Original Article

Systematic review and meta-analysis of the accuracy of 18F-FDG PET/CT for detection of regional lymph node metastasis in esophageal squamous cell carcinoma

Chenxue Jiang[#], Yun Chen[#], Yaoyao Zhu, Yapping Xu

Department of Radiation Oncology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai 200433, China

Contributions: (I) Conception and design: Y Xu; (II) Administrative support: Y Xu; (III) Provision of study materials or patients: C Jiang, Y Chen; (IV) Collection and assembly of data: C Jiang, Y Zhu; (V) Data analysis and interpretation: C Jiang, Y Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Yaping Xu, PhD. Department of Radiation Oncology, Shanghai Pulmonary Hospital, Tongji University School of Medicine, No. 507, Zhengmin Road, Shanghai 200433, China. Email: xuyaping1207@163.com.

Background: We performed a systematic review and meta-analysis to assess the accuracy of 18F-fluorodeoxyglucose positron emission tomography with computer tomography (18F-FDG PET/CT) for detection of regional lymph node metastasis in esophageal squamous cell carcinoma in per-patient and per-nodal station basis.

Methods: Electronic databases were researched for studies assessing the sensitivity and specificity of PET/CT to detect the regional lymph node metastasis published between January 2006 and December 2017 on esophageal squamous cell carcinoma. STATA software was performed to assess the sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odd ratio (DOR) and summary receiver operating characteristic (SROC) curve. The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) and Deeks' Funnel Plot Asymmetry Test were performed to evaluate the study quality and publication bias of included studies.

Results: Nineteen studies were eligible for meta-analysis, comprising 1,089 patients with esophageal cancer who underwent 18F-FDG PET/CT before surgery. According to the content of the article, we divided the selected studies into per-patient basis group and per-nodal basis group (one of the articles was involved in both groups). For the per-nodal station basis group (12 studies, 5,681 stations), the pooled sensitivity and specificity estimates of 18F-FDG PET/CT for detecting regional lymph node metastasis were 66% [95% confidence interval (CI): 51–78%] and 96% (95% CI: 92–98%), respectively. The corresponding values on a per-patient basis group (8 studies; 506 patients) were 65% (95% CI: 49–78%) and 81% (95% CI: 69–89%) in sensitivity and specificity, respectively.

Conclusions: Overall, 18F-FDG PET/CT have a moderate to low sensitivity and a high to moderate specificity for detection of regional nodal metastasis in esophageal cancer. Therefore, since the false rate is considerable, extending the extent of lymph node dissection or radiotherapy target volume is necessary after diagnosis of regional nodal metastasis by 18F-FDG PET/CT.

Keywords: Esophageal squamous cell carcinoma; positron emission tomography with computer tomography (PET/CT); regional lymphatic metastasis; meta-analysis

Submitted Apr 24, 2018. Accepted for publication Sep 30, 2018.
doi: 10.21037/jtd.2018.10.57
View this article at: http://dx.doi.org/10.21037/jtd.2018.10.57

Introduction

Esophageal squamous cell carcinoma, a highly aggressive malignant cancer that ranks sixth in cancer mortality and third in morbidity worldwide, is the most common type of esophageal cancer in China; it accounts for more than 90% of cases, while esophageal adenocarcinoma has a high incidence in Western countries (1). The majority of esophageal cancer patients are diagnosed with advanced disease due to unclear early symptoms. Lymph node metastasis is the main form of esophageal cancer metastasis. N staging determines the target volume of radiotherapy and the necessary extent of lymph node dissection in the resection of esophageal cancer and is related to the local control rate, recurrence and overall survival (OS).

Endoscopic ultrasound (EUS) is now considered the most accurate method available to assess esophageal carcinoma infiltration depth, with an accuracy of 89% (2). However, the sensitivity, specificity and accuracy of EUS for detecting N stage in esophageal cancer are 71%, 74% and 73%, respectively (3). Computer tomography (CT) is widely used to determine staging in thoracic malignancies, including esophageal cancer. However, the accuracy of CT in detecting regional lymph node metastasis in esophageal cancer is unsatisfactory. The accuracy of CT for detecting lymph nodes with a diameter less than 10 mm and for detecting para esophageal lymph nodes in esophageal carcinoma is only 16.78% and 9%, respectively (4,5). Another study reported the sensitivity and specificity of CT for the detection of lymph node metastasis in esophageal cancer as 38.57% and 93.93%, respectively (6).

With the development and improvement of diagnostic technology, the integration of 18F-fluorodeoxyglucose positron emission tomography with CT (18F-FDG PET/CT) has been used successfully with increasing frequency in the evaluation and clinical management of many malignant conditions. The aim of this systematic review and meta-analysis was to assess the accuracy of integrated 18F-FDG PET/CT for the detection of regional lymph node metastasis in esophageal cancer.

Methods

Literature search strategy and selection/exclusion criteria

PubMed, EMBASE and the Cochrane Library were systematically searched from January 2006 to December 2017, with the key words "esophageal squamous cell carcinoma", "PET/CT", "lymph node metastasis" and their synonyms. Two reviewers independently selected studies that examined the diagnostic value of 18F-FDG PET/CT, either in routine clinical practice or in symptomatic patients, in whom regional lymph node metastasis was suspected before surgery using data that could be extracted into a 2×2 contingency table. The reference standard for positive lymph node metastasis in each selected study must be pathology during or after surgery. Non-English language studies were excluded, except those in Chinese. Conference abstracts and letters to journal editors were excluded.

Quality evaluation

The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2, *Figure S1*) was performed to evaluate the diagnostic accuracy qualities of the 19 eligible articles. QUADAS-2 is a tool for systematic reviews of diagnostic studies developed from the QUADAS tool, and it is used to judge the risk of bias and applicability concerns, evaluating four key domains: patient selection, index text, reference standard, and flow and timing (7,8). QUADAS-2 evaluation was performed using Review manager software version 5.3.5 (The Nordic Cochrane Centre, The Cochrane Collaboration) and the full QUADAS-2 tool also could be found from the QUADAS website (www.quadas.org).

Statistical analysis

The data from the 19 selected studies was extracted and assembled into a 2×2 table, which consisted of true positive (TP), false-negative (FN), false-positive (FP) and true-negative (TN) values. Forest plots of sensitivity and specificity were generated using the forest command of the midas package for STATA version 14.0 (Stata Corporation, College Station, TX, USA). Summary receiver operating characteristic (SROC) curves were constructed to examine diagnostic accuracy. The inconsistency index (I^2) was calculated to assess the heterogeneity between studies. I^2 values greater than 50% were considered to indicate substantial heterogeneity. Deek's funnel plot was used to assess the publication bias in this meta-analysis (9,10). Meta-regression was performed to identify potential sources of bias. Statistical significance was defined as a P value less than 0.05.

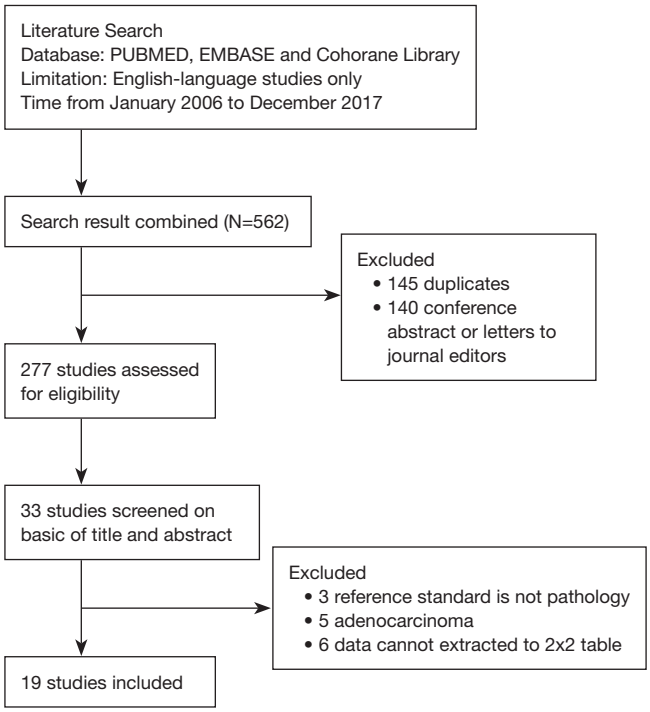

Figure 1 The flow chart of the search for eligible studies.

Results

Study selection and characteristics

A total of 19 studies were included in the review. The electronic search yielded 562 studies; after excluding 145 duplicates and 140 conference abstracts and letters to journal editors, 277 studies were assessed for eligibility. According to the content of their abstracts, 244 articles were excluded. Then, 33 articles were screened based on their full text and eventually we selected 19 articles (the flow chart of the screening of the literature is shown in *Figure 1*). *Table 1* summarizes the clinical characteristics and reported accuracy of the 19 selected eligible articles. Included studies were grouped according to whether the research unit was the patient or lymph nodes. *Table 2* summarizes the type of scanner, amount of tracer agent and the criteria for PET/CT positive detection of regional lymph nodes in the included studies in this meta-analysis.

Study quality and study design

Figure 2 summarizes the methodological quality of all

Table 1 Characteristics of the 19 eligible studies and diagnostic accuracy of 18F-FDG PET and PET/CT

Author	Year	Subgroup	Origin	No. pts	Median age [range]	Lymph node (group)	Design	No. yes in QUADAS-2	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	Accuracy (%)
Yuan (11)	2006	Station	China	45	57.5 [40–73]	397	Prosp.	11	77	25	5	290	93.90	92.60	92.44
Hsu (12)	2009	Patient	Taiwan (China)	45	60.8 [39–83]	–	Prosp.	11	12	4	9	20	57.10	83.30	71.11
Han (13)	2012	Station	China	22	60 [51–75]	424	Prosp.	11	39	14	8	363	82.98	96.29	94.81
Yano (14)	2012	Patient	Japan	81	63 [44–75]	–	Retro.	11	12	13	25	31	32.40	70.40	53.09
Wang (15)	2012	Station	China	26	Unmentined	119	Retro.	11	9	16	5	89	36.00	94.70	82.35
Tan (16)	2014	Station	China	115	57.9 [45–76]	946	Prosp.	11	141	6	46	75	75.40	98.68	92.02
Yamada (17)	2014	Station	Japan	258	66 [41–86]	1,231	Retro.	11	71	21	204	935	25.80	97.80	81.72
Sohda (18)	2010	Station	Japan	21	65.9 [43–80]	96	Prosp.	10	8	4	25	49	24.20	93.70	66.28
Shum (19)	2012	Patient	Taiwan (China)	26	60.4 [42–72]	–	Retro.	10	8	6	2	10	80.00	60.00	69.23
Yen (20)	2012	Patient	Taiwan (China)	36	Unmentioned	–	Retro.	10	3	1	4	28	42.86	96.55	86.11
Wang (21)	2016	Station	China	43	54.3	864	Retro.	10	107	51	47	641	69.48	92.71	88.42
Yu (22)	2011	Station	China	16	56.5 [46–70]	144	Prosp.	10	16	5	5	118	76.20	95.90	93.06
Manabe (23)	2013	Patient	Japan	156	61.4 [40–84]	–	Retro.	10	49	7	40	60	66.10	85.70	69.87
Kato (24)	2008	Station	Taiwan (China)	26	60.4 [42–72]	–	Retro.	10	8	6	2	10	80.00	60.00	69.23
Kim (25)	2015	Patient	Korea	51	69 [51–80]	–	Prosp.	9	20	13	3	15	82.60	53.50	68.63
Kim (26)	2012	Station	Korea	17	66.1 [52–75]	72	Retro.	9	10	14	7	41	58.80	90.90	70.83
Schreurs (27)	2008	Patient	Neterlands	61	63.4 [48–80]	–	Retro.	9	12	6	3	40	86.60	86.90	85.25
Bella (28)	2014	Station	China	27	64 [48–79]	117	Retro.	8	26	10	6	75	81.20	88.20	86.32
Okada (29)	2009	Station	Japan	18	68 [59–79]	210	Prosp.	8	15	1	10	184	60.00	99.46	94.76

18F-FDG PET, 18F-fluorodeoxyglucose positron emission tomography; Pts, patients; QUADAS-2, The Quality Assessment of Diagnostic Accuracy Studies 2.

Table 2 Criteria of positive regional lymph node by 18F-FDG PET/CT in included studies in this meta-analysis

Author	Year	Type of scanner	Amount of tracer agent	Slice thickness of CT	Criteria of positive regional lymph node by 18F-FDG PET/CT in included studies in this meta-analysis
Yuan (11)	2006	Discovery LS; GE Healthcare	370 MBq	4.25 mm/slice	18F-FDG uptake prominently compared with surrounding tissues and not related to normal physiologic uptake
Hsu (12)	2009	Discovery VCT; GE Healthcare, Waukesha, WI, USA	370 MBq	Unclear	SUVmax greater than 2.5 was considered positive
Han (13)	2012	MiniTrace; GE Healthcare, Piscataway, NJ, USA	300–400 MBq	4.25 mm/slice	The nodal accumulations with the intensity higher than that of the mediastinal blood pool were first visually detected on PET image and then precisely localized on PET/CT fusion image to determine whether they were LNs
Yano (14)	2012	Siemens-Asahi Medical Technologies, Tokyo, Japan	3.5 MBq/kg	Unclear	SUVmax value of above 1.8
Wang (15)	2012	Philips Gemini TF 16, Philips, The Netherlands	3.0–3.7 MBq/kg	3 mm/slice	Maximum standardized uptake values (SUVmax)
Tan (16)	2014	GE Discovery LS4 PET/CT, General Electrical Medical Systems	5.55 MBq/kg	5 mm/slice	Short diameter >10 mm; any sulcus oesophageal lymph node; enhanced lymph node with thin wall and ring shaped
Yamada (17)	2014	Nihon Medi-Physics Tokyo, Japan	Not mentioned	Unclear	SUVmax ≥5.0 in the tracheal bifurcation and pulmonary hilum or a value of 2.0 or more in other sites
Sohda (18)	2010	Discovery STE; GE Healthcare; Biograph 16; Siemens Medical Solutions	5-6 MBq/kg	Unclear	A faint uptake of 18 F-FDG
Shum (19)	2012	Discovery STE, GE Medical Systems, Milwaukee, WI, USA	370 MBq	3.75 mm/slice	Combined with early SUVmax ≥2.5 alone and retention index (RI)
Yen (20)	2012	GE Medical Systems, Milwaukee, WI, USA	370 MBq	4.25 mm/slice	SUVmax (not mentioned detail)
Wang (21)	2016	Amsterdam, The Netherlands	4.44 MBq/kg	Unclear	SUVmax was then computed, with cut-off values set at 2.5 and 5
Yu (22)	2011	BIOGRAPH 16HR, Siemens Molecular Imaging, Knoxville, TN, USA	7.4 MBq/kg	5 mm/slice	Short axis >10 mm or seen in the tracheoesophageal groove
Manabe (23)	2013	GEMINI GXL (Philips Healthcare)	241.3 6±70.0 MBq	Unclear	Increased 18F-FDG uptake greater than the background activity of the blood pool
Kato (24)	2008	GE Discovery ST8, GE, Milwaukee, USA	Not mentioned	10 mm/slice	Short axis >1 cm in CT. Not mentioned with PET/CT
Kim (25)	2015	Siemens Healthcare, Knoxville, TN, USA	5 MBq/kg	Unclear	Cut-of value was determined by ROC of 18F-FDG PET/CT parameters
Kim (26)	2012	Gemini Scanner; Philips, Bothell, WA, USA	5.6 MBq/kg	5 mm/slice	FDG uptake within structurally identifiable nodes that was focally prominent compared with background mediastinal activity (regardless of lymph node size)
Schreurs (27)	2008	Siemens/CTI, Knoxville, TN, USA	400–580 MBq	Unclear	Lymph nodes >1 cm on CT imaging with FDG-uptake on PET imaging
Bella (28)	2014	GE Discovery ST-16 PET/CT System; Wisconsin, USA	Not mentioned	Unclear	SUVmax for lymph nodes was 4.1, which was calculated by ROC curve
Okada (29)	2009	Biograph; Siemens Japan, Tokyo, Japan	3 MBq/kg	8 mm/slice	FDG uptake above the background

18F-FDG PET/CT, 18F-fluorodeoxyglucose positron emission tomography with computer tomography. SUV, standardized uptake value.

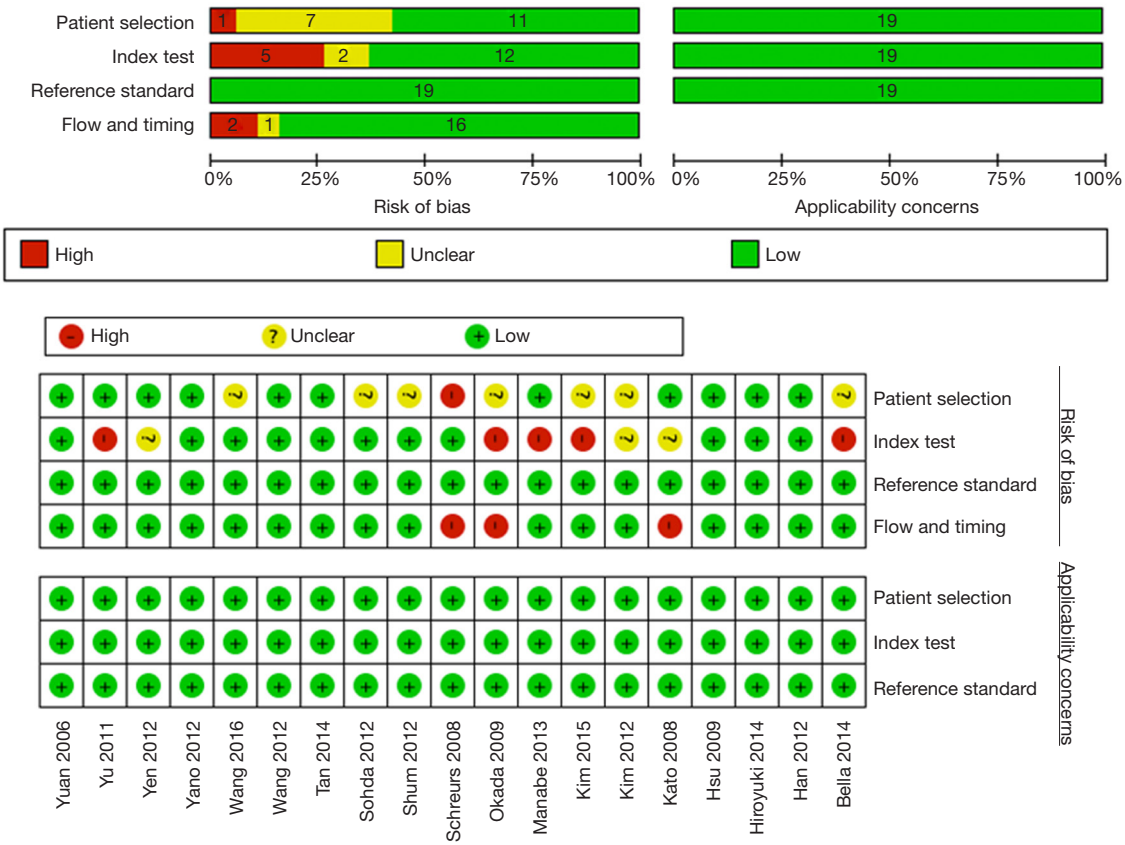

Figure 2 The quality assessment of 19 included studies by QUADAS-2 tool. QUADAS-2, The Quality Assessment of Diagnostic Accuracy Studies 2.

included studies after assessment by the QUADAS-2 tool. If the answers to all of the questions about a domain were judged as 'yes', indicating a low risk of bias, then this domain was judged to be at low risk of bias. In contrast, if one was judged as 'no', then that would indicate 'high risk', and a potential bias might exist. 'Unclear' indicated insufficient information to determine whether partial verification was present. In our study, seven studies (11-17) were rated as 'yes' for the 11 questions in the QUADAS-2 quality assessment tool. Four studies (18-21) were rated as 10 'yes' and 1 'unclear' on the 11 questions in the QUADAS-2 quality assessment tool, while two studies (22,23) were rated as 10 'yes' and 1 'no' in the 11 questions on the QUADAS-2 quality assessment tool. Two studies (24,25) were rated as 9 'yes', 1 'unclear' and 1 'no', while two studies were rated as 9 'yes' and two 'unclear' (26) or two 'no' (27). One study (28) was rated as 8 'yes', 2 'unclear' and 1 'no' and one study (29) was rated as 8 'yes', 1 'unclear' and 2 'no' in the QUADAS-2 quality assessment tool. In nine of nineteen studies, the study design was prospective.

Publication bias and sensitivity analysis

Deek's funnel plots of diagnostic odds ratio inverse of the square root of the effective sample size were constructed to assess the publication bias of the articles. The shape of the funnel plots revealed no asymmetry in both subgroups [t=0.48, P=0.65 on a per-patient basis (*Figure 3A*) and t=–0.05, P=0.96 on a per-nodal station basis (*Figure 3B*)]. A sensitivity analysis was performed to assess whether or not the meta analyses were stable by excluding studies one by one. The results showed that the data were stable and not significantly different on a per-patient basis (*Figure 3C*) or on a per-nodal station basis (*Figure 3D*).

Detection of lymph node metastasis on a per-patient basis

The paired forest plots of sensitivity and specificity for

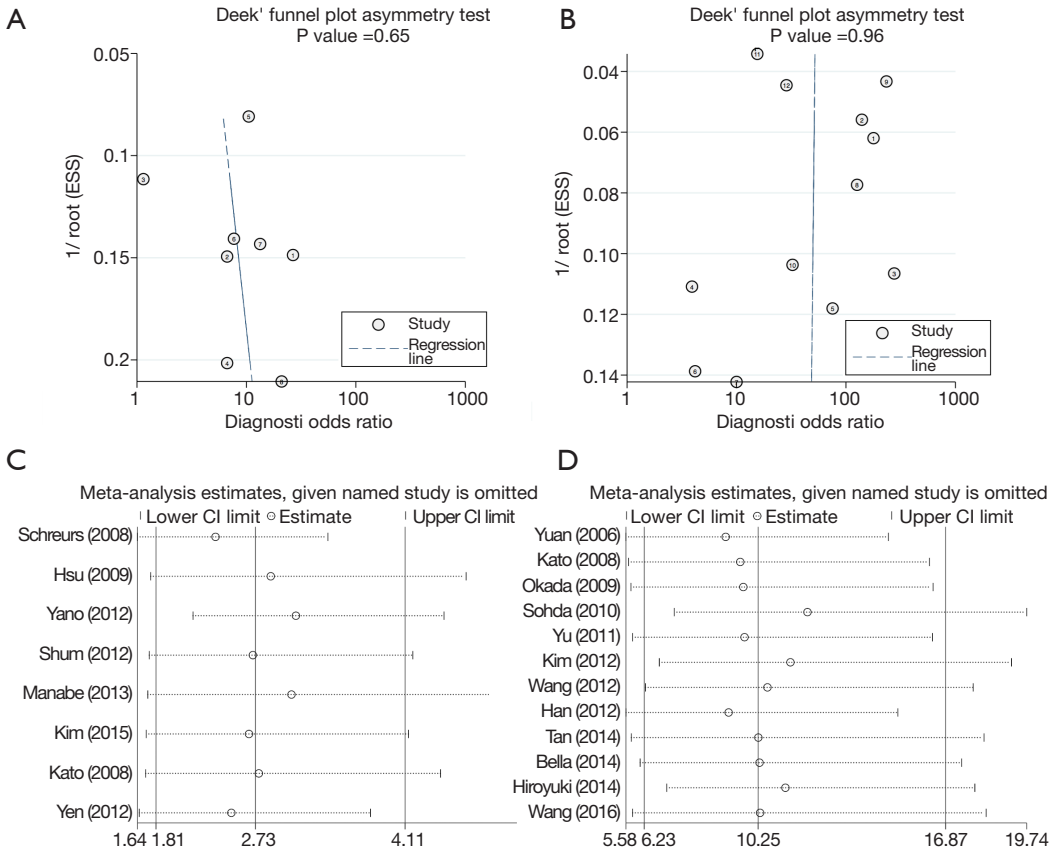

Figure 3 Deek's funnel plot and sensitivity analysis of 19 included studies. (A,C) On per-patient basis subgroup; (B,D) on per-nodal station basis subgroup.

the eight individual articles (a total of 506 patients) are presented in *Figure 4A* and indicate that 18-FDG PET/CT resulted in a low estimated sensitivity and moderate estimated specificity of 0.65 [95% confidence interval (CI): 0.49–0.78] and 0.81 (95% CI: 0.69–0.89), respectively. I^2-values were 75.26 (95% CI: 57.97–92.55, Cochrane's Q P=0.00) for sensitivity and 76.50 (95% CI: 60.28–92.72, Cochrane's Q P=0.00) for specificity and indicate substantial heterogeneity. However, no factor was caused the heterogeneity via meta-regression analysis. The positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odd ratio (DOR) were 3.4 (95% CI: 2.1–5.4), 0.44 (95% CI: 0.29–0.65) and 8 (95% CI: 4–16), respectively.

Figure 4B presents the SROC curve analysis (with prediction and confidence contours) of the ability of 18-FDG PET/CT to detect regional nodal metastasis in patients with esophageal cancer on a per-patient analysis in the eight eligible articles. The area under the SROC curve (AUC) was 0.80 (95% CI: 0.76–0.83).

Detection of lymph node metastasis on a per-nodal station basis

The paired forest plots of the sensitivity and specificity values reported in the 12 relevant individual articles are presented in *Figure 5A*. Of the total of 5,681 nodal stations analyzed, 18-FDG PET/CT had a low estimated sensitivity and a high estimated specificity of 0.66 (95% CI: 0.51–0.78) and 0.96 (95% CI: 0.92–0.98), respectively. I^2-values were 95.27 (95% CI: 93.61–96.94, Cochrane's Q P=0.00) for sensitivity and 94.66 (95% CI: 92.71–96.61, Cochrane's Q P=0.00) for specificity, which indicated substantial heterogeneity. Meta-regression showed the type of research (P=0.01) and origin (P=0.00) contributed to the high heterogeneity. The PLR, NLR, and DOR values were 15.2 (95% CI: 8.0–28.8), 0.36 (95% CI: 0.24–0.53), and 43 (95% CI: 19–96), respectively. *Figure 5B* illustrates the summary SROC (with prediction and confidence contours) for the

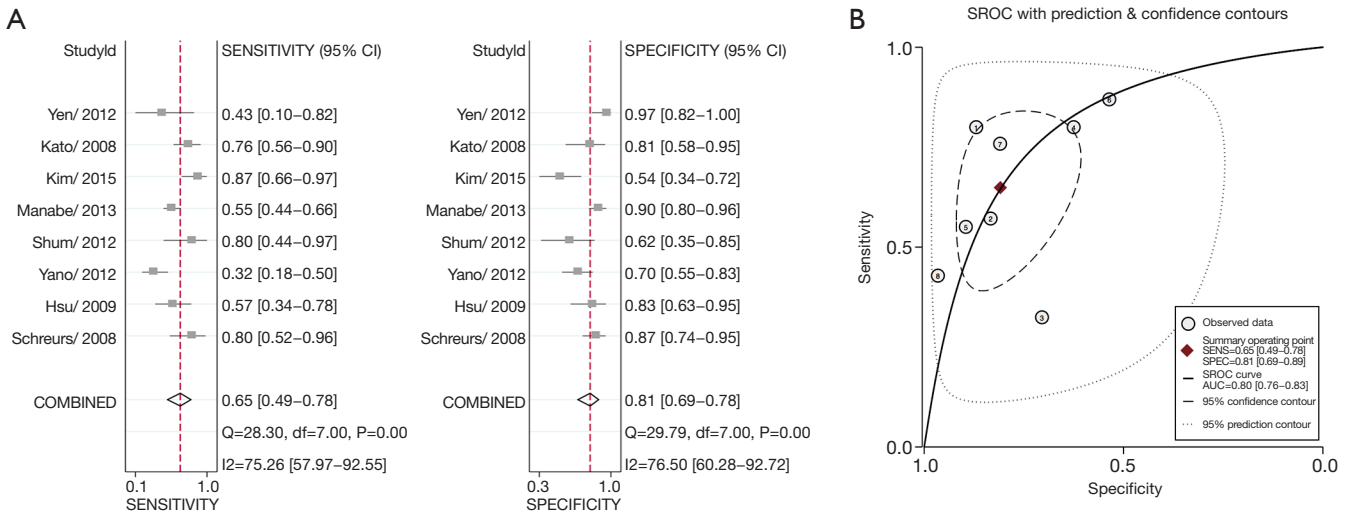

Figure 4 The forest plot of sensitivity and specificity (A) and SROC curve (B) for 18F-FDG PET/CT in the detection of regional lymph node metastasis in patients with esophageal cancer on per-patient basis. (Number in B represented included studies, sequence is shown in A). SROC, summary receiver operating characteristic; 18F-FDG PET/CT, 18F-fluorodeoxyglucose positron emission tomography with computer tomography.

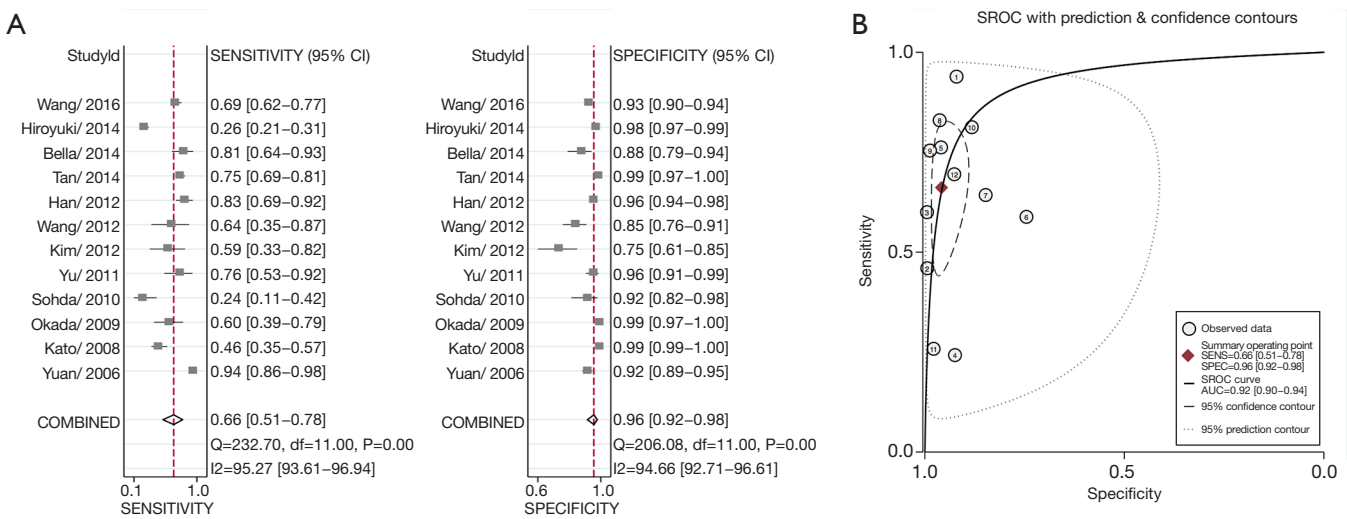

Figure 5 The forest plot of sensitivity and specificity (A) and SROC curve (B) for 18F-FDG PET/CT in the detection of regional lymph node metastasis in patients with esophageal cancer on per-nodal station basis. (Number in B represented included studies, sequence is shown in A). SROC, summary receiver operating characteristic; 18F-FDG PET/CT, 18F-fluorodeoxyglucose positron emission tomography with computer tomography.

ability of 18F-FDG PET/CT to detect regional nodal metastasis in patients with esophageal cancer on a per-station basis for the 12 eligible articles. The SROC AUC was 0.92 (95% CI: 0.90–0.94).

Discussion

As a result of the widespread application of 18F-FDG PET/CT, these techniques are now used to detect regional lymph node metastasis in a variety of malignant neoplasms (30-32).

The benefits and accuracy of 18F-FDG PET/CT remain controversial and inconclusive in esophageal squamous cell carcinoma. In this meta-analysis, the pooled sensitivity and specificity values for 18F-FDG PET/CT were 0.64 (95% CI: 0.47–0.78) and 0.78 (95% CI: 0.68–0.85) on a per-patient basis, respectively. On a per-nodal basis, the pooled sensitivity and specificity were 0.66 (95% CI: 0.51–0.78) and 0.96 (95% CI: 0.92–0.98), respectively, indicating that 18F-FDG PET/CT has a moderate/low sensitivity and high/moderate specificity for the detection of regional lymph node metastasis in esophageal squamous cell carcinoma.

There was high heterogeneity among studies in both subgroups on a per-patient basis and on a per-nodal station basis. Meta-regression showed that research type and origin or included studies led to a high heterogeneity in the subgroup on a per-nodal station basis. However, in the per-patient basis subgroup, no factor was found to be related to the high heterogeneity. The small number of studies included in this meta-analysis and the small sample size in each included study in the subgroup on a per-patient basis may have resulted in the high heterogeneity. Future studies should be designed to evaluate this heterogeneity.

The low sensitivity of PET/CT for regional lymph node metastasis may be related to Glut 1 expression. Glut 1 expression and tumor size are correlated with FDG accumulation and influence the sensitivity of PET scans in both primary tumors and metastatic lymph nodes of esophageal squamous cell carcinoma (33). The size of lymph node metastases is smaller in esophageal cancer than that in other cancers. Several studies have shown that small regional metastatic lymph nodes (range: 2–10 mm) could not be detected by FDG-PET in cases of esophageal carcinoma (34), and it might difficult to detect LN metastasis with a minimum size of 6–8 mm by FDG-PET near the cardiac-gastric region (35).

The DOR is an index of test accuracy that combines the sensitivity and specificity data into a single number. The DOR is the ratio of the odds of a positive test in a patient with the disease relative to the odds of a positive test in a patient without the disease, and it ranges from 0 to infinity, with higher values indicating better discriminatory test performance (36). There is no means to discriminate between patients with and without the disorder by the diagnostic test if the value of DOR is 1.0. In this meta-analysis, the pooled DOR values for 18F-FDG PET/CT in the per-patient and per-nodal station meta-analyses were 8 (95% CI: 4–16) and 43 (95% CI: 19–96), respectively, indicating that 18F-FDG PET and PET/CT have a low accuracy for the detection of regional lymph node metastasis in esophageal squamous cell carcinoma.

A similar result was reported in another study on esophageal cancer, and we also found that PET/CT had an overall high accuracy to detect regional nodal metastasis in primary head and neck cancer before treatment (37,38). We hypothesized that the low DOR value for PET/CT in detecting regional lymph nodal metastasis in esophageal squamous cell carcinoma is due to the common complications of esophageal squamous cell carcinoma such as esophagitis and infection.

Radiomics is a new field that extracts and analyzes large amounts of advanced quantitative imaging features with high throughput from medical images obtained with CT, PET or magnetic resonance imaging (MRI) (39). Radiomic analysis using density thresholds for FDG-PET/CT can improve the clinical value of 18F-FDG PET/CT, such as differentiating benign from malignant mediastinal and hilar lymph nodes and tumor subtypes in patients with lung cancer (40). PET/CT images that display the Haralick co-occurrence can identify and reveal the higher heterogeneity areas in lymph nodes in patients with metastatic breast cancer, which can be used to select suspicious lymph nodes for image-informed biopsy (41). The development of radiomics is promising to increase the PET/CT accuracy and precision in the detection of regional lymph node metastasis in patients with esophageal squamous cell carcinoma.

Since the DOR is not easy to interpret or use in clinical practice and likelihood ratios are considered more clinically meaningful, both the PLR and NLR were calculated as measures of diagnostic accuracy. PLR of >10 or NLR <0.1 are indicative of a high accuracy. The amalgamated PLR values for 18F-FDG PET/CT in the per-patient and per-nodal station meta-analyses were 3.4 (95% CI: 2.1–5.4) and 15.2 (95% CI: 8.0–28.8), respectively. The pooled per-nodal station PLR value indicated that 18F-FDG PET/CT is capable of determining nodal staging for patients with esophageal squamous cell carcinoma. However, the amalgamated per-patient value suggests that 18F-FDG PET/CT is not accurate enough to determine nodal staging for patients with esophageal squamous cell carcinoma. Moreover, the amalgamated NLR values for 18F-FDG PET/CT in the per-patient and per-nodal station meta-analysis were 0.44 (95% CI: 0.29–0.65) and 0.36 (95% CI: 0.24–0.53), respectively. These results suggest that we still need biopsy or other diagnostic tests to confirm the

diagnosis of negative but suspicious regional lymph nodes after PET/CT while other tomographic imaging methods (such as CT, MR or EUS) give a positive diagnosis in patients with esophageal cancer in clinical practice.

This meta-analysis possesses several limitations. First, the high heterogeneity between the individual studies had a limited impact on the meta-analysis. Meta-regression analysis showed that research type and origin or included studies led to the high heterogeneity in the subgroups on a per-nodal station basis. However, in the per-patient basis subgroups, meta-regression analysis did not detect potential sources of heterogeneity. The small number of included studies may have led to inaccurate estimates of heterogeneity. Second, the lack of clinical and imaging follow-up data may affect our assessment of the sensitivity and specificity of 18F-FDG PET/CT. Third, the spatial resolution of PET/CT increased the difficulty of identifying metastatic lymph nodes less than 5 mm in diameter; this difficulty might lead to underestimates of lymph node involvement. In addition, since the meta-analysis only included studies of esophageal squamous cell carcinoma, our results do not fully explore the role of PET/CT in detecting regional lymph nodes in esophageal adenocarcinoma. In addition, the discrepancies among different patient populations, types of scanners, the criteria for positive lymph nodes, and excluded articles including conference abstracts or letters to the editor may impact this evaluation of the accuracy of 18F-FDG PET/CT. Moreover, each of the abovementioned factors may affect the accuracy of 18F-FDG PET/CT for the detection of regional lymph node metastasis in patients with esophageal squamous cell carcinoma. With the developments in current research on radiomics, it is promising to improve the accuracy of PET/CT in the diagnosis of esophageal squamous cell carcinoma.

In conclusion, 18F-FDG PET/CT has a moderate/low sensitivity and high/moderate specificity for the detection of regional nodal metastasis in patients with esophageal squamous cell carcinoma. These results indicate that enlarging the extent of lymph node dissection or radiotherapy target volume in patients with a diagnosis of regional nodal metastasis based on 18F-FDG PET/CT may be necessary in esophageal squamous cell carcinoma, since 18F-FDG PET/CT has a considerable false negative rate for detection of regional nodal metastasis. In clinical practice, we still need pathologic or cytological examination to identify the suspected regional lymph nodes due to the high NLR of PET/CT for detection of regional lymph node metastasis in patients with esophageal squamous cell carcinoma.

Acknowledgements

This work was presented on September 19th to 21st at 15th International Society for Disease of the Esophagus World Congress (IDSE 2016) in Singapore (Abstrast ID: PS02.039).

Footnote

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

References

1. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin 2016;66:115-32.
2. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471-4.
3. Heeren PA, van Westreenen HL, Geersing GJ, et al. Influence of tumor characteristics on the accuracy of endoscopic ultrasonography in staging cancer of the esophagus and esophagogastric junction. Endoscopy 2004;36:966-71.
4. Sharma OP, Chandermohan, Mashankar AS, et al. Role of computed tomography in preoperative evaluation of esophageal carcinoma. Indian J Cancer 1994;31:12-8.
5. Wen-hai Y. The Clinical Value of 64-slices Spiral Computed Tomography in Detecting Abdominal Lymph Node Metastasis of Esophageal Carcinoma and Abdominal Cardiac Carcinoma. Chinese Journal of CT And MRI 2015;13:57-9.
6. Dai Y, Tianwen X. Analysis of diagnosis of lymph node metastasis of esophageal carcinoma by CT compared with pathology. Cancer Research and Clinic 2014;26:169-71.
7. 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.
8. Wu L, Zhang Y, Zeng X. The QUADAS-2 tool for the quality assessment of diagnostic accuracy study: an introduction. Journal of Hubei University of Medicine 2013;32:201-8.
9. He L, Jing-zhuang M. Graphing of Funnel in Meta-Analysis. Journal of Evidence-Based Medicine

2007;7:101-4.
10. Zhang TS, Zhong WZ, Xu TC. Drawing Funnel Plot and Testing for Funnel Plot Asymmetry in Stata. The Journal of Evidence-Based Medicine 2009;2:30.
11. Yuan S, Yu Y, Chao KC, et al. Additional value of PET/CT over PET in assessment of locoregional lymph nodes in thoracic esophageal squamous cell cancer. J Nucl Med 2006;47:1255-9.
12. Hsu WH, Hsu PK, Wang SJ, et al. Positron emission tomography-computed tomography in predicting locoregional invasion in esophageal squamous cell carcinoma. Ann Thorac Surg 2009;87:1564-8.
13. Han D, Yu J, Zhong X, et al. Comparison of the diagnostic value of 3-deoxy-3-18F-fluorothymidine and 18F-fluorodeoxyglucose positron emission tomography/computed tomography in the assessment of regional lymph node in thoracic esophageal squamous cell carcinoma: a pilot study. Dis Esophagus 2012;25:416-26.
14. Yano M, Motoori M, Tanaka K, et al. Preoperative staging of clinically node-negative esophageal cancer by the combination of 18F-fluorodeoxyglucose positron emission tomography and computed tomography (FDG-PET/CT). Esophagus 2012;9:210-6.
15. Wang F, Shen LY, Ma SH, et al. Advantages of positron emission tomography-computed tomography imaging in esophageal squamous cell carcinoma. Dis Esophagus 2013;26:832-7.
16. Tan R, Yao SZ, Huang ZQ, et al. Combination of FDG PET/CT and contrast-enhanced MSCT in detecting lymph node metastasis of esophageal cancer. Asian Pac J Cancer Prev 2014;15:7719-24.
17. Yamada H, Hosokawa M, Itoh K, et al. Diagnostic value of (1)(8)F-FDG PET/CT for lymph node metastasis of esophageal squamous cell carcinoma. Surg Today 2014;44:1258-65.
18. Sohda M, Kato H, Suzuki S, et al. 18F-FAMT-PET is useful for the diagnosis of lymph node metastasis in operable esophageal squamous cell carcinoma. Ann Surg Oncol 2010;17:3181-6.
19. Shum WY, Hsieh TC, Yeh JJ, et al. Clinical usefulness of dual-time FDG PET–CT in assessment of esophageal squamous cell carcinoma. Eur J Radiol 2012;81:1024-8.
20. Yen TJ, Chung CS, Wu YW, et al. Comparative study between endoscopic ultrasonography and positron emission tomography-computed tomography in staging patients with esophageal squamous cell carcinoma. Dis Esophagus 2012;25:40-7.
21. Wang GM, Liu DF, Xu YP, et al. PET/CT imaging in diagnosing lymph node metastasis of esophageal carcinoma and its comparison with pathological findings. Eur Rev Med Pharmacol Sci 2016;20:1495-500.
22. Yu W, Fu XL, Zhang YJ, et al. A prospective evaluation of staging and target volume definition of lymph nodes by 18 FDG PET/CT in patients with squamous cell carcinoma of thoracic esophagus. Int J Radiat Oncol Biol Phys 2011;81:e759-65.
23. Manabe O, Hattori N, Hirata K, et al. Diagnostic accuracy of lymph node metastasis depends on metabolic activity of the primary lesion in thoracic squamous esophageal cancer. J Nucl Med 2013;54:670-6.
24. Kato H, Kimura H, Nakajima M, et al. The additional value of integrated PET/CT over PET in initial lymph node staging of esophageal cancer. Oncol Rep 2008;20:857-62.
25. Kim SJ, Pak K, Chang S. Determination of regional lymph node status using 18F-FDG PET/CT parameters in oesophageal cancer patients: comparison of SUV, volumetric parameters and intratumoral heterogeneity. Br J Radiol 2016;89:20150673.
26. Kim SH, Lee KN, Kang EJ, et al. Hounsfield units upon PET/CT are useful in evaluating metastatic regional lymph nodes in patients with oesophageal squamous cell carcinoma. Br J Radiol 2012;85:606-12.
27. Schreurs LM, Pultrum BB, Koopmans KP, et al. Better assessment of nodal metastases by PET/CT fusion compared to side-by-side PET/CT in oesophageal cancer. Anticancer Res 2008;28:1867-73.
28. Bella AJE, Zhang YR, Fan W, et al. Maximum standardized uptake value on PET/CT in preoperative assessment of lymph node metastasis from thoracic esophageal squamous cell carcinoma. Chin J Cancer 2014;33:211-7.
29. Okada M, Murakami T, Kumano S, et al. Integrated FDG-PET/CT compared with intravenous contrast-enhanced CT for evaluation of metastatic regional lymph nodes in patients with resectable early stage esophageal cancer. Ann Nucl Med 2009;23:73-80.
30. Futamura M, Asano T, Kobayashi K, et al. Prediction of macrometastasis in axillary lymph nodes of patients with invasive breast cancer and the utility of the SUV lymph node/tumor ratio using FDG-PET/CT. World J Surg Oncol 2015;13:49.
31. Mattes MD, Moshchinsky AB, Ahsanuddin S, et al. Ratio of Lymph Node to Primary Tumor SUV on PET/CT Accurately Predicts Nodal Malignancy in Non-Small-Cell Lung Cancer. Clin Lung Cancer 2015;16:e253-8.
32. Payabvash S, Meric K, Cayci Z. Differentiation of benign

from malignant cervical lymph nodes in patients with head and neck cancer using PET/CT imaging. Clin Imaging 2016;40:101-5.
33. Hiyoshi Y, Watanabe M, Imamura Y, et al. The relationship between the glucose transporter type 1 expression and 18F-fluorodeoxyglucose uptake in esophageal squamous cell carcinoma. Oncology 2009;76:286-92.
34. Luketich JD, Schauer PR, Meltzer CC, et al. Role of positron emission tomography in staging esophageal cancer. Ann Thorac Surg 1997;64:765-9.
35. Kato H, Kuwano H, Nakajima M, et al. Comparison between positron emission tomography and computed tomography in the use of the assessment of esophageal carcinoma. Cancer 2002;94:921-8.
36. Glas AS, Lijmer JG, Prins MH, et al. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol 2003;56:1129-35.
37. 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.
38. Yongkui L, Jian L, Jingui L. 18 FDG-PET/CT for the detection of regional nodal metastasis in patients with primary head and neck cancer before treatment: A meta-analysis. Surg Oncol 2013;22:e11-6.
39. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234-48.
40. Flechsig P, Frank P, Kratochwil C, et al. Radiomic Analysis using Density Threshold for FDG-PET/CT-Based N-Staging in Lung Cancer Patients. Mol Imaging Biol 2017;19:315-22.
41. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.

Cite this article as: Jiang C, Chen Y, Zhu Y, Xu Y. Systematic review and meta-analysis of the accuracy of 18F-FDG PET/CT for detection of regional lymph node metastasis in esophageal squamous cell carcinoma. J Thorac Dis 2018;10(11):6066-6076. doi: 10.21037/jtd.2018.10.57

Supplementary

QUADAS-2

Phase 1: State the review question:

Patients (setting, intended use of index test, presentation, prior testing):
Index test(s):
Reference standard and target condition:

Phase 2: Draw a flow diagram for the primary study

Phase 3: Risk of bias and applicability judgments

QUADAS-2 is structured so that 4 key domains are each rated in terms of the risk of bias and the concern regarding applicability to the research question (as defined above). Each key domain has a set of signalling questions to help reach the judgments regarding bias and applicability.

DOMAIN 1: PATIENT SELECTION

A. Risk of Bias

Describe methods of patient selection:

- ❖ Was a consecutive or random sample of patients enrolled? Yes/No/Unclear
- ❖ Was a case-control design avoided? Yes/No/Unclear
- ❖ Did the study avoid inappropriate exclusions? Yes/No/Unclear

Could the selection of patients have introduced bias? **RISK: LOW/HIGH/UNCLEAR**

B. Concerns regarding applicability

Describe included patients (prior testing, presentation, intended use of index test and setting):

Is there concern that the included patients do not match the review question? **CONCERN: LOW/HIGH/UNCLEAR**

DOMAIN 2: INDEX TEST(S)

If more than one index test was used, please complete for each test.

A. Risk of Bias

Describe the index test and how it was conducted and interpreted:

- ❖ Were the index test results interpreted without knowledge of the results of the reference standard? Yes/No/Unclear
- ❖ If a threshold was used, was it pre-specified? Yes/No/Unclear

Could the conduct or interpretation of the index test have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the index test, its conduct, or interpretation differ from the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 3: REFERENCE STANDARD

A. Risk of Bias

Describe the reference standard and how it was conducted and interpreted:

- ❖ Is the reference standard likely to correctly classify the target condition? Yes/No/Unclear
- ❖ Were the reference standard results interpreted without knowledge of the results of the index test? Yes/No/Unclear

Could the reference standard, its conduct, or its interpretation have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

B. Concerns regarding applicability

Is there concern that the target condition as defined by the reference standard does not match the review question? **CONCERN: LOW /HIGH/UNCLEAR**

DOMAIN 4: FLOW AND TIMING

A. Risk of Bias

Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram):

Describe the time interval and any interventions between index test(s) and reference standard:

- ❖ Was there an appropriate interval between index test(s) and reference standard? Yes/No/Unclear
- ❖ Did all patients receive a reference standard? Yes/No/Unclear
- ❖ Did patients receive the same reference standard? Yes/No/Unclear
- ❖ Were all patients included in the analysis? Yes/No/Unclear

Could the patient flow have introduced bias? **RISK: LOW /HIGH/UNCLEAR**

Figure S1 The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2).