

Accuracy of ultrasound for the diagnosis of cervical lymph node metastasis in esophageal cancer: a systematic review and meta-analysis

Xue-Feng Leng^{1*}, Yi Zhu^{2*}, Ge-Ping Wang¹, Jian Jin¹, Lei Xian³, Yu-Hong Zhang⁴

¹Department of Cardiothoracic Surgery, the Affiliated Hospital of Chengdu University, Chengdu 610081, China; ²Department of Ultrasound, Sichuan Cancer Hospital, Chengdu 610041, China; ³Department of Thoracic Surgery, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China; ⁴Department of Gastroenterology, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China

Contributions: (I) Conception and design: XF Leng, Y Zhu, L Xian; (II) Administrative support: L Xian, J Jin; (III) Provision of study materials or patients: XF Leng, Y Zhu, GP Wang; (IV) Collection and assembly of data: XF Leng, Y Zhu, J Jin, GP Wang; (V) Data analysis and interpretation: XF Leng, Y Zhu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*These authors contributed equally to this work.

Correspondence to: Lei Xian, MD, PhD. Department of Thoracic Surgery, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China. Email: xianlei59@163.com.

Background: Esophageal cancer is considered a serious malignancy with respect to its prognosis and mortality rate. Cervical lymph node status is one of the keys to determining prognosis and treatment methods. However, published data vary regarding the accuracy of ultrasound in the diagnosis of cervical lymph node metastasis. We performed a meta-analysis to assess the efficacy of ultrasound for detecting cervical lymph node metastasis in patients with esophageal cancer.

Methods: The PubMed/MEDLINE, EMBASE, Web of Science, and Cochrane Library databases were searched to identify studies related to cervical lymph node metastasis, and 22 studies comprising 3,513 patients met our inclusion criteria. We used a bivariate meta-analysis following a random effects model to summarize the data. We also explored reasons for statistical heterogeneity using meta-regression, subgroup, and sensitivity analyses. Publication bias was assessed with a Deeks funnel plot.

Results: The area under the receiver operating characteristic curve was 0.97 [95% confidence interval (CI): 0.95–0.98], and the pooled diagnostic odds ratio was 121.00 (95% CI: 47.57–307.79). With cut-off values of 5 mm and >5 mm for cervical lymph node size, the sensitivities and specificities (95% confidence interval) for ultrasound detection of cervical lymph node metastasis were 84% (67–93%) and 93% (90–95%); and 94% (76–98%) and 98% (89–100%), respectively.

Conclusions: We show for the first time the diagnostic accuracy of ultrasound for predicting cervical lymph node-positive metastasis in esophageal cancer. Our analysis shows that ultrasonography may be an effective and reliable approach to detect cervical lymph node metastasis in esophageal cancer. However, to accommodate heterogeneity, high-quality studies are needed to further verify the efficacy of ultrasound detection.

Keywords: Esophageal cancer; cervical lymph node; ultrasonography

Submitted Apr 26, 2016. Accepted for publication Jun 12, 2016.

doi: 10.21037/jtd.2016.07.71

View this article at: <http://dx.doi.org/10.21037/jtd.2016.07.71>

Introduction

Esophageal cancer is considered a serious malignancy with respect to its prognosis and mortality rate. An estimated 455,800 new cases of esophageal cancer and 400,200 esophageal cancer-related deaths occurred in 2012 worldwide (1). High rates of recurrence and metastasis cause high mortality among patients with esophageal cancer (2).

Surgical resection has been well established as a part of multimodal treatment for esophageal cancer. The number of positive lymph nodes (LNs) removed is an independent predictor of survival in esophageal cancer after esophagectomy (3,4). Fortunately, the surgical dissection of cervical lymph nodes (CLNs) and three-field LNs contribute to a better prognosis and prolonged survival for patients with esophageal cancer (5-7). One of the most important methods to identify whether three-field lymphadenectomy should be performed is an examination of CLN status.

Ultrasound (US), computed tomography (CT), and 18-fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) are used clinically to detect regional LNs. Nevertheless, the value of CT and FDG-PET scanning is limited (8), and there is an associated financial burden and risk of high exposure to radiation. Comparatively speaking, US is more convenient and economical for the detection of CLNs.

Multiple studies have identified the potential benefits of US in detecting CLN metastasis; however, results regarding the extent of its benefits have been inconsistent. The purpose of this study was to review the literature and perform a meta-analysis to assess the efficacy of US in the detection of CLN status in patients with esophageal cancer.

Methods

Search strategy

In accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (*Figure 1*) (9), we systematically searched the PubMed/MEDLINE, EMBASE, Web of Science, and Cochrane Library databases for relevant citations published up to October 2014 and without language restrictions. The following medical subject headings and free-text words were chosen to be used alone or in combination: “esophageal cancer”, “oesophageal cancer”, “esophageal neoplasm”, “esophagus neoplasm”, “esophagus cancer”, “ultrasound”, “ultrasonography”, “sonography”, “ultrasonic”,

“echography,” and “echotomography”.

Study selection

Articles judged as suitable on the basis of the title and abstract were identified following the electronic search strategy. In all, 4,712 articles and abstracts were identified through the initial Web of Science search, 2,519 articles were identified through the EMBASE search, and 1,445 articles were identified through the PubMed/MEDLINE search. A search of the Cochrane Library databases identified three articles. Then, we combined and retrieved detailed full-text assessments of potentially relevant studies. Two investigators (X.F.L. and Y.Z.) further evaluated the data independently based on predefined selection criteria. Disagreements between the investigators were solved by discussion. *Figure 1* displays how the studies were selected.

Inclusion criteria

The selection criteria were:

- (I) Population: adult patients with confirmed esophageal cancer (all types of histopathology and lesion locations, and not a case-control design);
- (II) Intervention: US equipment without fine-needle aspiration (FNA) was used for the independent diagnosis of CLNs in esophageal cancer;
- (III) Comparison: the diagnostic accuracy of US in determining CLN status was evaluated according to the reference standard of histopathology obtained by fine needle aspiration (FNA) and/or a surgical procedure;
- (IV) Outcomes: results for true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs) were extracted to construct 2×2 tables for each study. Further calculation was performed if these data had not reported directly;
- (V) Studies published in English language.

Data extraction

The details were extracted through selected articles by two investigators independently. Following categories included: first author, year of publication, published regions, type of study, diagnostic criteria, histological reference standard and methodology, number of patients enrolled, sensitivity, specificity, and so on. Accurate TP, FP, FN, and TN results were extracted to construct a 2×2 table for qualified

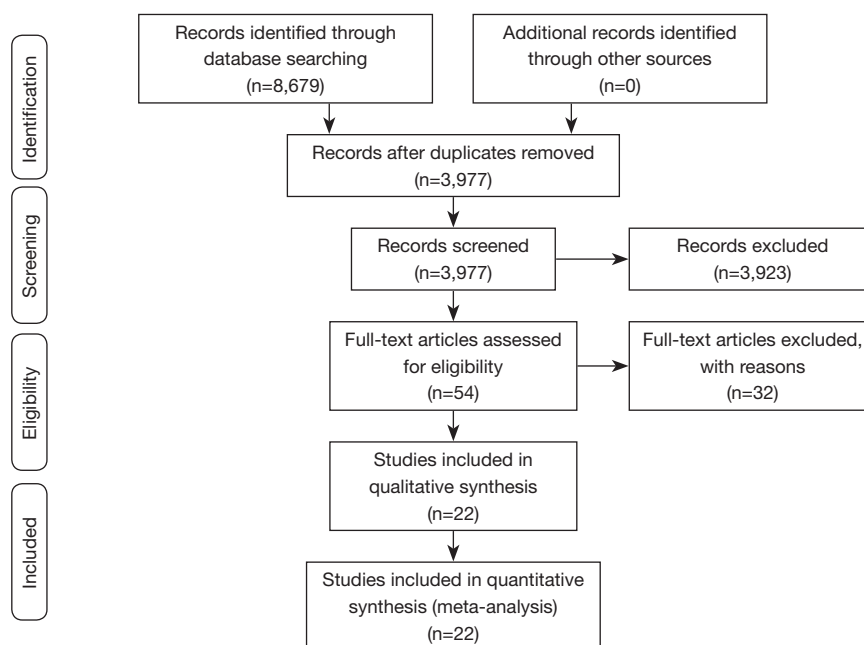


Figure 1 The flowchart of studies selection for the systematic reviews and meta-analysis (following PRISMA statement).

studies. RevMan 5.3 software (The Nordic Cochrane Centre, Rigshospitalet 2008) was used to calculate the not fully reported data. If these data were totally absent from the article, it was excluded from the meta-analysis. Any disagreements were resolved by mutual agreement.

Assessment of methodological quality

Two independent reviewers made the assessment by using the Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Reviews-2 (QUADAS-2) tool (10), which is recommended by the Cochrane Diagnostic Test Accuracy Working Group. There are two major parts of the QUADAS-2 tool, which are risks of bias and applicability concerns. The former further contains patient selection, index test, reference standard, and flow and timing. The latter covers contents of patient selection, index test, and reference standard. This procedure was performed through the RevMan 5.3 software (The Nordic Cochrane Centre, Rigshospitalet 2008).

Statistical analysis

We performed a bivariate meta-analysis with random effects model to calculate extracted data of sensitivity, specificity, likelihood ratios (LRs), and the diagnostic odds ratio (DOR).

Hierarchical summary receiver operating characteristic (HSROC) curve and the areas under the summary receiver operating characteristic (SROC) curve were constructed and calculated respectively to reflect synthesized diagnostic accuracy (11,12). Diagnostic accuracy was manifested by area under the curve (AUC) and indicator of Q^* (the point on the SROC curve at which sensitivity is equal to specificity). The more value of AUC is close to 100%, the more perfectly diagnostic accuracy it has (13).

The threshold effect can contribute to one of the primary causes of heterogeneity. In order to calculate and evaluate the threshold effect, the method of Spearman correlation coefficient between the logit of sensitivity and the logit of (1-specificity) was performed by Meta-DiSc version 1.4 (Hospital Ramón y Cajal, Madrid 2006) (14,15). Publication bias of diagnostic meta-analysis was assessed by Deeks funnel plot (16). Heterogeneity among the included studies was assessed using Cochran's Q statistics and the I^2 test. Normally, I^2 lies between 0% and 100%. If $I^2 > 50\%$, then there is greater heterogeneity among studies (17). A recommended method of meta-regression to detect heterogeneities can be done only when the number of included studies exceeded ten. A bivariate boxplot was used in a complementary fashion to evaluate heterogeneity. Subgroup analysis played a role to explore the potential affections of clinical and methodological heterogeneities.

Table 1 Characteristics of the eligible studies

| Studies | Year | Country | Cut-off (mm) | Lymph node location | Hospital level | Patients (n) | TP (n) | FP (n) | FN (n) | TN (n) |
|--------------------------------------|------|-------------------|--------------|---------------------|----------------|--------------|--------|--------|--------|--------|
| van Zoonen <i>et al.</i> (40) | 2012 | Netherlands | 6 | NL | Uni./Cen. | 193 | 14 | 1 | 0 | 178 |
| Blom <i>et al.</i> (39) | 2012 | Netherlands | 5 | NL and SCL | Uni./Cen. | 170 | 9 | 20 | 0 | 161 |
| Cwik <i>et al.</i> (A) (38) | 2011 | Poland | 5 | NL and SCL | Uni./Cen. | 26 | 17 | 4 | 0 | 9 |
| Cwik <i>et al.</i> (B) (37) | 2011 | Poland | 5 | NL and SCL | Uni./Cen. | 83 | 13 | 6 | 0 | 64 |
| Omluo <i>et al.</i> (36) | 2009 | Netherlands | 5 | NL and SCL | Uni./Cen. | 233 | 9 | 48 | 9 | 167 |
| Schreurs <i>et al.</i> (35) | 2008 | Netherlands | 5 | NL | Uni./Cen. | 136 | 11 | 3 | 1 | 121 |
| van Vliet <i>et al.</i> (A) (33) | 2007 | Netherlands | 7 | SCL | Uni./Cen. | 567 | 49 | 44 | 16 | 458 |
| van Vliet <i>et al.</i> (B) (34) | 2007 | Netherlands | 7 | SCL | Uni./Cen. | 546 | 49 | 9 | 4 | 484 |
| van Vliet <i>et al.</i> (32) | 2006 | Netherlands | 10 | SCL | Regional | 153 | 5 | 0 | 14 | 134 |
| Fang <i>et al.</i> (31) | 2003 | China | 5 | NL and SCL | Uni./Cen. | 42 | 9 | 0 | 2 | 31 |
| Griffith <i>et al.</i> (30) | 2000 | China (Hong Kong) | 7 | NL and SCL | Uni./Cen. | 121 | 38 | 15 | 0 | 68 |
| Nishimaki <i>et al.</i> (29) | 1999 | Japan | 5 | NL and SCL | Uni./Cen. | 248 | 46 | 6 | 71 | 125 |
| Natsugoe <i>et al.</i> (28) | 1999 | Japan | 6 | NL and SCL | Uni./Cen. | 153 | 38 | 6 | 13 | 96 |
| Shimizu <i>et al.</i> (27) | 1997 | Japan | 5 | NL | Regional | 31 | 1 | 0 | 2 | 28 |
| Bonvalot <i>et al.</i> (26) | 1996 | France | 5 | NL and SCL | Uni./Cen. | 50 | 13 | 1 | 6 | 30 |
| Matsubara <i>et al.</i> (25) | 1995 | Japan | 5 | NL | Regional | 179 | 2 | 13 | 11 | 153 |
| Tachimori <i>et al.</i> (24) | 1994 | Japan | 5 | NL and SCL | Uni./Cen. | 83 | 15 | 1 | 4 | 63 |
| Van Overhagen <i>et al.</i> (23) | 1993 | Netherlands | 5 | SCL | Uni./Cen. | 89 | 15 | 6 | 4 | 64 |
| van Overhagen <i>et al.</i> (22) | 1992 | Netherlands | 5 | SCL | Uni./Cen. | 135 | 22 | 12 | 0 | 101 |
| van Overhagen <i>et al.</i> (A) (20) | 1991 | Netherlands | 5 | SCL | Uni./Cen. | 37 | 7 | 3 | 0 | 27 |
| van Overhagen <i>et al.</i> (B) (21) | 1991 | Netherlands | 5 | SCL | Uni./Cen. | 100 | 14 | 8 | 2 | 76 |
| Tohnosu <i>et al.</i> (19) | 1989 | Japan | 5 | NL and SCL | Uni./Cen. | 156 | 29 | 8 | 8 | 111 |

NL, neck lymph node; SCL, subclavicular lymph node; Uni., medical university; Cen., medical center; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

Such factors included study region (Asia or Europe), research year (<2,000 or ≥2,000), research institute (a medical university/center or regional hospital), index test cut-off (5 mm or more than 5 mm for CLN size), involved CLN region (cervical or supraclavicular vs. cervical and supraclavicular), and the method of obtaining histology specimens (FNA or surgical resection). Sensitivity analysis was implemented to estimate the latent effect of each study.

Pre-test probabilities versus corresponding post-test probabilities were evaluated by using Fagan's nomogram analysis, which was underlay the summary sensitivity and specificity. And a LR scattergram was created to further evaluate the clinical utility (18). These methodologies mentioned above were analyzed by software of Stata version 12.0 (Stata CorpLP, USA) and Meta-DiSc version 1.4 (Hospital Ramón y Cajal, Madrid 2006).

Results

Characteristics of the included studies

The initial search identified 4,095 reference articles, and finally, 22 studies (N=3,513; range, 26–567) were met for criteria of this meta-analysis. All 22 studies were published as full-text articles in peer-reviewed journals (19–40). The flow diagram (PRISMA 2009) was summarized in *Figure 1*. All patients underwent US examination. The TP, FP, FN, and TN results and detailed characteristics of the eligible studies were shown in *Table 1*.

Quality assessment

The QUADAS-2 tool undertook the quality assessment. The qualities of eligible studies were summarized and

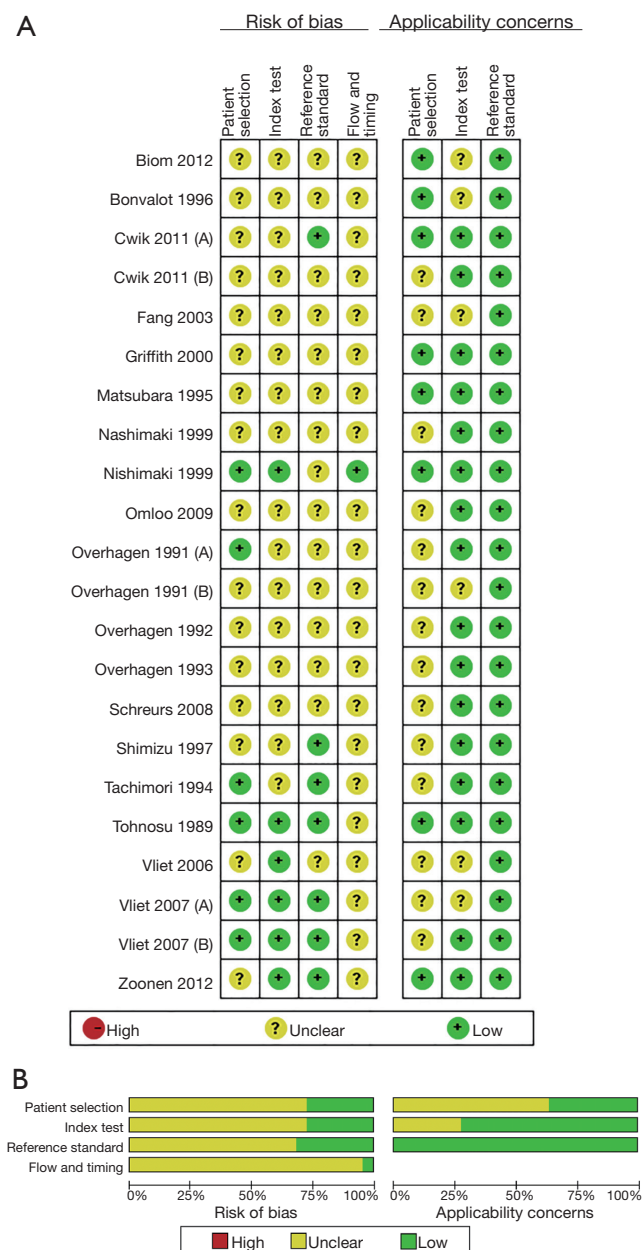


Figure 2 Risk of bias and applicability of included studies. (A) Summary of QUADAS-2 assessments of included studies; (B) proportion of studies with low, high, or unclear risk of bias.

displayed in *Figure 2*. And, on the whole, the included 22 studies met most of the quality criteria.

Diagnostic performance of US for CLNs

The corresponding SROC and HSROC curves were plotted

and are shown in *Figure 3*. The SROC curve illustrated an AUC of 0.97 [95% confidence interval (CI): 0.95–0.98] and the Q^* value was 0.9128. The results of the HSROC model displayed a beta = -0.509 (95% CI: -1.158 to 0.139), a Z value of -1.54 ($P=0.124$), and the manifested curve was symmetrical. In addition, the lambda value was 5.189 (95% CI: 4.033–6.346), which demonstrated that US had a high discrimination ability as a diagnostic methodology. An LR scattergram was plotted to further evaluate clinical application values. A diamond symbol located in the right upper quadrant (RUQ) indicates that a US examination could make a definite diagnosis [positive likelihood ratio (PLR) >10, negative likelihood ratio (NLR) >0.1] (*Figure 4A*).

We evaluated the pre-test probability and relevant post-test probability through a Fagan plot analysis underlay the pooled sensitivity and specificity (41). This reveals an estimation of the relationships among the pre-test probability, LR, and post-test probability. As *Figure 4B* shows, US examination had an 80% probability of correctly detecting a CLN following a positive test result when pre-test probabilities were 20%. This means that US could increase the probability of CLN detection by 60%, and it was enough to draw an accurate diagnosis when US results were positive.

The pooled DOR was 121.00 (95% CI: 47.57–307.79). Significant heterogeneity ($I^2=100.00\%$, $P<0.01$) was detected among all studies (*Figure 5*). Therefore, pooled sensitivities and specificities could not be assessed. Different diagnostic thresholds were analyzed instead. *Table 2* listed the following results, at a cut-off value of 5 mm for the estimated PLR and NLR of the test were 11.83 (95% CI: 8.06–17.36) and 0.18 (95% CI: 0.08–0.38), respectively. The estimated PLR and NLR at a cut-off value of >5 mm were 41.21 (95% CI: 8.08–210.14) and 0.07 (95% CI: 0.02–0.27), respectively.

Analysis of heterogeneity

Heterogeneity has played an important role to influence on the accuracy of the meta-analysis, and accordingly the exploration of heterogeneity is an indispensable procedure. First of all, we explored the threshold effect. The result manifested there was no statistic difference (Spearman's correlation coefficient = 0.303, $P=0.171$). Subsequently, we pooled a forest plot of the DOR to explore the non-threshold effect when threshold effect was not detected. Afterwards there was an obvious non-threshold effect (Cochran's $Q=91.60$, $P=0.00$). For further confirmation, a

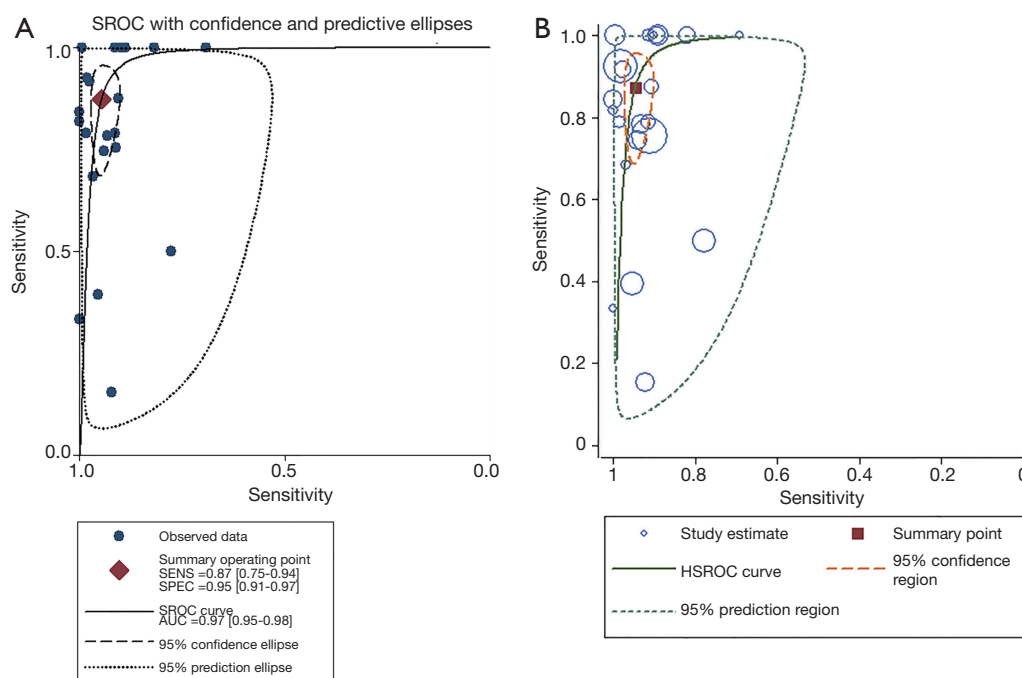


Figure 3 Summary receiver operating characteristic (SROC) and HSROC curves. (A) SROC graph with 95% confidence region and 95% prediction region for the diagnosis value. The solid circles represent the included studies. The square represents the summary estimate of sensitivity and specificity with the 95% confidence ellipse from the bivariate model; (B) hierarchical summary receiver operating characteristic (HSROC) curves for the diagnostic performance of ultrasound to predict CLN in ECA. The size of the blue circles indicate the weight in the studies. The summary sensitivity and specificity is shown with a square and the 95% confidence region is plotted in the short lines. The AUC was 0.97 (95% CI: 0.95–0.98). CI, confidence interval.

bivariate boxplot illustrated that some of these points were near the boundary of the loop and some were outliers. We also observed five points outside the fence. Further, it was demonstrated that a non-threshold effect existed (Figure 6A).

Under the meta-regression analysis in exploration of sources of heterogeneity, the relevant variations included study region, research year, research institute, index test cut-off, CLN regions, and method of obtaining histology specimens. On the basis of descending P values, which guided to remove and analyze the corresponding variables, the results implied that the primary cause of heterogeneity was the index test cut-off value ($P=0.0499$, 95% CI: 1.00–37.10). There were five studies (30,32–34,40) that used more than 5 mm as the cut-off value, and the remaining studies used 5 mm as the cut-off value. As the subgroup analysis demonstrated, the AUC was 0.99 (95% CI: 0.97–0.99) in the >5 mm group and 0.97 (95% CI: 0.95–0.98) in the 5 mm group. With cut-off values of 5 and >5 mm, the sensitivities and specificities (95% CI) for US detection of CLN metastasis were 84% (67–93%) and 93% (90–95%); and 94% (76–98%) and 98% (89–100%),

respectively. The use of diverse cut-off values could explain the observed heterogeneity.

A sensitivity analysis was implemented to estimate the latent influence upon diagnostic accuracy of US among each study. It revealed that the Vliet 2007 (B) (34) and Omloo 2009 (36) studies might be sources of heterogeneity in the meta-analysis. When these two outlier studies were removed, the AUC was still 0.97 (95% CI: 0.95–0.98) and the DOR was slightly elevated to 135.55 (95% CI: 51.35–357.82). This result inferred that the diagnostic accuracy was relatively stable. However, several design differences and factors among the studies might have produced the constant existence of heterogeneity.

There was no evidence of significant publication bias in this research followed by the description of Deeks funnel plot asymmetry test ($P=0.403$, Figure 6B).

Discussion

Esophagectomy has traditionally been the gold standard for

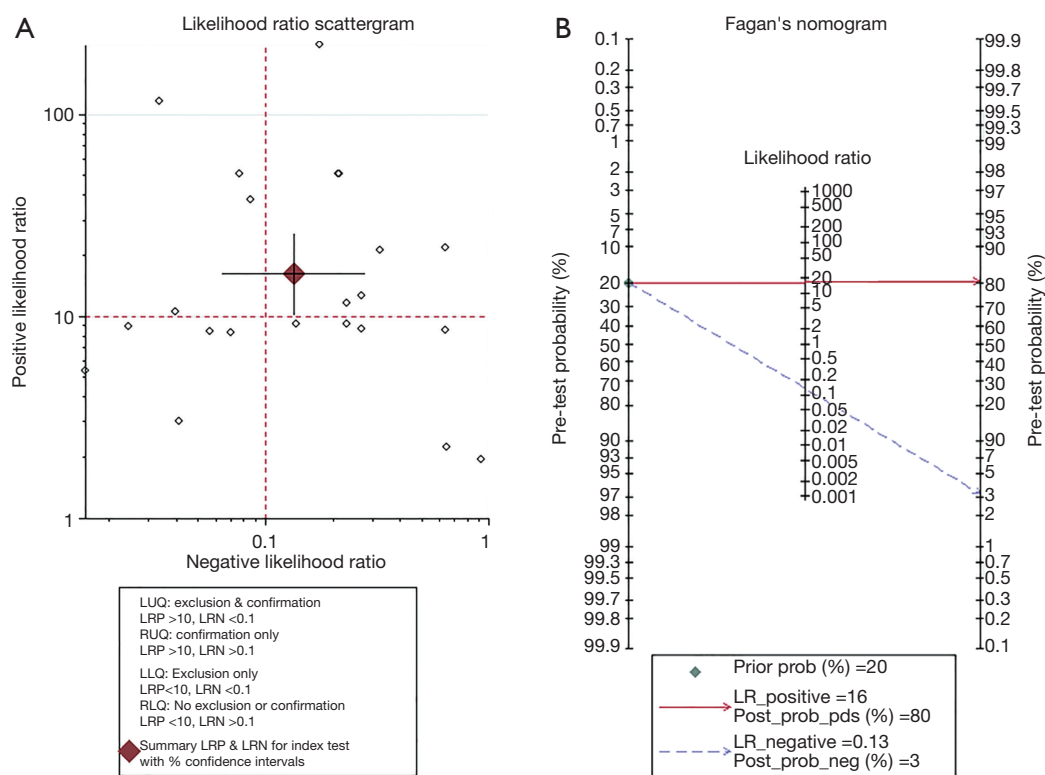


Figure 4 The likelihood ratio scattergram and Fagan plot. (A) The likelihood ratio scattergram to evaluate the clinical application values. The likelihood ratio scatters distribute to four quadrants. The diamond symbol located in RUQ (right upper quadrant) indicated the ultrasound can make a definite diagnosis (LRP>10, LRN>0.1); (B) fagan plot analysis to evaluate the clinical utility of ultrasonography for the detection of CLN in ECA. The Fagan plot consists of a vertical axis on the left with the pretest probability, an axis in the middle represents the likelihood ratio, and a vertical axis on the right represents the post-test probability. With a pretest probability of 20%, the post-test probability of CLN, given positive and negative ultrasound results, were 80% and 3%.

treating localized esophageal cancer. Radical esophagectomy and extensive node dissection have positive impacts on survival rates, particularly in patients with nodal metastases (42). The overall prevalence of CLN metastases has been documented as approximately 20% to 40%, regardless of the stage of the primary tumor (43,44). In the 7th edition of the esophageal AJCC (American Joint Committee on Cancer) staging guidelines, a regional LN has been redefined to include any paraesophageal node extending from the cervical nodes to the celiac nodes (45). Thus, the accurate staging of esophageal cancer is critical for both prognostic and therapeutic decisions, as well as in the evaluation of treatment results.

The sentinel node is the first lymphatic drainage area from the primary tumor, and could be the first site of micrometastasis. However, the lymphatic drainage system in the submucosa is very complex, and the presence of

nodal skip metastasis is common in thoracic esophageal carcinoma, especially in patients with tumors located in the middle and lower third of the esophagus (46). For patients who are diagnosed clinically as CLN metastasis-positive, three-field lymphadenectomy is recommended. However, the optimal extent of lymphadenectomy or nodal dissection for thoracic esophageal cancer is controversial. Until recently, two meta-analyses showed that three-field lymphadenectomy could provide better 5-year survival, and that the addition of CLN dissection could improve the long-term outcome over two-field lymphadenectomy alone, especially for esophageal cancer with metastasis-positive CLNs (47,48).

Palpation is unreliable as a diagnostic means. US, CT, and FDG-PET imaging are commonly used to detect regional LN metastasis. It has been known that endoscopic ultrasonography (EUS) has high sensitivity, and CT and

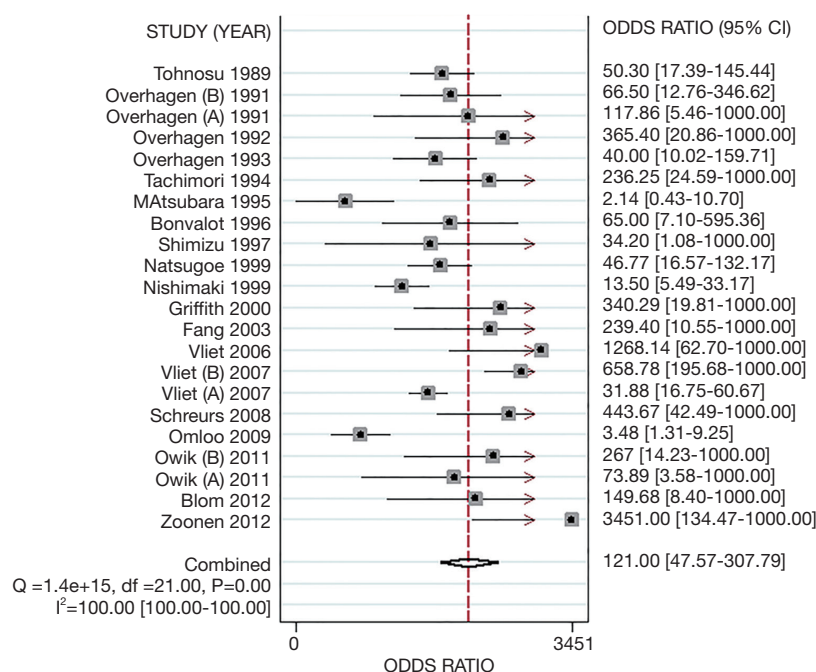


Figure 5 Forest plot of diagnostic odds ratio (DOR) estimates and 95% CI. Point estimates from each study are shown as solid squares. The solid lines represent the 95% confidence intervals (CI). The squares are proportional effect size for each study. The larger the square, the more weight it contributes to the results of the meta-analysis. The pooled estimate is denoted by the diamond at the bottom.

Table 2 Summary of the SEN, SPE, PLR, NLR, and AUC of US at different diagnostic thresholds

| Test threshold | Number of studies | AUC (95% CI) | Summary SEN (95% CI) | Summary SPE (95% CI) | Summary PLR (95% CI) | Summary NLR (95% CI) |
|----------------|-------------------|--------------------|----------------------|----------------------|-----------------------|----------------------|
| 5 mm | (17) | 0.95% (0.93–0.97%) | 84% (67–93%) | 93% (90–95%) | 11.83% (8.06–17.36%) | 0.18% (0.08–0.38%) |
| >5 mm | (5) | 0.99% (0.97–0.99%) | 94% (76–98%) | 98% (89–100%) | 41.21% (8.08–210.14%) | 0.07% (0.02–0.27%) |

AUC, area under the curve; SEN, sensitivity; SPE, specificity; PLR, positive likelihood ratio; NLR, negative likelihood ratio; US, ultrasound; CI, confidence interval.

FDG-PET have high specificity for the detection of regional LN metastases (8). However, the outcome of a recent meta-analysis was frustrating. The study reported that the sensitivity and specificity of EUS for N staging were 69% and 52%, respectively (49). Although the sensitivity and specificity can reach to 66% and 96% for EUS in CLNs detecting (50), and the accuracy of EUS in celiac axis LN metastasis can be improved through the use of FNA (51,52). According to the complexity of operation and financial burden of EUS, it could do no better than US and was not the first choice for the detection of CLN metastasis. Another study reported that PET-CT had lower sensitivity (55%) and specificity (76%) for the detection of regional nodal metastasis in patients with esophageal cancer

before surgery (53). Thus, FDG-PET or PET-CT alone is not recommended to be used for the detection of regional LN metastasis (54).

For more than two decades, US has been used as a highly accurate and cost-effective diagnostic tool for the differentiation of benign and malignant superficial LNs according to the indicators of size, shape, hilum, echogenicity, margins, and LN structural changes. The Japanese Ultrasound Guidelines for Clinical and Pathologic Studies on Carcinoma of the Esophagus suggest the following diagnostic criteria for CLN metastasis: (I) well-defined margins; (II) irregular inner echo images; (III) a ratio of the longitudinal diameter to the short axis diameter of over 0.5; and (IV) a longitudinal diameter over 5 mm (55).

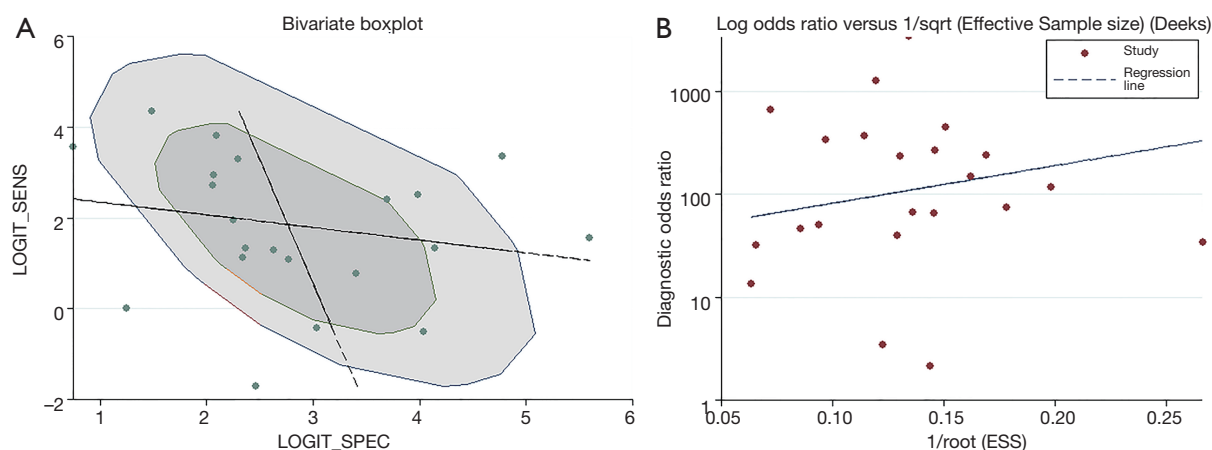


Figure 6 The Bivariate Boxplot and Deeks' funnel plot. (A) Bivariate Boxplot to assist verifying the potential heterogeneity. The gray part is a bag that contains 50% of the data points, a fence that separates inliers from outliers. Studies between the bag and the fence are marked by a light gray loop, whereas studies outside the fence are flagged as outliers. There are five observations outside the fence. Further demonstrates there was potential heterogeneity between studies; (B) Deeks' funnel plot to evaluate publication bias. The horizontal axis displays the inverse of the square root of the effective sample size [$1/\sqrt{\text{ESS}}$]. The vertical axis displays the diagnostic odds ratio (DOR). The funnel shape suggests publication bias is absent.

Therefore, a systematic evaluation of US efficacy for the detection of regional LN metastasis is meaningful for the selection of lymphadenectomy in patients with esophageal cancer.

In this study, our findings show that when combined, US had high SROC curve, AUC, and DOR values. Compared with a cut-off value of 5 mm, US had greater sensitivity and specificity at a threshold value of >5 mm. It is suggested that a cut-off value of >5 mm reflects the high diagnostic efficacy of US (an AUC of 0.99) and offers a helpful clinical utility in detecting CLN metastasis. However, this threshold could result in a higher FN rate when selected.

We demonstrated the diagnostic efficacy of US for predicting CLN-positive metastasis in esophageal cancer for the first time. The statistical methodology of bivariate meta-analysis following a random effects model could allow researchers to avert misleading conclusions (56). A comprehensive search strategy with various overlapping approaches was carried out to enable as many studies as possible to be retrieved. In addition, two reviewers (X.F.L. and Y.Z.) independently finished the searching, screening, filtering, and data extraction processes (a third reviewer intervened, if necessary). Thirdly, we picked and chose more relevant indicators comprising the DOR, AUC, LR and its scattergram, post-test probability, lambda value of the HSROC curve, and Fagan plot analysis, which were used to explore the clinical utility of US. Finally, publication

bias, threshold and non-threshold effect were explored. In order to diminish the effect of heterogeneity and to detect the origins of heterogeneity, the meta-regression, specific subgroup and sensitivity analyses were performed.

There are several limitations to our study that must be acknowledged. Despite searching several databases, we may have omitted studies unintentionally. Even though the assessments of two to three reviewers guarded against bias, there were still potential bias risks that presented in the included studies, which would affect credibility of the conclusions. Additionally, we could only extract data and details from full-text of English language studies. Some potential studies were researched by Japanese colleagues and published in Japanese (57-59). This could have introduced a selection bias in the outcomes because of the disadvantage of language obstacles. It should be noted that the full text of the study by Fang *et al.* (31) was published in Chinese. We included it because of the high research quality and our language advantage. Thirdly, the experience level of the sonographer was not mentioned in most studies. The only information provided was that most studies were performed in high-volume medical centers or at universities. Therefore, we presumed that most physicians were experienced and skilled. The investigators who were not blinded to the reference standard results might have overestimated the DOR and other results. Regardless of whether FP or FN results were exaggerated, both would

result in misleading clinicopathologic staging. Lastly, despite using meta-regression methods and subgroup and sensitivity analyses, some variables still could not satisfy the requirements of detection. These involved the diagnostic capabilities of the radiologists, different types of equipment, unreported clinical variables, and some latent factors which could influence upon the diagnostic accuracy of US detection. The deficiencies and imperfections in both methodology and research quality that we found in our review could contribute to further investigation in this field.

Conclusions

We conclude that US could be helpful in the detection of CLN metastasis to assess LN status for staging and therapeutic procedures and to improve the prognosis and guide the follow-up of patients with esophageal cancer. A large-scale study to verify the diagnostic efficacy of US in the detection of CLN metastasis in esophageal cancer patients is needed for scientific evaluation in clinical settings.

Acknowledgements

We appreciate the invaluable assistance of Jian-Ming Huang M.D. (Dept. of Biochemistry & molecular Biology, Sichuan Cancer Hospital & Institute) for the review.

Funding: This study was supported by the Natural Science Foundation of Guangxi Zhuang Autonomous Region (2010GXNSFB013061) and the Scientific Research Projects in Universities of Guangxi Zhuang Autonomous Region (2013LX032).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Torre LA, Bray F, Siegel RL, et al. Global cancer statistics, 2012. *CA Cancer J Clin* 2015;65:87-108.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015;65:5-29.
3. Peyre CG, Hagen JA, DeMeester SR, et al. The number of lymph nodes removed predicts survival in esophageal cancer: an international study on the impact of extent of surgical resection. *Ann Surg* 2008;248:549-56.
4. Korst RJ, Rusch VW, Venkatraman E, et al. Proposed revision of the staging classification for esophageal cancer. *J Thorac Cardiovasc Surg* 1998;115:660-69; discussion 669-70.
5. Chen J, Zhu K, Zheng X, et al. Prognostic analysis of cervical lymph node metastasis in patients with thoracic esophageal squamous cell carcinoma. *Zhonghua Zhong Liu Za Zhi* 2014;36:612-6.
6. Kawahara K, Maekawa T, Okabayashi K, et al. The number of lymph node metastases influences survival in esophageal cancer. *J Surg Oncol* 1998;67:160-3.
7. Isono K, Sato H, Nakayama K. Results of a nationwide study on the three-field lymph node dissection of esophageal cancer. *Oncology* 1991;48:411-20.
8. van Vliet EP, Heijenbrok-Kal MH, Hunink MG, et al. Staging investigations for oesophageal cancer: a meta-analysis. *Br J Cancer* 2008;98:547-57.
9. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
10. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36.
11. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-90.
12. Macaskill P. Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. *J Clin Epidemiol* 2004;57:925-32.
13. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;13:313-21.
14. Zamora J, Abraira V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006;6:31.
15. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; 21:1539-58.
16. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882-93.
17. Vamvakas EC. Meta-analyses of studies of the diagnostic accuracy of laboratory tests: a review of the concepts and methods. *Arch Pathol Lab Med* 1998;122:675-86.
18. Hellmich M, Lehmacher W. A ruler for interpreting diagnostic test results. *Methods Inf Med* 2005;44:124-6.

19. Tohnosu N, Onoda S, Isono K. Ultrasonographic evaluation of cervical lymph node metastases in esophageal cancer with special reference to the relationship between the short to long axis ratio (S/L) and the cancer content. *J Clin Ultrasound* 1989;17:101-6.
20. van Overhagen H, Laméris JS, Zonderland HM, et al. Ultrasound and ultrasound-guided fine needle aspiration biopsy of supraclavicular lymph nodes in patients with esophageal carcinoma. *Cancer* 1991;67:585-7.
21. van Overhagen H, Laméris JS, Berger MY, et al. Supraclavicular lymph node metastases in carcinoma of the esophagus and gastroesophageal junction: assessment with CT, US, and US-guided fine-needle aspiration biopsy. *Radiology* 1991;179:155-8.
22. van Overhagen H, Laméris JS, Berger MY, et al. Assessment of distant metastases with ultrasound-guided fine-needle aspiration biopsy and cytologic study in carcinoma of the esophagus and gastroesophageal junction. *Gastrointest Radiol* 1992;17:305-10.
23. Van Overhagen H, Laméris JS, Berger MY, et al. Improved assessment of supraclavicular and abdominal metastases in oesophageal and gastro-oesophageal junction carcinoma with the combination of ultrasound and computed tomography. *Br J Radiol* 1993;66:203-8.
24. Tachimori Y, Kato H, Watanabe H, et al. Neck ultrasonography for thoracic esophageal carcinoma. *Ann Thorac Surg* 1994;57:1180-3.
25. Matsubara T, Ueda M, Nakajima T. Preoperative assessment of lymph nodes in the prediction of disease spread and outcome in cancer of the thoracic oesophagus. *Br J Surg* 1995;82:356-9.
26. Bonvalot S, Bouvard N, Lothaire P, et al. Contribution of cervical ultrasound and ultrasound fine-needle aspiration biopsy to the staging of thoracic oesophageal carcinoma. *Eur J Cancer* 1996;32A:893-5.
27. Shimizu Y, Mera K, Tsukagoshi H, et al. Endoscopic Ultrasonography for the Detection of Lymph Node Metastasis in Superficial Esophageal Carcinoma. *Dig Endosc* 1997;9:178-82.
28. Natsugoe S, Yoshinaka H, Shimada M, et al. Assessment of cervical lymph node metastasis in esophageal carcinoma using ultrasonography. *Ann Surg* 1999;229:62-6.
29. Nishimaki T, Tanaka O, Ando N, et al. Evaluation of the accuracy of preoperative staging in thoracic esophageal cancer. *Ann Thorac Surg* 1999;68:2059-64.
30. Griffith JF, Chan AC, Ahuja AT, et al. Neck ultrasound in staging squamous oesophageal carcinoma-a high yield technique. *Clin Radiol* 2000;55:696-701.
31. Fang WT, Zhang ZH, Chen WH, et al. Ultrasound surveillance of cervical lymph node metastasis in thoracic esophageal carcinoma. *Zhonghua Wai Ke Za Zhi* 2003;41:523-5.
32. van Vliet EP, Eijkemans MJ, Kuipers EJ, et al. A comparison between low-volume referring regional centers and a high-volume referral center in quality of preoperative metastasis detection in esophageal carcinoma. *Am J Gastroenterol* 2006;101:234-42.
33. van Vliet EP, van der Lugt A, Kuipers EJ, et al. Ultrasound, computed tomography, or the combination for the detection of supraclavicular lymph nodes in patients with esophageal or gastric cardia cancer: a comparative study. *J Surg Oncol* 2007;96:200-6.
34. van Vliet EP, Steyerberg EW, Eijkemans MJ, et al. Detection of distant metastases in patients with oesophageal or gastric cardia cancer: a diagnostic decision analysis. *Br J Cancer* 2007;97:868-76.
35. Schreurs LM, Verhoef CC, van der Jagt EJ, et al. Current relevance of cervical ultrasonography in staging cancer of the esophagus and gastroesophageal junction. *Eur J Radiol* 2008;67:105-11.
36. Omloo JM, van Heijl M, Smits NJ, et al. Additional value of external ultrasonography of the neck after CT and PET scanning in the preoperative assessment of patients with esophageal cancer. *Dig Surg* 2009;26:43-9.
37. Cwik G, Dąbrowski A, Skoczylas T, et al. The value of ultrasound in the assessment of cervical and abdominal lymph node metastases and selecting surgical strategy in patients with squamous cell carcinoma of the thoracic esophagus treated with neoadjuvant therapy. *Adv Med Sci* 2011;56:291-8.
38. Cwik G, Dąbrowski A, Skoczylas T, et al. Imaging of the cervical and abdominal lymph nodes in a combined treatment of squamous cell oesophageal carcinoma. *Pol Przegl Chir* 2011;83:95-101.
39. Blom RL, Vliegen RF, Schreurs WM, et al. External ultrasonography of the neck does not add diagnostic value to integrated positron emission tomography-computed tomography (PET-CT) scanning in the diagnosis of cervical lymph node metastases in patients with esophageal carcinoma. *Dis Esophagus* 2012;25:555-9.
40. van Zoonen M, van Oijen MG, van Leeuwen MS, et al. Low impact of staging EUS for determining surgical resectability in esophageal cancer. *Surg Endosc* 2012;26:2828-34.
41. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N*

- Engl J Med 1979;300:1350-8.
42. Altorki N, Skinner D. Should en bloc esophagectomy be the standard of care for esophageal carcinoma? *Ann Surg* 2001;234:581-7.
 43. Altorki N, Kent M, Ferrara C, et al. Three-field lymph node dissection for squamous cell and adenocarcinoma of the esophagus. *Ann Surg* 2002;236:177-83.
 44. Udagawa H, Ueno M, Shinohara H, et al. The importance of grouping of lymph node stations and rationale of three-field lymphadenectomy for thoracic esophageal cancer. *J Surg Oncol* 2012;106:742-7.
 45. Rice TW, Blackstone EH, Rusch VW, et al. 7th Edition of the AJCC Cancer Staging Manual: Esophagus and Esophagogastric Junction. *Ann Surg Oncol* 2010;17:1721-4.
 46. Zhu Z, Yu W, Li H, et al. Nodal skip metastasis is not a predictor of survival in thoracic esophageal squamous cell carcinoma. *Ann Surg Oncol* 2013;20:3052-8.
 47. Ye T, Sun Y, Zhang Y, et al. Three-field or two-field resection for thoracic esophageal cancer: a meta-analysis. *Ann Thorac Surg* 2013;96:1933-41.
 48. Ma GW, Situ DR, Ma QL, et al. Three-field vs two-field lymph node dissection for esophageal cancer: a meta-analysis. *World J Gastroenterol* 2014;20:18022-30.
 49. Sun F, Chen T, Han J, et al. Staging accuracy of endoscopic ultrasound for esophageal cancer after neoadjuvant chemotherapy: a meta-analysis and systematic review. *Dis Esophagus* 2015;28:757-71.
 50. Chandawarkar RY, Kakegawa T, Fujita H, et al. Endosonography for Preoperative Staging of Specific Nodal Groups Associated with Esophageal Cancer. *World J Surg* 1996;20:700-2.
 51. Puli SR, Reddy JB, Bechtold ML, et al. Accuracy of endoscopic ultrasound in the diagnosis of distal and celiac axis lymph node metastasis in esophageal cancer: a meta-analysis and systematic review. *Dig Dis Sci* 2008;53:2405-14.
 52. Puli SR, Reddy JB, Bechtold ML, et al. Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review. *World J Gastroenterol* 2008;14:1479-90.
 53. Shi W, Wang W, Wang J, et al. Meta-analysis of 18FDG PET-CT for nodal staging in patients with esophageal cancer. *Surg Oncol* 2013;22:112-6.
 54. Shimizu S, Hosokawa M, Itoh K, et al. Can hybrid FDG-PET/CT detect subclinical lymph node metastasis of esophageal cancer appropriately and contribute to radiation treatment planning? A comparison of image-based and pathological findings. *Int J Clin Oncol* 2009;14:421-5.
 55. Asakura S, Nabeya KI, Hanaoka T, et al. The effectiveness of ultrasonography in diagnosis of cervical lymph node metastasis in preoperative esophageal cancer. In: Kin-ichi N, Tateo H, Hiroshi N, editors. *Recent advances in diseases of the esophagus*. Kyoto: Springer;1993:580-91.
 56. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;20:2865-84.
 57. Murata Y, Muroi M, Ide H. Ultrasonic diagnosis of esophageal cancer. *Kango Gijutsu* 1986;32:1073-5.
 58. Ide H, Hanyu F, Murata Y, et al. Multidisciplinary treatment of thoracic esophagus carcinoma based on preoperative staging. *Gan To Kagaku Ryoho* 1988;15:589-96.
 59. Yoshinaka H, Shimazu H, Morifuji H, et al. Lymph node metastases in the thoraco-cervical transitional region in thoracic esophageal cancer--with ultrasonic detection and a comment on the guide lines. *Nihon Geka Gakkai Zasshi* 1989;90:496-503.

Cite this article as: Leng XF, Zhu Y, Wang GP, Jin J, Xian L, Zhang YH. Accuracy of ultrasound for the diagnosis of cervical lymph node metastasis in esophageal cancer: a systematic review and meta-analysis. *J Thorac Dis* 2016;8(8):2146-2157. doi: 10.21037/jtd.2016.07.71