 sentences to improve the clarity of the argument.

Reply 6: Thanks so much for your comments. We have corrected this inappropriate expression.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 42-44).

Comment 7: -Line 144 - unnecessary "in",

Reply 7: Thanks so much for your comments. We have corrected this mistake.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 157).

Comment 8: -Line 157-161 - is this analysis necessary in the discussion here? In my opinion, this was not a significant problem for this study. Either I would shorten it or delete this paragraph.

Reply 8: Thanks so much for your comments. We have deleted this paragraph.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 171-174).

Comment 9: Patients characteristics and Table 1- the Table 1 summarizes all the most essential features. I believe that there is no point in duplicating some of the information in the Materials and Methods section.

Reply 9: Thanks so much for your comments. Indeed, there is a lot of information duplicated between these two parts. The information in Table 1 is more comprehensive, and all the variables in Table 1 were taken into univariate Cox analysis (Table 2). So, we deleted and modified some repetitive information in the Patients and Methods section.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 124-132).


Comment 10: What does extended pneumonectomy mean in your paper? Resection of the superior vena cava, chest wall resection, tracheal carina, left atrium, thoracic resection, diaphragm resection etc. - it should be explained in the text.

Reply 10: Thanks so much for your comments. According to the SEER variables dictionary, radical pneumonectomy was defined as pneumonectomy with mediastinal lymph node dissection, and extended pneumonectomy was radical pneumonectomy with the dissection of surrounding structures such as the diaphragm, pleura, and chest wall.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 78-81).


Comment 11: - Figure 1A - This flow diagram is very appreciated, but in my opinion, the criteria for inclusion in the study are simple and can be summarized in one sentence in Materials and Methods. This figure is unnecessary.

Reply 11: Thanks so much for your comments. The authors summarize the contents of the flow chart in part of Patients and Methods, and indicate the number of patients included and excluded of the study cohort in part of Results.

Changes in the text: We have deleted the figure and modified our text as advised (Marked version without title page: Line 120-124).


Comment 12: - Figure 1B - in my opinion, this figure is not needed. It is well known that the number of pneumonectomies worldwide is decreasing at the expense of other minimally invasive methods. Moreover, it is not the subject of this paper.

Reply 12: Thanks so much for your comments. The authors also hold the view that the number of pneumonectomies worldwide is decreasing. We have reviewed relevant literature, but did not find direct data to support this opinion, besides the latest data on the proportion of pneumonectomy is about 8-15% (6,7), which seems to be much higher than clinical practice. Therefore, we hope to further demonstrate the above

point of view through this nationwide database analysis. The study found that the proportion of pneumonectomy gradually decreased from 2011 to 2015, while the average proportion was only 3.9%, which is much lower than the data reported in the literature.

Changes in the text: We have deleted the figure and modified our text as advised (Marked version without title page: Line 120-124).

**Reviewer D**


Comment 1: Though this is a retrospective study, it seems unusual that the training set data (2011-2015) occurred chronologically after the validation cohort (2010). Can the authors comment on the rationale for this and provide other examples / literature to support this approach?

Reply 1: Thanks so much for your comments and advice. Initially, we randomly assigned the whole group into the training and verification set at 7:3 ratio, but we modified the data allocation approach according to the Statement of Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) (8). According to the TRIPOD Statement, the random allocation can be only regarded as internal verification (Type 2a), while the data nonrandomly being split (e.g., by location or time) into 2 groups was defined as Type 2b. Type 2b is a stronger design for evaluating model performance than type 2a, because allows for nonrandom variation between the 2 data sets. Type 2b is arguably referred to as "external validation studies". We also found that the checklist of TRIPOD Statement is recommended as a standardized format file for clinical predictive model articles in the *Journal of Thoracic Disease*.
Changes in the text: We have added a relevant explanation on this concern. (Marked version without title page: Line 88-89).


Comment 2: Lines 28 of the Introduction suggests a pneumonectomy rate far higher than most would expect to see in everyday practice. This is likely because the reference cited is from 2002. I suggest revising these estimates of pneumonectomy rates using more recent multi-institutional data.

Reply 2: Thanks so much for your comments. We have reviewed the literature and found the latest data was that pneumonectomy accounted for about less than 15% of all lung cancer resections (7,9). Meanwhile, one purposes of this study was to summarize the latest SEER database to provide more accurate data to show the current proportion of pneumonectomy in lung cancer surgery. During the period of this study, the proportion of pneumonectomy was only 3.9% among all the lung cancer resections with a reduction tendency in 2011-2015, which was lower than the data in the literature. This study also analyzed the changing trend of the pneumonectomy proportion year by year, and demonstrated the view from the real-world data that the proportion of pneumonectomy was gradually decreasing with which most of the scholars were agree.

Changes in the text: We have updated the data and references. (Marked version without title page: Line 29).


Comment 3: Additionally, as sleeve lobectomy and pulmonary arterioplasty are increasingly performed with similar oncologic benefits as pneumonectomy, it

warrants mention in the Introduction that alternative treatment options for central tumors are commonly done (Line 31).

Reply 3: Thanks so much for your comments. A number of studies recently have demonstrated that sleeve lobectomy and pulmonary arterioplasty are valuable and oncologically safe treatments for NSCLC patients, with lower short-term mortality and morbidity, without affecting long-term oncological results (10-12). These surgery techniques are more and more widely used in clinical practice and become favorable and effective alternatives to pneumonectomy.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 27-29).

Comment 4: References are needed in the Study Population of the Methods section to support the percentage of the US population included within SEER.

Reply 4: Thanks so much for your comments. We have inserted the data reference into the manuscript. The data was from the official website of the SEER database.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 58).

Comment 5: The methodology as described is concerning for a guaranteed time bias in how overall survival was calculated. Defining OS from the time of diagnosis until the time of death then selecting only patients who underwent pneumonectomy essentially guarantees some variable amount of survival time from the period from when the patient was diagnosed until when they actually underwent surgery. If the authors goals are to create a predictive model of survival AFTER pneumonectomy, then it would appear to me that survival time should only be counted from the date of surgery onward.

Reply 5: Thanks so much for your comments. We highly agree with your opinion and carefully assess the impact of this time bias. The authors carefully checked the SEER Variables Dictionary, and the definition of OS is indeed from the diagnosis to death. Considering the current situation, we very regret that we are unable to get detailed information about the specific time between diagnosis and surgery for each patient to make the survival time more accurate. But as far as this study is concerned, all the patients involved were pathologically diagnosed with lung cancer after surgery under microscope, and postoperative pathological examination is an important criterion for diagnosis. According to the pathological diagnosis techniques after 2010, the time from operation to microscopic diagnosis may only take about 7 or 10 working days, and we believe that the impact of this time on the long-term survival of the NSCLC is limited. We reviewed some literatures on long-term survival after surgical treatment of NSCLC based on SEER data. The definition of OS in these studies was also from the time of diagnosis (4,5). We also discuss the natural bias of this retrospective database study in the limitation part.

Changes in the text: We discuss the bias in the limitations section (Marked version without title page: Line 236-237).

Comment 6: Terms such as radical pneumonectomy and extended pneumonectomy require clarification.

Reply 6: Thanks so much for your comments. According to the SEER variables dictionary, radical pneumonectomy was defined as pneumonectomy with mediastinal lymph node dissection and extended pneumonectomy was radical pneumonectomy with the dissection of surrounding structures such as the diaphragm, pleura, and chest wall.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 78-81).

Comment 7: Table 1 lists "differentiation" though I believe the authors mean Stage. Additionally, stage III/IV is combined though I thought their criteria for inclusion eliminated stage IV patients. Please clarify.

Reply 7: Thanks so much for your comments. The variable of pathological grading in SEER database was included in the analysis. Differentiation here means pathological grading. This word may indeed be confusing and we modified it. Patients with stage IV have been excluded from screening.

Changes in the text: We have modified our text as advised (Marked version without title page: Line 76, 141).

Comment 8: I found the description of the Risk Classification System in the Results. It is difficult to discern the importance or meaning of comparing high and low risk patients within different stages all of whom underwent pneumonectomy is challenging to interpret or draw meaning from. Clearly the authors aren't advocating for revising staging based on this risk classification system so I'm not sure the relevance of explaining how "high risk" lower staged patients fared similarly (or worse) than "low risk" higher staged patients.

Reply 8: Thanks so much for your comments. Indeed, it seems that the significance of Risk Classification System in the prognosis evaluation is not as important as Stage, but we think the meaning of risk group can be listed as follows.

1. One purpose of the Risk Classification System is to confirm the effectiveness of the nomogram model. There is a significant difference in prognosis between the high- and low-risk groups in the training, verification, and total cohort.

2. Risk classification system can be used as an auxiliary tool for the staging system. We truly believe that the staging criteria currently used are more effective, but the risk group system involved in this study can be further evaluation tool for stage system of

patients. For patients undergoing pneumonectomy at the same stage, there are significant differences in OS among low- and high-risk groups. Physicians can identify patients with different risk levels by this risk classification system, especially high-risk group patients, and give these patients closer follow-up and medical support in clinical practice, which may be conducive to the long-term survival of patients.

3. Risk classification system can be used as a simplified application of the nomogram predictive model, through the risk grouping tool to give a general evaluation of patient scores and enhance clinical practicability.

Changes in the text: We added a brief explanation on this concern. (Marked version without title page: Line 26-213). If needed, we will explain more in the next version of manuscript.

Comment 9: Ultimately, I think the challenge with this study is that as the authors point out, pneumonectomy perioperative mortality rates have decreased considerably in recent years and that the long-term survival of these patients is very much related to tumor biology, such as stage, nodal metastasis, etc. I don't have a clear understanding from this analysis as to why pneumonectomy in particular should be viewed differently than sleeve or bilobectomy with respect to long-term survival. Moreover, the c-statistic is quite modest in terms of prediction accuracy.

Reply 9: Thanks so much for your comments. As you mentioned, it is true that this approach of thoracic surgery is less performed recently, but it is still an inevitable treatment. In particular, many studies have taken pneumonectomy as an independent risk factor for the prognosis of NSCLC patients after thoracic surgery (13). Therefore, although the proportion and perioperative mortality of pneumonectomy has decreased, it is still a kind of operation that has a great impact on the long-term survival of patients and more attention should be paid to those patients undergoing pneumonectomy.

We reviewed the relevant literature and found that both short-term and long-term outcomes of patients undergoing pneumonectomy have significant disadvantages compared with sleeve lobectomy (14,15). A meta-analysis found a similar conclusion (11). Most studies have found that compared with sleeve or bilobectomy, pneumonectomy has higher perioperative mortality, long-term cardiopulmonary complications, and lower quality of life, which is also a high-risk factor for poor postoperative survival. So, some scholars believe that "pneumonectomy is a disease" (16).

We reviewed the relevant literature at present and there are no randomized controlled trials on pneumonectomy, some scholars also believe that it is difficult to conduct large-scale randomized clinical research for these patients undergoing pneumonectomy (3). There is only one retrospective nomogram study of prognostic survival analysis in patients undergoing pneumonectomy up to date. In that research, the sample size of discovery group was only 100, and its accuracy in model verification was quite moderate (17).

The purpose of our research is to study the prognostic factors of these patients through the retrospective analysis of the national database and to construct a nomogram and risk classification system, in order to give more specific evaluation to these patients in clinical practice. We extracted and comprehensively analyzed all the variables that can be obtained from the SEER database, and obtained several factors significantly related to prognosis. Considering the actual situation, we cannot obtain all the perioperative clinicopathological factors from the SEER database, but we believe that these identified prognostic factors are indeed related to long-term survival and also have been demonstrated in several other literatures. In addition, through the quantitative analysis of nomogram, the influence of prognostic factors on long-term prognosis in this special population has also been visualized.

It is true that the C-index of this study is quite modest, but we have reviewed some other nomogram studies, and there are many literatures with C-index or AUC between 0.65 and 0.70 (18-21). Compared with those studies, the nomogram in our research was developed by a rich training cohort and validated by an external verification group with large sample size, and the evaluation accuracy of 1, 3, and 5 years is more than 0.7, which is relatively acceptable. In addition, we look forward to more multicenter studies in the future to give a more detailed and comprehensive assessment and research on this special population.

Changes in the text: We added a brief explanation on this concern. (Marked version without title page: Line 243-248). If needed, we will explain more in the next version of manuscript.