**Peer review file**

**Article information: http://dx.doi.org/10.21037/jtd-20-3611**

**Reviewer A**

**Comment 1**: I am a bit confused about the inclusion criteria. Were only pathologically node negative patients included? If so, I would think you'd perhaps want to do clinical T1a-2bN0M0 patients instead. If the number of N1 nodal stations examined was important, one would think that more accurate pathologic staging and more appropriate use of adjuvant therapy was the underlying mechanism. This cannot be shown if everyone is node negative.

**Reply 1**: Thank you for your kind suggestion! In this study, all included patients is node-negative. The hypotheses of study of Osarogiagbon et al, because the number of N1 nodes examined will be directly proportional to the likelihood of correctly identifying patients with pN0 disease, there will be a sequential improvement in survival of patients with pN0 as more N1 lymph nodes are examined; pN0 patients with more non-hilar N1 (stations 11 to 14) lymph nodes examined will have a more favorable prognosis(1), inspired us to design this study. Different from their study, considering the possibility of lymph nodes being clipped during dissection, which make the number of N1 LNs is increased, we analysis the number of N1 station examined in pN0 patients.

**Comment 2:** It would be helpful if you commented on the pattern of recurrence seen in this cohort and how it was impacted by the number of nodes/nodal stations examined.

**Reply 2:** Thank you for your kind suggestion! In our center, many patients receive the treatment

although it is far from their residence. When the treatment is finished, part of patients likely to realize follow up in the local hospital as a matter of convenience. The follow-up office in our hospital would regularly carry out telephone follow-up for every patient. However, the results are not satisfactory, many patients can't report recurrence site accurately. The information of recurrence site is unavailable for a large proportion of patients. Therefore, we only evaluate OS and DFS in this study. This limitation has been clarified in the revised manuscript (see Page16-17, Line 344-347).

**Comment 3:** You mentioned in your discussion that inadequate examination of the N1 nodes could result in incorrect staging and receipt of adjuvant therapy. However, if your analysis only included node negative patients that's not what you have shown here. In that case what do you believe is the behind the differences in survival based on the nodes examined? If you did include some node positive patients, you should clarify the wording in your methods section. It would also be helpful to provide information of pathologic nodal upstaging based on the nodal stations examined and perform an analysis looking at receipt of adjuvant therapy based on this as well.

**Reply 3:** Thank you for your constructive idea! For cN0 patients, removing more LNs has two effect. One is that examining more LNs can provide a more accurate N stage. Another is that examining more LNs may remove metastatic lymph nodes, which has a therapeutical effect. Patients with node metastasis would receive adjuvant therapy, which also has a therapeutical effect. Therefore, I thought examining more LNs can improve prognosis of cN0 patients. For pN0 patients, examining more doesn't have therapeutical effect. So we used "associate with better DFS" instead of "improve DFS" in conclusion. In line with view of **reviewer E**,

pN0 need dissect sufficient LNs of N1 station and N2 station and evaluate. Node-negative patients with insufficient LNs of N1 station and N2 station examined may be a “fake pN0”. Fake pN0 is seen as true pN0, which is the behind the differences in survival based on the nodes examined.

**Comment 4:** The discussion on sublobar resection seems a bit out of place considering they were excluded from your analysis.

**Reply 4:** Thank you for your king suggestion! As you said in comment, there is no data about sublobar resection in this study. It is unsuitable to give any conclusion or recommendation about sublobar resection. We have rewritten this part to make it relevant with this artcle. (see Page15-16, Line 309-335)

**Reviewer B**

**Comment 1**: Unfortunately, one major issue with your manuscript is the English language. There are many typographic and grammatical errors throughout the manuscript. There are parts of this interesting manuscript that are hardly understandable due to wrong using of words and prepositions. I strongly suggest having the manuscript revised by a native speaker or someone with similar skills of the English language. This editing should also yield a clearer and more focused language throughout the manuscript.

**Reply 1**: Thank you for your kind suggestion! Because the first language of us is not English, we have consulted an English language editing service from American Journal Experts to improve readability. The language editing number is C3WRK5RB. The language editing

certificate will also be submitted if needed.

**Comment 2**: The authors describe that they performed different surgical approaches like sublobular resections (segmentectomy, wedge resection), lobectomy or probably pneumectomy. However, they provide no data on the surgical approach, apart from VATS or thoracotomy. The surgical approach may have a bigger impact on survival than LN dissection. The authors have to provide this basic data.

**Reply 2**: We are very sorry for our negligence of missing the data of surgical approach. In our study, patients with sublobular resection were excluded. Most of included patients received lobectomy. In all 1935 patients, 1752 patients received lobectomy (90.5%), 118 patients received lobectomy combined with sublobectomy in other lobe (6.1%), 35 patients received lobectomy in two different lobe (1.8%) and 30 patients received pneumectomy (1.6%). These details of surgical approach can be seen in (see Page 9, line 181-184, Table 1 and 2).

**Comment 3**: The authors explain that older patients tended to have less LN resected probably due to a different surgical approach. What about other patients' characteristics like preoperative lung function, comorbidities? Did they have an influence on the surgeon's choice for a surgical approach?

**Reply 3**: Thank you for your rigorous attitude! In this study, we have already excluded patients with sublobectomy and 90.5% of patients received lobectomy. Lung function or comorbidities did not have an influence on the surgeon's choice for a surgical approach.

**Comment 4:** Lines 63ff: The authors quote Subramanian et al. who analyzed lobectomy vs. sublobular resection for early stage NSCLC. The authors conclude that the differences in their collective was based on the different number of resected lymph nodes. However, it was due to an unappropriated surgical approach in form of the sublobular resection. The authors tend to rip of facts from the references, which suit their results, but which were clearly not the major conclusion of the underlying papers throughout this manuscript.

**Reply 4:** Thank you for your rigorous attitude! We carefully read the literature again and rewrote this part of background to make it more in line with the original meaning of the reference (see Page 5-6, line 10-105).

**Comment 5:** Lines 84ff: Surgeons chose LN sampling, selective LN dissection "…" according to general condition? What kind of condition? Patient's condition? Which would be incomprehensible, condition of the surgeon? Condition of the lymph nodes? The authors have to provide more information. This is clearly a bias.

**Reply 5:** Thank you for your rigorous and scientific attitude! Because this is a retrospective study, we can't request the extent of mediastinal nodal remove. The selection of mediastinal lymph node examination is determined by the attending surgeon. Different surgeon may have different choice when facing same patients. In addition, surgical technique and concepts changed a lot during in the decade. One surgeon may have different choice at different times when facing same patients. So, it is hard to answer this question clearly. In general, surgeon would choose thorough mediastinal nodal retrieval for patients with a large primary tumor (T2a-T2b), solid nodules or radiographic pleural invasion.

**Comment 6:** Statistical analysis lines 100ff: The authors do not describe their propensity model. A propensity matching is not calculated by a multinomial regression. Those two terms do not necessarily have anything to do with each other. One must assume that the authors have not understood the statistical tests in their entirety.

**Reply 6:** Thank you for your kind suggestion! The propensity score was calculated by a logistic regression model, we rewrote this part to make it clearly (see Page 8, line 149-151).

**Comment 7:** Please use multivariate models only if $p<0.05$ in univariate analysis. Otherwise, multinomial regressions will not yield significant results.

**Reply 7:** Thank you for your kind suggestion! We reperformed the statistical analysis according to your suggestion. No apparent change can be seen in the results (see Page 8, Line 163, Table 2 and Supplemental Table).

**Comment 8:** Lines 123-125: Histology: 145 patients were diagnosed with histology other than Adeno or SCC. What about SCLC? Did you include those patients? Please provide data.

**Reply 8:** Thank you for your rigorous attitude! According to the inclusion criteria, only NSCLC was included in this study. 145 patients with other histology included 49 adenosquamous carcinomas, 41 lymphoepithelioma-like carcinomas, 5 large cell carcinomas, 5 carcinoid, 6 mixed neuroendocrine carcinomas, 6 sarcomatoid carcinomas and 33 other rare histological type or undifferentiation (see Page9, Line 178-181).

**Comment 9:** Lines 129-130: Patients with only one LN examined had also less N2 LN examined. This result is not surprising, although I do not even know the cohort. Who would skip dissecting N1 LN but only dissect N2 lymph nodes apart from a diagnostic operation like mediastinoscopy? Please delete this segment.

**Reply 9:** Thank you for your kind suggestion! As you say in the comment, it is unusual to examine only N2 LN but without N1 LN examined in lobectomy. Therefore only 2.1% of patients had no N1 LN examined. One condition is surgeon tried to dissect N1 LN but have no found. The number of patients with different N1 station examined is an important baseline information, so we keep the information of number of patients with one, two, three, four and five N1 station examined (see Page9-10, Line 190-191).

**Comment 10:** Lines 228-231: The authors discuss that segmentectomy is a more suitable surgical approach for LN resection than wedge resection. Why would that be? Perhaps because most surgeons simply do not perform a sufficient lymphadenectomy when performing a wedge resection? this would lend itself very well to discussion if the authors have data on the proportion of their surgical accesses.

**Reply 10:** Thank you for your king suggestion! Consider there is no data of sublobar resection in this study. We revised this part of discussion considerably and added 2 references to make it more relevant with this article (see Page15-16, Line 309-335).

**Reviewer C**

**Comment:** The authors present a well-designed retrospective analysis of a large group of

patients with early stage lung cancer (IA-IIA). The authors make a valid attempt to perform an analysis on the number of N1 stations postoperatively examined. After propensity score matching the number of N1 station examined was an independent prognostic factor for DFS. I think that this study is highly valuable for thoracic surgeons. and this article is worth reading. There are however a few points that should be addressed prior to final consideration for publication.

Have all patients been staged with PET/CT?

Where the location of the primary tumor relevant and was the number of resected station correlated?

However, I congratulate the authors on a well written article of an important topic for thoracic surgeons and one in which we are faced with a complex clinical problem frequently.

**Reply:** Thank you for your recognition of this study. In this study, few patients underwent PET-CT before surgery. On the one hand, this study reviewed the patients from 2008 to 2018. PET-CT has become popular in recent years in our region. On the other hand, we would not routinely recommend PET-CT for early stage NSCLC. We usually recommend PET-CT when patients have mediastinal lymph node metastasis in enhanced CT scanning.

In this study, tumor location was relevant with the number of resected N1 station ($p<0.001$). The median number of N1 station examined in all 1935 patients is 3 and the median number of N1 LN examined is 8. Tumor in right middle lobe had significant less median N1 station examined (2 N1 stations) and N1 LN examined (5 N1 LNs). However, tumor location did not have apparent impact on OS (before PSM: log-rank $p=0.077$; after PSM: log-rank $p=0.276$) and DFS (before PSM: log-rank $p=0.312$; after PSM: log-rank $p=0.664$). So, we didn’t consider

tumor location as a covariable in PSM and COX analysis.

**Reviewer D**

**Comment:** The authors present a single-centre retrospective series examining the prognostic association of ex-vivo lymph node examination in lobectomy specimens. I have several concerns about the clarity of the writing and the statistical methodology, but I would consider these irrelevant given that the central premise of the study seems to be fundamentally flawed. The central premise of the paper seems to be that looking at more lymph nodes in the removed lobe, after it has been removed from the patient, has some impact on the survival outcomes for the patient. This seems nonsensical to me. How can looking at nodes in a post-resection patient change whether the patient lives or dies or whether the cancer recurs? All that this study demonstrates is that stage migration occurs when we examine more nodes, and that better node examination assigns a more accurate prognosis, which is already well known. The authors can not make any claims or conclusions beyond the fact that stage migration is likely present. Any recommendations about the utility of more extensive N1 dissection or the role of different sublobar resections can not be made based on the data in this paper.

**Reply:** Thank you for your comment! I believe most of researcher would agree that the station (or number) of LNs examined indicates thoroughness of pathologic nodal examination and accuracy of node staging. Many studies supported that variation in the thoroughness and accuracy of pathologic nodal staging is a cause of variation in survival of patients after curative-intent resection of non-small cell lung cancer(2-4). It

cannot be denied that examining more N1 station is not directly responsible for the improved prognosis. Therefore, we used “associate with prolonged DFS” instead of “improve DFS” in conclusion. In addition, because there is no data about wedge resection versus segmentectomy. The conclusions did not recommend wedge resection or segmentectomy. In discussion, we rewrote this part to make the tone down.

**Reviewer E**

**Comment 1**: In the Methods section, Surgeons chose systemic LND, LN sampling or selective MLND. Is there any difference of OS or DFS among 3 different groups?

**Reply 1**: Thank you for your careful attitude! Although receiving varying degrees of lymph node resected, systemic LND and LN sampling had similar number of N2 station examined. It is difficult to clearly distinguish between systemic LND and LN sampling in this study. But both number of N2 LN examined (OS: p=0.956; DFS: P=0.233) and number of N2 station examined (OS: p=0.258; DFS: P=0.748) did not show statistical significance in survival analysis, so we assume that the survival had no difference between the three groups.

**Comment 2**: In all patients, were Hilar LN and Interlobar LN resected?

**Reply 2**: Thank you for your careful attitude! This study did not require all patients examined hilar and interlobar LN and some patients did not have group 10 and 11 lymph nodes examined.

**Comment 3**: How many stations or number of N2 LN resected?

**Reply 3**: Thank you for your careful attitude! This study required at least one station N2 LN

examined. Before PSM, the median number of N2 station examined is 3 and the median number of N2 LN examined is 11. The mean number of N2 station examined is 3.3 (range from 1 to 8) and the mean number of N2 LN examined is 13.2 (range from 1 to 104). After PSM, the median number of N2 station examined is 3 and the median number of N2 LN examined is 11. The mean number of N2 station examined is 3.2 (range from 1 to 8) and the mean number of N2 LN examined is 11.9 (range from 1 to 48) (see Page 9, Line 187-189).

**Comment 4:** In your practice, what is the indications of adjuvant chemotherapy? even N0 patients, big tumor or visceral pleural invasion, LVI are sometimes indicated for adjuvant therapy.

**Reply 4:** Thank you for your careful attitude! In this study, 342 patients received adjuvant chemotherapy (ACT). Whether a patient receives ACT or not is at the discretion of the attending doctor. According to the NCCN guideline, poorly differentiated tumors, vascular invasion (VI), wedge resection, size of tumor > 4 cm, visceral pleural involvement (VPI), and Nx were defined as high-risk factors. The attending doctor intended to recommend ACT for patients with high-risk factor. In addition, ACT did not significantly affect the OS (log-rank p=0.008) and DFS (log-rank p=0.130) in stage IB-IIA patients in this study. So didn't consider ACT as a covariable in PSM and COX analysis (see Page 7, Line 136-137 and Page 9, Line 186).

**Comment 5:** Based on your data, I felt there was no survival benefit of aggressive LN dissection. but I thought that nodal dissection is survival benefit because staging migration and dissection itself. How about your opinion?

**Reply 5:** Thank you for your scientific attitude! The data of this study showed that DFS improved with an increase in the number of N1 station examined after PSM. Before PSM, OS also benefited from the increased number of N1 station examined. In my opinion, for patients with lobectomy, all Intrapulmonary lymph nodes are removed along with lobe. Thorough N1 nodal dissection is survival benefit because staging migration. For patients sublobectomy, increasing N1 nodal dissection is survival benefit because staging migration and dissection itself. Thorough N1 nodal dissection is survival benefit because staging migration and dissection itself but increases the time of operation and the risk of complications.

**Comment 6:** From viewpoint of pathologic N0, I thought that pathologist had to dissect sufficient LNs of N1 station and N2 station and evaluate.

**Reply 6:** Thank you for your constructive ideas! Although the results JCOG0804 showed that sublobectomy was not inferior to lobectomy and did not require LN dissection (79.4% patients without LN dissection)(5), we should realize that almost all patients in JCOG0804 are AIS or MIA. For invasive adenocarcinoma or squamous carcinoma, it is necessary to dissect sufficient LNs of N1 station and N2 station and evaluate.