Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-24-1419

Reviewer A

The authors spend a lot of effort in preparation of this manuscript and fulfilled the criteria of the TRIPOD Checklist.
Discussion is well structured and comprehensible.
Unfortunately, grammar and syntax should be improved in the whole manuscript. Preferably with corrections of a native-speaker.
There are several annotations:

- The authors mention “sensitive” treatments. What is the definition of “sensitive” in this case?
Reply: Thank you for your thoughtful question. In this case, by “sensitive” treatments, we are referring to the initial effectiveness of radiotherapy and chemotherapy in treating SCLC. SCLC typically shows a good initial response to these therapies, meaning the cancer cells are more susceptible to being controlled or killed by these treatments. However, we acknowledge that this responsiveness is often temporary, as resistance can develop over time. we hope this explanation clarifies the intended meaning.

- There are several words and whole sentences which are not comprehensible in almost all sections of the manuscript.

- Do the authors use cTNM or pTNM for the calculation of their model?
- Reply: Thank you for your question. According to the Variable & Recode Definitions section of the SEER database, the T, N, M, and Stage data were collected using the Collaborative Stage (CS) system and processed with the AJCC 8th edition CS algorithm. These data may reflect either clinical or pathologic staging, as the CS system uses a "best stage" approach, combining the most informative elements from both clinical and pathologic sources. Consequently, the TNM staging data we extracted do not explicitly differentiate between clinical and pathologic staging but instead integrate information from both. We apologize for not making this clearer, and appreciate your understanding of the limitations this integration may pose for the predictive accuracy of our model. Thank you for your insightful critique and suggestions.

- Is it possible to use the Karnofsky or ECOG Index for the propensity score matching as well?
- Reply: Thank you for your thoughtful question. The SEER database does not typically include direct data on Karnofsky Performance Status or ECOG Performance Status. SEER primarily contains clinical, pathological, treatment, and survival data but does not systematically collect performance status scores such as KPS or ECOG.

While these performance status scores are not available, we can indirectly reflect a patient's health status using variables such as age, stage, treatment, and survival status. For example, older patients with advanced cancer are often associated with lower KPS or ECOG scores, but these data do not directly represent performance status. Alternatively, by considering related treatment data or clinical factors (such as whether the patient underwent surgery, chemotherapy, or radiation therapy), along with information on the treatment plan and prognosis, we may be able to make reasonable assumptions about their KPS or ECOG scores. However, these methods can only provide approximations.

If a more detailed and accurate measurement of these indices is needed for propensity score matching, it would be necessary to access other specialized databases that collect this data or obtain it from hospital records (such as clinical trial data). We appreciate your suggestion and hope to incorporate external validation and include these indices in our future research.

- Are the authors able to present a p-value for the differences in OS and CSS?
- Reply: We compared the differences in OS and CSS within the surgery and non-surgery groups after propensity score matching (PSM). The results showed that in the non-surgery group, OS was significantly lower than CSS (P = 6e-04), suggesting that non-surgical patients might have lower OS due to comorbidities or non-cancer-related causes. This indicates that other health factors, apart from the cancer itself, significantly impact the survival time of non-surgical patients. In the surger y group, despite receiving surgical treatment, OS was still significantly lower than CSS (P = 9e-05), implying that postoperative complications, the recovery process, or other non-cancer-related factors might affect survival. This highlights the need for special attention to postoperative care and long-term health management to reduce non-cancer-related mortality.

- The authors explain lymph node metastases where not part of the model, but lymph node status was. The lymph node status displays the occurrence of lymph node metastases, doesn’t it?
- Reply: Thank you for your thoughtful question. Yes, lymph node status does indeed indicate the presence or absence of lymph node metastases. However, when we mentioned that lymph node metastases were not included in our model, we were referring to the lack of more detailed information, such as the lymph node metastasis ratio. While our model included lymph node status as a general indicator of metastasis, it did not account for specific or more quantifiable details of lymph node involvement, which can be challenging to assess accurately before surgery. We apologize for any confusion and appreciate your understanding of this distinction.

In conclusion: The authors cover an interesting topic with good results and your discussion is well thought out, but you must review the language and the content.

Reviewer B

The authors reported their work named “The screening of optimal primary tumor resection candidates in patients with small cell lung cancer: a population-based predictive model” and concluded “We built a preoperative predictive model for small cell lung cancer patients to screen optimal surgery candidates. This model has the potential to help clinicians determine whether it is beneficial to operate on patients with SCLC.”. I have the following comments:

- Abstract: Please rephrase this sentence and specify the chosen cutoff time “Among patients undergoing surgery, we assumed that patients who had longer median cancer specific survival (CSS) time than non-surgical patients could benefit from surgery. “.
- Reply: Thank you for your insightful comment. To clarify, we set the cutoff time for the analysis between 2014 and 2018. Based on this time frame, we hypothesized that patients undergoing surgery during this period might benefit from the procedure if their median cancer-specific survival (CSS) time was longer than that of non-surgical patients. we have modified our text as advised (see Page 1, lines 13-16).
- Changes in the text: We assumed that patients undergoing surgery between 2014 and 2018 would benefit from the procedure if their median cancer-specific survival (CSS) time was longer than that of non-surgical patients. (see Page 1, lines 13-16).

- Line 23: Please specify the compared groups in this line “median CSS time 37.00 vs 16.00 months”.
-Reply: Thank you for your comment. The compared groups are the surgery group vs the non-surgery group. We have made the revisions in the text to clarify this.
- Changes in the text: with a median CSS time of 37.00 months for the surgery group compared to 16.00 months for the non-surgery group (see Page 1, lines 23-24).

- Line 25: Please edit this line “which were used to establish the predictive model with a nice stable.”.
-Reply: Thank you for your comment. We have revised the sentence for clarity and accuracy. The edited sentence is: “which were used to establish a stable predictive model.”
- Changes in the text: which were used to establish a stable predictive model. (see page 1, Line 26)

- Please try to do a dynamic nomogram, that can be one using shiny package in r.
-Reply: Thank you for your constructive feedback. We have built a dynamic nomogram based on the nomogram model using the Shiny package in R. The dynamic nomogram can be accessed at the URL: http://127.0.0.1:7290. Additionally, we have uploaded the

figure of dynamic nomogram m for easier access and visualization. We greatly appreciate your review and hope this addition enhances our model's clarity and utility.

- Line 90: Please report median follow up time, that can be done using reversed Kaplan Meier method.
-Reply: Thank you for your comment. According to the reversed Kaplan-Meier method, the median follow-up time was calculated to be 9 months, and we have included this result in the manuscript. (see Page 4, Line 117)
- Changes in the text: with a median follow-up time of 9 months.

- Please add reference to “rms” package.
-Reply: Thank you for your suggestion. We have added the reference to the 'rms' package in the manuscript as requested. (see Page 4, Line 106)

- Please report standard mean difference (SMD) for the matched groups. It is better to add “love plot” that shows SMD before and after matching.
-Reply: Thank you for your valuable suggestions on our work. We have detailed the standard mean difference (SMD) for the matched groups in the uploaded Table S2. Additionally, we have provided the “love plot” showing SMD before and after matching as Figure S2 to visually illustrate the matching effect.
- Changes in the text: The standard mean difference (SMD) for the matched groups was presented in the Table S2. Additionally, the “love plot” showed the SMD before and after matching, visually illustrating the matching effect (Figure S2). (see Page 5, Lines 128-130)

- Please specify if you included “Sequence number 0 or 1” only i.e.: those with only one cancer type during their lifetime.
-Reply: Thank you for your comment. To address your concern, we have only included patients with a Sequence number 0 or 1 in the analysis, which corresponds to those who had only one cancer type during their lifetime. Specifically, Sequence number 0 refers to patients who were diagnosed with a single cancer type, and Sequence number 1 refers to those diagnosed with only one cancer type following a previous diagnosis, without any other cancer type being present. We excluded patients with multiple cancer types in our analysis to maintain the focus on those with a single cancer type.

- Table 1 footnote: Please try to mention details on “Surgery of other sites”. You can do subgroup analysis excluding those with other surgery.
- Reply: Thank you for your suggestion. By “Surgery of other sites,” we are referring to patients who have undergone surgeries for conditions other than lung cancer. We performed a subgroup analysis, and in the non-surgery group, there was no significant difference in OS and CSS between patients who had not undergone other surgeries and those who had, after PSM (Figure_subgroup.a and b). Similarly, the same results were

observed in the surgery group (Figure_subgroup.c and d). Detailed results are presented in the Figure_subgroup.

- Table 2: Please add headings specifying OS and CSS.
- Reply: Thank you for your valuable suggestion. We have added the headings 'Univariate/Multivariate Analysis of OS' and 'Univariate/Multivariate Analysis of CSS' to clearly distinguish the analyses, as requested. We hope these modifications improve the clarity of the table.

- Table 1 and 2: Please specify the percentages of missing values and how you managed that.
- Reply: We have managed missing values in our data as follows: (1) Variables with unclear or unknown TNM stages, survival times of 0 or unknown, and cases where age, race, or tumor position were unspecified were identified. These accounted for 55.2% of the overall data and were subsequently removed. (2) For the categorical variable 'Grade', missing values were treated as a separate category to ensure comprehensive analysis.

- Table S1: Please add p value and make OS and CSS bold in font.
- Reply: Thank you for your valuable suggestion. We have added the p-values in the Table S1. Additionally, we have made the headings for OS and CSS bold for better clarity and emphasis, as per your recommendation.

- Figure 3: Please provide the used R code.

- Reply: Here is the R code：

```
library(data.table)
mydata=as.data.frame(fread("Train.csv"))
str(mydata)
mydata$Marital= as.factor(mydata$Marital)
mydata$Agegroup=as.factor(mydata$Agegroup)
mydata$Race= as.factor(mydata$Race)
mydata$Sex=as.factor(mydata$Sex)
mydata$Year=as.factor(mydata$Year)
mydata$Site=as.factor(mydata$Site)
mydata$Grade= as.factor(mydata$Grade)
mydata$Histology= as.factor(mydata$Histology)
mydata$TNM=as.factor(mydata$TNM)
mydata$Tstage=as.factor(mydata$Tstage)
mydata$Nstage=as.factor(mydata$Nstage)
mydata$Mstage=as.factor(mydata$Mstage)
mydata$Age=as.numeric(mydata$Age)
mydata$Radiation=as.factor(mydata$Radiation)
mydata$Chemotherapy=as.factor(mydata$Chemotherapy)
```

```
mydata$Stage=as.factor(mydata$Stage)
mydata$Surgery= as.factor(mydata$Surgery)
mydata$SM= as.factor(mydata$SM)
mydata$Benefit=as.factor(mydata$Benefit)
mydata$Survival.months=as.numeric(mydata$Survival.months)
mydata$CSS=as.factor(mydata$CSS)

library(rms)
LR1<-lrm(Benefit ~ Sex+Site+Agegroup+Tstage+Nstage+Mstage, mydata,x=T,y=T)
dd<-datadist(mydata)
options(datadist="dd")
Nomo_LR<-nomogram(LR1,
                  fun = plogis,
                  lp=F,
                  fun.at = c(0,1,seq(0.2,0.9,by=0.1),0.9),
                  funlabel = "Probabilities of Benefit")
plot (Nomo_LR)
```

- Please add more details on this “The surgical patients were randomly divided into two sets at a ratio of 7:3, training and testing set.” And the benefit from this.
- Reply: Thank you for your comment. We have added more details regarding the division of surgical patients into the training and testing sets. Specifically, we used a 7:3 ratio to split the patients into two groups: 70% of the patients were allocated to the training set, and 30% to the testing set. This approach allows for the development of a robust predictive model in the training set, while the testing set serves to independently evaluate the model's performance and generalizability.
- Changes in the text: The surgical patients were randomly divided into two sets at a 7:3 ratio: 70% were allocated to the training set and 30% to the testing set. (See Page 4, Lines 96-98)

- Figure 4: Please expand the legend
- Reply: Thank you for your suggestion. We have revised the text as requested. The updated description now reads: The internal and external validation of this nomogram. The ROC curve in training set (a) and testing set (b); The calibration plots in training set (c) and testing set (d); The decision curve analysis curve of nomogram in training set (e) and testing set (f).

Additionally, we also noticed errors in the legend in the top-right corner of Figures 4e and 4f. The legend "LCSS" has been corrected to "Nomogram".