Peer Review File

Article information: https://dx.doi.org/10.21037/tlcr-24-213

Reviewer A

I congratulate the authors for this elegant study on a conflicting topic in clinical practice. I have some comments and suggestions and I thank the authors for reading and considering them.

Comment 1. As correctly discussed in the manuscript the SEER database is strongly biased due to the particularities of the health care system in the Country were the data are coming from. SEER data must be interpreted carefully as there are some important limitations which have been discussed in Park HS, et al. Limitations and biases of the Surveillance, Epidemiology, and End Results database. Curr Probl Cancer. 2012 Jul-Aug;36(4):216-24. doi: 10.1016/j.currproblcancer.2012.03.011. You could include this reference in your manuscript.

Reply1: thank you for your guidance. We have updated and added relevant reference as per your suggestion.

Changes in the text: page 16, line 524.

Comment 2. After psm, you are missing more than 75% of cases from the first cohort of T1N2-3 cohort. Thus, your conclusions and recommendations cannot be exported to the universe of T1N2-3 NSCLC cases and that should be clearly underlined in your report. Advantages and inconvenients of psm process have been analysed in many published papers and you should include at least one of them in your reference list, maybe Shiba K, Kawahara T. Using Propensity Scores for Causal Inference: Pitfalls and Tips. J Epidemiol. 2021 Aug 5;31(8):457-463. doi: 10.2188/jea.JE20210145.

Reply 2: Thank you for your review and your professional advice, which make our work more complete. We have made the corresponding modifications in accordance with your suggestion. And we have consulted the references you provided and incorporated them into our research.

Changes in the text: page 14, line 476-478; page 6, line 178.

Comment 3. Instead of psm you could adjust for confounders by iptw to avoid missing such a high number of cases.

Reply 3: Thank you for your professional suggestions, and we agree with the reviewer's opinion. However, the methods that adjusting for confounders by IPTW to expand the sample for constructing our models will rewrite our work, which is not feasible given the costs involved. And this work would not significantly support our argument because of the dependence of the SEER database. For this reason, we chose to use the PSM method. After that, we add some sentence in page 16, line 541-546; page 14, line 476-478 in response to this issue.

Changes in the text: No changes.

Comment 4. The usefulness and accuracy of nomograms must be proved in prospective series of cases not included in their construction. That should also be clarified in your manuscript before recommending its use in clinical practice.

Reply 4: Thanks for your precious suggestion and we agree with your opinion. We have clarified related contents in page 16, line 541-546. The concrete text was that to further validate our results and provide more robust clinical guidance before being recommended for clinical use, prospective large-sample RCTs with comprehensive data on clinicopathological variables, performance status, and detailed treatment regimens should be conducted. These trials would offer a more precise assessment of the outcomes observed in our study and enhance the reliability of clinical recommendations.
Changes in the text: Page 16, line 541-546.

Comment 5. You must be congratulated for making your nomograms accessible for free on the Internet. Nevertheless, your website should include a warning to the user clarifying that the accuracy of predictions have not been prospectively validated and that predictive models are constructed on a specific population and their predictive capacity on different populations has not been validated.
Reply 5: We thank the reviewer for pointing this out and We add the related sentences. The new sentence reads as follows: It is important to note that this tool may offer clinicians instructions for survival counseling and treatment strategy making conveniently, but we should apply it to clinical practice cautiously before its predictive capacity has been validated in prospective, large-sample RCTs.
Changes in the text: page 11, line 366-369.

Finally, we would like to thank the referee again for taking the time to review our manuscript.

Reviewer B

This paper tackles a difficult and complex subject: a subgroup of stage 3 patients, specifically cases where the primary tumor is rather small (T1) in spite of an advanced lymph node stage.

Firstly, the authors mention the heterogeneity of stage 3 cancers, and in this they are quite correct. I think it would have been useful to further probe what these differences are, and why the authors came to focus specifically on T1 cancers, since this is not immediately obvious.

Now, in order to address such a specific subcategory, one would need access to very precise, focused data; I am not convinced that the SEER database is able to provide such data, and so right from the beginning there seems to be a fundamental (maybe insurmountable?) methodological hurdle, that goes far beyond just the retrospective nature of the research. Essential information is missing, including the nature of N disease that is not accounted for (microscopic, bulky, single station, multistation), and the sequence of treatment modalities, that is also missing. This is critical, since it has been fairly well-documented that in N2 disease, the effectiveness of adjuvant (compared to neo-adjuvant) treatment is extremely limited. The authors present visually appealing survival graphs, but, in light of the above limitations, I think it is critical to make sure that they are not over (or mis-) interpreted.

I am not an expert in the techniques of nomogram development and the other statistical methods used by the authors, and therefore cannot comment on these specifically. However, I would say that there is rather a flood of data that are presented in the results section, which is quite overwhelming and makes this section extremely difficult to process. In my opinion including both N2 and N3 disease adds to the confusion. An interesting take-away from the paper is the possibility of performing sublobar resections in certain cases of advanced lymph node disease, specifically in cases of small primary tumours. (But then again, if a patient was fit enough for chemotherapy, radiotherapy, and then surgery, are we really saying that performing a lobectomy instead of a more limited resection makes any appreciable difference in overall morbidity?)

Perhaps it would be helpful if the authors laid out this problem of lobar vs sublobar resection from the very beginning of their paper, and focused the subsequent results and discussion on this issue. I think that having such a consistent thread throughout the paper would help ensure continuity and make the text much easier to follow.

And I do think this is certainly an interesting question: indeed why would we expect a difference (or not) between lobectomy and sublobar resection in this context? Why, conceptually, would addressing local parenchymal disease affect survival in cases of advanced nodal disease? In order to have more credence, these results would benefit from an attempt at an underlying theoretical rationale.

I did not quite understand if the authors compared surgery + chemotherapy to chemotherapy + radiotherapy? (since radiotherapy was used as a factor to match their experimental and control groups).

In conclusion, I think that a lot of work went into this paper; but I am not sure that, in its current form, it is able to answer any specific research question.

Reply: Thanks for reviewer's meticulous and professional suggestions.
Firstly, we focus on the subgroup of T1N2-3 NSCLC patients, not because of its inherent specificity compared to other patients with stage III lung cancer. But as health awareness among individuals improves and low-dose spiral CT screening gains popularity among long-term smokers, the likelihood of detecting smaller lung cancers has significantly increased in clinical practice. However, some patients present with lymph node metastasis at the time of diagnosis. Although there is existing previous literature focusing on the outcome of stage III lung cancer, there remains a controversial topic regarding optimal surgical approach for this special subset of patients. Thus, our study aimed to obtain a preliminary conclusion regarding survival advantages of different surgical approach in T1N2–3M0 NSCLC patients, utilizing the SEER database.

The SEER database did exist some weakness, and as the reviewer refers, the lack of relevant pivotal information such as lymph node status and the sequence of treatment modalities, which may make the results of our study lack credibility. However, there is no specific

research on this issue to date. Our study could provide preliminary results and visually represent the benefits of surgical intervention through the nomogram regarding OS and CSS. It was important that our nomogram models exhibited certain prognostic ability, reliability, and clinical applicability, and hence may offer clinicians instructions for survival counseling, treatment strategy making, and clinical trial design. Certainly, our results need to be further validated through conducting prospective large-sample RCTs with comprehensive data on clinicopathological variables, performance status, and detailed treatment regimens.

We apologize for any inconvenience caused by the large amount of data in the results section. Our study can be broadly divided into two parts. One part consisted of comparing the survival prognosis of the two surgical approaches, and subgroup analyses were performed to further compare the prognosis of the different surgical approaches. The other part was to construct a model for the visualization of survival benefit among different surgical approaches. Therefore, the overall amount of data is relatively large, for which we apologize again. For the N2 and N3 disease included in our manuscript, exited prospective study indicated that patients with N2 and N3 disease exhibited similar rates of mediastinal node sterilization, which suggests that patients with N3 disease may have tumor behavior aligned with N2 and could potentially benefit from multimodal therapy, including surgery. So we included the N3 disease into analysis and conducted subgroup analysis based on different N stage to further explore the optimal surgical approach in N3 stage. After that, we also find that there was no significant difference on the effects of survival prognosis between N2 stage and N3 stage in patients with T1 cancer (Table 4). Currently, for patient receiving chemotherapy or radiotherapy before surgery, the optimal surgical approach remains unclear, which need to be further explored in subsequent prospective clinical trial.

For special patients with limited cardiopulmonary function and could not receive the lobectomy, such as elder patients, whether the sublobectomy could provide survival benefit to T1N2-3 NSCLC patients was unclear. Our study found that sublobectomy confers a survival benefit when compared to the non-surgery group. Sublobectomy may still be a viable consideration for stage T1N2–3M0 NSCLC patients, particularly those of advanced age or with compromised cardiopulmonary function. And existing literature also support segmentectomy has been shown to be oncologically equivalent to lobectomy while offering the advantage of better preservation of pulmonary function (PMID: 27823756). The above-mentioned content was the reason why we compared the outcome of lobectomy and sub lobectomy.

We appreciate the reviewer’s insightful suggestion and agree that it would be useful. But we did not compare surgery + chemotherapy with chemotherapy + radiotherapy in our study. Because our study aimed to determine the optimal surgical approach in T1N2–3M0 NSCLC patients and conducted the subgroup analysis based on different treatment regimens. The comparison of surgery + chemotherapy with chemotherapy + radiotherapy was beyond the scope of our paper, which aims to show that the optimal treatment for patients with T1N2-3M0 NSCLC.

Finally, we would like to thank the referee again for taking the time to review our manuscript.

Changes in the text: No change.

Reviewer C

Review of the paper entitled: "Identifying optimal surgical approach among T1N2–3M0 non-small cell lung cancer patients: a population-based analysis". The authors, analyzing the SEER database, reported the overall survival (OS) and cancer specific survival (CSS) of patients with NSCLC in stage IIIA, t1N2-3. They analyzed the impact of surgery compared with other kind of treatment and also analyzed the outcomes of lobectomy in comparison with sublobectomy in this setting. The paper is complex, articulated, well written, but it suffers from several issues coming from the SEER database and in particular: the long study period and the absence of some clinical details (DFS, timing of treatment). Other issues:

Comment 1. in the paper is not completely specified if patients are classified as clinical or pathological T1N2-3 and for patients in cT1N2-3 what was the staging procedure performed (EBUS, mediastinoscopy)?
Reply1: Thanks for your precious suggestion. We collect the information of T stage, N stage, and M stage from the SEER database according to the SEER program Coding and Staging Manual 2018. The database does not specify that patients are classified as clinical or pathological staging concretely, for which we are so sorry. For the specific performed staging procedure, we were unable to obtain the relevant specific information from the SEER database. Thus, A prospective large-sample RCTs with comprehensive data on clinicopathological variables, performance status, and detailed treatment regimens should be conducted to further validate our results, as we mentioned in the limitation section of the article.
Changes in the text: page 16, line 541-546**.**

Comment 2. the disease-free survival and the details of recurrence (distant or local) are basilar to understand the behavior of an heterogeneous stage as stage III. Unfortunately this issues could not be solved.
Reply 2: Thanks for your precious and meticulous suggestion. The related detailed information about recurrence and the received treatment after the patients had recurrence could not obtained in the SEER database, for which we are very sorry. But we believe that our study was conducted using clinically relevant factors that are accessible in the SEER database and benefited from a large sample size, which could offer valuable insights for clinical practice. However, the related conclusion may be further verified through conducting prospective large-sample RCTs.
Changes in the text: No change.

Comment 3. another issue that could not be solved is the absence of detail regarding the timing of the treatments, for example neoadjuvant or adjuvant.

Reply 3: Thanks for your advice and we agreed with your opinion. the sequence of chemotherapy and radiotherapy in relation to surgery was lacking in the SEER database, which could potentially impact the reliability of our study and the performance of our predictive model. Due to the dependence on the SEER database, we were unable to include these parameters in our analysis. But our study could obtain a preliminary conclusion regarding survival advantages of different surgical approach in T1N2–3M0 NSCLC patients.
Changes in the text: No change.

Comment 4. the introduction should be shortened.
Reply 4: Thanks for your precious and professional suggestion. We have shortened the introduction.
Changes in the text: Page 3, line 78-102.

Comment 5. the comparison between surgery and no-surgery has several biases. In the clinical practice, only patients with low burden of N disease and who had a clinical response of neoadjuvant treatment could be operated on. This is a huge bias that affect the comparison between the two groups.
Reply 5: Thanks for your advice and we agreed with your opinion. However, the sequence of chemotherapy and radiotherapy in relation to surgery couldn't be obtained from the SEER database, so that we couldn't distinguish the patients who received neoadjuvant or adjuvant treatment. We're very sorry about this. But we believed our study could provide a preliminary conclusion through a large sample size and offer certain construction for informing survival discussions between patients and clinicians, guiding the design and monitoring of clinical trials, and facilitating the development of more personalized treatment strategies.
Changes in the text: No changes.