Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-22-1077

Reviewer A

Comment
This is a retrospective cohort study utilizing the NCDB to investigate treatment patterns and outcomes in patients with Pancoast tumors. The authors conclude that neoadjuvant treatment is superior to upfront surgery – a finding that is already well documented. They conclude that a multimodality regimen is associated with improved outcomes in the adjuvant setting – again a finding that is already well documented through prospective data. The strengths include a relatively large sample size, a credible national data source, multivariable adjusted analyses (although no tables were provided documenting these results), and a pertinent surgical question. The weaknesses include the retrospective design and an inability to accurately identify patients with Pancoast tumors. My comments:

Why were patients diagnosed in 2018 and 2019 not included? More recent data would add value to the study.

Reply
Thank you for those very helpful comments. The most recent iteration of the NCDB at the time of this study went to 2017. That was the most updated database available when we submitted our query.

Comment
Please include the results tables for the multivariable logistic and Cox regression analyses. Also please include the variables included in these models under 'statistical analysis' in the methods section.

Reply
We agree that including the multivariable logistic results for the table will help. Since it is a very large table, we have included it as a supplemental table. We have also referenced it in the manuscript. In the *Demographics* subsection of the Results, we have added to the text "Supplemental Table 1 shows the multivariate logistic regression analysis of these results."

Comment
You may want to consider propensity score matching to more effectively compare the two study groups.

Reply
Thank you for this great suggestion. Although we did consider propensity analysis during our initial statistical work, we did not feel that we had enough patients to perform propensity analysis

adequately, given the number of covariates we would have had to include. This problem highlights the fact that Pancoast tumors are not as prevalent as other disease processes where propensity matching could be performed more easily.

Comment
The study period spans 13 years with significant advancements in neoadjuvant regimens and surgical processes from 2004 to 2017. This heterogeneity is largely unaccounted for and should be acknowledged in the limitations section. Please include a limitations section.

Reply
I completely agree with this comment and it is a great point. We have included this idea in the limitations section. In the fourth paragraph of the Discussion section we have now added "During this time, there were significant advancements in neoadjuvant regimens and surgical conduct. We feel that this study may prompt future studies to examine whether practice patterns have changed in the last few years as the database is updated."

Comment
Please provide some detail with regard to how volume was calculated for each facility. Were total cases per facility divided by number of years (13)?

Reply
We did indeed divide the total number of cases by that institution by the total number of years. Since the study went from the beginning of 2004 to the end of 2017, we divided by 14 years. In the *Variables* subsection of the Methods section of the manuscript, we have added "To calculate yearly volume, we divided the total number of cases at each institution by the number of years of the study."

Comment
The discussion section requires more references.

Reply
We have significantly enhanced and increased the reference list, both in the discussion and throughout the manuscript. In addition, we have added more recent references and have looked for references within the last year. There is a paucity of literature, however. As such, we feel that this analysis will be a welcome addition to the literature and will be useful to readers.

Comment
25% of patients receiving neoadjuvant treatment is intriguing. You could check to see if this percentage has changed with time and add some discussion perhaps.

Reply
The NCDB had data available only until 2017. We agree that future studies should investigate whether the practice patterns have changed in recent years. In the Conclusion section of the

manuscript we have added “It will be beneficial to see whether neoadjuvant treatment for Pancoast tumors has increased in recent years.”

Reviewer B

Comment
Good paper. The findings of the authors are interesting and useful for the readers.

Reply
Thank you very much for your kind comments. Yes, we do feel that the results will be of interest to a broad group of readers.

Reviewer C

Comment
The authors report a patterns-of-care and outcomes analysis for patients with node-negative Pancoast tumors using the National Cancer Database Data. They found there was a remarkably low rate of utilization of standard-of-care neoadjuvant chemoradiation and, further, that this led to potential survival detriment.

The statistical analysis, in its current state, lacks the depth the support the conclusions. For example, if the goal was to assess patterns of care, there should be univariate and multivariate regression analyses reported. Both the methods section and abstract indicate that there was a logistic regression analysis performed. This is the most important data to present, but this is missing.

Reply
We apologize for the miscommunication. Multivariate regression analyses were used throughout the manuscript as you have correctly stated. We agree that without multivariate analyses the conclusions would not be useful. In the original version, we said that “multivariate Cox proportional hazard models were used to estimate hazard ratios adjusted for demographics and facility characteristics.” But to clarify further, we have added a supplemental table with the entire multivariate analyses, odds ratios and 95 percent confidence intervals. We did not include this in the original version because it is a very busy table. But we now understand that readers will not realize that multivariate regression analyses were performed without the supplemental table. We believe that addition of this table and clarification in the methods will resolve this miscommunication. Thank you for bringing this to our attention. In the *Demographics* subsection of the Results section, we have added “Supplemental Table 1 shows the multivariate

logistic regression analyses of these results. In this supplemental table we have included the results of the multivariate analyses with corresponding 95 percent confidence intervals."

Comment
Similarly, in the absence of multivariate logistic regression, this manuscript does not explore how variables such as insurance status, facility type, and Charlson-Deyo Score could actually impact treatment decisions.

Reply
Thank you for pointing out this miscommunication again. As mentioned above, we believe that the supplemental table and addition to the methods section shows the multivariate logistic regression analyses that we performed throughout the study.

Comment
In addition, making conclusions about survival endpoints from NCDB data without a multivariate Cox regression analysis is strife with confounders that limit one's ability to draw a conclusion from the data. This should be included when discussing NCDB survival data.

Reply
We agree with this statement and again apologize for the miscommunication. In the original version in the *Statistical methods* subsection of the Methods section, we said that "Multivariate Cox proportional hazard models were used to estimate hazard ratios adjusted for demographics and facility characteristics."

Comment
One variable in particular, the year of diagnosis, is grouped into two disproportionate groups 2013 and prior and then 2014-2017, but it is unclear why this was done. In a patterns-of-care analysis such as this one, I'd anticipate more exploration of the receipt of neoadjuvant CRT over time. I'd recommend regrouping this into either more logical or equally distributed groups for the above multivariate analysis.

Reply
We agree with your statement about this arbitrary division. We have removed this time grouping. In addition, there was no significant change in the rate of receipt of neoadjuvant treatment over the time period of the study. But we have added in the second to last paragraph of the Discussion section "We feel that this study may prompt future studies to examine whether practice patterns have changed in the last few years as the database is updated."

Comment
For the survival analysis performed in Figure 3, I believe the upfront surgery group would be constituted by any patient, regardless of adjuvant treatment. However, generally by virtue of their disease stage, these patients would have indications for adjuvant therapy or therapies (in particular chemotherapy). I would recommend performing a sensitivity analysis making

comparisons between patients treated with neoadjuvant CRT vs. adjuvant chemo or chemoradiation.

Reply
We appreciate this comment. We felt that it was a very understandable division between neoadjuvant chemoradiation vs. no treatment. It becomes a bit more difficult to discuss neoadjuvant chemoradiation vs surgery + some treatment. With that comparison we would not be able to account for the implications of neoadjuvant treatment on surgery itself and would not be able to control for some of those variables. In addition, it is likely that the amount of treatment that patients would get in the adjuvant setting would be more likely to be incomplete compared to the neoadjuvant setting as patients are recovering from surgery. Once again, we would not be able to control for that given the limitations of the NCDB. As such, we did not perform that sensitivity analysis although we think it is a great suggestion.

Comment
The authors state that the primary goal was to identify treatment patterns and outcomes in patients with Pancoast tumors. In doing so, the text should clearly state what the primary (presumably receipt of neoadjuvant chemoradiation) and secondary endpoints are (presumably OS).

Reply
Thank you for this comment. It will help readers to clarify our endpoints. In the *Variables* subsection of the Methods section we have added "The primary endpoint was the receipt of neoadjuvant chemoradiation. The secondary endpoints were overall survival in both the neoadjuvant and upfront surgery groups.

Comment
T stages should be reported.

Reply
We agree with this comment and we have added the T stages to the manuscript.

Comment
It is inappropriate to present great circle distance as a mean. It should instead be reported as at least a median or by quartiles. I am surprised that this would be unassociated with receipt of SOC.

Reply
We have removed the mean value and left the median values, with first and third quartiles.

Comment
As it relates to the reporting of race, the authors should be defining what is constituted by the "other" category and offer explanation (presumably small N) in a footnote.

Reply

In the *Statistical methods* subsection of the Methods section we have added "Race was reported as Caucasian, African American and "other" since the "other" category was only 2.9 percent of the cohort. The "other" category consisted of Hispanic, Asian, Pacific Islander and Native American."

Comment
Using data with as many cases as the NCDB, p-values should be reported to at least 3 decimal places (i.e., $p<0.001$ rather than $p<0.01$)

Reply
Thank you for this suggestion. We have made every p value at 3 decimal points.

Comment
The authors present a variable called income. It is important to note that income is not a personal income at the patient-level. Rather, it is based on the median income of the zip code in which a patient resides. This variable should reflect that or at least warrants clarification and explanation in the methods section.

Reply
We agree that this is an important distinction. In the *Variables* subsection we have added "Income was measured as the median income of the zip code in which the patient lived."

Comment
Line 178: the authors note that 90 -day and 14 year survival were improved in patients who received neoadjuvant chemoradiotherapy. 14-year survival is somewhat misleading, the majority of patients would be either censored or deceased by that timepoint. It would be more appropriate to say 10 or 14 year survival rate estimates were XX for the neoadjuvant CRT group and YY for the upfront surgery group. The Kaplan-Meier curve could then be referred to as "long-term follow up" rather than 14 year.

Reply
We agree with this distinction. To report the long term survival we have now said "At the 14-year period survival estimates were worse for the upfront surgery group compared to the neoadjuvant group."

Comment
The tables should be reformatted so that each row has a line; these are hard to follow. Their current form and it appears that there is data incorrectly placed in certain rows (see page 19; primary payor).

Reply
Thank you for pointing this out. We have added hard lines to every row.

Comment

Are the numbers below the KM curve in Figure 4 corresponding to # at risk? This should be clearly stated. Consider calling the group labeled “surgery” either surgery only or “no adjuvant therapy.”

Reply

Thank you for this suggestion. We have added to the figure legend to clarify that the surgery group was surgery alone without adjuvant treatment. We have also clarified that the numbers represent the remaining patients at risk.