Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-24-971

Reviewer A

Thank you for the opportunity to review the manuscript entitled "Impact of lymph node evaluation standard in patients undergoing lung resection for clinical stage IA NSCLC". The authors retrospectively reviewed 654 patients who underwent lung resection for clinical stage IA NSCLC and compared patients who met and did not meet the CoC standard. Left-sided resections, open technique, and lobectomy were associated with meeting the CoC standard. Nodal upstaging was more common in patients who met the CoC standard. Time to recurrence and overall survival did not differ between the CoC standard groups. However, meeting the CoC standard appeared to have a more pronounced effect in patients classified as pN0.

I think the most interesting finding of this study is that meeting the CoC standard seemed to have a more pronounced effect in patients classified as pN0, which may be due to more occult lymph node metastasis in patients who did not meet the CoC standard, as the authors suggested in the discussion. Therefore, the authors should focus more on this finding in the survival analyses.

Reply: We agree with the reviewer that further analysis is warranted. However, our data merely suggests a trend towards significance. We have made this clearer in our results (see page 11 line 264-247). In the discussion we hypothesize that a larger sample size and/or longer follow-up may further clarify this difference which would need to be obtained in a future study.

1) Why did the authors exclude patients with other than adenocarcinoma or squamous cell carcinoma when the title of the manuscript is "Impact of lymph node evaluation standard in patients undergoing lung resection for clinical stage IA NSCLC"?

Reply: The study was initially designed to only include adenocarcinoma and squamous cell carcinoma because these account for the vast majority of non-small cell lung carcinoma cases at our institution. Adenocarcinoma and squamous cell carcinoma also account for the majority of NSCLC worldwide. In our study, it was unclear whether or not additional histologies such as “large cell carcinoma” or “non-small cell carcinoma, not otherwise specified” would be powered to identify true differences when comparing groups. Therefore, those cases were excluded prior to statistical analysis. We have adjusted the title to reflect only inclusion of adenocarcinoma and squamous cell carcinoma (See Title Page).

2) Because this study is a retrospective analysis comparing patients who met and did not meet the CoC standard who underwent lung resection for clinical stage IA NSCLC, the authors' institutional policy for the extent of lymph node removal during the study period should be presented. In addition, the institutional policy for preoperative lymph node assessment, such as PET/CT and invasive mediastinal staging, can be described.

Reply: We thank the reviewer for this comment. The point is extremely valid. Our study incuded patients from 2005-2014. PET/CT was not routinely used in the earlier years of the study period; however, it quickly became the standard for all patients with suspected or diagnosed lung cancer. Because of the variability inherent to the study period, we clearly describe the use of preoperative staging methods in Table 1. Invasive mediastinal staging in clinical T1 disease was more frequently used in cases of central tumors, but ultimately left to the discretion of the treating surgeon. We excluded patients with suspected N+ disease based on preoperative imaging, as they would not be classified as clinical stage I.
The CoC standard 5.8 came into effect only in 2021. During the study period (2005-2014) lymph node dissection was left to the discretion and best clinical judgement of each individual surgeon, as there were no clear recommendations on the extent of lymph node removal for patients with stage I lung cancer. To note, our institution participated in ACOSOG Z0030 and some patients included in this retrospective analysis were subjected to randomization between sampling vs. dissection. ACOSOG Z0030 was published in 2011, therefore, its results did not impact a large part of the study period. Overall, our institutional policy since then has been to emphasize nodal stations rather than lymph node count. This has been clarified in the methods section (see page 7 line 142-147). The objective of the present analysis was to retrospectively evaluate if lymph node removal meeting the current CoC standard 5.8 had any long-term impact on patients undergoing curative-intent surgical treatment of clinical IA lung adenocarcinoma or squamous cell carcinoma. This older cohort provided sufficient follow up to assess long-term effects. Since the CoC standard 5.8 only came into effect in 2021, cohort studies after that date would not have sufficient follow up to evaluate outcomes.

3) As you know, the cT factor is decided based on the solid component size on CT, not the whole tumor size, and the solid component size should have more impact on postoperative outcomes than the whole tumor size. Therefore, the solid component size on CT should be presented and compared between patients who met the CoC standard and those who did not.

Reply: The reviewer brings to attention a very important point. This is an important concept for patients with subsolid tumors, usually representing adenocarcinoma. However, this concept only came into effect after recommendatiosn by the IASLC in 2016 (Travis WD et al. J Thorac Oncol 2016:11(8):1204-23). Therefore, this was not applicable when patient was treated during the study period (2005-2014). Because of this, we do not have

measurements of solid components on preoperative CT scans or measurements of true invasive component on pathology reports of patients with subsolid nodules. Unfortunately, this data is not available. Whole tumor size was used to identify and classify patients as having clinical T1 disease. We understand the limitations of this and have added this to the discussion section (page 16, lines 354-355). Even without this information, we think that our findings are important. Currently, there is a paucity of data supporting the validity of the CoC standard 5.8 recommendations. Our study helps bridging the gap and identifying areas of future reasearch.

4) Why did the authors include the pN factor for adjustment in the postoperative outcome analyses? This might minimize the effect of occult lymph node metastasis in patients who did not meet the CoC standard. The pN factor should be removed from the adjustment factors. In addition, because this study is a retrospective analysis, only preoperative factors, including the solid component size on CT, can be used for adjustment.

Reply: In page 11, under the subtitle "Compliance with the CoC standards and oncologic outcomes" we present the unadjusted and adjusted effects of meeting the CoC standard on tumor recurrence and overall survival. As clearly shown, meeting the CoC standard was not significantly associated with recurrence or overall survival on univariable or multivariable analyses. This remained true even when only including preoperative factors in the model, as suggested by the reviewer.

- **For recurrence, adjusted HR: 0.88, CI 95% 0.64-1.21, p=0.420 (only adjusting for type of resection).**
- **For overall survival, adjusted HR: 0.82, CI 95% 0.61-1.10, p=0.184) (only adjusting for age, gender, smoking history, histology and type of resection).**

As observed, results are not significantly changed.
When analyzing long-term oncologic outcomes, it is important to also evaluate known strong prognostic factors. Many patients with clinical stage IA lung cancer are upstaged on both the T and N descriptors after surgery. It is well known that pathologic staging is one of the strongest predictors of long-term outcomes in patients with lung cancer. In our dataset, pathologic N staging was significantly associated with recurrence (HR 2.01, CI 95% 1.63-2.50, p<0.001) and overall survival (HR 1.55, CI 95% 1.24-1.92, p<0.001). As such, we do believe that it is important to adjust for its effects in this type of survival analyses. Even though we did not find that meeting the CoC standard was independently associated with improved long-term outcomes, we did find that it was significantly associated with nodal upstaging (cN0→pN+). Since pathologic N staging is significantly associated with long term outcomes, better identification of patients with pN+ disease becomes important. That is where the CoC standard 5.8 could help. As shown in Figures 3

and 4, not meeting the CoC standard in patients classified as pN0 appeared to have an impact, possibly because of undetected nodal disease.

5) Median follow-up times should be reported in any postoperative outcome analysis.

Reply: We thank the reviewer for pointing this out. These data have been included in the manuscript in page 12, lines 255-256 and 259-260.

6) The authors mentioned in the discussion that nodal upstaging in the patients who met the CoC standard has downstream implications on prognosis and candidacy for adjuvant therapies. Therefore, the information regarding postoperative adjuvant therapies should be presented

Reply: We have indicated how our practice changes with nodal upstaging and how this impacts patient's receiving adjuvant therapy in pages 14-15 and added additional references to support those statements.

Reviewer B

This is a very nice study and very timely. The ACS has recently stood up two cancer research committees, one of which is studying standards as to their validity. There is very little data on standard 5.8 and this is a nice contribution to the literature. I have very few comments as I think this is a very well-done study. Some minor points and comments

1) intro - line 37 it states that ACOSOG z0030 described removal of nodes from "each accessible station." This is somewhat misleading as the enrollment criteria was for 2,4,7,10 on the right and 5,6,7,10 on the left, rather than "each accessible station." Patients randomized to sampling had no further nodes removed. Only those randomized to MLND had each accessible station. Perhaps this is what the authors mean, that the study defined what a MLND is, but lines 35-38 were a bit confusing.

Reply: We appreciate the reviewer's help in clarifying our point. We were trying to make reference to the definition of complete mediastinal lymph node dissection used in the trial. We have changed the wording accordingly (see page 5, lines 99-103).

2. The authors do a nice job of describing patients who missed the COC standard by a single station as opposed to multiple stations. Do the authors have data on the nodal upstaging rates for these patients? Was there any "dose dependent" response such that patients who "just missed" the standard compared to those with zero stations had different upstaging rates? Were the patients who missed by a single station closer (or even equal) to those that met the standard?

Reply: We thank the reviewer for an excellent suggestion to strengthen the paper. We have included additional data in the manuscript in page 10, lines 218-226. This includes rates of nodal upstaging when only one component of the standard (hilar or mediastinal) was met or when no component was met. Since the hilar component of the CoC standard only requires 1 lymph node, further analysis is not possible, as all patients not meeting the standard because of this would have missed it by 1 lymph node station. On the other hand, we added data on nodal upstaging and rates of pN2 disease in patients that met the hilar standard but missed the mediastinal component of the standard by 1, 2 or 3 lymph node stations. We hope this is helpful and satisfactory.

3) the authors have a lot of excellent granular data. Did additional N1 stations beyond one also increase nodal upstaging. Figure 1C highlights the impact on CN0-pN2, based on number of mediastinal nodal stations, why not include similar figure for N1 upstaging and N1 stations ?

Reply: We did not identify the same phenomenon for pN1 upstaging based on the number of hilar lymph nodes. We have added a Supplementary Figure to show these results. To note, the CoC standard 5.8 only allows named or numbered hilar stations to satisfy this component of the standard. However, there are unnumbered more distal intrapulmonary lymph nodes that can also result in upstaging, even if numbered hilar lymph nodes were not collected. We have included those in the Supplementary Figure as "Unnumbered LN only".

4. The authors sort of gloss over the difference in chyle leak. Yes, the p value was 0.052 but the point values showed a 3.5x fold increase in the rate. This should be acknowledged a little more explicitly. This is probably real as is prolonged air leak, chest tube days, and length of stay).

Reply: We thank the reviewer for pushing us on this point. We have added a paragraph to the discussion acknowledging the observation of a trend toward significance in postoperative complications suggesting that meeting the COC standard may have a downside and may factor into the reason for low rates of meeting the COC standard, particularly for perceived "lower risk" lung nodules. (See page 15-16, lines 338-346).

5. Figure 3 and figure 4 are fascinating and really the key figures.

Reply: We thank the reviewer for these comments. Although out study has limitations, we do believe it will contribute significantly to our understanding of the role of current lymph node resection recommendations.

Nice work.