Peer review file

Article information: http://dx.doi.org/10.21037/jtd-20-3393

Reviewer A

The authors analyzed the resected T2-4aN0M0 EC from the Surveillance, Epidemiology, and End Results (SEER) database and their center, for the purpose of identifying the prognostic factors and developing a nomogram to select the patients with an enhanced likelihood of poor survival. They demonstrated that age, sex, histology, chemotherapy, lymph node harvested (LNH), and T stage were identified as the independent prognostic factor. Furthermore, they divided the patients into two groups: the low-risk group and the high-risk group and found that OS of the low-risk group was significantly better than the high-risk subgroup in the entire cohort. Previously, there were other reports that used SEER database and described predictive nomograms, however, this study included the larger case number with long-term follow-up and both internal and external validation to make the nomogram more reliable. This study may be valuable for the clinicians. However, I believe that it could be further enhanced after several corrections. Following are my comments.

Reply:

Thank the reviewer very much for reading our manuscript very carefully. The comments and suggestions are very professional and constructive for us.

Comments:

Comment 1. Which timing did chemotherapy and radiotherapy perform, preoperative or postoperative?

Reply 1: Due to the limited information recorded in the SEER program, Chemotherapy information recorded in the database only includes: performed or did not perform chemotherapy, but does not include any information about the timing. Regarding Radiotherapy, the related information recorded in the database includes: no radiation, radiation prior to surgery, radiation after surgery, radiation before and after surgery, intraoperative radiation, intraoperative radiation with radiation before/after surgery and sequence unknown but both were given. Therefore, it is tough and confusing to separate it into preoperative category and postoperative category.

We do agree with you that it is necessary to separate perioperative therapy into preoperative category and postoperative category, because these two therapies may have a different effect on survival. We hoped that our future studies could overcome the limitation.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line19-20).

Comment 2. The authors used pathological prognostic group, not pathological stage. The authors should describe it clearly.

Reply 2: We have described it more clearly in the text.

Changes in the text: We have modified our text as advised (see Page 3, Line 5-7; Page 11, Line 4-5; Page 13, Line 23-24).

Comment 3. Group IC was not assigned for squamous cell carcinoma, but only for adenocarcinoma. I suggest that pathological TNM should be used in the study, because the staging classification was identical in both squamous and adeno.

Reply 3: Yes, according to the current TNM staging system, group IC is only applied

for adenocarcinoma but not for squamous cell carcinoma. As we all know, the current TNM staging system does have some known drawbacks. The purpose of the TNM staging system is to create sub-groups through different combinations of various categories of prognostic factors such as histology, depth of invasion, grade, location, nodal status, and metastasis status. The TNM staging system is applied for all the early stage tumor and advanced tumor. From our perspective, prognostic factors that constitute the current TNM staging might not be identical between early stage tumor and advanced tumor. Our study confirmed that the TNM stage factors such as T stage and histology are important prognostic factors, but not grade and location, and four other prognostic factors (age, sex, lymph node harvested and chemotherapy) might should be considered.

In this article, we actually conducted the risk stratification of patients with stage pT2-4aN0M0 esophageal carcinoma based on the TNM staging manual, and what we have done is to improve the performance of the current TNM staging system. Although the TNM staging only assigns group IC to adenocarcinoma but not to squamous cell carcinoma, and the staging classification of our nomogram was identical between squamous cell carcinoma and adenocarcinoma, the C-index and decision curve analysis (DCA) of our model showed greater performance than the TNM staging system. This article might indicate that our novel staging classification was more efficient than the current TNM staging classification for stage pT2-4aN0M0 esophageal carcinoma. It putted forward a good beginning for further research on improving the TNM staging performance of this population subset in the future.

Changes in the text: None.

Comment 4. LNH could be associated with surgical type. The authors did not assess

or discuss this point. Could the authors mention about it?

Reply 4: We have added information of surgical type (information of surgical type recorded in the SEER program only includes partial or total esophagectomy, not for three-incision radical resection, Ivor-Lewis surgery, Mckeown surgery or radical surgery for left thoracic esophageal carcinoma) in the Table 1 (revised version) and analyzed the correlation between LNH and surgical type through Pearson's χ^2 test. The result showed that these two variables were not correlated (P=0.127).

Changes in the text: We have added information of surgical type in the Table 1 (see Page 25).

Comment 5. The authors transformed the continuous variables into categorial variables using cutoff values. Why did the authors divide age, LHN, and tumor size into three categories? Please discuss it.

Reply 5: In our article, X-tile software was applied to transform the continuous variables into categorial variables, and the software could offer us one or two cutoff values. As we all know, categorization is associated with loss of information. Thus, three categories were more comprehensive than two categories. In addition, the reason that why we did not use continue variables is that it is more practical and convenient for clinicians to use categorial variables rather than continue variables in clinical practice.

Changes in the text: None.

Comment 6. The authors divided the patients into the low risk group and the high risk group. Which variables did the authors selected to classify into the two groups?

Reply 6: Variables with P value less than 0.05 in the multivariate Cox analyses were

entered into the nomogram. Each number/category of the independent prognostic variables in the nomogram was assigned a score on the point scale and finally added up to a total score. After that, a cutoff value, achieved by X-tile software, was applied to transform the continuous score into2 categories (low-risk and high-risk groups).

Changes in the text: We have added it in the method section of revised version (see Page 7, Line 10-15).

Comment 7. If age, LHN, and tumor size were included in the risk classification, which data among three did the authors use for statistical analysis?

Reply 7: Actually, age and LNH were included in the risk classification, and tumor size was not included. Data of these three variables were all involved in the statistical analysis. Univariate Cox analysis revealed that age ($P < 0.001$) and LNH ($P < 0.001$) were significant prognostic factors, and tumor size was not a correlated factor ($P = 0.357$). Multivariate Cox analysis confirmed that age ($P < 0.001$) and LNH ($P < 0.001$) were independent prognostic factors. Variables with P value less than 0.05 in the multivariate Cox analysis (age, LNH, T stage, sex, histology and chemotherapy) were entered into the nomogram, and a risk classifying system was formulated.

Changes in the text: None.

Comment 8. LNH and T category have already been known as risk factor for the esophageal cancer patients. The authors should stratify the patients by LNH or T category and perform additional subgroup analysis.

Reply 8: As you suggest, we have performed additional subgroup analysis according to T stage and LNH categories. The details are as follows: in the T stage subsets analysis, low-risk cases had better OS rate compared with high-risk cases in all the three T

substages (T2, $P<0.001$, Figure 3G; T3, $P<0.001$, Figure 3H; T4a, $P=0.020$, Figure 3I). Regarding LNH, the OS of low-risk cases was also superior than that of high-risk cases in all the three categories (LNH≤6, $P<0.001$, Figure 3J; 6<LNH≤17, $P<0.001$, Figure 3K; LNH>17, $P<0.001$, Figure 3L). The results above shown an excellent risk classifying efficacy of the nomogram.

Changes in the text: We have added additional subgroup analysis according to T stage and LNH categories in the result section of revised version (see Page 10, Line 7-11)

Reviewer B

Title:

Comment 1. I would suggest authors to clarify if the TNM refers to the clinical (cT2-4N0M0) or pathological (pT2-4N0M0) evaluation.

Reply 1: All cases included in our research were pathologically diagnosed. As you suggest, we have revised "T2-4N0M0" into "pT2-4N0M0" in the whole manuscript.

Changes in the text: We have modified our text as advised (see Page 1, Line 2; Page 2, Line 4; Page 2, Line 7; Page 3, Line 1; Page 3, Line 6; Page 4, Line 7; Page 4, Line 9; Page 4, Line 20; Page 6, Line 12; Page 6, Line 16; Page 6, Line 19; Page 6, Line 23; Page 11, Line 5; Page 11, Line 10; Page 11, Line 12; Page 11, Line 14; Page 12, Line 1; Page 12, Line 15; Page 13, Line 23; Page 17, Line 8; Page 17, Line 10; Page 17, Line 16; Figure 1).

Abstract: Well written, concise.

Introduction: Well written

Methods:

Comment 2. The TNM classification refers to the pTNM or cTNM?

Reply 2: All cases included in our research were pathologically diagnosed, and we have revised "TNM" into "pTNM" as you suggest.

Changes in the text: We have modified our text as advised (see Page 6, Line 23).

Comment 3. How clinical staging was performed?

Reply 3: All cases included in our research were pathologically diagnosed, and information of clinical staging is unavailable in the SEER database.

Changes in the text: None.

Comment 4. Did authors exclude the mortality related to surgery (post-operative mortality)?

Reply 4: Yes, the perioperative death patients (survival month recorded in the SEER database is 0 and death within 30 days after the operation or any time after the operation if the patient did not leave the hospital alive in the SYSUCC database) were excluded in our cohort. We have added it as one of our exclusion criteria.

Changes in the text: We have added the post-operative death as one of our exclusion criteria in the method section of the revised version (see Page 6, Line 15; Figure 1).

Results:

Comment 5. What was the perioperative mortality in the studied population?

Reply 5: The perioperative death patients were excluded in our cohort.

Changes in the text: None.

Comment 6. What was the radiation dose?

Reply 6: Due to the limited information recorded in the SEER program, radiotherapy information recorded in the database only includes: no radiation, radiation prior to surgery, radiation after surgery, radiation before and after surgery, intraoperative radiation, intraoperative radiation with radiation before/after surgery and sequence unknown but both were given, but does not contain information of radiation dose.

We totally agree with you that radiation dosage is an important predictor. We have

added it as one of our research limitations in the discussion section of revised version. We hoped that our future studies could overcome the limitation.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line 20).

Comment 7. What was the chemotherapy regimen adopted?

Reply 7: The information of chemotherapy recorded in the database only includes: performed or did not perform chemotherapy, but does not include any information about regimen. We totally agree with you that chemotherapy regimens may confer a different effect on patients' survival. We have added it as one of our research limitations in the discussion section of revised version.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line 19).

Discussion

Comment 8. Authors could discuss more about how the nomogram would influence the clinical practice, and how clinicians should use the nomogram to make decisions in esophageal cancer patients.

Reply 8: We have described the clinical significance of this nomogram more clearly in the discussion section. The details are as follow: in this study, only nodal negative patients were included. When compared with advanced tumors, the clinical manifestation of relapse or progression of these patients might be less obvious. As we all know, there are still many developing countries in the world. Esophageal carcinoma mostly occurs in the group of people with relatively poor economic situation, and most

of them live in rural area far away. The expensive follow-up examinations and long distances form major obstacles for these patients to get scheduled follow-up examinations. Based on our classifying system, we strongly recommend that the high-risk patients should return to hospital for a scheduled follow-up examination or even postoperative therapy, irrespective of cost.

Changes in the text: We have added the clinical significance of the nomogram in the discussion section of the revised version (see Page 11, Line 16-24)

Reviewer C

Study Design:

Administrative database retrospective review (SEER [n=2441] + Sun Yat-sen University Cancer Center in China [n=1323])

Study period = 2001-2015

Derivation and validation cohorts within SEER database

Sun Yat-sen University Cancer Center patients served as external validation cohort

Objective:

Create nomogram to identify patients at high risk of postoperative

Outcome:

Survival (all cause mortality)

Comments to the Authors:

Comment 1. I commend the authors on their efforts to use large databases to improve the care of esophageal cancer patients. The manuscript reads relatively well. However, it would still benefit from further grammatical and syntax review. In general, it also probably best to avoid the use of acronyms that will not be easily recognized by most readers. The following comments are organized according to the section of the manuscript to which they apply.

Reply 1: We have sent the wordings of the main text to the medical writing service (AME Editing Service), and we have revised our manuscript according to your comments. A detailed list of revision in a point-by-point fashion is shown below.

Changes in the text: None.

Abstract

No specific comment

Introduction

Comment 2. Although the objective is stated, there is no statement of the research hypothesis being tested.

Reply 2: Thank you very much for reading our paper so carefully. We have revised the statement in the last paragraph of the introduction section as you suggest. The details are as follows: in the current study, we analyzed resected cases of pT2-4aN0M0 EC from the Surveillance, Epidemiology, and End Results (SEER) database and Sun Yat-sen University Cancer Center (SYSUCC). The independent prognostic factors of overall survival (OS) and cancer specific survival (CSS) were identified, and prognostic nomograms were formulated. Risk-stratifying systems, based on the nomograms, were established to select the patients with an enhanced likelihood of poor survival. Different statistical methods were carried out to validate the models. Using the models, clinicians could refine treatment strategies and subsequently improve patient prognosis.

Changes in the text: We have revised the statement in the last paragraph of the introduction section as you suggest (see Page 4, Line 22-25 and Page 5, Line 1)

Methods

Comment 3. Some patients have received treatment modalities other than surgery:

The sequence of treatment should be specified as neoadjuvant treatment may affect pathologic stage.

Due to the unpredictable nature of the effect of neoadjuvant therapy on each tumor, the analysis should be stratified by treatment regimen i.e. surgery alone, neoadj chemo + surgery, neoadj chemoxrt + surgery, surgery + adjuvant.

If neoadjuvant patients are compared, it would be ideal to stratify by tumor regression grade (if this pathologic descriptor is available). Otherwise, the unknown treatment effect on pathologic stage is a problematic confounding factor.

Reply 3: Due to the limited information recorded in the SEER program, Chemotherapy information recorded in the database only includes: performed or did not perform chemotherapy, but does not include any information about the timing. Regarding Radiotherapy, the related information recorded in the database includes: no radiation, radiation prior to surgery, radiation after surgery, radiation before and after surgery, intraoperative radiation, intraoperative radiation with radiation before/after surgery and sequence unknown but both were given. Thus, it is tough and confusing to separate it into preoperative category and postoperative category. Also, information of tumor regression grade is unavailable in the SEER database.

We totally agree with you that it is necessary to separate perioperative therapy into preoperative category and postoperative category, because the sequence of treatment do have a different effect on survival. We have added it as one of our research limitations in the discussion section of revised version. We hoped that our future studies could overcome the limitation.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line 19-20).

Comment 4. There is a wide variation in lymph node yield within the study group. How do the authors explain this? What type of resection did patients undergo?

Reply 4: The SEER program of the National Cancer Institute is a population-based cancer registry system that provides data on cancer incidence and survival from 18 registries among 14 states across the United States and covers 30% of the population. Owing to various tumor stages, various tumor sizes, different skilled surgeons and various surgical types, a wide variety of lymph mode harvested may occur. In this research, we have analyzed the correlation between lymph node harvested and other clinical variables through Pearson's $\chi 2$ test or Fisher's exact test, and the result revealed that advanced tumor stage ($P = 0.032$) and larger tumor size ($P < 0.001$) were correlated with more lymph nodes harvested. In current clinical practice, as for advanced disease, more lymph nodes shall be harvested in order to remove metastasis or potential metastasis and stage the disease more accurately, which may benefit patients.

Due to the limited information recorded in the SEER program, related information about surgical type recorded in the database only includes partial or total esophagectomy. 973 (87.0%) patients received total esophagectomy and 145 (13.0%) patients received partial esophagectomy (please see Table 1).

Changes in the text: We have added information of surgical type in the Table 1 (see Page 25).

Comment 5. The term "high-risk" is used throughout but never clearly defined. What is the definition of high and low risk? At least, the methodology by which the binary risk stratification is achieved should be more clearly explained. Defining a high and low risk patient profile and trying to predict 3, 5, and 10-year survival probability are 2 different objectives. In the former, a score threshold would dichotomize patients in risk groups and in the latter, a nomogram such as the one presented could be used.

Reply 5: We have described the definition of high-risk more clearly in the methodology.

We also have added a table (Table 4 in the revised version) which listed the score of different categories of variables. The details are as follows: Variables with *P* value less than 0.05 in the multivariate Cox analyses were entered into the nomogram. Each number/category of these independent prognostic variables in the nomogram is assigned a score on the points scale and finally added up to a total score. After that, a cutoff value, achieved by X-tile software, was applied to transform the continuous score into two categories (low-risk and high-risk group).

We totally agree with you that defining different risk patient profiles and estimating individual prognosis are two different objectives. In this research, a prognostic nomogram was developed to predict 3, 5, and 10-year survival probability of pT2-4aN0M0 EC patients. Based on the nomogram, a risk classifying system was formulated to define the poor survival patients.

Changes in the text: We have described the definition of high-risk more clearly in the methodology (see Page 7, Line 10-15). We also have added a table which listed the score of different categories of prognostic variables (Table 4 in the revised version).

Comment 6. How were points assigned to each statistically significant prognostic factor? This is not well explained in the nomogram section of the results. This explanation should be in the Methods section rather than in the results.

Reply 6: As revised in Comment 3, we have described the definition of high-risk more clearly in the methodology. We also have added a table (Table 4 in the revised version) which listed the score of different categories of variables.

Changes in the text: We have described the definition of high-risk more clearly in the methodology (see Page 7, Line 10-15). We also have added a table which listed the score of different categories of prognostic variables (Table 4 in the revised version).

Comment 7. It would have been more appropriate to use cancer-specific survival since the authors are proposing to use their clinical decision tool to identify patients who may benefit from additional cancer therapy. These therapies are toxic and have no effect on the long-term risk of non-cancer related mortality.

Reply 7: As you suggest, we have added cancer specific survival (CSS) as the other endpoints in this article.

Changes in the text: We have added cancer specific survival as the other endpoints in this article (see Page 2, Line 12; Page 2, Line 19; Page 2, Line 21-23; Page 3, Line 2; Page 4, Line 23; Page 7, Line 8-9; Page 8, Line 20-24; Page 9, Line 2; Page 9, Line 4-6; Page 9, Line 9; Page 9, Line 14-16; Page 9, Line 18; Page 9, Line 20; Page 10, Line 12-18; Page 11, Line 4; Figure 2; Figure 4; Figure S2; Figure S3)

Comment 8. Completeness of resection is an important determinant of survival. What was the impact of R0 resection on survival?

Reply 8: We totally agree with you that resection margin is a critical prognostic factor for survival. Due to the limited information recorded in the SEER program, the related information about completeness of resection is unavailable in the database. We have added it as one of the limitations of our paper. We hoped that our future work could overcome the limitation.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line 20)

Comment 9. I am assuming that TMM stage refers to pathological stage rather than clinical stage. This distinction needs to be specified in this section.

Reply 9: All cases included in our research were pathologically diagnosed. As you suggest, we have revised "T2-4N0M0" into "pT2-4N0M0" in the whole manuscript.

Changes in the text: We have modified our text as advised (see Page 1, Line 2; Page 2, Line 4; Page 2, Line 7; Page 3, Line 1; Page 3, Line 6; Page 4, Line 7; Page 4, Line 9; Page 4, Line 20; Page 6, Line 12; Page 6, Line 16; Page 6, Line 19; Page 6, Line 23; Page 11, Line 5; Page 11, Line 10; Page 11, Line 12; Page 11, Line 14; Page 12, Line 1; Page 12, Line 15; Page 13, Line 23; Page 17, Line 8; Page 17, Line 10; Page 17, Line 16; Figure 1).

Comment 10. The authors should consider including a performance comparison of the nomogram by Wu and colleagues versus their own nomogram in the Sun Yat-sen University Cancer Center external database.

Reply 10: As you suggest, we have made a performance comparison of the nomogram by Wu et. al versus our nomogram. We have to admitted that the performance of their nomogram is superior than ours. Because complete data analyses were performed in our research, which might lead to selection bias, and it is associated with loss of information. Also, as mentioned in the limitation, the insufficiency of variables included in our nomogram may contribute to the little pale C-index of our nomogram. The details of discussion are as follows. Wu et. al (1) also presented a large study of 20,623 EC adenocarcinoma from the SEER dataset and established a nomogram to predict OS and cancer-specific survival. Grade, T stage, N stage, M stage, performed surgery or not, insurance record and marital status were entered into their nomogram which showed great performance (C index: 0.720-0.733) when compared that with our nomogram. However, chemotherapy, a critical prognostic factor on survival, was lacking in their study. In addition, tumor stage was diagnosed either pathologically or

clinically due to the inclusion criteria in their research, which may lead to bias.

1. Wu XX, Chen RP, Chen RC, Gong HP, Wang BF, Li YL, et al. Nomogram predicting cancer-specific mortality in patients with esophageal adenocarcinoma: a competing risk analysis. J Thorac Dis. 2019; 11:2990-3003.

Changes in the text: We have modified our text as advised (see Page 12, Line 21-23).

Result

Comment 11. The authors need to specify the sequence of non-surgical treatment (neoadjuvant vs adjuvant).

Reply 11: Due to the limited information recorded in the SEER program, Chemotherapy information recorded in the database only includes: performed or did not perform chemotherapy, but does not include any information about the timing. Regarding Radiotherapy, the related information recorded in the database includes: no radiation, radiation prior to surgery, radiation after surgery, radiation before and after surgery, intraoperative radiation, intraoperative radiation with radiation before/after surgery and sequence unknown but both were given. Thus, it is tough and confusing to separate it into preoperative category and postoperative category.

We do agree with you that it is necessary to separate perioperative therapy into preoperative category and postoperative category, because these two therapies may have a different effect on survival. We hoped that our future studies could overcome the limitation.

Changes in the text: We have added it as one of our research limitations in the discussion section of revised version (see Page 13, Line 19-20).

Comment 12. A minimum number of lymph nodes harvested for this stage range of esophageal cancer would be 20. The reported 70% of patients who had 17 lymph nodes

or less harvested is casting doubt on the accuracy of nodal staging in this study population. Since the focus is on node-negative esophageal cancer, it would make sense to eliminate cases where <7 lymph nodes were analyzed.

Reply 12: We totally agree with you that insufficiency of lymph node harvested may lead to stage migration. 315 (28.2%) patients were suffered from insufficiency of lymph node harvested (< 7) in our cohort, and these patients account for a slightly large proportion. From our perspectives, eliminating these cases may lead to selection bias and loss of more information. Several our studies (in the revised stage and is not published yet) using SEER data also confirmed that both in esophageal carcinoma and non-small cell lung cancer, the phenomenon of insufficiency of lymph node harvested occurred. In our opinion, it might be a limitation of the database itself.

Changes in the text: None.

Comment 13. The survival curves do not contribute additional information (Figure 3). It is clear that patients with 1 or more statistically significant risk factors for mortality will have a lower survival.

Reply 13: Yes, we totally agree with you that the more risk factors patients had, the lower survival rate they would be. Figure 3 demonstrates that the OS curves of these two subgroups (low-risk group and high-risk group), in the full analysis set or stratified by TNM stage and histology, showed significant distinctions. In this article, the performance of the nomogram was evaluated by concordance index (C-index), calibration plots, decision curve analysis (DCA) and survival curves, and Figure 3 is a symbol of good performance of our model. Also, Figure 3 might help readers master a better understanding of the nomogram and risk classifying system. Moreover, another reviewer suggests me to do more subgroup analysis except for TNM stage and histology.

So, we would keep Figure 3 in this manuscript.

Changes in the text: None.

Comment 14. The T4N0 subgroup is extremely rare and even in this large database, the number are too small to be able to state conclusively that the model can be applied for these patients. These cases should be excluded. The same would apply to upper third tumors. These small subgroups probably create noise in the data more than contribute to the results.

Reply 14: The reason why we chose pT2-4aN0M0 esophageal carcinoma patients was that the NCCN guideline recommends this population subset either surveillance or adjuvant chemoradiation. However, the survival rate of these patients differs substantially across different cases. Herein, we developed a nomogram and risk classifying system with purpose of screening out the patients who may need closer follow-up or even adjuvant therapy. Thus, in our opinion this population subset is special cohort, and it is also the standing point of clinical significance of this article. In addition, on pT4a subgroup analysis (Figure 3F), the OS curve of these two different risk groups shows statistically significant distinction. So, we would keep Figure 3 in this manuscript. We have to admit that upper third tumors were relative rare in our cohort. However, in clinical practice. many upper third esophageal carcinoma patients also receive neoadjuvant radiochemotherapy/neoadjuvant radiotherapy + surgery resection. In order to avoid loss of more information and selection bias, in our perspectives, it is essential to integrate these patients into the entire cohort, and it makes our cohort more representative.

Changes in the text: None.

Discussion

Comment 15. Although extent of lymphadenectomy is an important quality indicator in esophageal cancer resection, it has not been conclusively proven to yield a survival benefit in randomized trial.

Reply 15: Yes, many preview clinical trials indicated that the extent of lymphadenectomy is not a strong prognostic predicator for resected esophageal carcinoma. However, the most recent clinical trial NEOCRETC5010 of our center demonstrated that greater number of lymph node harvested was associated with significantly better overall survival and disease-free survival of esophageal carcinoma patients who received neoadjuvant chemoradiatherapy + surgery, but without any negative impact on postoperative complications (2). In addition, less lymph node harvested (< 20 vs ≥ 20) was significantly associated with increased local recurrence and total recurrence rates (2). Also, several retrospective researches (3-5) also revealed that lymph node harvested is an important prognostic factor. From our perspectives, lymphadenectomy might eliminate the potential micrometastasis of lymph nodes and lead to more accurate staging, which may benefit these patients. We hope more future studies could verify our assumption.

2. Guo X, Wang Z, Yang H, Mao T, Chen Y, Zhu C, et al. Impact of Lymph Node Dissection on Survival after Neoadjuvant Chemoradiotherapy for Locally Advanced Esophageal Squamous Cell Carcinoma: From the Results of NEOCRTEC5010, a Randomized Multicenter Study. Ann Surg. 2021.

3. Du F, Sun Z, Jia J, Yang Y, Yu J, Shi Y, et al. Development and Validation of an Individualized Nomogram for Predicting Survival in Patients with Esophageal Carcinoma after Resection. J Cancer. 2020; 11:4023-9.

4. Deng W, Zhang W, Yang J, Ni W, Yu S, Li C, et al. Nomogram to Predict Overall Survival for Thoracic Esophageal Squamous Cell Carcinoma Patients After Radical Esophagectomy. Ann Surg Oncol. 2019; 26:2890-8.

5. Zheng YZ, Li XQ, Wang JY, Yang H, Wen J, Zhai WY, et al. Impact of Examined Lymph Node

Count for Esophageal Squamous Cell Carcinoma in Patients who Underwent Right Transthoracic Esophagectomy. Ann Surg Oncol. 2020.

Changes in the text: None.

Comment 16. What is the clinical value of closer follow-up in higher risk patients? Most recurrences will be symptomatic and thus detected irrespective of the frequency of oncologic outpatient follow-up.

Reply 16: In this study, only nodal negative patients were included. When compared with advanced tumors, the clinical manifestation of relapse or progression of these patients might be less obvious. As we all know, there are still many developing countries in the world. Esophageal carcinoma mostly occurs in the group of people with relatively poor economic situation, and most of them live in rural area far away. The expensive follow-up examinations and long distances form major obstacles for these patients to get scheduled follow-up examinations. Based on our classifying system, we strongly recommend that the high-risk patients should return to hospital for a scheduled follow-up examination irrespective of cost.

Changes in the text: We have added the clinical significance of the nomogram in the discussion section of the revised version (see Page 11, Line 16-24).

A few important limitations not being addressed properly. For instance:

Comment 17. It is very likely that collinearity exists between number of lymph node harvested and N0 status. The lesser the number of lymph nodes removed (LNH) the higher the probability of N0. and nodal stage are likely to be correlated. This concern about the data could potentially be resolved by stratifying LNH by T stage and showing that there was a preponderance of lower T-stage tumors in the lower LNH groups. This would be no means be grounds to support the practice of limiting LNH based on T stage.

However, it could be a measure of the procedural quality in the database.

Reply 17: We have analyzed the correlation between T stage and lymph nodes harvested (categorized) through chi-squared test and correlation between T stage and lymph nodes harvested (continued) through one-way ANOVA test, and all the results were nonsignificant (chi-squared test: $P = 0.771$; one-way ANOVA test: $P = 0.155$). Considering the relatively poor data quality of SEER database, we have added the limitation of insufficient lymph node harvested in the discussion section. The details are as follows, the quality of several data in the SEER database is relatively poor, for example, the median number of lymph node harvested in the SEER database is only 12, which is less than the number (at least 15 examined lymph nodes) recommended by the NCCN guideline (6), and it might lead to stage migration.

6. National Comprehensive Cancer Network. Esophageal and Esophagogastric Junction Cancers. Version 1.2019. Available at: https://www.nccn.org/professionals/physician_gls/pdf/esophageal.pdf.

Changes in the text: We have added the limitation of insufficient lymph node harvested in the discussion section of the revised version (see Page 13, Line 11-15).

Comment 18. Lymph node yield is highly dependent on methodology used to process specimens in pathology. I suspect there is no detail provided to address this in the SEER database but perhaps this information is available for the Sun Yat-sen University Cancer Center database.

Reply 18: In our center, lymph nodes dissection is completed both by attending surgeons and assistants. According to the location and extent of the tumor, the McKeown, Ivor Lewis, left transthoracic approach or cervical-thoracic-abdominal triple incision esophagectomy are implemented. Standard two-field (abdominal and thoracic) lymph nodes resection is performed for almost all the patients in our center. Three-field lymph nodes dissection is a less common practice but the lymph nodes

along the bilateral recurrent nerves are always being removed. After operation, the assistants finish the remaining works including excising the proximal and distal resection margins for pathological examination, dissecting the remaining lymph nodes that can be visible or palpable and assigning them into different stations according to the locations. All resected specimens are carefully examined and recorded by expert pathologists in a standard manner.

Changes in the text: None.