Peer Review File

Article information: https://dx.doi.org/10.21037/tlcr-22-752

Reviewer A

The analysis is well conducted and the manuscript is well and clearly written.

I have some questions/comments:

Comment 1: Chemotherapy/radiotherapy was adjuvant or neoadjuvant? Which percentage received neoadjuvant if any?

Reply 1: You have raised an important point here. However, in our study we focused on surgical treatment. Unfortunately, we do not have such information. In Poland after surgical treatment patients are referred to different oncology centers with their own databases. Reaching full data about such treatment would be very difficult. We did not include neoadjuvant treatment in our study.

Changes in the text: none.

Comment 2: Patient with single metastasis cM1 (brain for example) were excluded from the analysis? because cM status is not reported

Reply 2: Yes, that is correct, they were excluded from the analysis. We do not include such patients because of the different management - it would be too heterogeneous group.

Changes in the text: none.

Comment 3: You report a difference in the post operative complications rate...it would be helpful to specify the different grades of complications (according to Clavien Dindo for example) in order to have a meaningful comparison

Reply 3: As far as the severity of symptoms is concerned - we do not use the same grading in our base as in oncology. There is no rigid framework of specific complications, so we are not able to order complications according to a scale. However, we are of the opinion that the complication table is sufficiently comprehensive to correctly identify differences between centres.

Changes in the text: none.

Comment 4: You reported in the results section "In the non-ACA group, the mean number of days from surgery to discharge and the entire period of hospitalization was significantly higher than those in the ACA group (9.58 and 14.0 days vs. 7.94 and 12.8 days, respectively, $p<0.001$). In the table 9.58 is the SD and the mean is 9.26.

Reply 4: We are very sorry for this typo, which we did not catch. Of course you are right, instead of 9.58 it should be 9.26.

Changes in the text: In the aforementioned sentence, the data has been amended as shown in Table 3. The value of “9.58” has been replaced by “9.26”.

Reviewer B

Thanks to the authors for this interesting work comparing academic vs non-academic centers for lung cancer surgical treatment and its outcomes. The findings presented can help to understand different results from single-center studies in relation to health disparities.

My comments and suggestions are:

Comment 1: When reporting epidemiological information (75-77) use the 2020 GLOBOCAN study as it`s more recent than the current reference.

Reply 1: Thank you for paying attention to this aspect. Following your recommendation, we have used the latest data from the GLOBOCAN 2020 studies, thus amending the data in the introduction to our paper.

Changes in the text: We have replaced the sentence on the number of deaths worldwide from lung cancer with the following: "Each year, approximately 1.8 million people die from lung cancer accounting for 18% of all cancers worldwide." Reference 1 has been changed to: "https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf".

Comment 2: Please add the hypothesis tested by this study in the introduction.

Reply 2: Of course, according to your comment, we have added a hypothesis to the introduction.

Changes in the text: We have added a new sentence in the Introduction section: "Before starting the study and after analysing the available literature, we hypothesized that in academic centers there is better preoperative diagnosis, fewer complications and thus the 5-year survival achieved is higher than in non-academic centers."

Comment 3: When reporting population description statistics for continuous variables please use mean + SD just for data with normal distribution. For data without normal distribution use median + IQR. You don`t need to use both.

Reply 3: For such large sample size the normality tests have very high power and hence any small difference between distribution and the Normal distribution is meaningful and leads to the rejection of the null hypothesis. for example Shapiro-Wilk test gives p-value<0.001. That's why we use both description statistics.

Changes in the text: none.

Comment 4: Mann-Whytney test was used for comparing continuous variables. This is a non-parametric test used after checking data distribution. Was the distribution tested? How? Please describe in the methods.

Reply 4: Thank you for that remark. We decided on Mann-Whitney test is free-distribution test, and his power is more robust against gross errors than that of the t-tests, and their efficiency loss is quite small even in the rare case of not holding normality assumption [3]. This is also procedure advised by Frank Harrell
(see
https://stats.stackexchange.com/questions/19675/what-normality-assumptions-are-required-for-an-unpaired-t-test-and-when-are-the/19677#19677)

Changes in the text: We have added new sentences and one reference in the Methods

section.

Comment 5: Line 173: please expand on what FDR stands for.
Reply 5: FDR stands for False Discovery Rate.
Changes in the text: We have expanded the acronym FDR in the text.

Comment 6: Lines 173-174: please add references for R and its packages.
Reply 6: Thank you for this comment, of course references should be added to the above mentioned data.
Changes in the text: We have added 2 new references.

Comment 7: Line 216: Just reporting that survival was not different with a p-value is not sufficient. Since it was the result of a survival analysis please report HRs and its 95% confidence intervals.
Reply 7: Thank you for this comment, however, we do not fully understand it. We have not duplicated the information in Table 4 in the text. Please, have a look on Table 4 once again. Could you specify exactly what data you are referring to and we will certainly add it.
Changes in the text: none.

Comment 8: The discussion needs to be entirely reformated, the actual structure looks like a lot of information without a good flow for the reader. I suggest the following structure:

-1st paragraph: summarize the key questions and findings. Why is this an important topic, what were your main findings, and why are they relevant.
-2nd paragraph: Go over additional findings of importance. Focus on secondary outcomes, those from subgroups or sensitivity analysis.
-3rd-5th paragraph: Put your findings in context (it`s what you have done in your current discussion). But organize this.
-4th-5th paragraph: Describe whether your findings are in line with others. If not, explain why not.
-6th-7th paragraph: Limitations and strengths.
-Final paragraph: Conclusion with clinical, research, and policy implications from your study. Focus for future studies on this topic. Remaining questions.

Reply 8: We have changed the discussion section to make it more readable for researchers. We have tried to do as you recommended in your comment. We hope that the current form is better than the previous version. However, we realise that we have not been able to follow your advice on all issues. We have deliberately combined several points into one in order to describe the problem in one paragraph, to describe the available literature, and finally to describe what results we have come up with and how they differ from the others, which is why paragraphs 3-5 as you indicated are jumbled up, but in our opinion this is advisable so that the reader can know exactly the whole context of the problem point by point.
Changes in the text: The Discussion section has been changed.