

# Modified Nodal Stage of esophageal cancer based on the evaluation of the hazard rate of the negative and positive lymph node

**Jinling Zhang**

Linyi People's Hospital

**Hongyan Li**

Linyi People's Hospital

**Liangjian Zhou**

Linyi People's Hospital

**Lianling Yu**

Shandong University

**Fengyuan Che**

Linyi People's Hospital

**Xueyuan Heng** (✉ [xueyuanheng@yahoo.com](mailto:xueyuanheng@yahoo.com))

Linyi People's Hospital <https://orcid.org/0000-0002-8174-0347>

---

## Research article

**Keywords:** Esophageal carcinoma, SEER, Tumor metastasis, TNM

**Posted Date:** August 19th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-54021/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on December 7th, 2020. See the published version at <https://doi.org/10.1186/s12885-020-07664-w>.

# Abstract

## Objective:

The study aimed to propose a modified Nodal stage of esophageal cancer (EC) on basis of the number of positive lymph node (PLN) and the number of negative lymph node (NLN) simultaneously.

## Method:

Data from 13,491 patients with EC registered in the SEER database were reviewed. The parameters related to prognosis were investigated using a Cox proportional hazards regression model. A modified N stage was proposed based on the cut-off number of the re-adjusted ratio of the number of PLN ( $_{\text{number}}\text{PLN}$ ) to the number of NLN ( $_{\text{number}}\text{NLN}$ ), which were derived from the comparison of the hazard rate (HR) of  $_{\text{number}}\text{PLN}$  and  $_{\text{number}}\text{NLN}$ . The modified N stage was confirmed using the cross-validation method with the training and validation cohort, and it was also compared to the N stage from the American Joint Committee on Cancer (AJCC) staging system (7th edition) using Receiver Operating Characteristic (ROC) curve analysis.

## Results:

The  $_{\text{number}}\text{PLN}$  on prognosis was 1.064, while  $_{\text{number}}\text{NLN}$  was 0.962. The modified N stage was defined as follows: N1 stage: the ratio range was from 0 to 0.08; N2 stage: more than 0.08, but no more than 0.63; N3 stage: more than 0.63. Cross-validation method within the cohort identified the predictive accuracy of this modified N stage, and ROC curve analysis demonstrated the relative superiority of the modified N stage over that of the AJCC N stage.

## Conclusion:

The modified N stage based on the re-adjusted ratio of  $_{\text{number}}\text{PLN}$  to  $_{\text{number}}\text{NLN}$  can evaluate tumor stage relative accurately than the traditional N stage.

# Background

Esophageal cancer (EC) is one of the common malignant tumors of the digestive tract, which has poor prognosis[1]. Lymph node (LN) metastasis usually occurs in the beginning of diagnosis, and accurate evaluation of the tumor stage is a key step in determining post-operation treatment[2]. However, at present, the definition of N stage is controversial. The N stage has typically been defined by the American Joint Committee on Cancer (AJCC) as the number of positive lymph node (PLN),but a new N stage was proposed by the Japanese Society for Esophageal Diseases (JSED) [3, 4]. The JSED N stage is defined according to the site of PLN. The site of PLN has been demonstrated to play an important role in the prognosis of patients with EC, while the key role of number of PLN on the prognosis of patients with EC was repeatedly confirmed and widely received by researchers[5-7]. Furthermore, recent research revealed

that the site of PLN was weaker than the number of PLN ( $_{\text{number}}\text{PLN}$ ) in the multiple-parameters analysis using a survival model of EC[8]. Nevertheless, neither of them considered the influence of the number of negative lymph node (NLN) on the prognosis of patients with EC.

Greenstein first proposed the impact of the number of NLN ( $_{\text{number}}\text{NLN}$ ) on the outcome of patients with EC[9]. He suggested that the higher the NLN resected in during surgery would be associated with better post-operative outcome for patients. Hsu confirmed the above identification again in EC[10]. In another study,  $_{\text{number}}\text{NLN}$  was included in a scoring system for determining the prognosis of EC[11]. In other words, it was accepted that  $_{\text{number}}\text{NLN}$  counted in the operation could increase the accuracy of identifying the N stage of AJCC. It was also inferred that  $_{\text{number}}\text{NLN}$  could represent site information for tumor metastasis to some extent.

Because of the significant impact  $_{\text{number}}\text{PLN}$  and NLN has on the prognosis of patients with EC, a modified N stage that consists of both PLN and NLN might provide a more accurate representation of the extent of tumor metastasis in the regional LN station. However, it might not be accurate to define the modified N stage using the ratio of  $_{\text{number}}\text{PLN}$  and NLN directly.

This study investigates the feasibility of a modified N stage which is based on combination analysis including the number of positive LN and negative LN in the meantime. The combined analysis refer to the result of the Cox proportional hazard regression model.

## Methods

### Data Source

The Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute is a comprehensive source of population-based cancer information in the United States. The SEER database collects disease incidence, patient treatments, and survival data from population-based cancer registries covering around 28% of the country's population. SEER data comes primarily from hospital medical records as well as records from outpatient surgical, pathology, and radiology centers. The routine data collected in SEER database includes detailed information on demographics, diagnosis, and tumor characteristics. The work team engaged in active follow-up on the cases included in the SEER.

### Inclusion Criteria for Patients

This study reviewed patient information collected from 2004 to 2011. Data was downloaded using the SEER\*Stat software (8.3.5, The Surveillance Research Program of the Division of Cancer Control and Population Sciences, National Cancer Institute.). The inclusion criteria in this study are as follows: 1) All patients should have experienced radical lymphadenectomy; 2) The LN number collected during the operation was clear; 3) EC is the specific reason of patient death.

A wide range of patient information was obtained from the SEER database. More specifically, the following variables and covariates were collected for this study: age, gender, tumor size, tumor extension, regional nodes positive, regional nodes negative, race (white, black and other), primary site of tumor (cervical segment, chest segment, abdominal segment and cross-section(refers to esophageal cancer across at least one segment)), grade classification of tumor (I, II,III, and IV), AJCC Group (I, II,III, and IV), radiation sequence with surgery (no radiation and/or cancer-directed surgery, radiation prior to surgery, radiation after surgery, radiation before and after surgery, surgery both before and after radiation, sequence unknown, but both were given), tumor metastasis to bone, tumor metastasis to brain, tumor metastasis to liver, tumor metastasis to lung, the survival time and the status of patients.

## Statistical Analysis

The total population was divided into two groups using a random number table. One group was the training population, and the other was the validation population. The cross-validation method was used between the training population and the validation population.

Cox proportional regression model was used to build a prediction function for time event data. The prediction function including the HR of PLN and NLN provided the coefficient to calculate the re-adjusted number of PLN and NLN for proposing the modified N stage of EC. The cutoff number for the ratio of the PLN count to NLN count was investigated using the method of the minimum of P values. This was performed using the software X-tile (2.0, University of Chicago).

Differences in survival rates between subgroups categorized by N stage were analyzed using the Kaplan–Meier analysis and log-rank test. Receiver Operating Characteristic (ROC) Curve Analysis was used to investigate whether the modified N stage proposed by this study was more effective than the previous N stage definition.

All analyses were performed using IBM SPSS version 21.0(SPSS Inc. Chicago, Illinois, USA). Continuous variables were presented as the mean  $\pm$  standard deviation (SD) or when the data exhibited a skewed distribution, as the median and interquartile range (IQR). P values of 0.05(two-tailed) were established as the threshold for statistical significance.

## Results

### Baseline Characteristics and Outcomes

The data for around 100,000 EC patients were reviewed through SEER statistical software, but only the medical records of 13,491 patients were collected under the inclusion criteria. The 13,491 patients with EC were classified into two groups according to the random number method. The two groups were the training population (n = 6698) and the validation population (n = 6793).

The mean age of the total population was  $66.70 \pm 11.12$  years, and 10,776(79.9%) of the patients were male. The median size of tumors was  $104.45 \pm 223.14$  millimeters, and the average tumor size was

366.58 ± 179.92. The proportions of white, black, and other races were 84.9%, 10.1%, 4.7%, respectively. Follow-up data revealed that 5,327(39.5%) of patients survived, and 8,164(60.5%) of patients had died. Average survival time was 11.31 ± 11.35 months.

In the comparison of the results from the training population with that from the validation population, no significant differences were observed according to sex, age, tumor site, tumor size, the organ metastasis, number<sup>number</sup>PLN and NLN (Table 1).

Table 1  
Baseline Characteristics of Subjects in the Training and Validation Data.

<b>Prognostic factors</b>	<b>Total data (n = 13492)</b>	<b>Training data (n = 6746)</b>	<b>Validation data (n = 6746)</b>	<b>t/<math>\chi^2</math></b>	<b>P</b>
Age in years (years)	66.70 ± 11.14	66.69 ± 11.04	66.70 ± 11.24	-0.036	0.971
CS tumor size(mm)	65.72 ± 184.05	63.24 ± 177.98	68.20 ± 189.90	-1.563	0.118
CS tumor extension(mm)	366.58 ± 179.92	368.29 ± 178.75	364.89 ± 181.08	1.108	0.309
Regional nodes positive	0.26 ± 1.29	0.25 ± 1.25	0.27 ± 1.32	-0.811	0.417
Regional nodes negative	3.01 ± 7.50	3.06 ± 7.62	2.97 ± 7.38	0.692	0.489
Male sex (n,%)	10776 (79.9)	5400(80.0)	5376 (79.7)	0.266	0.606
<b>Race</b>				0.4697	0.925
White (n,%)	11449 (84.9)	5719 (84.8)	5730 (84.9)		
Black (n,%)	1301(9.7)	649 (9.6)	652(9.7)		
Other (n,%)	694 (5.1)	355 (5.3)	339 (5.0)		
Unknown (n,%)	48 (0.7)	23(0.3)	25(0.4)		
<b>Primary Site</b>				0.735	0.865
Cervical segment (n,%)	280 (2.3)	145 (2.4)	135(2.2)		
Chest segment (n,%)	11413 (92.6)	5714(92.6)	5699(92.5)		
Abdominal segment (n,%)	83 (0.7)	41(0.7)	42 (0.7)		
Cross-section (n,%)	554 (4.5)	270 (4.4)	284 (4.6)		
<b>Grade classification</b>				2.851	0.583
Grade I (n,%)	662 (4.9)	319 (4.7)	343(5.1)		
Grade II (n,%)	4361 (32.3)	2179 (32.3)	2182 (32.3)		
Grade III (n,%)	5532 (41.0)	2797(41.5)	2735 (40.5)		
Grade IV (n,%)	182 (1.3)	96 (1.4)	86 (1.3)		
<b>AJCC Group</b>				4.804	0.187
I (n,%)	2728 (20.2)	1351 (19.5)	1413 (20.9)		
II (n,%)	2414 (17.9)	1205 (17.9)	1209 (17.9)		

<b>Prognostic factors</b>	<b>Total data (n = 13492)</b>	<b>Training data (n = 6746)</b>	<b>Validation data (n = 6746)</b>	<b>t/<math>\chi^2</math></b>	<b>P</b>
III (n,%)	3531 (26.2)	1789(26.5)	1742 (25.8)		
IV (n,%)	4818 (35.7)	2437 (36.1)	2318 (35.3)		
<b>Radiation sequence with surgery</b>				1.901	0.863
No radiation and/or cancer-directed surgery (n,%)	11049 (81.9)	5517 (81.8)	5532(82.0)		
Radiation prior to surgery (n,%)	1764 (13.1)	883(13.1)	881 (13.1)		
Radiation after surgery (n,%)	598 (4.4)	304 (4.5)	294 (4.4)		
Radiation before and after surgery (n,%)	60 (0.4)	30 (0.4)	30(0.4)		
Surgery both before and after radiation (n,%)	10 (0.1)	7(0.1)	3 (0.1)		
Sequence unknown, but both were given (n,%)	9 (0.1)	4 (0.1)	5 (0.1)		
<b>CS mets at DX-bone</b>				1.358	0.244
No (n,%)	12067 (91.4)	6007(91.2)	6060 (91.7)		
Yes (n,%)	1130 (8.6)	583 (8.8)	547 (8.3)		
<b>CS mets at DX-brain</b>				1.386	0.239
No (n,%)	12920 (98.2)	6447 (98.0)	6473 (98.3)		
Yes (n,%)	242 (1.8)	130 (2.0)	112 (1.7)		
<b>CS mets at DX-liver</b>				1.589	0.207
No (n,%)	11021 (83.4)	5476 (82.9)	5545 (83.8)		
Yes (n,%)	2201 (16.6)	1126 (17.1)	1075 (16.2)		
<b>CS mets at DX-lung</b>				0.494	0.482
No (n,%)	11751 (89.4)	5860(89.3)	5891 (89.6)		
Yes (n,%)	1386 (10.6)	705(10.7)	681 (10.4)		
<b>Outcome</b>				2.157	0.142
Srvvival	5327 (39.5)	2622 (38.9)	2705 (40.1)		
Death	8164 (60.5)	4124 (61.1)	4040 (59.9)		

Prognostic factors	Total data (n = 13492)	Training data (n = 6746)	Validation data (n = 6746)	$t/\chi^2$	$P$
Survival time (month)	11.31 ± 11.35	11.26 ± 11.31	11.35 ± 11.38	-0.456	0.648

### The Parameters Identified in the Results of Cox Proportional Hazard Regression Analysis

The univariate analysis revealed that sex, race, age, tumor site, tumor size, tumor length, pathological grade of tumor, AJCC stage, the post-operation treatment, the extent of esophageal tumor metastasis, and  $\text{number}_{\text{NLN}}$  were all independent prognostic factors. The result of the multivariate analysis demonstrated that  $\text{number}_{\text{PLN}}$  was also an independent prognostic factor in addition to the above parameters (Table 2).

Table 2  
Univariate and multivariate for overall survival: Cox proportional hazard regression model

<b>Prognostic factors</b>	<b>p value (univariate)</b>	<b>p value (multivariate)</b>	<b>HR (95%CI)</b>
CS extension(continuous)	0.000	0.000	1.001(1.001 ~ 1.002)
Regional nodes positive(continuous)	0.005	0.000	1.064(1.036 ~ 1.094)
Regional nodes negative(continuous)	0.000	0.000	0.962(0.955 ~ 0.970)
Grade(I, II, III,IV)	0.000	0.000	1.276 (1.193 ~ 1.364)
AJCC Group(I, II, III,IV)	0.000	0.000	1.225(1.159 ~ 1.284)
Radiation sequence with surgery	0.000	0.000	
Sequence unknown, but both were given			1.000
No radiation and/or cancer-directed surgery			1.880(0.264 ~ 13.360)
Radiation prior to surgery			0.990(0.525 ~ 0.897)
Radiation after surgery			0.813(0.139 ~ 7.069)
Radiation before and after surgery			1.203(0.167 ~ 8.645)
Surgery both before and after radiation			1.355(0.175 ~ 21.798)
CS mets at DX-bone	0.000	0.000	
No			1.000
Yes			1.506(1.283 ~ 1.767)
CS mets at DX-brain	0.000	0.000	
No			1.000
Yes			1.422(1.013 ~ 1.997)
CS mets at DX- liver	0.000	0.000	
No			1.000
Yes			1.357(1.187 ~ 1.552)
CS mets at DX-lung	0.000	0.000	
No			1.000
Yes			1.405(1.206 ~ 1.636)

## The Modified N Stage Proposed In This Study

According to the results of the Cox proportional hazard model, the HR of PLN was 1.064, and the HR of NLN was 0.962. The distance between the HR of PLN and the statistical standard point (which was 1) was 0.064 ( $\Delta\text{HP}_{\text{positive}}$ ), and the distance between the HR of NLN and the statistical point was 0.038 ( $\Delta\text{HP}_{\text{negative}}$ ). The ratio of  $\Delta\text{HP}_{\text{positive}}$  to  $\Delta\text{HP}_{\text{negative}}$  (N ratio) was used as a coefficient to produce the re-adjusted number of PLN and NLN.

The analysis result of the minimum P value method indicated that the following ranges for the modified N stage were an appropriate solution:

for N0 stage, the re-adjusted N ratio = 0;

for N1 stage, the re-adjusted N ratio = (0-0.08];

for N2 stage, the range of rate was (0.08–0.63];

for N3, the re-adjusted N ratio (0.63,+∞]

In order to calculate the ratio of the re-adjusted number of PLN to NLN in special situations, the authors settled on the following two definitions of the ratio:

- 1) When  $\text{number}_{\text{PLN}}$  and NLN were both 0, the ratio of the readjusted number of PLN to NLN (re-adjusted N ratio) was defined as 0;
- 2) When one of  $\text{number}_{\text{PLN}}$  and  $\text{number}_{\text{NLN}}$  is 0, 0 was defined as 0.0001.

## The Feasibility And Superiority Of The Modified N Stage

A cross-validation study was performed on the modified N stage. The modified N stage was developed from the training population and validated using the validation population. The log-rank test indicated that significant survival differences were confirmed among the N1, N2 and N3 sub-groups of patients in the training population, and the survival difference could be replicated in the validation population using the Kaplan-Meier analysis ( $P < 0.05$ , Fig. 1).

The survival analysis on all the patients using the method of N stage coming from AJCC tumor stage system, the log-rank test indicated that significant survival differences were confirmed among the N1, N2 and N3 sub-groups of all patients. Log-rank value of modified N stage is 304.7, the traditional AJCC N stage is 17.03, the K-M curve interval of modified N stage N1, N2, N3 is wider than the traditional AJCC N stage (Fig. 2).

The comparison between the modified N stage and the AJCC N stage was performed using the ROC method. The result of ROC analysis revealed that the area under AJCC N stage curve was 0.934, and the area under modified N stage curve was 0.956, which indicated that the modified N stage could represent the N stage of EC more accurately (Fig. 3). The results show that the area under modified N stage curve

was 0.956, the log-rank value was 304.7, which was greater than 0.934 of the area under AJCC N stage curve and log-rank value 17.03 of AJCC N stage, and the K-M curve interval of modified N stage N1, N2, N3 is wider than the traditional AJCC N stage, it is considered that all the patients using the modified N stage method were relative accurate in predicting patients outcome compare with AJCC N stage.

## Discussion

This study investigates the feasibility of a modified N stage which is based on combination analysis including the number of positive LN and negative LN in the meantime. The results show that the area under modified N stage curve was 0.956, the log-rank value was 304.7, which was greater than 0.934 of the area under AJCC N stage curve and log-rank value 17.03 of AJCC N stage, and the K-M curve interval of modified N stage N1, N2, N3 is wider than the traditional AJCC N stage, it is considered that all the patients using the modified N stage method were relative accurate in predicting patients outcome compare with AJCC N stage.

The results of recent studies have shown that patients with EC might not benefit from the extensive lymphadenectomy because the parameters of the tumor metastasis site may not be stable [12, 13]. The site would very likely be affected by the extent of operation, whereas  $\text{number}_{\text{PLN}}$  collected in lymphadenectomy would not be. Because the PLN was usually enlarged, the surgeon was likely to notice it during lymphadenectomy and would remove it. This means  $\text{number}_{\text{PLN}}$  would remain stable regardless of the individual extent of lymphadenectomy and the practice of different surgeons.

Research has indicated that the NLN number could represent the extent of lymphadenectomy in patients with EC, therefore, more NLN removed meant a better prognosis [14]. However, it was controversial how many NLN should be removed in the lymphadenectomy to achieve a better prognosis. Greenstein advised that 18 NLN should be removed in the lymphadenectomy to obtain a better outcome for patients with EC, while 19 NLN was suggested by another study [9, 10]. Baba suggested that 31 NLN should be resected at least in the lymphadenectomy, however, this study found that 31 NLN should be resected only on patients who experienced the three-field dissection for the lymphadenectomy [15]. This finding indicated that  $\text{number}_{\text{NLN}}$  was an important factor in the prognosis of patients with EC.

However, the result of the above studies on the advised resected number of NLN was slightly different. The reason for this might be that  $\text{number}_{\text{NLN}}$  was easily influenced by the confounders than others. Although  $\text{number}_{\text{NLN}}$  could indicate the extent of lymphadenectomy and could reflect the site of tumor metastasis to some extent, a stable cut-off number of NLN removed in lymphadenectomy indicating a better prognosis were not replicated in this study.

The ratio of  $\text{number}_{\text{PLN}}$  to  $\text{number}_{\text{NLN}}$  or the ratio of  $\text{number}_{\text{PLN}}$  to the total number of LN removed in lymphadenectomy could be used to explore the cut-off number which differentiates patients into sub groups with different outcome. This cut-off number could be used for proposing the modified N stage. Dhar first reported the ratio of  $\text{number}_{\text{PLN}}$  to the total number of LN as a prognostic factor in EC in

2000[16]. Mariette showed that the ratio of  $\frac{\text{number of PLN}}{\text{total number of LN}}$  to the total number of LN was a strong independent prognosis factor[17]. The above studies demonstrated that the ratio  $\frac{\text{number of PLN}}{\text{total number of LN}}$  to the total number of LN was as important as  $\frac{\text{number of PLN}}{\text{total number of LN}}$  regardless of the extent of the lymphadenectomy and the application of neoadjuvant chemoradiation. However, what the best cut-off number was for the ratio of  $\frac{\text{number of PLN}}{\text{total number of LN}}$  to the total number of LN remained controversial. Several studies proposed 0.2 as the cut-off number for the ratio in their modified N stage regimens, while other studies concluded that the cut-off number for the ratio should be 0.3[18–21]. In the meantime, Tan suggested that 0.25 might be a more appropriate cut-off number for the ratio than 0.35, which was identified in Shao's research [22, 23]. The above results indicated that the ratio of  $\frac{\text{number of PLN}}{\text{total number of LN}}$  to the total number of LN failed to consistently predict the prognosis for patients with EC. The reason might be that none of the above research compared the relative impact of PLN, NLN, and total LN removed in lymphadenectomy on the general prognosis, but they simply used the ratio between them to explore the modified N stage. The criterion of the above-modified N stage would be affected by the research cohort or the proportion of patients.

This study proposed the cut-off ratio of the PLN count to the NLN count based on the results of the Cox proportional hazard model. The procedure in this study was more reasonable than those procedures which directly explored the ratio of the PLN count to the NLN count. Furthermore, the cut-off ratio proposed in this study has been further confirmed using the cross-validation method on cohort data from the SEER database.

The N stage introduced by the 7th AJCC was a regular criterion. In predicting prognoses, the priority 7th AJCC was usually compared by procedures of modified N stage recently. The survival analysis in this study confirmed that the survival line of subgroups from the modified N stage separated more significantly than that of the N stage of 7th AJCC. Furthermore, the result of the ROC curve analysis demonstrated the superiority of this modified N stage system, which in turn supported the assumption of this study: the relative impact of PLN and NLN on the prognosis based on the results of the Cox proportional hazard model should be considered in the modified N stage.

It was widely accepted that tumor differentiation,  $\frac{\text{number of PLN}}{\text{total number of LN}}$  and NLN, the tumor stage of 7th AJCC, and organ metastasis were all independent prognostic factors [24, 25]. This study confirmed that finding. Studies also showed that the age of patients was also a prognostic factor [26, 27]. This study confirmed the finding as well. However, based on the analysis result of the Cox proportional hazard model, age only had a small impact on a patient's prognosis. This result implied that the impact of age on prognosis would only be noticed when using a big cohort. This finding was consistent with our previous research[28].

A recent report revealed that EC patients with organ metastasis or distant metastasis in bone had the worst prognosis than others[5]. In this study, patients with tumor metastasis in bone had a worse outcome than those with tumor metastasis in the lungs or liver, but had a better outcome than patients with tumor metastasis in the brain. This finding was consistent with the report which showed that brain

metastatic tumors with a primary tumor located in the esophagus only had a mean survival time of six months[29].

Although the current study proposed a reasonable modified N stage for EC, it has several limitations. First, the study was retrospective; its results may be affected by confounding factors that were not controlled for. Second, although the SEER database was prepared according to strict criteria, the data was collected from multiple research centers with different operational habits. As a result, some differences in findings may be due to differences in research center practices. Third, the sample in this study consisted of different pathological types of EC. Because the data did not include which type of EC, this study could not determine whether there was a difference between the application of this modified N stage system or in the ESCC and SCC sub-cohorts.

In summary, based on the results of the Cox proportional hazard model, the study proposed a modified N stage derive from the N stage system of 7th AJCC for EC. The study also identified the reasonability and superiority of the modified N stage using the cross-validation method comparing to the N stage system of 7th AJCC. This modified N stage system is a promising step toward more accurately identifying the N stage of EC and in turn, providing more effective treatment for this devastating disease.

## **Declarations**

### **Ethical Approval and Consent to participate**

Because the data was collected from the Surveillance, Epidemiology, and End Results (SEER) , database, therefore it doesn't involve ethical approval and consent to participate.

### **Consent for publication**

We are all consent for publication

### **Funding sources**

This study was supported by Shandong Provincial Medical and Health Development plan (2013WSA13018), Natural Science Foundation of Shandong Province (ZR2014HL062).

### **Availability of supporting data**

The data sets supporting the results of this article are included within the article and its additional files.

### **Conflicts of interest**

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

### **Author Contributions**

Jinling Zhang: conceived, design the study, and write the paper

Hongyan Li: analyzed the data and write the paper

Liangjian Zhou: assist in the research

Lianling Yu: assist in the research

Fengyuan Che: assist in the research

Xueyuan Heng: overall command the research

### **Authors' information**

Jinling Zhang, Cancer center of Linyi people's hospital, Shandong University, School of medicine, Linyi 276000, Shandong Province, People's Republic of China. Email: [jinlingzhang\\_931@yahoo.com](mailto:jinlingzhang_931@yahoo.com)

Hongyan Li, Department of Central Laboratory, Linyi people's hospital, Shandong University, School of medicine,, Linyi 276000, Shandong Province, People's Republic of China. Email: 15153913680@qq.com

Liangjian Zhou, Cancer center of Linyi people's hospital, Shandong University, School of medicine,, Linyi 276000, Shandong Province, People's Republic of China. Email:2100997128@qq.com

Lianling Yu, Cancer center of Linyi people's hospital, Shandong University, School of medicine,, Linyi 276000, Shandong Province, People's Republic of China. Email: [yll66\\_@126.com](mailto:yll66_@126.com)

Author Contributions: Assist in research

Fengyuan Che, Department of Central Laboratory, Linyi people's hospital, Shandong University, School of medicine,, Linyi 276000, Shandong Province, People's Republic of China. Email:che1971@126.com

Author Contributions: Assist in research

Xueyuan Heng , Cancer center of Linyi people's hospital, Shandong University, School of medicine,

Linyi 276000, Shandong Province, People's Republic of China. Tel./fax: 0539-8071229. Email: [xueyuanheng@yahoo.com](mailto:xueyuanheng@yahoo.com)

Author Contributions: Overall command the research.

### **Acknowledgements**

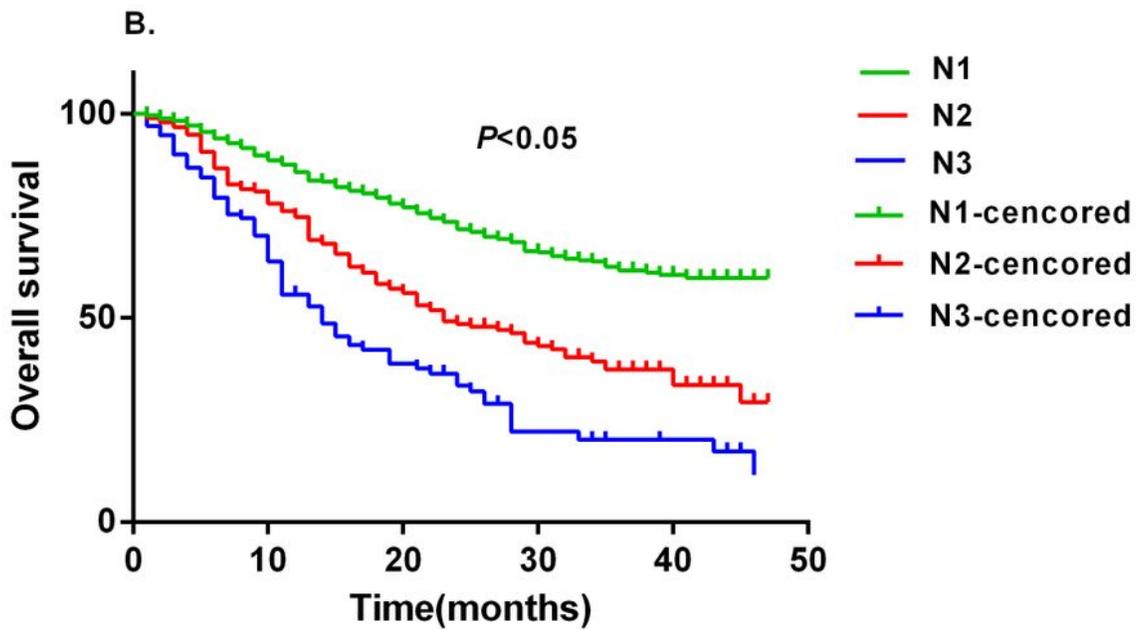
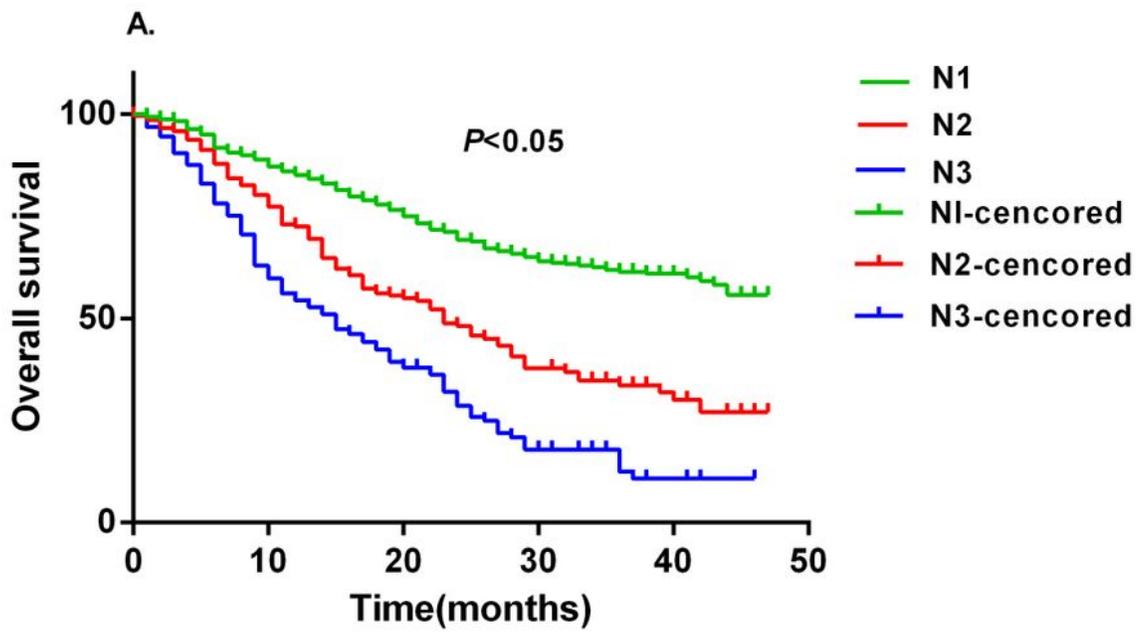
This study was supported by Shandong Provincial Medical and Health Development plan (2013WSA13018), Natural Science Foundation of Shandong Province (ZR2014HL062).

## **References**

1. Sohda M, Kuwano H. Current Status and Future Prospects for Esophageal Cancer Treatment. *Annals of thoracic and cardiovascular surgery: official journal of the Association of Thoracic Cardiovascular Surgeons of Asia*. 2017;23:1–11.
2. Rice TW, et al. Esophageal Cancer: Associations With (pN+) Lymph Node Metastases. *Annals of surgery*. 2017;265:122–9.
3. Kajiyama Y [New Japanese Classification of Esophageal Cancer (11th Edition)]. *Gan to kagaku ryoho Cancer & chemotherapy* 2016;43:1049-52.
4. Oweira H, et al. Validation of the eighth clinical American Joint Committee on Cancer stage grouping for esophageal cancer. *Future oncology*. 2018;14:65–75.
5. Wu SG, et al. Sites of metastasis and overall survival in esophageal cancer: a population-based study.
6. Chen J, et al. A new clinical staging system for esophageal cancer to predict survival after definitive chemoradiation or radiotherapy. *Diseases of the esophagus: official journal of the International Society for Diseases of the Esophagus* 2018.
7. Rice TW, et al. Recommendations for clinical staging (cTNM) of cancer of the esophagus and esophagogastric junction for the 8th edition AJCC/UICC staging manuals.
8. Kunisaki C, et al. Impact of lymph-node metastasis site in patients with thoracic esophageal cancer. *Journal of surgical oncology*. 2010;101:36–42.
9. Greenstein AJ, et al. Effect of the number of lymph nodes sampled on postoperative survival of lymph node-negative esophageal cancer. *Cancer*, 2008, 112(6):1239–1246.
10. Hsu PK, et al. The prognostic value of the number of negative lymph nodes in esophageal cancer patients after transthoracic resection. *Ann Thorac Surg*. 2013;96(3):995–1001.
11. Ma G, et al. A novel multivariate scoring system for determining the prognosis of lymph node-negative esophageal squamous cell carcinoma following surgical therapy: an observational study. *Eur J Surg Oncol*. 2015;41(4):541–7.
12. Hagens ERC, et al. The extent of lymphadenectomy in esophageal resection for cancer should be standardized. *J Thorac Dis*. 2017;9(S8):713.
13. Phillips AW, et al. Impact of Extent of Lymphadenectomy on Survival, Post Neoadjuvant Chemotherapy and Transthoracic Esophagectomy. *Ann Surg*. 2017;265(4):750–6.
14. Yang HX, et al. An evaluation of the number of lymph nodes examined and survival for node-negative esophageal carcinoma: data from China. *annals of surgical oncology*. 2010;17(7):1901–11.
15. Baba Y, et al. Negative lymph-node count is associated with survival in patients with resected esophageal squamous cell carcinoma. *Surgery*. 2013;153(2):234–41.
16. Dhar DK, et al. Appraisal of a revised lymph node classification system for esophageal squamous cell cancer. *annals of thoracic surgery*. 2007;83(4):1265–72.
17. Mariette C, et al. The number of metastatic lymph nodes and the ratio between metastatic and examined lymph nodes are independent prognostic factors in esophageal cancer regardless of

- neoadjuvant chemoradiation or lymphadenectomy extent. *Ann Surg.* 2008;247(2):365–71.
18. Bhamidipati CM, et al. Pathologic lymph node ratio is a predictor of survival in esophageal cancer. *Ann Thorac Surg.* 2012;94(5):1643–51.
  19. Hsu WH, et al. The metastatic lymph node number and ratio are independent prognostic factors in esophageal cancer. *J Gastrointest Surg.* 2009;13(11):1913.
  20. Feng Jf F, - Huang Y, et al. Prognostic analysis of esophageal cancer in elderly patients: metastatic lymph node ratio versus 2010 AJCC classification by lymph nodes. *World J Surg Oncol.* 2013;11:162.
  21. He Z, et al. Use of the metastatic lymph node ratio to evaluate the prognosis of esophageal cancer patients with node metastasis following radical esophagectomy. *PLoS One.* 2013;8(9):e73446.
  22. Tan Z, et al. Can lymph node ratio replace pn categories in the tumor-node-metastasis classification system for esophageal cancer? *Journal of Thoracic Oncology.* 2014;9(8):1214–21.
  23. Shao Y, et al. Assessment of Lymph Node Ratio to Replace the pN Categories System of Classification of the TNM System in Esophageal Squamous Cell Carcinoma. *Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer.* 2016;11(10):1774–84.
  24. Akutsu Y, et al. The Prevalence of Overall and Initial Lymph Node Metastases in Clinical T1N0 Thoracic Esophageal Cancer: From the Results of JCOG0502, a Prospective Multicenter Study. *Annals of surgery.* 2016;264:1009–15.
  25. Talsma K, et al. Impact of Neoadjuvant Chemoradiation on Lymph Node Status in Esophageal Cancer: Post hoc Analysis of a Randomized Controlled Trial. *Annals of surgery.* 2017;266:e52-e3.
  26. Cheng YFA-Ohoo, et al. Esophageal squamous cell carcinoma and prognosis in Taiwan. *LID – 10.1002/cam4.1499.*
  27. Miyata H, et al. Clinical Outcome of Esophagectomy in Elderly Patients With and Without Neoadjuvant Therapy for Thoracic Esophageal Cancer. *Annals of Surgical Oncology,* 2015:794–801.
  28. Cheng J, et al. Explore the radiotherapeutic clinical target volume delineation for thoracic esophageal squamous cell carcinoma from the pattern of lymphatic metastases. *J Thorac Oncol.* 2013;8(3):359–65.
  29. Rades D, et al. Predicting overall survival in patients with brain metastases from esophageal cancer. *Anticancer Res.* 2014;34(11):6763–5.

## Figures



**Figure 1**

The comparison of overall survival between training group and validation group. A: The survival difference among N1, N2 and N3 sub group in training data were significant( $P \leq 0.05$ ). B: The survival difference among N1, N2 and N3 sub group in validation data were significant( $P \leq 0.05$ ).

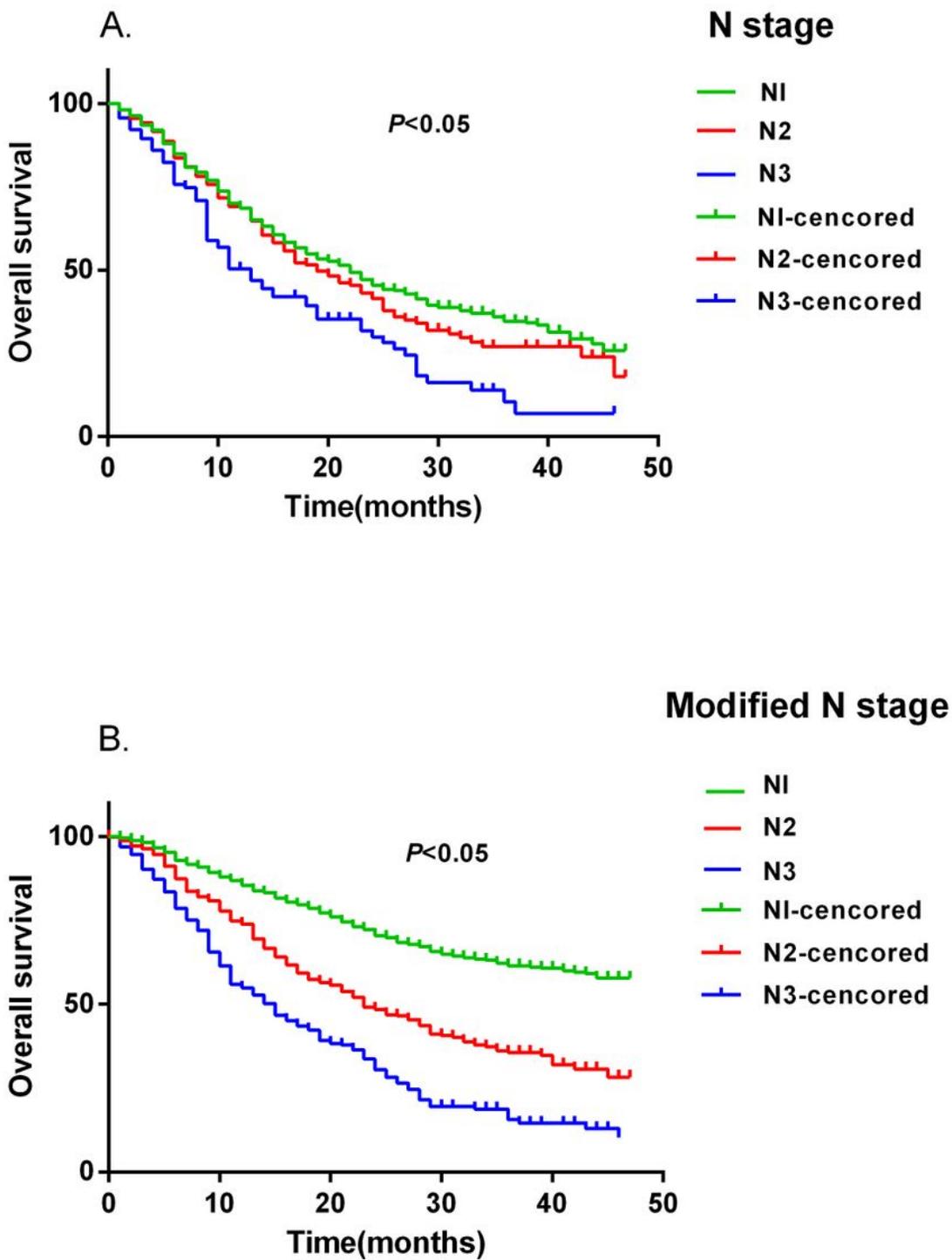
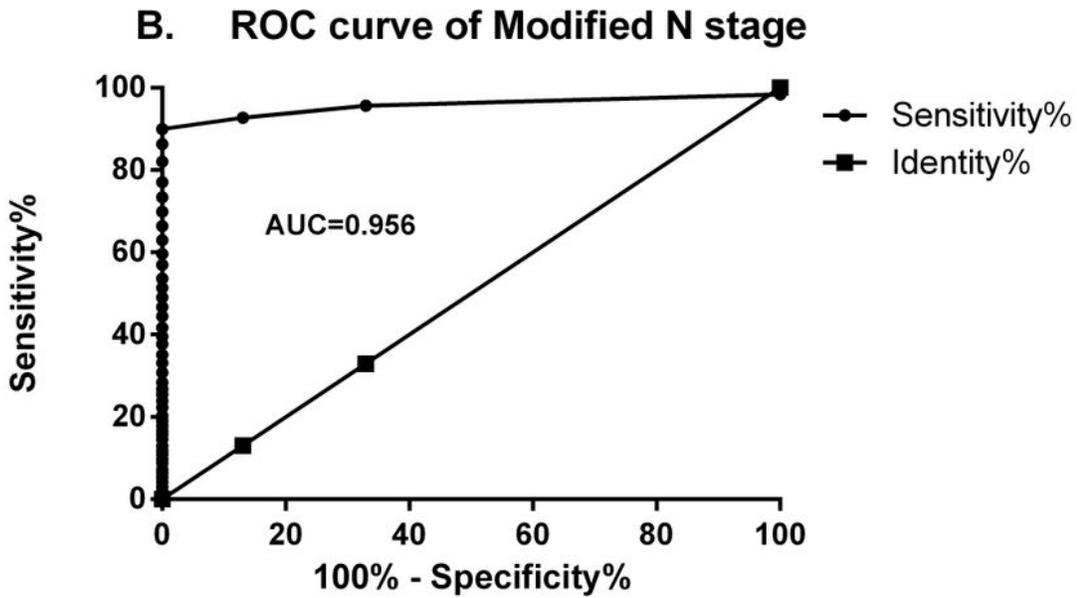
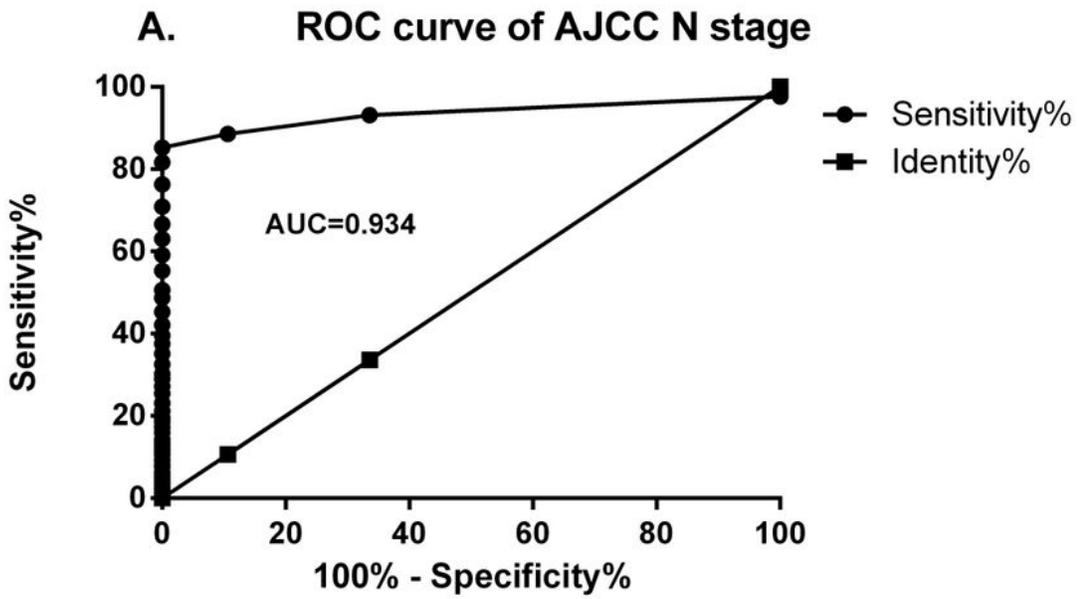


Figure 2

The comparison of survival analysis between the method using N stage of AJCC and the modified N stage, respectively. A: The survival analysis on all the patients using the method of N stage coming from AJCC tumor stage system. B: The survival analysis on all the patients using the modified N stage method.



**Figure 3**

The ROC analysis on the AJCC N stage and modified N stage. The result of ROC analysis revealed that the area under AJCC N stage curve was 0.934, the area under modified N stage curve was 0.956 .