BMJ Open

Cervical lymph node metastasis in papillary thyroid cancer: a meta-analysis of ultrasound and CT for diagnosis and management

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051568
Article Type:	Original research
Date Submitted by the Author:	09-Apr-2021
Complete List of Authors:	Yang, Jian ; Xishan Coal Electricity Group Workers General Hospital, Department of Radiology Zhang, Fengyan; First Clinical Medical College, Shanxi Medical University, Department of Radiology Qiao, Ying; First Clinical Medical College, Shanxi Medical University, Department of Radiology
Keywords:	Ultrasound < RADIOLOGY & IMAGING, Computed tomography < RADIOLOGY & IMAGING, Computed tomography < RADIOTHERAPY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Cervical lymph node metastasis in papillary thyroid cancer: a meta-analysis of ultrasound and CT for diagnosis and management

Jian Yang[1,†], Fengyan Zhang[2,†], Ying Qiao[2,*]

[1]Department of Radiology, Xishan Coal Electricity Group Workers General Hospital, Taiyuan, 030053, Shanxi Province, China

[2]Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

[†]These authors contributed equally to this work.

[*]**Correspondence:**

Ying Qiao

Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

Tel: +86-15103462912

E-mail: qiaoying910723@163.com

Running Head: US and CT for CLNM for PTC

Abstract

Objectives: To determine the diagnostic accuracy of ultrasound (US), computed tomography (CT) and their combination in detecting cervical lymph node metastasis (CLNM) in patients with papillary thyroid cancer (PTC).

Design: This is a meta analysis study.

Setting: Not applicable.

Participants: Patients with PTC level-by-level.

Interventions: Not applicable.

Primary and secondary outcome: Studies that reported the absolute numbers of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) results directly.

Measures: Medline (via PubMed), Web of Science and Embase were searched to identify studies that used both US and CT to detect CLNM in patients with PTC. Primary outcomes were sensitivity, specificity, and diagnostic odds ratios (DORs) in level-by-level or patient-based analysis. Secondary outcomes were sensitivity, specificity and DORs in central and lateral compartments.

Results: Fourteen studies involving 6167 patients with 11601 neck lymph nodes met the inclusion criteria. Based on level-by-level analysis, the pooled sensitivity, specificity, DORs for US were 0.35(95% confidence interval [(CI) 0.34-0.37], 0.95(95% CI 0.94-0.95), 13.94(95% CI 9.34-20.82), for CT were 0.46(95% CI 0.44-0.47), 0.88(95% CI 0.87-0.89), 7.24(95% CI 5.46-9.62), for the combination of US and CT were 0.51(95% CI 0.49-0.52), 0.85(95% CI 0.84-0.86), 6.01(95% CI 3.84-9.40), respectively. The pooled estimates of sensitivity, specificity, and diagnostic OR of US were 0.41(95% CI 0.36-

0.46), 0.92(95% CI 0.89-0.94), 7.56(95% CI 4.08-14.01), of CT were 0.49(0.44-0.54), 0.91(0.89-0.94), 9.40(5.79-15.27), of combination of US and CT were 0.64(95% CI 0.57-0.71), 0.83(95% CI 0.77-0.88), 8.59(95% CI 5.37-13.76) on patient-based analysis, respectively.

Conclusions: These findings suggest that US, with a DOR of almost twice that for CT on level-by-level analysis, was superior for CT in detecting CLNM in patients with PTC, especially at lateral compartment, and the combination of US and CT increased sensitivity by 30-50% on patient-based analysis.

Strengths and limitations of this study

1. Fourteen studies involving 6167 patients with 11601 neck lymph nodes met the inclusion criteria.

2. Based on level-by-level analysis, the pooled sensitivity, specificity, DORs for US were 0.35, 0.95, 13.94, for CT were 0.46, 0.88, 7.24, for the combination of US and CT were 0.51, 0.85, 6.01, respectively

3.The literature included is limited due to the study design and timing of imaging

Keywords: ultrasound; computed tomography; cervical lymph node metastasis; papillary thyroid cancer; meta-analysis

Introduction

Papillary thyroid carcinoma (PTC) is an endocrine neoplasia with high incidence of lymphatic metastasis and is associated with regional recurrence[1-3] . The incidence of cervical lymph node metastasis (CLNM) in patients with thyroid cancer has been reported to be 20–90%[4]. The presence of CLNM might increase the risk of locoregional recurrence after surgery [5, 6], decreasing the survival rate in patients[7]. Therefore, it is of great clinical importance to accurately evaluate CLNM and determine appropriate extent of neck dissection[8]. Although prophylactic central compartment neck dissection (ipsilateral or bilateral) is recommended by American Thyroid Association (ATA) guidelines in patients with clinically involved central nodes, especially for those with advanced primary tumors, the information regarding the prophylactic lateral compartment neck dissection has not been clearly stated[8]. Thus, the indications for neck dissection, especially the lateral compartment, should be carefully assessed as it might lead to severe postoperative complications[9].

Preoperative staging with ultrasound (US) for cervical lymph nodes, including both central and lateral neck compartments, is the most widely accepted first imaging technique for patients with thyroid or suspicious malignancies cytologic or molecular findings, while preoperative use of computed tomography (CT) with intravenous (IV) contrast is complementary to US in patients with advanced disease[8, 10, 11]. Although several studies have failed to prove the benefit of CT over US in detecting lateral lymph node metastasis[12-14], few studies have suggested superior diagnostic performance of preoperative combination of US with CT over US alone[13, 15-18].

Meta-analyses examining the diagnostic accuracy of US and CT in detecting cervical

CLNM in patients with PTC have been previously conducted[19-23]. However, these meta-analyses studies have integrated the findings of US and CT from different studies and populations.

To our knowledge, no previous meta-analysis included studies that evaluated cervical CLNM in patients with PTC using both US and CT, which could minimize the confounding effect of operator in interpreting the diagnostic accuracy of preoperative imaging. This meta-analysis aimed to evaluate the sensitivity, specificity, and diagnostic odds ratios of US, CT, and the combination of both in assessing cervical CLNM in patients with PTC based on the central and lateral neck levels and by using level-by-level and patient-based analyses.

Methods

Institutional Review Board approval was not required because this article is a meta-analysis. The data comes from published articles and does not require ethical approval.

Systematic Literature Research

This meta-analysis was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines[24]. Ethical approval was waived off due to secondary data acquisition from previously published papers that are available in the public domain. A systematic search of the Medline (via PubMed), Web of Science and Embase till June 1, 2020 was conducted to identify studies that assessed the accuracy of US and CT in detecting CLNM in patients with PTC. The search strategy was developed in collaboration with a hospital librarian and included subject headings and text words, which were as follows: ("thyroid cancer" OR "thyroid cancinoma" OR "thyroid tumor"

OR “papillary thyroid cancer” OR “thyroid neoplasm”) AND (“cervical lymph node” OR “neck lymph node”) AND (“metastasis” or “metastatic”) AND (“ultrasonography” or “ultrasound” or “US”) AND (computed tomography” or “CT”). The studies were initially screened by examining their titles and abstracts, and the full-texts of potentially eligible studies were retrieved for further review. No language restriction was applied. A manual search of additional records and reference lists was also performed to include more relevant studies.

Study selection

The inclusion criteria of the studies were as follows: (a) prospective or retrospective studies that evaluated the diagnostic accuracy of both US and CT for detecting CLNM in patients with PTC level-by-level or patient-based analysis; (b) studies with population >10 patients; (c) studies with reference standard of histopathology or cytology (the diagnostic gold standard was the pathological diagnosis of the resected lymph nodes); (d) studies that reported the absolute numbers of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) results directly or derived from the reported data or communicated by authors in response to our request; and (e) studies published in English.

Exclusion criteria were as follows: (i) case reports, case series, review articles, pictorial essays, letters to editors, unpublished data, conference abstracts, and proceedings on the topic of interest; (ii) studies that used only US or only CT; (iii) insufficient data regarding TP, FP, FN, and TN; (iv) duplicate publications in different databases and studies; (v) if the patient population of one article is overlapping with the patient population of other or multiple articles, then the article with larger sample size

was included; and (vi) studies with less than 10 cases as confirmed by reference standard. One reader reviewed the full-texts of candidate articles and selected those that met the inclusion criteria. A second reader reviewed the process of inclusion of articles in the meta-analysis. No inter-reader disagreements were observed.

Primary and secondary outcomes

Primary outcomes were sensitivity, specificity, and diagnostic odds ratios (DORs) in level-by-level or patient-based analysis. Secondary outcomes were sensitivity, specificity and DORs in central and lateral compartments in level-by-level or patient-based analysis.

Data Extraction and quality assessment

Two reviewers independently performed data extraction. Data such as study characteristics, clinical and patient characteristics, reference standard or standards, cervical lymph node compartment, technical characteristics of CT and US and contrast enhancement, the definition of CLNM according to CT and US image findings; and the diagnostic performance of CT and US, such as TP, FP, FN, TN were obtained from each study.

Two reviewers who were not blinded to the journal names, author names, and year of publication assessed the methodologic and reporting quality of each study by using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) [25]. Each study was independently assessed by two reviewers after a tutorial meeting on the guidelines for interpreting the items. Any disagreements were resolved by discussion with an experienced third reviewer.

Statistical analysis

The pooled sensitivity, specificity, diagnostic odds ratio (OR), positive likelihood ratio (LR+) and negative likelihood ratio (LR-) were calculated for US and CT in level-by-level analysis (at neck level, central neck level and lateral neck level) and patient-based analysis (patient level, central patient level and lateral level). The heterogeneity of pooled sensitivity, specificity, diagnostic OR, LR+ and LR- was measured by the inconsistency (I^2). Heterogeneity in the included articles was defined as small I^2<25%, moderate I^2 25–50%, and obvious I^2> 50%. If heterogeneity was detected (P-value < 0.10 or I^2≥50 %), then a random-effects model was applied; otherwise, a fixed-effects model was used. Bivariate logistic regression model was used for meta-analysis of diagnostic test accuracy[26] and forest plots were created. Pooling of sensitivity, specificity, diagnostic OR, LR+ and LR-, was performed with Meta-Disc software (version 1.4, Madrid, Spain). Foreplot, and summary receiver operating characteristic (SROC) curve were produced using RevMan 5.3. A P value of less than 0.05 was used as the threshold to indicate statistical significance.

Results

Characteristics of included studies

Initial literature search yielded 1135 potential studies for this meta-analysis. A total of 449 articles were screened after removing duplications. Of these, 372 studies were excluded by reviewing the titles and abstracts, and 63 articles were excluded after reviewing the full-texts (Figure 1). Fourteen studies were ultimately selected for inclusion[12, 13, 15-18, 27-34]: 10 studies based on level-by-level analysis, 2 articles based on

patient-based analysis, and 2 articles based on both level-by-level and patient-based analyses. Five studies have reported diagnostic performance by combining both US and CT[13, 15-18]. A total of 6167 patients with 11601 neck lymph nodes were included, and all patients were diagnosed with PTC except one who was diagnosed with medullary thyroid cancer. The earliest study was started in 1997, whereas the latest one was started in 2012. The median number of patients per study was 171 (range 20–3668), while the median number of lymph nodes per study was 331 (range 107–6557). Eleven were retrospective studies and 3 were prospective studies, 13 studies were performed preoperatively and 1 study was performed postoperatively. Twelve, one and one were conducted in Korea, the United states and Japan, respectively (Table 1). The studies included in this meta-analysis was of moderate quality (supplementary Figure 1[sFigure 1], sFigure 2).

The diagnostic accuracy of US and CT in level-by-level analysis

There were 11 studies that included both CT and US for detecting CLNM in patients with PTC and 5 of them assessed the diagnostic accuracy of the combination of CT and US. The pooled sensitivity, specificity, diagnostic OR, LR+ and LR- for US were 0.35(95% CI 0.34-0.37), 0.95(95% CI 0.94-0.95), 13.94(95% CI 9.34-20.82), 6.79(95% CI 4.79-9.63), 0.50(95% CI 0.41-0.60), for CT were 0.46(95% CI 0.44-0.47), 0.88(95% CI 0.87-0.89), 7.24(95% CI 5.46-9.62), 3.77(95% CI 2.08-6.84), 0.52(95% CI 0.45-0.61), and for the combination of US and CT were 0.51(95% CI 0.49-0.52), 0.85(95% CI 0.84-0.86), 6.01(95% CI 3.84-9.40), 3.04(95% CI 1.93-4.80), 0.52(95% CI 0.45-0.60) with marked heterogeneity (Table 2, Figure 2, Figure 3).

Subgroup analysis of central and lateral neck level were performed to investigate the effects of cervical lymph node compartment based on the diagnostic accuracy of US and

CT. Subgroup analysis of central neck level revealed that the pooled sensitivity, specificity, and diagnostic OR of US were 0.28(95% CI 0.24-0.32), 0.97(95% CI 0.96-0.98), 14.07(95% CI 6.66-29.71) from 4 studies, of CT were 0.32(95% CI 0.28-0.36), 0.89(95% CI 0.86-0.91), 5.48(95% CI 2.15-13.98) from 4 studies, and of the combination of US and CT were 0.40(95% CI 0.35-0.45), 0.85(95% CI 0.82-0.88), 4.32(95% CI 2.09-8.92) from 3 studies, respectively (Table 2, sFigure 3, sFigure 4).

In contrast, subgroup analysis of lateral neck level revealed that the pooled sensitivity, specificity, and diagnostic OR of US were 0.74(95% CI 0.69-0.78), 0.92(95% CI 0.90-0.94), 24.41(95% CI 11.16 -53.42) from 6 studies, of CT were 0.73(95% CI 0.68-0.77) , 0.89(95% CI 0.87-0.91), 15.55(95% CI 7.98 -30.32) from 6 studies, and of the combination of US and CT were 0.88(95% CI 0.83-0.91), 0.79(95% CI 0.73-0.84), 22.59(95% CI 11.29 -45.19) from 4 studies, respectively (Table 2, sFigure 5, sFigure 6).

The diagnostic accuracy of US and CT in patient-based analysis

There were 4 studies that included both US and CT in detecting CLNM in patients with PTC and 2 of them assessed the diagnostic accuracy by combining both CT and US. The pooled estimates of sensitivity, specificity, and diagnostic OR of US were 0.41(95% CI 0.36-0.46), 0.92(95% CI 0.89-0.94), 7.56(95% CI 4.08-14.01), of CT were 0.49(0.44-0.54), 0.91(0.89-0.94), 9.40(5.79-15.27), and of the combination of US and CT were 0.64(95% CI 0.57-0.71), 0.83(95% CI 0.77-0.88), 8.59(95% CI 5.37-13.76) , respectively (Table 2, Figure 4, Figure 5).

There are only two studies that assessed the diagnostic accuracy of US, CT, and the combination of both on patient-based analysis. On central patient level, the pooled estimates of sensitivity, specificity, and diagnostic OR of US were 0.21(95% CI 0.16-

0.28), 0.95(95% CI 0.91-0.97), 4.53(95% CI 2.34-8.77), of CT were 0.38(95% CI 0.32-0.46), 0.90(95% CI 0.85-0.93), 5.02(95% CI 0.46-54.54), and of the combination of CT and US were 0.47(95% CI 0.39-0.54), 0.85(95% CI 0.80-0.89), 4.88(95% CI 2.58-9.23), respectively (Table 2, sFigure 7, sFigure 8).

In contrast, the pooled estimates of sensitivity, specificity, and diagnostic OR of US were 0.87(95% CI 0.74-0.95), 0.89(95% CI 0.83-0.93), 20.11(95% CI 6.77-59.70), of CT were 0.92(95% CI 0.81-0.98), 0.88(95% CI 0.83-0.93), 36.88(95% CI 11.40 -119.35), and of the combination of US and CT were 0.98(95% CI 0.89-0.99), 0.92(95% CI 0.87-0.96), 78.10(95% CI 2.82 -2160.4) from central patient level (Table 2, sFigure 9, sFigure 10).

Discussion

The main findings of this meta-analysis demonstrated that the DORs of US by level-by-level analysis was higher than CT or the combination of both on central, lateral and neck levels. Differentiated thyroid carcinoma, particularly PTC, involves CLNM in 20%–50% of patients [35-38], which could prevent small and intrathyroidal primary tumors[39]. However, the clinical implications of macro-metastases (≥2mm) are more significant when compared to micro-metastases, in which 90% of patients might reach according to the sensitivity of the imaging methods[40, 41]. The combination of US features might increase the likelihood for detecting CLNM as several US features are suggestive of metastatic lymph nodes, including enlargement, loss of fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, peripheral vascularity, calcifications, etc[42]. Preoperative US identifies lymph node or soft-tissue metastases in up to 39% of patients

who had no physical examination[43], and changed the operative management in 23% patients[44].

Our data found that the DORs of CT was higher than US and the combination and the DORs of the combination remained higher than US and CT by patient-based analysis. This was reasonable because the sensitivity of CT on patient-based analysis was higher than that of the US on central, lateral and patient level analysis, respectively. This result might still need further investigation because of the inclusion of small number of studies in the subgroup analysis. The operator independent CT could be used as an adjunct in imaging deep anatomic structures, including the mediastinum, infraclavicular, retropharyngeal, and parapharyngeal regions and those structures that are acoustically shadowed by bone or air. In addition, preoperative knowledge on the extent of laryngeal, tracheal, esophageal involvement, as well as bulky nodal disease from neck CT with contrast significantly influences the surgical plan by indicating the need for sternotomy, tracheal or laryngeal resection and reconstruction[45].

Our results suggested that the sensitivity on the lateral compartment tend to be higher than that on the central compartment regardless of the use of US, CT or the combination of both by level-by-level and patient-based analysis, respectively. The location of the lymph nodes helps in decision-making as most of the metastatic nodes situated in the lower third of the neck and reactive enlarged lymph nodes occurred in the upper part of the neck[46]. Besides, the lateral compartment should be carefully evaluated for skip metastases that are located in the upper pole, or ≤1cm in diameter[47]. For patients who had preoperative CT and US and subsequently underwent total thyroidectomy and neck dissection, the sensitivity of CT was shown to be much better than US for evaluating

CLNM on neck level, but the sensitivity, specificity, and DORs for lateral neck level tended to be higher than those of the central neck level for both CT and US[12]. Dual-energy computed tomography (DECT) for assessing CLNM in patients with PTC was not included in this meta-analysis as it can generate iodine-based material decomposition (MD) images and spectral HU curve[48-50]. In accordance with the findings from CT, combined gemstone spectral image (GSI) parameters from DECT also demonstrated better diagnostic accuracy of CLNM in patients with PTC when compared to those that are obtained by combining the US morphological parameters, especially in the lateral compartment[50].

Our findings revealed that compared to US or CT alone, the combination of both US and CT demonstrated higher sensitivity, i.e., a meta-analytic summary sensitivity of 0.51(0.49-0.52) and 0.64(0.57-0.71), and a lower specificity, i.e., a meta-analytic summary specificity of 0.85(0.84-0.86) and 0.83(0.77-0.88) for evaluating CLNM in patients with PTC by level-by-level and patient-based analysis, respectively. In patients undergoing primary and revision surgical treatment for PTC, combined preoperative mapping with US and CT yielded significantly higher sensitivity for detecting macroscopic lymph-nodes in both lateral and central neck, especially in the central neck[32].

It should be noted that the study has strengths. Firstly, Boolean operatives of AND rather than OR were used for combined datasets for all studies. Namely, only studies of direct head-to-head comparison by US, CT, and combination of both in the same patient population were included in this meta-analysis, avoiding bias due to differences in patient and institutional factors. Secondly, meta-analysis of the included studies was performed by using level-by-level and patient-based analyses and on all, central and lateral neck

levels. Lastly, our data suggested that future follow-up study should be performed to determine the comparative role of US and CT in identifying false negative nodes which are not biopsied or excised.

Despite great clinical significance, there are several limitations in the current meta-analysis that are mostly associated with the available data and heterogeneity of design, interpretation of results, and reporting of data in primary studies. Firstly, the sources of heterogeneity among primary studies in meta-analysis studies have been reported by several previous studies, which included contrast amount, scan phase, and reconstruction slice thickness for CT[20], and the criteria of lymph node diameter and vascular flow for US[23]. Secondly, the literature included is limited due to the study design and timing of imaging. Eleven of 14 studies (78.6%) were retrospective and 1 of the 14 studies was imaging postoperative study. Large proportion of retrospective studies might increase the sensitivity of CT and US. Twelve of the 14 studies were conducted in Korea, and so ethnic factor might affect the results of this meta-analysis. Thus, the complementary use CT may be routine in Korea but not necessarily applicable to other parts of the world, especially in lesser developed countries. Thirdly, modern high resolution US transducers have a lateral resolution of 2mm which is not feasible for CT, allowing for the detection of small nodes and the presence of microcalcification. The included CT studies may not be comparable from one study to another, particularly over the decade as it depends on the equipment, slice thickness, amount of contrast injected etc. Fourthly, 4 of the included 14 studies were with patient-based results and 12 of 14 studies were of suboptimal quality, and no definite recommendation could be drawn from the present study. Finally, MRI, US-guided FNA and PET-CT were not included in the meta-analysis

in order to directly compare CT and US although they are also paly complementary role in the management of cervical lymph node metastasis in papillary thyroid cancer.

Despite these potential drawbacks, this meta-analysis demonstrated the unique complementary value of CT secondary to US in detecting CLNM in patients with PTC by patient-based analysis. More importantly, the choice of diagnostic test should be tailored to have feasible access to these imaging modalities at individual healthcare centers.

Conclusion

These findings suggest that US, with a DOR of almost twice that for CT on level-by-level analysis, was superior for CT in detecting CLNM in patients with PTC, especially at lateral compartment, and the combination of US and CT increased sensitivity by 30-50% on patient-based analysis. The CT might be valid as candidate imaging techniques secondary to US in the management of CLNM in patients with PTC.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship

1. guarantor of integrity of the entire study: Jian Yang

2. study concepts and design: Ying Qiao

3. literature research: Jian Yang

4. clinical studies: Ying Qiao

5. experimental studies / data analysis: Jian Yang

6. statistical analysis: Fengyan Zhang

7. manuscript preparation: Ying Qiao

8. manuscript editing: Fengyan Zhang

Acknowledgements

None.

Funding

None.

Data Sharing Statement

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet* 2016;**388**(10061):2783-95.

2. Gubbi S, Thakur S, Avadhanula S*, et al.* Comprehensive guidance on the diagnosis and management of primary mesenchymal tumours of the thyroid gland. *Lancet Oncol* 2020;**21**(11):e528-e37.

3. Durante C, Grani G, Lamartina L*, et al.* The Diagnosis and Management of Thyroid Nodules: A Review. *JAMA* 2018;**319**(9):914-24.

4. Wang TS, Dubner S, Sznyter LA*, et al.* Incidence of metastatic well-differentiated thyroid cancer in cervical lymph nodes. *Arch Otolaryngol Head Neck Surg* 2004;**130**(1):110-3.

5. Mazzaferri EL, Kloos RT. Clinical review 128: Current approaches to primary therapy for papillary and follicular thyroid cancer. *J Clin Endocrinol Metab* 2001;**86**(4):1447-63.

6. Lee BJ, Wang SG, Lee JC*, et al.* Level IIb lymph node metastasis in neck dissection for papillary thyroid carcinoma. *Arch Otolaryngol Head Neck Surg* 2007;**133**(10):1028-30.

7. Lundgren CI, Hall P, Dickman PW*, et al.* Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. *Cancer* 2006;**106**(3):524-31.

8. Haugen BR, Alexander EK, Bible KC*, et al.* 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines

Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* 2016;**26**(1):1-133.

9. Cheah WK, Arici C, Ituarte PH, *et al.* Complications of neck dissection for thyroid cancer. *World J Surg* 2002;**26**(8):1013-6.
10. American Thyroid Association Guidelines Taskforce on Thyroid N, Differentiated Thyroid C, Cooper DS, *et al.* Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009;**19**(11):1167-214.
11. Stack BC, Jr., Ferris RL, Goldenberg D, *et al.* American Thyroid Association consensus review and statement regarding the anatomy, terminology, and rationale for lateral neck dissection in differentiated thyroid cancer. *Thyroid* 2012;**22**(5):501-8.
12. Ahn JE, Lee JH, Yi JS, *et al.* Diagnostic accuracy of CT and ultrasonography for evaluating metastatic cervical lymph nodes in patients with thyroid cancer. *World J Surg* 2008;**32**(7):1552-8.
13. Choi JS, Kim J, Kwak JY, *et al.* Preoperative staging of papillary thyroid carcinoma: comparison of ultrasound imaging and CT. *AJR Am J Roentgenol* 2009;**193**(3):871-8.
14. Liu X, Ouyang D, Li H, *et al.* Papillary thyroid cancer: dual-energy spectral CT quantitative parameters for preoperative diagnosis of metastasis to the cervical lymph nodes. *Radiology* 2015;**275**(1):167-76.
15. Kim E, Park JS, Son KR, *et al.* Preoperative diagnosis of cervical metastatic lymph nodes in papillary thyroid carcinoma: comparison of ultrasound, computed

tomography, and combined ultrasound with computed tomography. *Thyroid* 2008;**18**(4):411-8.

16. Lee DW, Ji YB, Sung ES, *et al.* Roles of ultrasonography and computed tomography in the surgical management of cervical lymph node metastases in papillary thyroid carcinoma. *Eur J Surg Oncol* 2013;**39**(2):191-6.
17. Lee Y, Kim JH, Baek JH, *et al.* Value of CT added to ultrasonography for the diagnosis of lymph node metastasis in patients with thyroid cancer. *Head Neck* 2018;**40**(10):2137-48.
18. Na DK, Choi YJ, Choi SH, *et al.* Evaluation of cervical lymph node metastasis in thyroid cancer patients using real-time CT-navigated ultrasonography: preliminary study. *Ultrasonography* 2015;**34**(1):39-44.
19. Suh CH, Baek JH, Choi YJ, *et al.* Performance of CT in the Preoperative Diagnosis of Cervical Lymph Node Metastasis in Patients with Papillary Thyroid Cancer: A Systematic Review and Meta-Analysis. *AJNR Am J Neuroradiol* 2017;**38**(1):154-61.
20. Cho SJ, Suh CH, Baek JH, *et al.* Diagnostic performance of CT in detection of metastatic cervical lymph nodes in patients with thyroid cancer: a systematic review and meta-analysis. *Eur Radiol* 2019;**29**(9):4635-47.
21. Raijmakers PG, Paul MA, Lips P. Sentinel node detection in patients with thyroid carcinoma: a meta-analysis. *World J Surg* 2008;**32**(9):1961-7.
22. Wu LM, Gu HY, Qu XH, *et al.* The accuracy of ultrasonography in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma: A meta-analysis. *Eur J Radiol* 2012;**81**(8):1798-805.

23. Zhao H, Li H. Meta-analysis of ultrasound for cervical lymph nodes in papillary thyroid cancer: Diagnosis of central and lateral compartment nodal metastases. *Eur J Radiol* 2019;**112**:14-21.

24. Liberati A, Altman DG, Tetzlaff J*, et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;**6**(7):e1000100.

25. 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**(8):529-36.

26. 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**(10):982-90.

27. Jeong HS, Baek CH, Son YI*, et al.* Integrated 18F-FDG PET/CT for the initial evaluation of cervical node level of patients with papillary thyroid carcinoma: comparison with ultrasound and contrast-enhanced CT. *Clin Endocrinol (Oxf)* 2006;**65**(3):402-7.

28. Choi YJ, Yun JS, Kook SH*, et al.* Clinical and imaging assessment of cervical lymph node metastasis in papillary thyroid carcinomas. *World J Surg* 2010;**34**(7):1494-9.

29. Morita S, Mizoguchi K, Suzuki M*, et al.* The accuracy of (18)[F]-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography, ultrasonography, and enhanced computed tomography alone in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma. *World J Surg* 2010;**34**(11):2564-9.

30. Yoon JH, Kim JY, Moon HJ, *et al.* Contribution of computed tomography to ultrasound in predicting lateral lymph node metastasis in patients with papillary thyroid carcinoma. *Ann Surg Oncol* 2011;**18**(6):1734-41.

31. Seo YL, Yoon DY, Baek S, *et al.* Detection of neck recurrence in patients with differentiated thyroid cancer: comparison of ultrasound, contrast-enhanced CT and (18)F-FDG PET/CT using surgical pathology as a reference standard: (ultrasound vs. CT vs. (18)F-FDG PET/CT in recurrent thyroid cancer). *Eur Radiol* 2012;**22**(10):2246-54.

32. Lesnik D, Cunnane ME, Zurakowski D, *et al.* Papillary thyroid carcinoma nodal surgery directed by a preoperative radiographic map utilizing CT scan and ultrasound in all primary and reoperative patients. *Head Neck* 2014;**36**(2):191-202.

33. Kim SK, Woo JW, Park I, *et al.* Computed Tomography-Detected Central Lymph Node Metastasis in Ultrasonography Node-Negative Papillary Thyroid Carcinoma: Is It Really Significant? *Ann Surg Oncol* 2017;**24**(2):442-9.

34. Eun NL, Son EJ, Kim JA, Gweon HM,et al. Comparison of the diagnostic performances of ultrasonography, CT and fine needle aspiration cytology for the prediction of lymph node metastasis in patients with lymph node dissection of papillary thyroid carcinoma: A retrospective cohort study. Int J Surg 2018; 51:145-150.

35. Scheumann GF, Gimm O, Wegener G, *et al.* Prognostic significance and surgical management of locoregional lymph node metastases in papillary thyroid cancer. *World J Surg* 1994;**18**(4):559-67; discussion 67-8.

36. Grebe SK, Hay ID. Thyroid cancer nodal metastases: biologic significance and therapeutic considerations. *Surg Oncol Clin N Am* 1996;**5**(1):43-63.

37. Nam-Goong IS, Kim HY, Gong G, *et al.* Ultrasonography-guided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. *Clin Endocrinol (Oxf)* 2004;**60**(1):21-8.

38. Whiting PF, Weswood ME, Rutjes AW, *et al.* Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006;**6**:9.

39. Hay ID, Grant CS, van Heerden JA, *et al.* Papillary thyroid microcarcinoma: a study of 535 cases observed in a 50-year period. *Surgery* 1992;**112**(6):1139-46; discussion 46-7.

40. Qubain SW, Nakano S, Baba M, *et al.* Distribution of lymph node micrometastasis in pN0 well-differentiated thyroid carcinoma. *Surgery* 2002;**131**(3):249-56.

41. Arturi F, Russo D, Giuffrida D, *et al.* Early diagnosis by genetic analysis of differentiated thyroid cancer metastases in small lymph nodes. *J Clin Endocrinol Metab* 1997;**82**(5):1638-41.

42. Leboulleux S, Girard E, Rose M, *et al.* Ultrasound criteria of malignancy for cervical lymph nodes in patients followed up for differentiated thyroid cancer. *J Clin Endocrinol Metab* 2007;**92**(9):3590-4.

43. Kouvaraki MA, Shapiro SE, Fornage BD, *et al.* Role of preoperative ultrasonography in the surgical management of patients with thyroid cancer. *Surgery* 2003;**134**(6):946-54; discussion 54-5.

44. O'Connell K, Yen TW, Quiroz F, *et al.* The utility of routine preoperative cervical ultrasonography in patients undergoing thyroidectomy for differentiated thyroid cancer. *Surgery* 2013;**154**(4):697-701; discussion -3.

45. Yeh MW, Bauer AJ, Bernet VA, *et al.* American Thyroid Association statement on preoperative imaging for thyroid cancer surgery. *Thyroid* 2015;**25**(1):3-14.

46. Kuna SK, Bracic I, Tesic V, *et al.* Ultrasonographic differentiation of benign from malignant neck lymphadenopathy in thyroid cancer. *J Ultrasound Med* 2006;**25**(12):1531-7; quiz 8-40.

47. Park JH, Lee YS, Kim BW, *et al.* Skip lateral neck node metastases in papillary thyroid carcinoma. *World J Surg* 2012;**36**(4):743-7.

48. Zhao Y, Li X, Li L, *et al.* Preliminary study on the diagnostic value of single-source dual-energy CT in diagnosing cervical lymph node metastasis of thyroid carcinoma. *J Thorac Dis* 2017;**9**(11):4758-66.

49. He M, Lin C, Yin L, *et al.* Value of Dual-Energy Computed Tomography for Diagnosing Cervical Lymph Node Metastasis in Patients With Papillary Thyroid Cancer. *J Comput Assist Tomogr* 2019;**43**(6):970-5.

50. Li L, Cheng SN, Zhao YF, *et al.* Diagnostic accuracy of single-source dual-energy computed tomography and ultrasonography for detection of lateral cervical lymph node metastases of papillary thyroid carcinoma. *J Thorac Dis* 2019;**11**(12):5032-41.

Figure legend

Figure 1. Flowchart of literature search process

Figure 2. Forest plots for the sensitivities and specificities of US, CT, and combination in level-by-level analysis

Figure 3. SROC of US, CT, and combination in detecting CLNM in patients with PTC in level-by-level analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer.

Figure 4. Forest plots for the sensitivities and specificities of US, CT, and combination in patient-based analysis

Figure 5. SROC of US, CT, and combination in detecting CLNM in patients with PTC in patient-based analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer

Table 1. Characteristics of included studies

Study	Country	Study design	Timing of imaging	Duration of patient recruitment	Sample (n) (Male/Female)	Age (range)	No of lymph nodes	Diagnosis	Analysis methods
Jeong HS 2006[27]	Korea	Retrospective	Preoperative	July 2004-March 2005	26(7/19)	44 (17-73)	31	All PTC	L
Kim E 2008[15]	Korea	Retrospective	Preoperative	April 2006–October 2006	165 (25/140)	48 (16-78)	27	All PTC	L+P
Ahn JE 2008[12]	Korea	Retrospective	Preoperative	January 2005–December 2005	37(7/30)	47(20-68)	181	All PTC	L
Choi JS 2009[13]	Korea	Retrospective	Preoperative	February 2006–April 2007	299 (44/255)	45 (20-74)	352	All PTC	L
Choi YJ 2010[28]	Korea	Retrospective	Preoperative	January 2007–December 2008	589(121/468)	46	589	All PTC	P
Morita S 2010[29]	Japan	Prospective	Preoperative	January 2007–December 2009	74 (12/62)	66 (16-84)	349	All PTC	L

Yoon JH 2011[30]	Korea	Retrospective	Preoperative	February 2007-December 2007	113(16/97)	46	(15-83)	122	All PTC	L
Seo YL 2012[31]	Korea	Retrospective	Postoperative	August 2008–August 2011	20(4:16)	49.8		197	19 PTC, 1 MTC	L
Lee DW 2013[16]	Korea	Retrospective	Preoperative	January 2007–May 2010	252(45/207)	49	(15-82)	512	All PTC	L+P
Lesnik D 2014[32]	USA	Prospective	Preoperative	2003–2008	95(NA)	NA		1998	All PTC	L
Na DK 2015[18]	Korea	Retrospective	Preoperative	March 2011–February 2012	176 (44/132)	43	(23-74)	352	All PTC	P
Kim SK 2017[33]	Korea	Retrospective	Preoperative	January 1997–June 2015	3668(NA)	NA		6257	All PTC	L
Eun NL 2018[34]	Korea	Retrospective	Preoperative	January 2013–December 2015	302(76:226)	44		398	All PTC	L
Lee Y 2018[17]	Korea	Prospective	Preoperative	November 2011–December	351(78:273)	47.1		801	All PTC	L

2012

NA not available; PTC papillary thyroid carcinoma; MTC medullary thyroid carcinoma; L, Level-by-Level analysis; P, Patient-based analysis

Table 2. Pooled estimates of sensitivity, specificity, diagnostic OR, LR+, LR-

	Studies, n	Sensitivity (95% CI)	Specificity (95% CI)	Diagnostic OR (95% CI)	LR+ (95% CI)	LR- (95% CI)
Diagnostic accuracy of CT or US on neck level						
All neck level						
US	11	0.35(0.34-0.37, 97.5)	0.95(0.94-0.95, 90.8)	13.94(9.34-20.82, 81.0)	6.79(4.79-9.63, 86.)	0.50(0.41-0.60, 95.2)
CT	11	0.46(0.44-0.47, 97.6)	0.88(0.87-0.89, 97.9)	7.24(5.46-9.62, 72.2)	3.77(2.08-6.84, 98.5)	0.52(0.45-0.61,89.7)
US/CT	5	0.51(0.49-0.52, 92.8)	0.85(0.84-0.86, 97.8)	6.01(3.84-9.40, 89.2)	3.04(1.93-4.80, 96.5)	0.52(0.45-0.60, 78.8)
Central neck level						
US	4	0.28(0.24-0.32, 94.3)	0.97(0.96-0.98, 53.0)	14.07(6.66-29.71, 53.1)	14.07(6.66-29.71, 53.1)	14.07(6.66-29.71,53.1)
CT	4	0.32(0.28-0.36, 88.2)	0.89(0.86-0.91, 84.6)	5.48(2.15-13.98, 84.3)	3.71(1.79-7.66, 83.2)	0.74(0.62-0.89, 86.7)

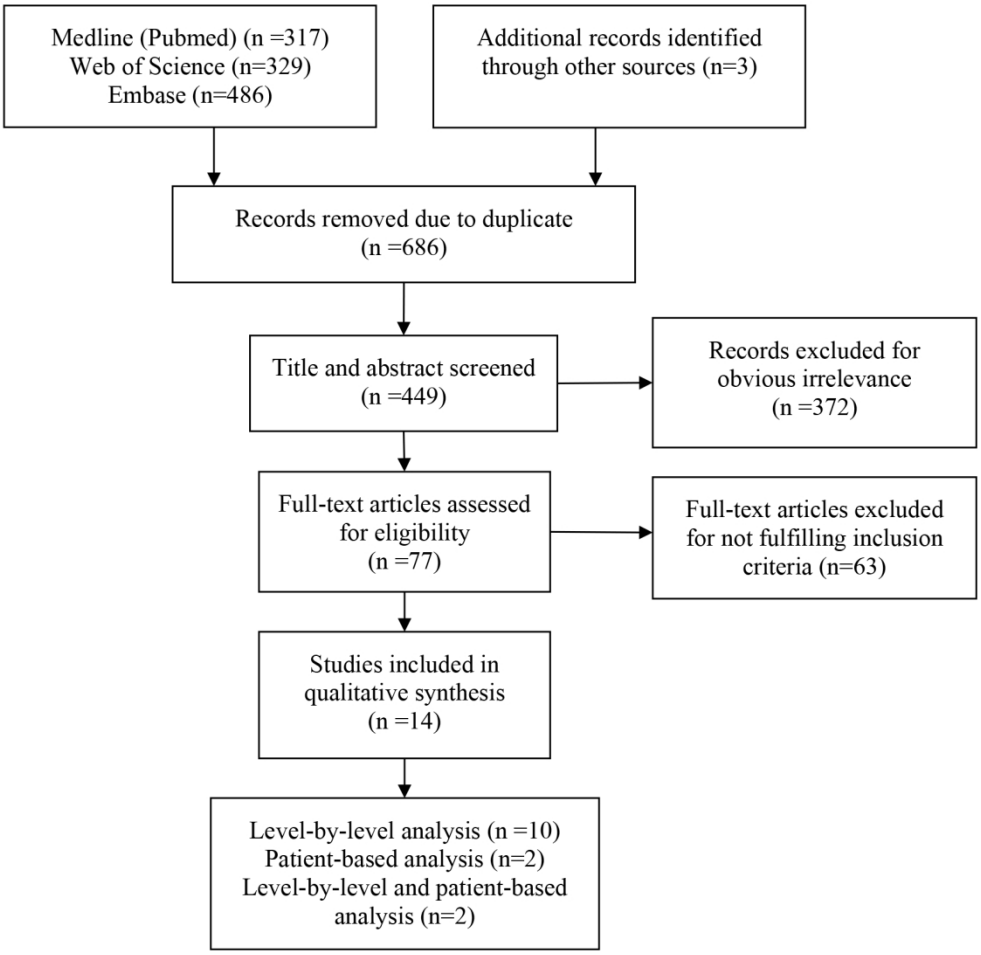
US/CT	3	0.40(0.35-0.45, 82.3)	0.85(0.82-0.88, 37.0)	4.32(2.09-8.92, 81.0)	2.85(1.75-4.65, 77.5)	0.67(0.52-0.86, 83.9)
Lateral neck level						
US	6	0.74(0.69-0.78, 78.1)	0.92(0.90-0.94, 89.6)	24.41(11.16 -53.42, 71.9)	6.67(2.91-15.30, 90.5)	0.35(0.28-0.43, 30.1)
CT	6	0.73(0.68-0.77, 90.7)	0.89(0.87-0.91, 91.8)	15.55(7.98 -30.32, 64.6)	5.54(2.95-10.39, 82.9)	0.35(0.21 -0.59, 91.4)
US/CT	4	0.88(0.83-0.91, 64.1)	0.79(0.73-0.84, 89.6)	22.59(11.29 -45.19, 46.6)	3.31(1.53-7.17, 91.1)	0.19(0.14-0.25, 0)
Diagnostic accuracy of CT or US on patient level						
All patient level						
US	4	0.41(0.36-0.46, 95.5)	0.92(0.89-0.94, 0)	7.56(4.08-14.01, 51.3)	4.48(3.31 -6.05, 0)	0.65(0.53-0.80, 75.5)
CT	4	0.49(0.44-0.54, 95.9)	0.91(0.89-0.94, 66.8)	9.40(5.79-15.27, 34.4)	4.84(3.66-6.39, 0)	0.53(0.37-0.75, 85.6)
US/CT	2	0.64(0.57-0.71, 0)	0.83(0.77-0.88, 0)	8.59(5.37-13.76, 0)	3.71(2.72 -5.08, 0)	0.43(0.36-0.53, 0)

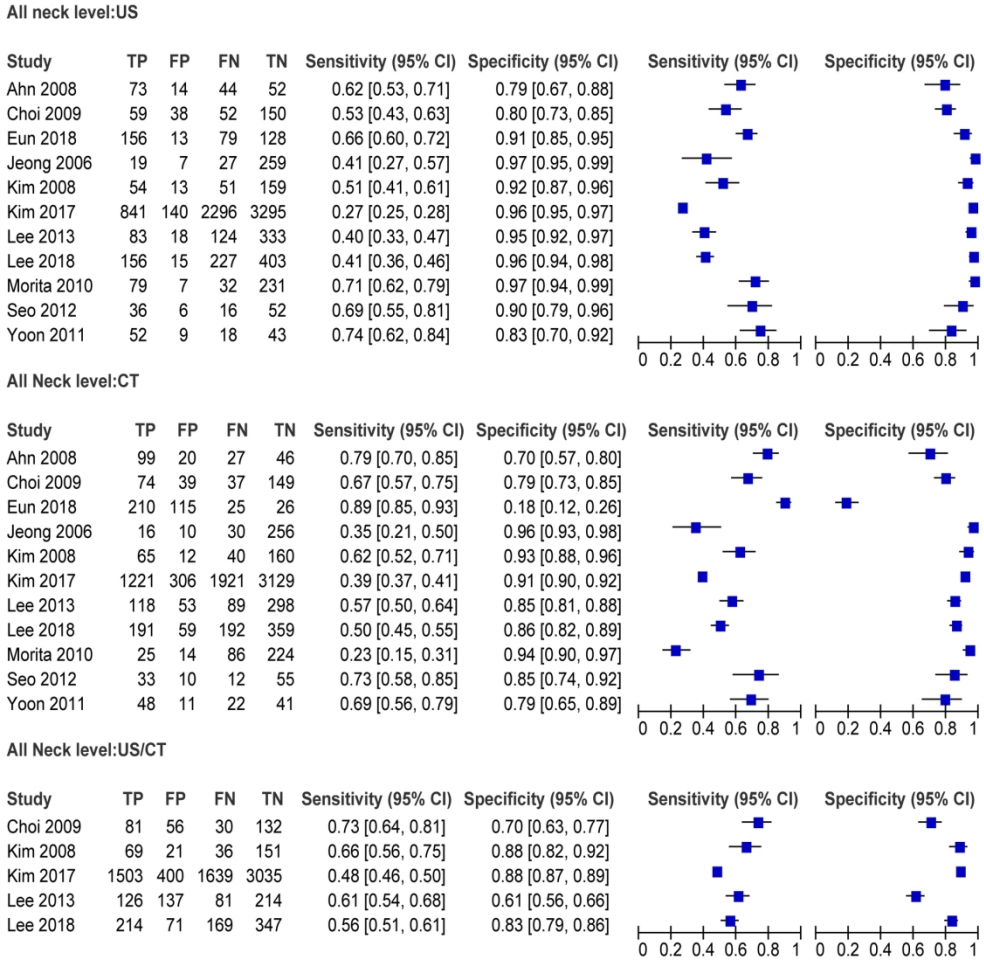
Central patient level						
US	2	0.21(0.16-0.28, 24.8)	0.95(0.91-0.97, 0)	4.53(2.34-8.77, 0)	3.78(2.08-6.86, 0)	0.84(0.76 -0.93, 36.4)
CT	2	0.38(0.32-0.46, 92.6)	0.90(0.85-0.93, 87.2)	5.02(0.46-54.54, 94.4)	3.52(0.52-23.84, 93.8)	0.71(0.41-1.22, 95.3)
US/CT	2	0.47(0.39-0.54, 80.8)	0.85(0.80-0.89, 0)	4.88(2.58-9.23, 46.4)	3.14(2.23-4.41, 0)	0.64(0.47-0.86, 78.4)
Lateral patient level						
US	2	0.87(0.74-0.95, 88.2)	0.89(0.83-0.93, 91.4)	20.11(6.77-59.70, 0)	3.58(0.85-15.16, 91.0)	0.22(0.05-1.08, 58.0)
CT	2	0.92(0.81-0.98, 88.6)	0.88(0.83-0.93, 96.8)	36.88(11.40 -119.35, 0)	3.44(0.29-40.77, 98.3)	0.23(0.10-0.52, 0)
US/CT	2	0.98(0.89-0.99, 58.9)	0.92(0.87-0.96, 97.5)	78.10(2.82 -2160.4, 65.2)	5.30(0.15-186.11, 98.6)	0.08(0.02-0.41, 0)

Results are presented as n (95% CI; I2, %); OR, odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio.

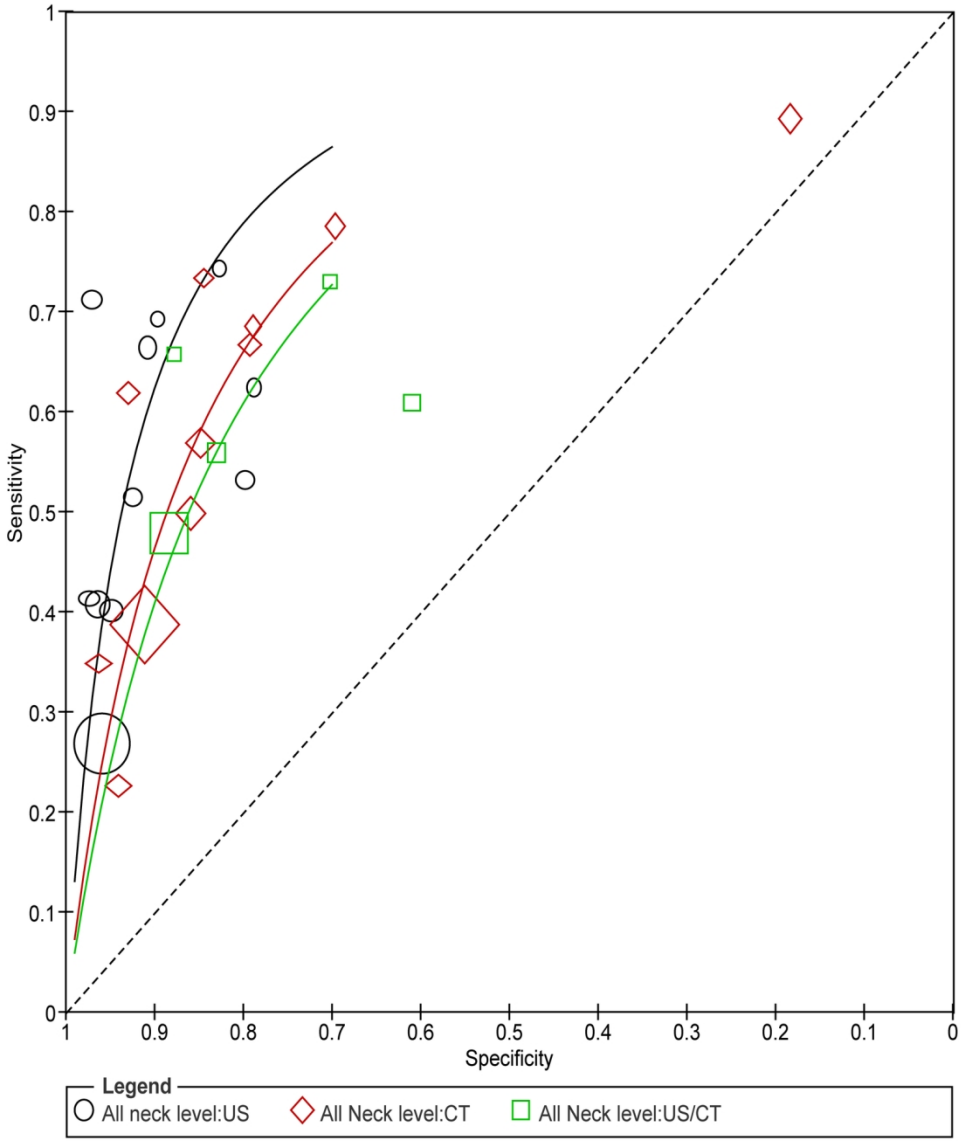
136/bmjopen-2021-051568 on 4 July 2022. Downloaded from

Supplementary materials





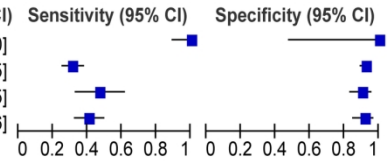
171x177mm (300 x 300 DPI)



140x170mm (300 x 300 DPI)

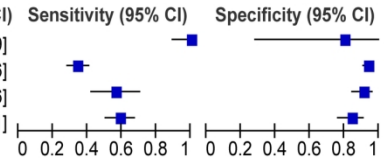
All patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	0	0	5	1.00 [0.89, 1.00]	1.00 [0.48, 1.00]
Choi 2010	75	27	163	324	0.32 [0.26, 0.38]	0.92 [0.89, 0.95]
Kim 2008	25	11	28	101	0.47 [0.33, 0.61]	0.90 [0.83, 0.95]
Lee 2013	59	9	85	99	0.41 [0.33, 0.49]	0.92 [0.85, 0.96]



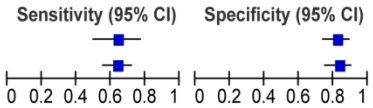
All patient level:CT

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	1	0	4	1.00 [0.89, 1.00]	0.80 [0.28, 0.99]
Choi 2010	82	22	156	329	0.34 [0.28, 0.41]	0.94 [0.91, 0.96]
Kim 2008	30	10	23	102	0.57 [0.42, 0.70]	0.91 [0.84, 0.96]
Lee 2013	85	17	59	91	0.59 [0.51, 0.67]	0.84 [0.76, 0.91]

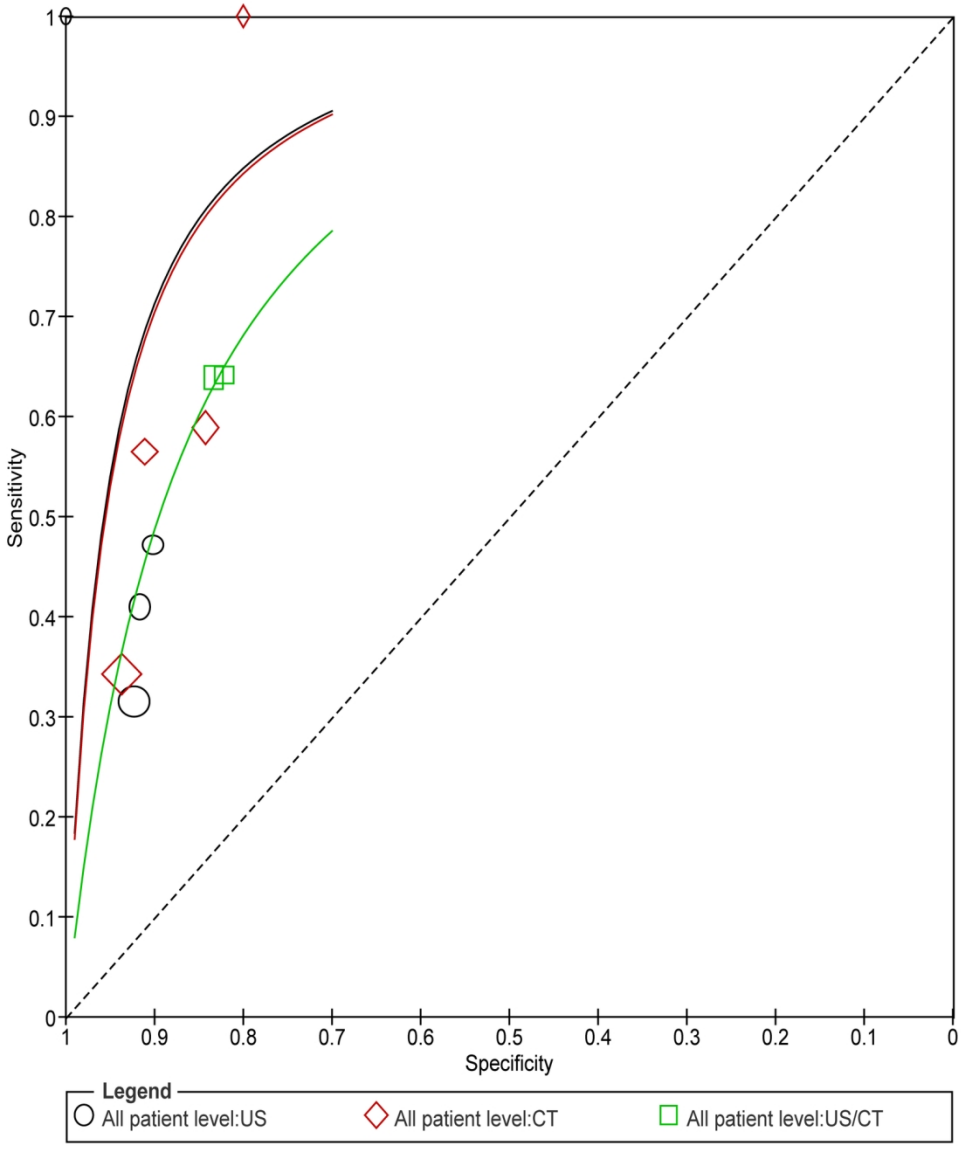


All patient level:US/CT

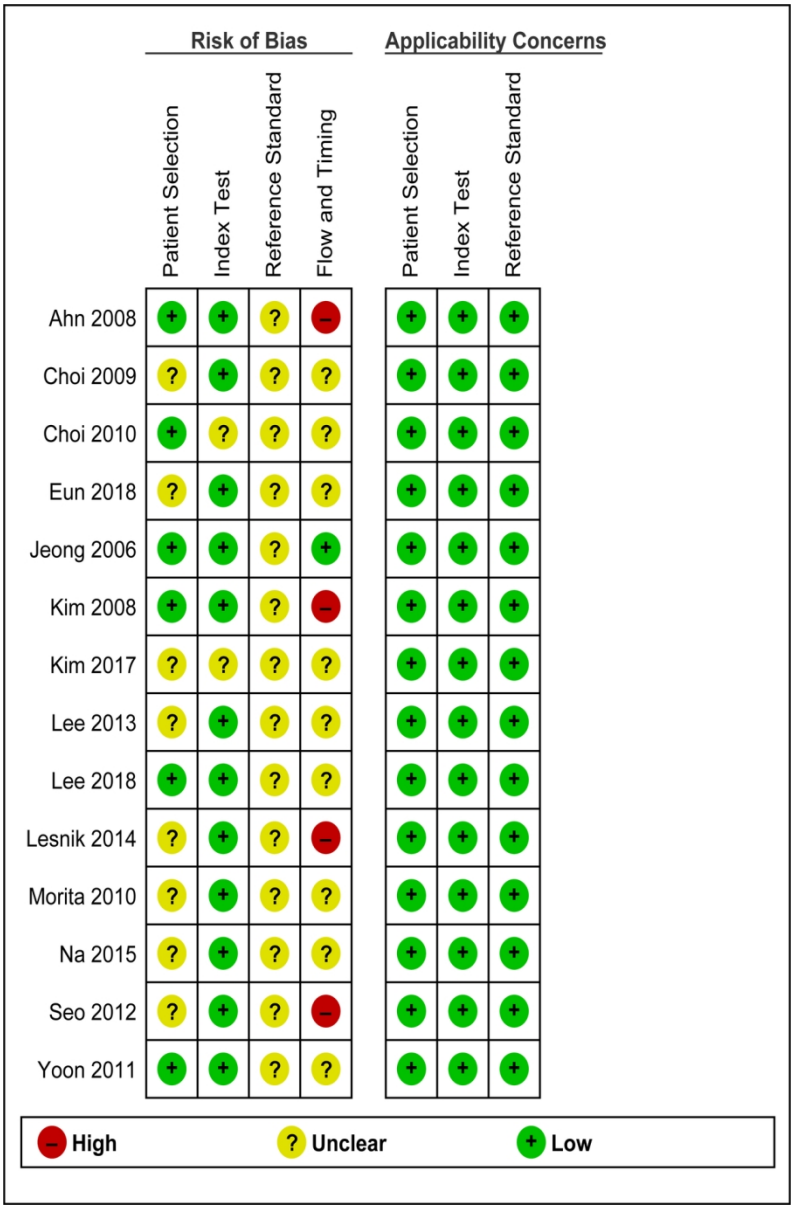
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	34	20	19	92	0.64 [0.50, 0.77]	0.82 [0.74, 0.89]
Lee 2013	92	18	52	90	0.64 [0.55, 0.72]	0.83 [0.75, 0.90]



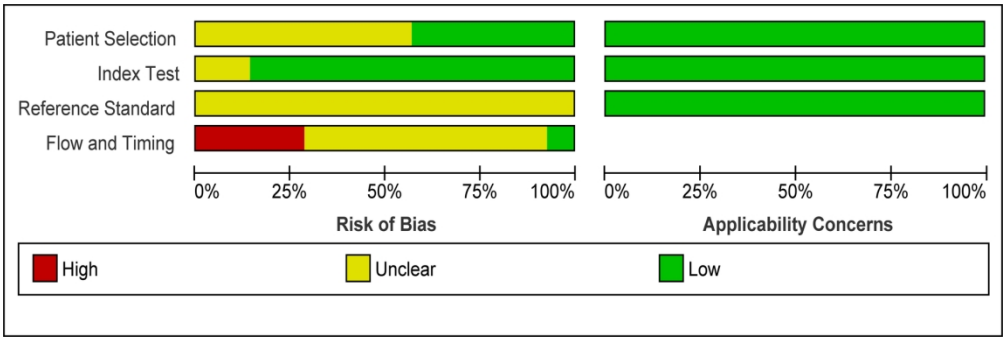
161x106mm (300 x 300 DPI)



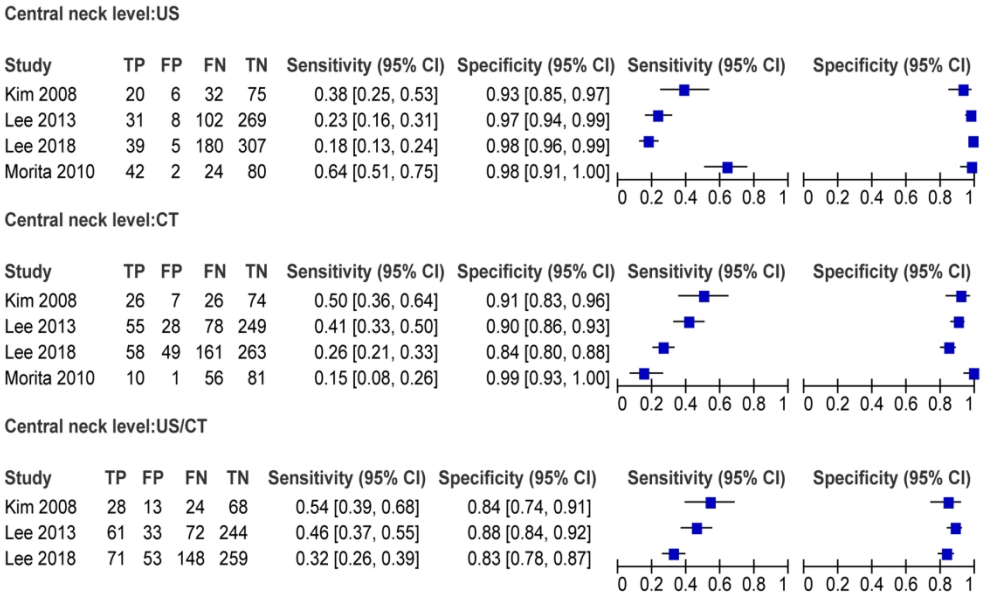
140x170mm (300 x 300 DPI)



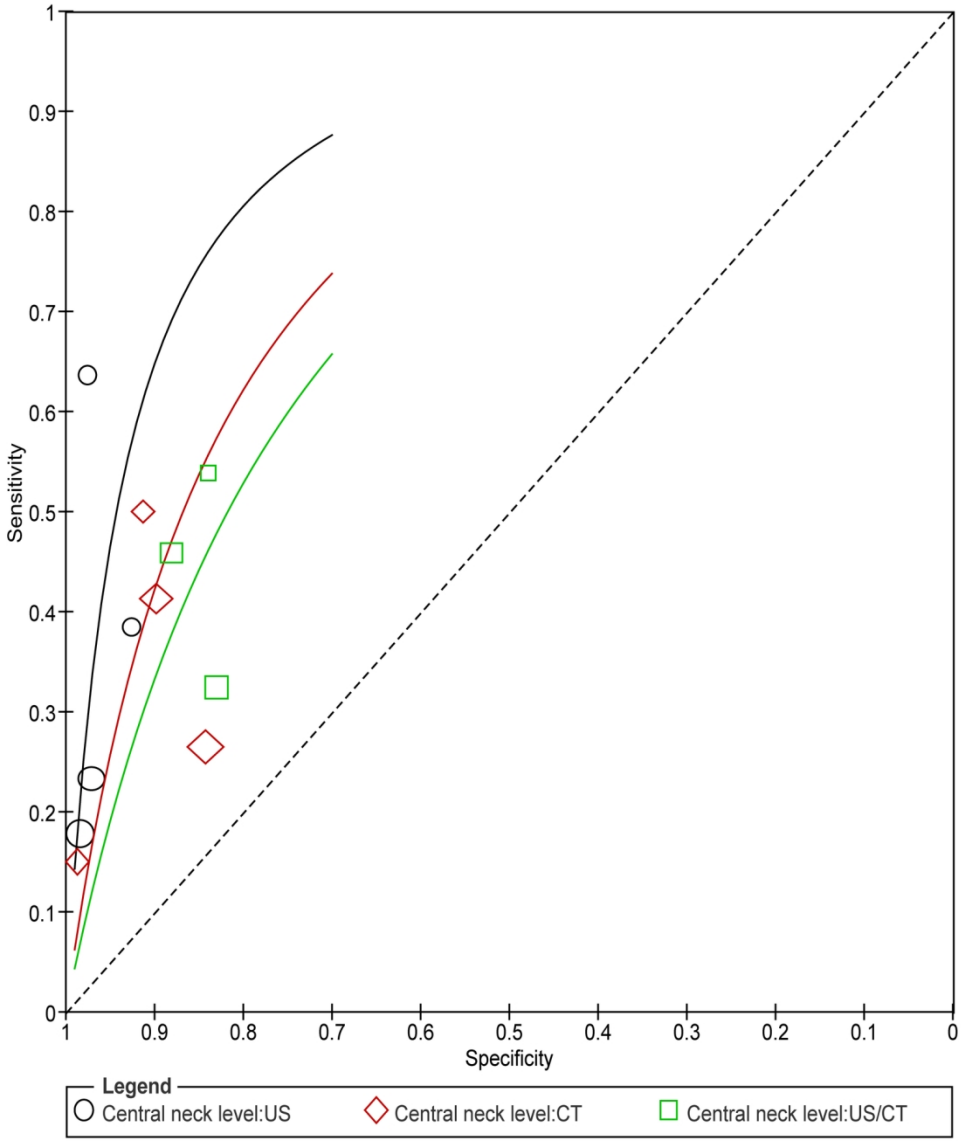
107x163mm (300 x 300 DPI)



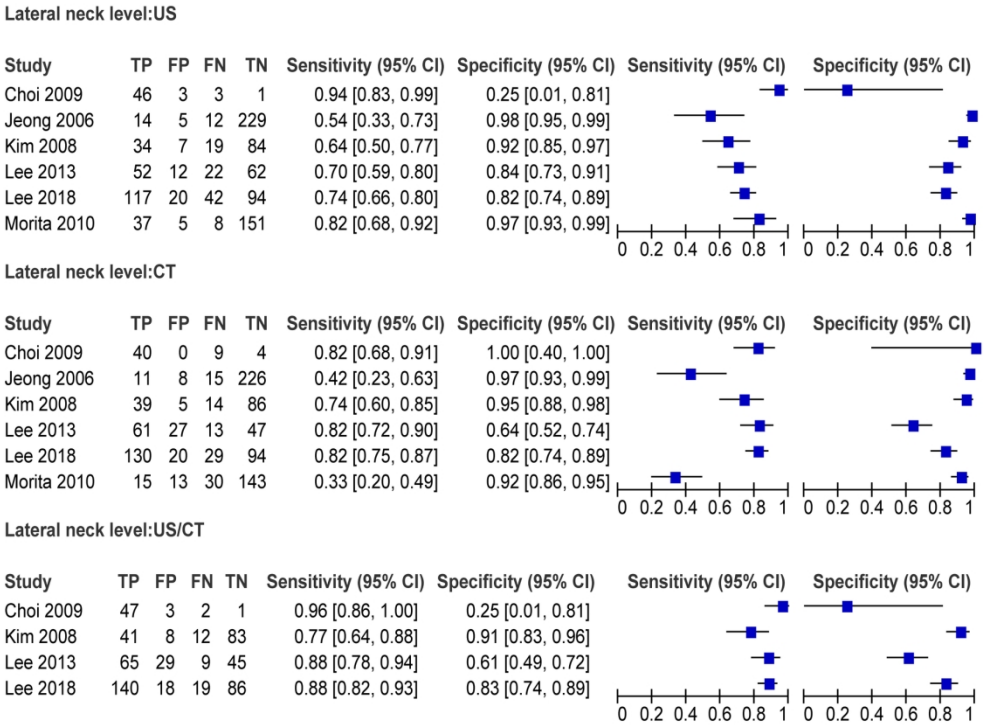
159x52mm (300 x 300 DPI)



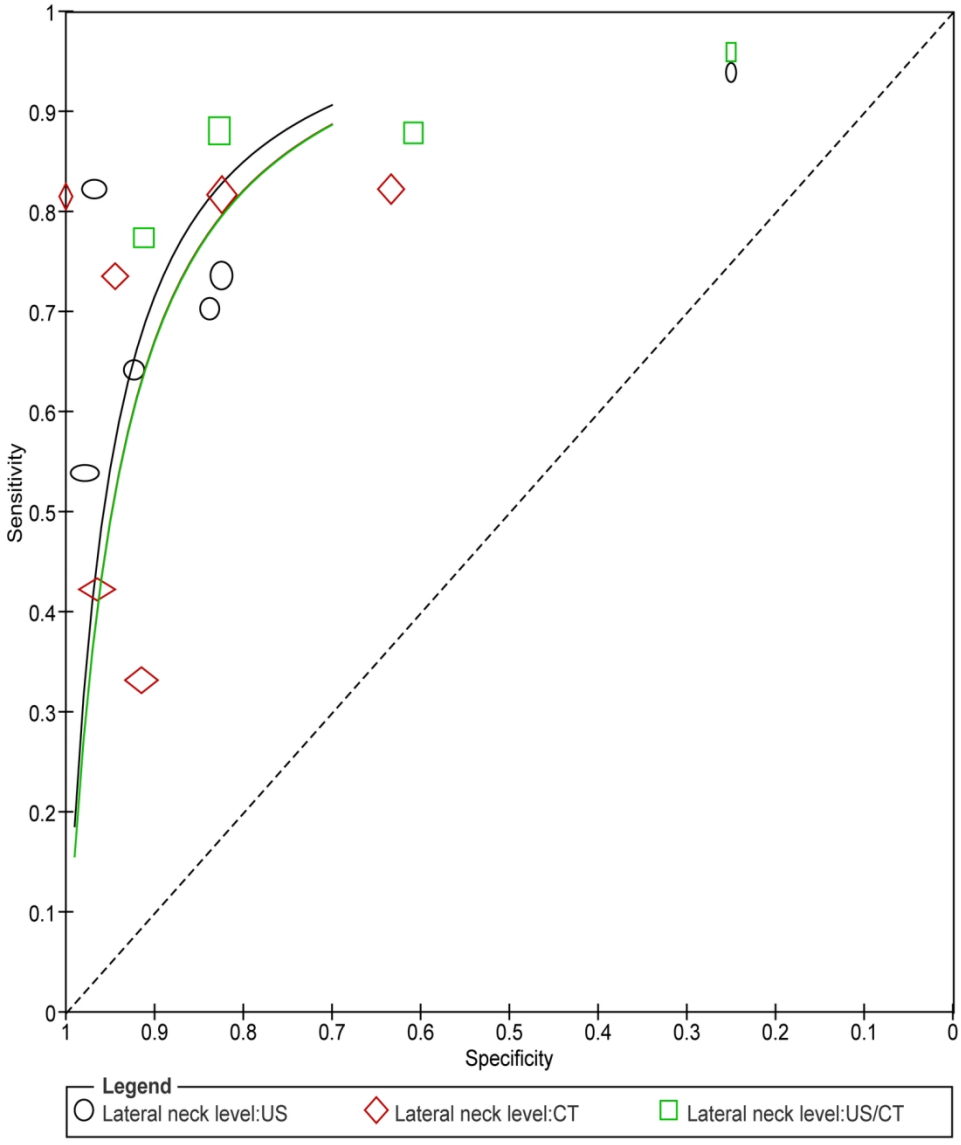
163x110mm (300 x 300 DPI)



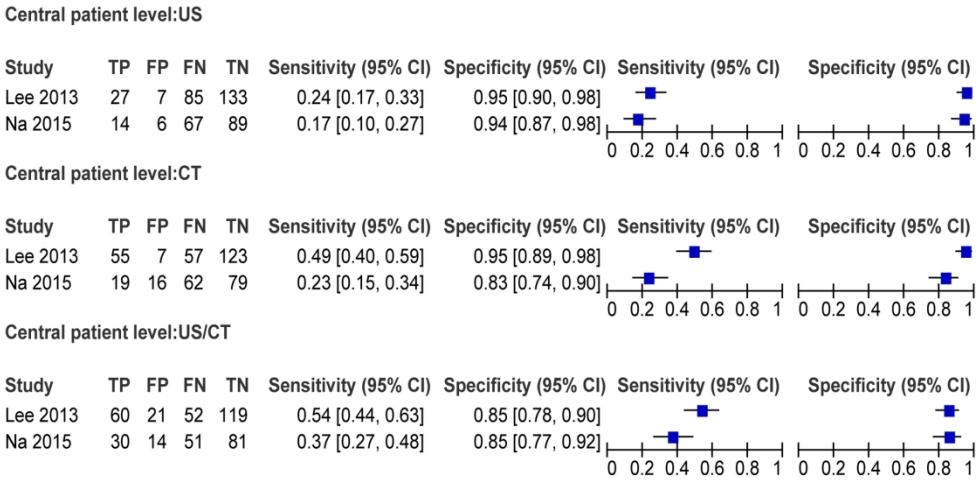
140x170mm (300 x 300 DPI)



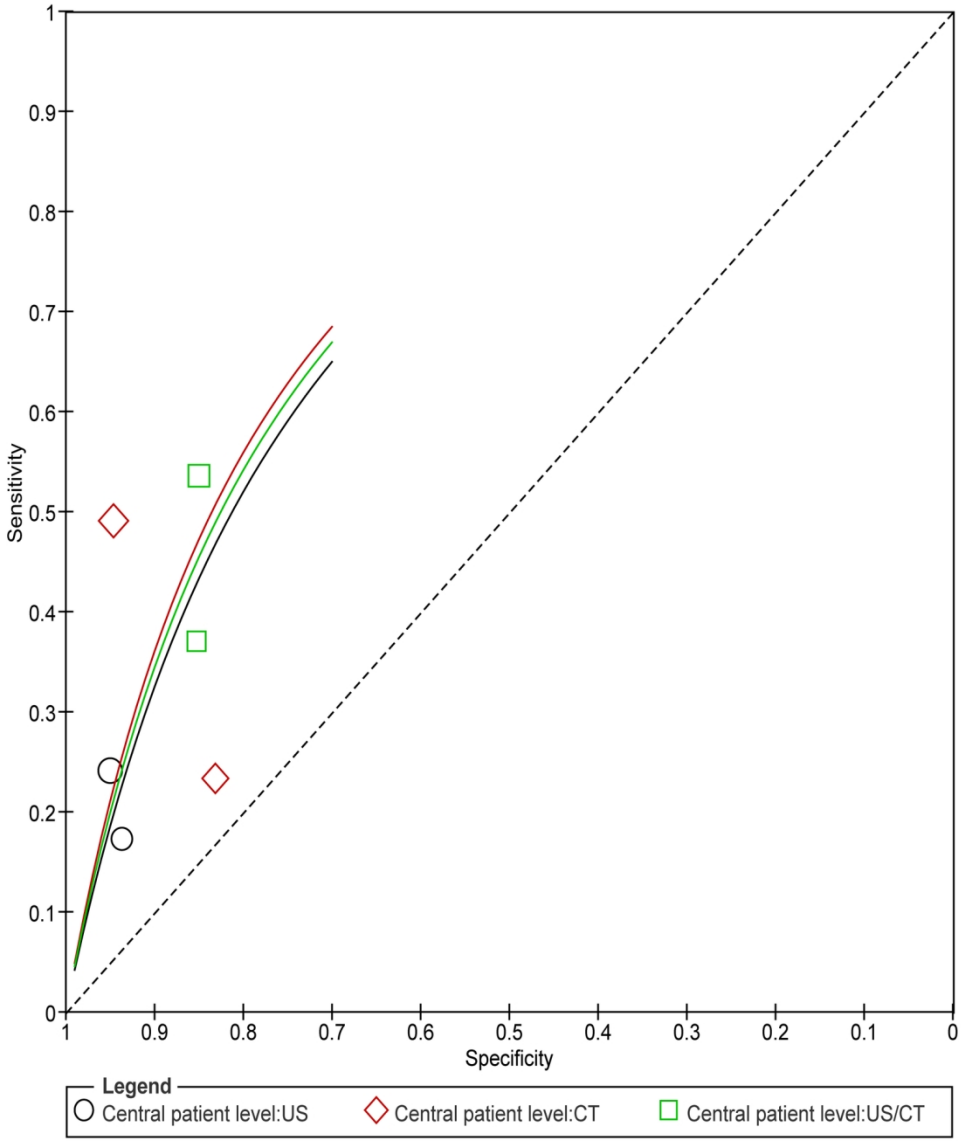
163x131mm (300 x 300 DPI)



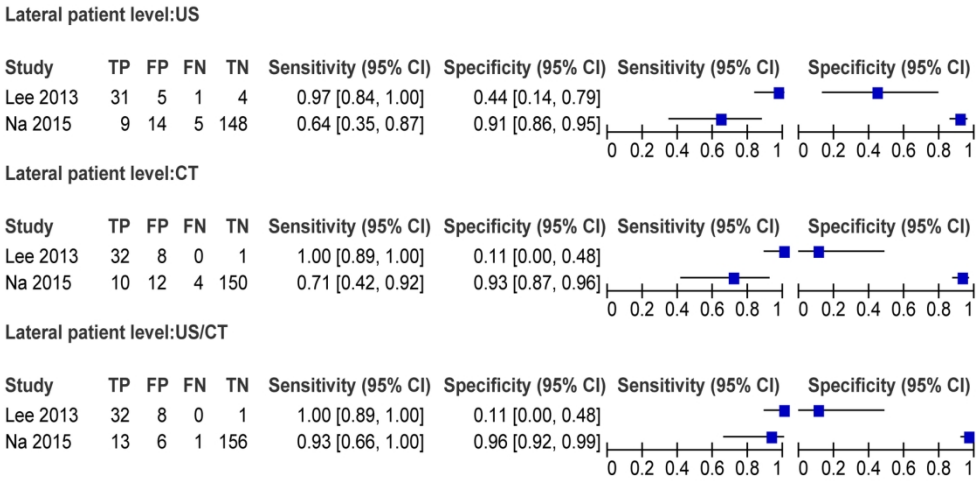
140x170mm (300 x 300 DPI)



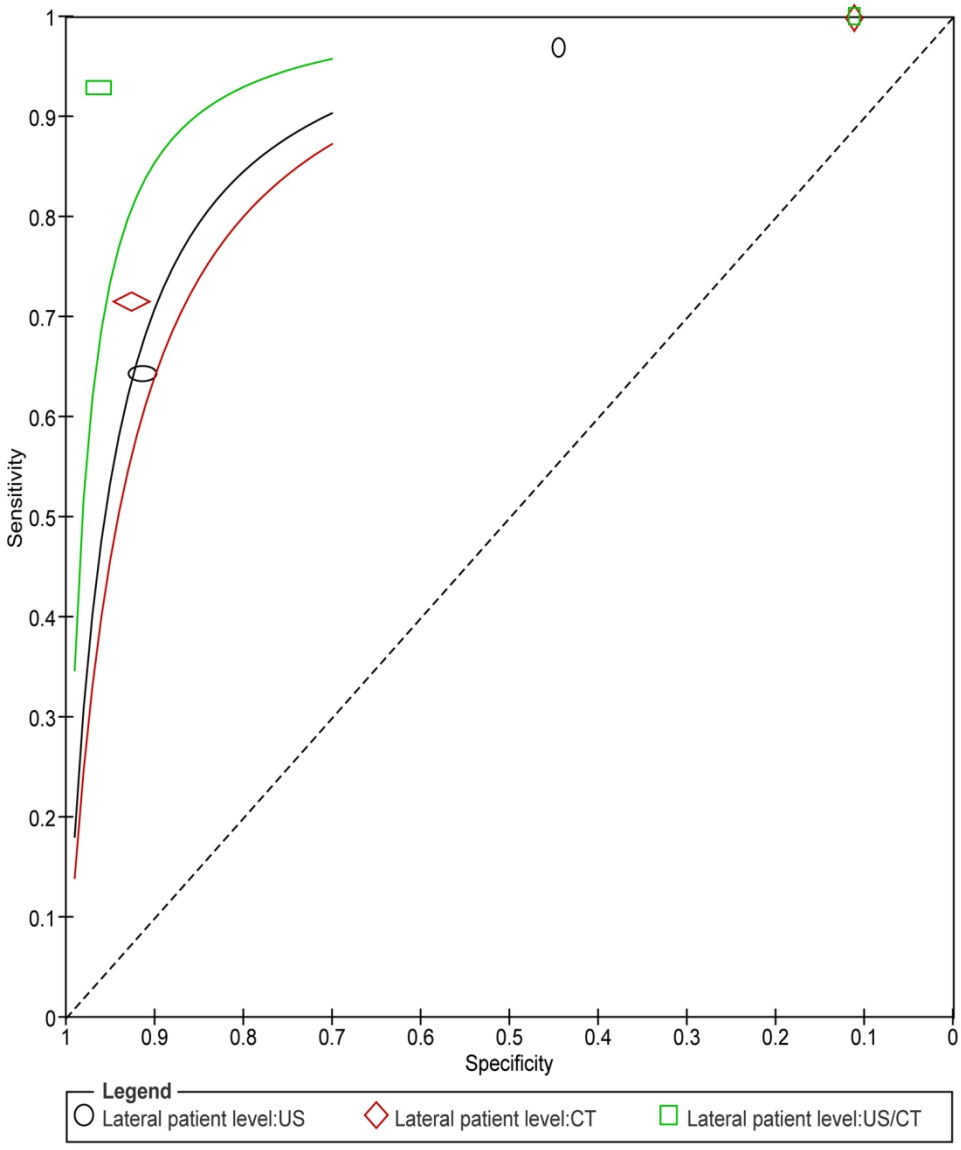
158x89mm (300 x 300 DPI)



140x170mm (300 x 300 DPI)



158x89mm (300 x 300 DPI)



140x170mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5-8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5-8
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5-8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5-8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5-8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5-8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5-8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8-11
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8-11
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	8-11
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	None

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Diagnostic accuracy of ultrasound, computed tomography, and their combination in detecting cervical lymph node metastasis in patients with papillary thyroid cancer: a systematic review and meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051568.R1
Article Type:	Original research
Date Submitted by the Author:	07-Jan-2022
Complete List of Authors:	Yang, Jian ; Xishan Coal Electricity Group Workers General Hospital, Department of Radiology Zhang, Fengyan; First Clinical Medical College, Shanxi Medical University, Department of Radiology Qiao, Ying; First Clinical Medical College, Shanxi Medical University, Department of Radiology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Radiology and imaging
Keywords:	Ultrasound < RADIOLOGY & IMAGING, Computed tomography < RADIOLOGY & IMAGING, Computed tomography < RADIOTHERAPY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Diagnostic accuracy of ultrasound, computed tomography, and their combination in detecting cervical lymph node metastasis in patients with papillary thyroid cancer: a systematic review and meta-analysis

Jian Yang[1,†], Fengyan Zhang[2,†], Ying Qiao[2,*]

[1]Department of Radiology, Xishan Coal Electricity Group Workers General Hospital, Taiyuan, 030053, Shanxi Province, China

[2]Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

[†]These authors contributed equally to this work.

[*]**Correspondence:**

Ying Qiao

Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

Tel: +86-15103462912

E-mail: qiaoying910723@163.com

Running Head: US and CT for CLNM for PTC

Abstract

Objectives: To determine the diagnostic accuracy of ultrasound (US), computed tomography (CT) and their combination in detecting cervical lymph node metastasis (CLNM) in patients with papillary thyroid cancer (PTC).

Methods: Medline (via PubMed), Web of Science, Embase were searched to identify studies published till December 5, 2021, that used US and CT to detect CLNM in patients with PTC. The primary outcomes were sensitivity, specificity, and diagnostic odds ratios (DORs) in neck level-based (lymph nodes are analyzed by neck level) or patient-based (lymph nodes are analyzed by patient) analysis. Secondary outcomes were sensitivity, specificity, and DORs in the central and lateral compartments.

Results: Fourteen studies (6167 patients with 11,601 neck lymph nodes) met the inclusion criteria. Based on the neck level-based analysis, the pooled sensitivity, specificity and DORs were 0.35 (95% confidence interval [(CI) 0.34-0.37], 0.95 (95% CI 0.94-0.95), and 13.94 (95% CI 9.34-20.82) for US, were 0.46 (95% CI 0.44-0.47), 0.88 (95% CI 0.87-0.89), and 7.24 (95% CI 5.46-9.62) for CT, were 0.51 (95% CI 0.49-0.52), 0.85 (95% CI 0.84-0.86), 6.01 (95% CI 3.84-9.40) for the combination of US and CT. In the patient-based analysis, the pooled estimates of sensitivity, specificity, and DOR were 0.41 (95% CI 0.36-0.46), 0.92 (95% CI 0.89-0.94), and 7.56 (95% CI 4.08-14.01) for US, were 0.49 (0.44-0.54), 0.91 (0.89-0.94), 9.40 (5.79-15.27) for CT, and were 0.64 (95% CI 0.57-0.71), 0.83 (95% CI 0.77-0.88), 8.59 (95% CI 5.37-13.76) for the combination of US and CT.

Discussion: These findings suggest US, with a DOR almost twice that of CT in the neck level-based analysis, was superior to CT in detecting CLNM in patients with PTC, especially in the lateral compartment. The combination of US and CT increased the

sensitivity from 41%-49% for the individual modalities to 64% for combined modalities in the patient-based analysis.

Keywords: ultrasound; computed tomography; cervical lymph node metastasis; papillary thyroid cancer; meta-analysis.

Strengths and limitations

- Only studies that analyzed CT and US were included.
- The analyses were performed based on the neck level and the patient level.
- Heterogeneity was observed due to study design and timing of the examinations.
- The use of CT for CLNM screening is not recognized everywhere globally.

Introduction

Papillary thyroid carcinoma (PTC) is an endocrine neoplasia with a high incidence of lymphatic metastasis and is associated with regional recurrence [1-3]. The incidence of cervical lymph node metastasis (CLNM) in patients with thyroid cancer has been reported to be 20%-90% [4]. The presence of CLNM might increase the risk of locoregional recurrence after surgery [5 6], worsening prognosis and survival [7]. Therefore, it is of great clinical importance to accurately evaluate CLNM and determine the extent of neck dissection [8]. Although prophylactic central compartment neck (groups VI and VII) dissection (ipsilateral or bilateral) is recommended by the American Thyroid Association (ATA) guidelines in patients with clinically positive central nodes, especially for those with advanced primary tumors, the information regarding prophylactic lateral compartment (groups I-V) neck dissection has not been clearly stated [8]. Thus, the indications for neck dissection, especially the lateral compartment, should be carefully assessed as it might lead to severe postoperative complications [9].

Preoperative staging with ultrasound (US) for cervical lymph nodes, including both central and lateral neck compartments, is the most widely accepted first imaging technique for patients with thyroid or suspicious malignancies cytologic or molecular findings. It can observe node enlargement, loss of fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, peripheral vascularity, and calcifications, which are all indicators of malignant invasion [10]. In addition, US is inexpensive, widely available, can be carried out bedside, and provide real-time imaging. Still, US is operator-dependent, and the images will vary depending on the angle and pressure of the probe on the neck. Computed tomography (CT) provides three-dimensional neck images that avoid operator-

dependency issues. On the other hand, the analysis of each layer takes time, and the use of contrast carries a risk of kidney injury. The preoperative use of computed tomography (CT) with intravenous (IV) contrast is complementary to US in patients with advanced disease [8 11 12]. A suspicious node on US can be confirmed by CT, and CT can detect nodes that were not visible because they were behind solid or air-containing structures or were not considered suspicious for various reasons. Although several studies have failed to prove the benefit of CT over US in detecting lateral lymph node metastasis [13-15], some studies suggested a superior diagnostic performance of the combination of preoperative US with CT over US alone [14 16-19].

Some meta-analyses examining the diagnostic accuracy of US and CT in detecting CLNM in patients with PTC have been previously conducted [20-24]. However, these meta-analyses studies integrated the findings of US and CT from different studies and populations.

To the best of our knowledge, no previous meta-analysis included studies that evaluated CLNM in patients with PTC using both US and CT, which could minimize the confounding effect of an operator in interpreting the diagnostic accuracy of preoperative imaging. This meta-analysis aimed to evaluate the sensitivity, specificity, and diagnostic odds ratios (DORs) of US, CT, and their combination in detecting positive CLNM in patients with PTC based on the central and lateral neck levels and by using neck level-based (lymph nodes are analyzed by neck level) and patient-based (lymph nodes are analyzed by patient, irrespective of the level) analyses.

Methods

Systematic literature research

This meta-analysis was reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines [25]. Ethical approval was waived due to the secondary data acquisition from previously published papers available in the public domain. A systematic search of Medline (via PubMed), Web of Science, and Embase was conducted to identify studies published up to December 5, 2021, that assessed the accuracy of US and CT in detecting CLNM in patients with PTC. The search strategy was developed in collaboration with a hospital librarian and included subject headings and text words: ("thyroid cancer" OR "thyroid carcinoma" OR "thyroid tumor" OR "papillary thyroid cancer" OR "thyroid neoplasm") AND ("cervical lymph node" OR "neck lymph node") AND ("metastasis" or "metastatic") AND ("ultrasonography" or "ultrasound" or "US") AND (computed tomography" or "CT") (Supplementary Table S1). The studies were initially screened by examining their titles and abstracts, and the full texts of potentially eligible studies were retrieved for further review. No language restriction was applied. A manual search of additional records and reference lists was also performed to include more relevant studies.

Study selection

The inclusion criteria of the studies were (a) prospective or retrospective studies that evaluated the diagnostic accuracy of both US and CT for detecting CLNM in patients with PTC, using neck level-based or patient-based analysis; (b) studies with >10 patients; (c) studies with a reference standard of histopathology or cytology (the diagnostic gold

standard was the pathological diagnosis of the resected lymph nodes); (d) studies that reported the absolute numbers of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) results directly or derived from the reported data or communicated by the authors in response to our request; (e) studies published in English.

The exclusion criteria were (i) case reports, case series, review articles, pictorial essays, letters to editors, unpublished data, conference abstracts, and proceedings on the topic of interest; (ii) studies that used only US or only CT; (iii) insufficient data regarding TP, FP, FN, and TN; (iv) duplicate publications using the same databases and studies; (v) if the patient population of one article is overlapping with the patient population of other or multiple articles, then the article with the largest sample size was included; (vi) studies with less than 10 cases confirmed by the reference standard. One reader reviewed the full texts of the candidate articles and selected those that met the inclusion criteria. A second reader reviewed the process of the inclusion of articles in the meta-analysis. No inter-reader disagreements were observed.

Primary and secondary outcomes

The primary outcomes were sensitivity, specificity, and DORs in a neck level-based or patient-based analysis. Secondary outcomes were sensitivity, specificity, and DORs in central and lateral compartments in neck level-based or patient-based analysis.

Data extraction and quality assessment

Two reviewers independently performed the data extraction. Data such as study characteristics, clinical and patient characteristics, reference standard or standards, cervical lymph node compartment, technical characteristics of CT and US and contrast enhancement, the definition of CLNM according to CT and US image findings, and the

diagnostic performance of CT and US, such as TP, FP, FN, TN were obtained from each study.

Two reviewers who were not blinded to the journal names, author names, and year of publication assessed the methodologic and reporting quality of each study by using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) [26]. Each study was independently assessed by two reviewers after a tutorial meeting on the guidelines for interpreting the items. Any disagreements were resolved by discussion with an experienced third reviewer.

Statistical analysis

The pooled sensitivity, specificity, diagnostic odds ratio (OR), positive likelihood ratio (LR+), and negative likelihood ratio (LR-) were calculated for US and CT in a neck level-based analysis (at neck level, central neck level, and lateral neck level) and a patient-based analysis (patient level, central patient level, and lateral level). The heterogeneity of pooled sensitivity, specificity, diagnostic OR, LR+, and LR- was measured by the inconsistency (I^2). Heterogeneity in the included articles was defined as small I^2 <25%, moderate I^2 25%-50%, and obvious I^2 >50%. If heterogeneity was detected (P-value <0.10 or $I^2 \geq 50\%$), a random-effects model was applied; otherwise, a fixed-effects model was used. A bivariate logistic regression model was used for meta-analysis of diagnostic test accuracy [27], and forest plots were created. The pooling of sensitivity, specificity, diagnostic OR, LR+, and LR- was performed using the Meta-Disc software (version 1.4, Madrid, Spain). Forest plots and summary receiver operating characteristic (SROC) curves were obtained using RevMan 5.3. A P-value of <0.05 was considered statistically significant.

Patient and public involvement

The patients and the public were not involved in this study.

Results

Characteristics of included studies

The initial literature search yielded 1135 potential studies for this meta-analysis. A total of 449 articles were screened after removing the duplicates. Of these, 372 studies were excluded after reviewing the titles and abstracts, and 63 articles were excluded after reviewing the full texts (Figure 1). Fourteen studies were ultimately selected for inclusion [13 14 16-19 28-35]: 10 studies used a neck level-based analysis, two studies used a patient-based analysis, and two studies used both. Five studies reported the diagnostic performance by combining both US and CT [14 16-19]. A total of 6167 patients with 11,601 neck lymph nodes were included, and all patients were diagnosed with PTC except one who was diagnosed with medullary thyroid cancer. The earliest study was started in 1997, whereas the latest one was started in 2012. The median number of patients per study was 171 (range 20-3668), while the median number of lymph nodes per study was 331 (range 107-6557). Eleven were retrospective studies, and three were prospective studies; 13 studies were performed preoperatively, and 1 study was performed postoperatively. Twelve, one, and one were conducted in Korea, the United States, and Japan, respectively (Table 1). The studies included in this meta-analysis were of moderate quality (supplementary Figure 1[sFigure 1], sFigure 2).

Neck level-based diagnostic accuracy of US and CT

Eleven studies used both CT and US for detecting CLNM in patients with PTC, and five of them assessed the diagnostic accuracy of the combination of CT and US. The pooled

sensitivity, specificity, diagnostic OR, LR+, and LR- were 0.35 (95% CI 0.34-0.37), 0.95 (95% CI 0.94-0.95), 13.94 (95% CI 9.34-20.82), 6.79 (95% CI 4.79-9.63), and 0.50 (95% CI 0.41-0.60) for US, were 0.46 (95% CI 0.44-0.47), 0.88 (95% CI 0.87-0.89), 7.24 (95% CI 5.46-9.62), 3.77 (95% CI 2.08-6.84), and 0.52 (95% CI 0.45-0.61) for CT, and were 0.51 (95% CI 0.49-0.52), 0.85 (95% CI 0.84-0.86), 6.01 (95% CI 3.84-9.40), 3.04 (95% CI 1.93-4.80), and 0.52 (95% CI 0.45-0.60) for the combination of US and CT, with marked heterogeneity (Table 2, Figure 2, Figure 3).

Subgroup analyses of central and lateral neck levels were performed to investigate the effects of cervical lymph node compartment based on the diagnostic accuracy of US and CT. The subgroup analysis of the central neck level revealed that the pooled sensitivity, specificity, and DOR of US were 0.28 (95% CI 0.24-0.32), 0.97 (95% CI 0.96-0.98), and 14.07 (95% CI 6.66-29.71) from four studies. For CT, the pooled sensitivity, specificity, and DOR were 0.32 (95% CI 0.28-0.36), 0.89 (95% CI 0.86-0.91), and 5.48 (95% CI 2.15-13.98) from four studies. The pooled sensitivity, specificity, and DOR of the combination of US and CT were 0.40 (95% CI 0.35-0.45), 0.85 (95% CI 0.82-0.88), and 4.32 (95% CI 2.09-8.92) from three studies (Table 2, sFigure 3, sFigure 4).

In contrast, the subgroup analysis of the lateral neck level revealed that the pooled sensitivity, specificity, and DOR of US were 0.74 (95% CI 0.69-0.78), 0.92 (95% CI 0.90-0.94), and 24.41 (95% CI 11.16 -53.42) from six studies; the values for CT were 0.73 (95% CI 0.68-0.77), 0.89 (95% CI 0.87-0.91), and 15.55 (95% CI 7.98-30.32) from six studies; the values for the combination of US and CT were 0.88 (95% CI 0.83-0.91), 0.79 (95% CI 0.73-0.84), and 22.59 (95% CI 11.29-45.19) from four studies (Table 2, sFigure 5, sFigure 6).

Patient-based diagnostic accuracy of US and CT

Four studies included both US and CT in detecting CLNM in patients with PTC, and two of them assessed the diagnostic accuracy by combining both CT and US. The pooled estimates of sensitivity, specificity, and DOR of US were 0.41 (95% CI 0.36-0.46), 0.92 (95% CI 0.89-0.94), and 7.56 (95% CI 4.08-14.01); the values for CT were 0.49 (0.44-0.54), 0.91 (0.89-0.94), and 9.40 (5.79-15.27); the values for the combination of US and CT were 0.64 (95% CI 0.57-0.71), 0.83 (95% CI 0.77-0.88), and 8.59 (95% CI 5.37-13.76) (Table 2, Figure 4, Figure 5).

Only two studies assessed the diagnostic accuracy of US, CT, and their combination on a patient basis. On the patient level, the pooled estimates of sensitivity, specificity, and DOR were 0.21 (95% CI 0.16-0.28), 0.95 (95% CI 0.91-0.97), and 4.53 (95% CI 2.34-8.77) for US, were 0.38 (95% CI 0.32-0.46), 0.90 (95% CI 0.85-0.93), and 5.02 (95% CI 0.46-54.54) for CT, and were 0.47 (95% CI 0.39-0.54), 0.85 (95% CI 0.80-0.89), and 4.88 (95% CI 2.58-9.23) for the combination of CT and US (Table 2, sFigure 7, sFigure 8).

In contrast, the pooled estimates of sensitivity, specificity, and DOR of US were 0.87 (95% CI 0.74-0.95), 0.89 (95% CI 0.83-0.93), and 20.11 (95% CI 6.77-59.70); the values for CT were 0.92 (95% CI 0.81-0.98), 0.88 (95% CI 0.83-0.93), and 36.88 (95% CI 11.40-119.35); the values for the combination of US and CT were 0.98 (95% CI 0.89-0.99), 0.92 (95% CI 0.87-0.96), and 78.10 (95% CI 2.82-2160.4) (Table 2, sFigure 9, sFigure 10).

Discussion

This meta-analysis revealed that the DORs of US in the neck level-based analysis was higher than for CT or their combination on the central, lateral, and neck levels.

Differentiated thyroid carcinoma, particularly PTC, involves CLNMs in 20%-50% of the patients [36-39], which could prevent small and intrathyroidal primary tumors[40]. Still, the clinical implications of macrometastases (≥2 mm) are more significant than micrometastases, in which 90% of patients might reach the criteria according to the sensitivity of the imaging methods [41 42]. The combination of US features might increase the likelihood of detecting CLNM as several US features are suggestive of metastatic lymph nodes, including enlargement, loss of fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, peripheral vascularity, and calcifications [10]. The preoperative US identifies lymph node or soft-tissue metastases in up to 39% of patients who had no physical examination [43] and changed the operative management in 23% of patients [44].

Previous meta-analyses examined CT and US. Suh et al. [20] and Cho et al. [21] demonstrated the value of CT for CNLM but did not include US. Raijmakers et al. [22] only examined the detection of the sentinel lymph node. Wu et al. [23] and Zhao et al. [24] examined the value of US for CLNMs but did not include CT. Therefore, these studies did not examine CT and US simultaneously. Our data found that the DORs of CT were higher than US and the combination, and the DORs of the combination remained higher than US and CT by patient-based analysis. This was reasonable because the sensitivity of CT in the patient-based analysis was higher than that of US in the central, lateral, and patient analyses. This result might still need further investigation because of the inclusion of a small number of studies in the subgroup analysis. The operator-independent CT could be used as an adjunct in imaging deep anatomic structures, including the mediastinum, infraclavicular, retropharyngeal, and parapharyngeal regions and the structures that are acoustically

shadowed by bone or air. In addition, preoperative knowledge on the extent of laryngeal, tracheal, and esophageal involvement, as well as bulky nodal disease from neck CT with contrast, significantly influences the surgical plan by indicating the need for sternotomy, tracheal or laryngeal resection, and reconstruction [45].

The results suggested that the sensitivity on the lateral compartment tended to be higher than for the central compartment regardless of the use of US, CT, or their combination in the neck level-based and patient-based analyses. The location of the lymph nodes helps in decision-making as most of the metastatic nodes are found in the lower third of the neck, and reactive enlarged lymph nodes are found in the upper part of the neck [46]. Besides, the lateral compartment should be carefully evaluated for skip metastases located in the upper pole or are ≤1 cm in diameter [47]. For patients who had preoperative CT and US and subsequently underwent total thyroidectomy and neck dissection, the sensitivity of CT was much better than US for evaluating CLNM on the neck level, but the sensitivity, specificity, and DORs for the lateral neck level tended to be higher than those of the central neck level for both CT and US[13]. Dual-energy CT (DECT) for assessing CLNM in patients with PTC was not included in this meta-analysis as it can generate iodine-based material decomposition (MD) images and spectral HU curve [48-50]. In accordance with the findings from CT, combined gemstone spectral image (GSI) parameters from DECT also demonstrated better diagnostic accuracy of CLNM in patients with PTC when compared to those that are obtained by combining the US morphological parameters especially in the lateral compartment [50].

Our findings revealed that compared to US or CT alone, the combination of both US and CT demonstrated higher sensitivity, i.e., a meta-analytic summary sensitivity of 0.51

(0.49-0.52) and 0.64 (0.57-0.71), and a lower specificity, i.e., a meta-analytic summary specificity of 0.85 (0.84-0.86) and 0.83 (0.77-0.88) for evaluating CLNM in patients with PTC using neck level-based and patient-based analyses, respectively. In patients undergoing primary and revision surgical treatment for PTC, combined preoperative mapping with US and CT yielded significantly higher sensitivity for detecting macroscopic lymph nodes in both lateral and central neck, especially in the central neck [33].

It should be noted that the study has strengths. Firstly, Boolean operatives of "AND" rather than "OR" were used for combined datasets for all studies. Namely, only studies of direct head-to-head comparison by US, CT, and combination of both in the same patient population were included in this meta-analysis, avoiding bias due to differences in patient and institutional factors. Secondly, a meta-analysis of the included studies was performed by using neck level-based and patient-based analyses and on all, central, and lateral neck levels. Lastly, our data suggested that future follow-up studies should be performed to determine the comparative role of US and CT in identifying false-negative nodes that are not biopsied or excised.

Despite great clinical significance, there are several limitations in the current meta-analysis that are mostly associated with the available data and heterogeneity of design, interpretation of results, and reporting of data in primary studies. Firstly, the sources of heterogeneity among primary studies in meta-analyses have been reported by several previous studies, which included contrast amount, scan phase, and reconstruction slice thickness for CT [21], and the criteria of lymph node diameter and vascular flow for US [24]. Secondly, the literature included is limited due to the study design and timing of imaging. Eleven of the 14 studies (78.6%) were retrospective, and one of the 14 studies

was a postoperative imaging study. A large proportion of retrospective studies might increase the sensitivity of CT and US. Twelve of the 14 studies were conducted in Korea, and so ethnic factors might affect the results of this meta-analysis. Thus, the complementary use of CT might be routine in Korea but not necessarily applicable to other parts of the world, especially in developing countries. Thirdly, modern high-resolution US transducers have a lateral resolution of 2 mm, which is not feasible for CT, allowing for the detection of small nodes and the presence of microcalcification. The included CT studies might not be comparable from one study to another, particularly over the decade, depending on the equipment, slice thickness, amount of contrast injected, etc. Fourthly, four of the 14 included studies were with patient-based results, and 12 of 14 studies were of suboptimal quality, and no definite recommendation could be drawn from the present study. Finally, MRI, US-guided FNA, and PET-CT were not included in the meta-analysis to directly compare CT and US, although they also play complementary roles in managing CLNMs in PTC.

Despite these potential drawbacks, this meta-analysis demonstrated the unique complementary value of CT secondary to US in detecting CLNMs in patients with PTC in the patient-based analysis. More importantly, the choice of a diagnostic test should be tailored to have feasible access to these imaging modalities at individual healthcare centers.

Conclusion

These findings suggest that US, with a DOR of almost twice that for CT in the neck level-based analysis, was superior to CT in detecting CLNM in patients with PTC, especially in the lateral compartment. The combination of US and CT increased the

sensitivity from 41%-49% for the individual modalities to 64% for combined modalities in the patient-based analysis. CT might be valid a candidate imaging technique secondary to US in the management of CLNM in patients with PTC.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship

Jian Yang is the guarantor of the integrity of the entire study. Ying Qiao designed and conceptualized the study concepts and design, Jian Yang performed the literature search. Ying Qiao analyzed the literature. Jian Yang analyzed the data. Gengyan Zhang performed the statistical analyses, Ying Qiao prepared the manuscript. Fengyan Zhang edited the manuscript.

Acknowledgments

None.

Funding

None.

Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. Lancet 2016;**388**(10061):2783-95 doi: 10.1016/S0140-6736(16)30172-6[published Online First: Epub Date]|.
2. Gubbi S, Thakur S, Avadhanula S, et al. Comprehensive guidance on the diagnosis and management of primary mesenchymal tumours of the thyroid gland. Lancet Oncol 2020;**21**(11):e528-e37 doi: 10.1016/S1470-2045(20)30332-6[published Online First: Epub Date]|.
3. Durante C, Grani G, Lamartina L, et al. The Diagnosis and Management of Thyroid Nodules: A Review. JAMA 2018;**319**(9):914-24 doi: 10.1001/jama.2018.0898[published Online First: Epub Date]|.
4. Wang TS, Dubner S, Sznyter LA, et al. Incidence of metastatic well-differentiated thyroid cancer in cervical lymph nodes. Arch Otolaryngol Head Neck Surg 2004;**130**(1):110-3 doi: 10.1001/archotol.130.1.110[published Online First: Epub Date]|.
5. Mazzaferri EL, Kloos RT. Clinical review 128: Current approaches to primary therapy for papillary and follicular thyroid cancer. J Clin Endocrinol Metab 2001;**86**(4):1447-63 doi: 10.1210/jcem.86.4.7407[published Online First: Epub Date]|.
6. Lee BJ, Wang SG, Lee JC, et al. Level IIb lymph node metastasis in neck dissection for papillary thyroid carcinoma. Arch Otolaryngol Head Neck Surg 2007;**133**(10):1028-30 doi: 10.1001/archotol.133.10.1028[published Online First: Epub Date]|.

7. Lundgren CI, Hall P, Dickman PW, et al. Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. Cancer 2006;**106**(3):524-31 doi: 10.1002/cncr.21653[published Online First: Epub Date]|.

8. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;**26**(1):1-133 doi: 10.1089/thy.2015.0020[published Online First: Epub Date]|.

9. Cheah WK, Arici C, Ituarte PH, et al. Complications of neck dissection for thyroid cancer. World J Surg 2002;**26**(8):1013-6 doi: 10.1007/s00268-002-6670-4[published Online First: Epub Date]|.

10. Leboulleux S, Girard E, Rose M, et al. Ultrasound criteria of malignancy for cervical lymph nodes in patients followed up for differentiated thyroid cancer. J Clin Endocrinol Metab 2007;**92**(9):3590-4 doi: 10.1210/jc.2007-0444[published Online First: Epub Date]|.

11. American Thyroid Association Guidelines Taskforce on Thyroid N, Differentiated Thyroid C, Cooper DS, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2009;**19**(11):1167-214 doi: 10.1089/thy.2009.0110[published Online First: Epub Date]|.

12. Stack BC, Jr., Ferris RL, Goldenberg D, et al. American Thyroid Association consensus review and statement regarding the anatomy, terminology, and rationale for lateral

neck dissection in differentiated thyroid cancer. Thyroid 2012;**22**(5):501-8 doi: 10.1089/thy.2011.0312[published Online First: Epub Date]|.

13. Ahn JE, Lee JH, Yi JS, et al. Diagnostic accuracy of CT and ultrasonography for evaluating metastatic cervical lymph nodes in patients with thyroid cancer. World J Surg 2008;**32**(7):1552-8 doi: 10.1007/s00268-008-9588-7[published Online First: Epub Date]|.
14. Choi JS, Kim J, Kwak JY, et al. Preoperative staging of papillary thyroid carcinoma: comparison of ultrasound imaging and CT. AJR Am J Roentgenol 2009;**193**(3):871-8 doi: 10.2214/AJR.09.2386[published Online First: Epub Date]|.
15. Liu X, Ouyang D, Li H, et al. Papillary thyroid cancer: dual-energy spectral CT quantitative parameters for preoperative diagnosis of metastasis to the cervical lymph nodes. Radiology 2015;**275**(1):167-76 doi: 10.1148/radiol.14140481[published Online First: Epub Date]|.
16. Kim E, Park JS, Son KR, et al. Preoperative diagnosis of cervical metastatic lymph nodes in papillary thyroid carcinoma: comparison of ultrasound, computed tomography, and combined ultrasound with computed tomography. Thyroid 2008;**18**(4):411-8 doi: 10.1089/thy.2007.0269[published Online First: Epub Date]|.
17. Lee DW, Ji YB, Sung ES, et al. Roles of ultrasonography and computed tomography in the surgical management of cervical lymph node metastases in papillary thyroid carcinoma. Eur J Surg Oncol 2013;**39**(2):191-6 doi: 10.1016/j.ejso.2012.07.119[published Online First: Epub Date]|.

18. Lee Y, Kim JH, Baek JH, et al. Value of CT added to ultrasonography for the diagnosis of lymph node metastasis in patients with thyroid cancer. Head Neck 2018;**40**(10):2137-48 doi: 10.1002/hed.25202[published Online First: Epub Date]|.

19. Na DK, Choi YJ, Choi SH, et al. Evaluation of cervical lymph node metastasis in thyroid cancer patients using real-time CT-navigated ultrasonography: preliminary study. Ultrasonography 2015;**34**(1):39-44 doi: 10.14366/usg.14030[published Online First: Epub Date]|.

20. Suh CH, Baek JH, Choi YJ, et al. Performance of CT in the Preoperative Diagnosis of Cervical Lymph Node Metastasis in Patients with Papillary Thyroid Cancer: A Systematic Review and Meta-Analysis. AJNR Am J Neuroradiol 2017;**38**(1):154-61 doi: 10.3174/ajnr.A4967[published Online First: Epub Date]|.

21. Cho SJ, Suh CH, Baek JH, et al. Diagnostic performance of CT in detection of metastatic cervical lymph nodes in patients with thyroid cancer: a systematic review and meta-analysis. Eur Radiol 2019;**29**(9):4635-47 doi: 10.1007/s00330-019-06036-8[published Online First: Epub Date]|.

22. Raijmakers PG, Paul MA, Lips P. Sentinel node detection in patients with thyroid carcinoma: a meta-analysis. World J Surg 2008;**32**(9):1961-7 doi: 10.1007/s00268-008-9657-y[published Online First: Epub Date]|.

23. Wu LM, Gu HY, Qu XH, et al. The accuracy of ultrasonography in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma: A meta-analysis. Eur J Radiol 2012;**81**(8):1798-805 doi: 10.1016/j.ejrad.2011.04.028[published Online First: Epub Date]|.

24. Zhao H, Li H. Meta-analysis of ultrasound for cervical lymph nodes in papillary thyroid cancer: Diagnosis of central and lateral compartment nodal metastases. Eur J Radiol 2019;**112**:14-21 doi: 10.1016/j.ejrad.2019.01.006[published Online First: Epub Date]|.

25. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;**6**(7):e1000100 doi: 10.1371/journal.pmed.1000100[published Online First: Epub Date]|.

26. 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**(8):529-36 doi: 10.7326/0003-4819-155-8-201110180-00009[published Online First: Epub Date]|.

27. 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**(10):982-90 doi: 10.1016/j.jclinepi.2005.02.022[published Online First: Epub Date]|.

28. Jeong HS, Baek CH, Son YI, et al. Integrated 18F-FDG PET/CT for the initial evaluation of cervical node level of patients with papillary thyroid carcinoma: comparison with ultrasound and contrast-enhanced CT. Clin Endocrinol (Oxf) 2006;**65**(3):402-7 doi: 10.1111/j.1365-2265.2006.02612.x[published Online First: Epub Date]|.

29. Choi YJ, Yun JS, Kook SH, et al. Clinical and imaging assessment of cervical lymph node metastasis in papillary thyroid carcinomas. World J Surg 2010;**34**(7):1494-9 doi: 10.1007/s00268-010-0541-1[published Online First: Epub Date]|.

30. Morita S, Mizoguchi K, Suzuki M, et al. The accuracy of (18)[F]-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography, ultrasonography, and enhanced computed tomography alone in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma. World J Surg 2010;**34**(11):2564-9 doi: 10.1007/s00268-010-0733-8[published Online First: Epub Date]|.

31. Yoon JH, Kim JY, Moon HJ, et al. Contribution of computed tomography to ultrasound in predicting lateral lymph node metastasis in patients with papillary thyroid carcinoma. Ann Surg Oncol 2011;**18**(6):1734-41 doi: 10.1245/s10434-010-1527-9[published Online First: Epub Date]|.

32. Seo YL, Yoon DY, Baek S, et al. Detection of neck recurrence in patients with differentiated thyroid cancer: comparison of ultrasound, contrast-enhanced CT and (18)F-FDG PET/CT using surgical pathology as a reference standard: (ultrasound vs. CT vs. (18)F-FDG PET/CT in recurrent thyroid cancer). Eur Radiol 2012;**22**(10):2246-54 doi: 10.1007/s00330-012-2470-x[published Online First: Epub Date]|.

33. Lesnik D, Cunnane ME, Zurakowski D, et al. Papillary thyroid carcinoma nodal surgery directed by a preoperative radiographic map utilizing CT scan and ultrasound in all primary and reoperative patients. Head Neck 2014;**36**(2):191-202 doi: 10.1002/hed.23277[published Online First: Epub Date]|.

34. Kim SK, Woo JW, Park I, et al. Computed Tomography-Detected Central Lymph Node Metastasis in Ultrasonography Node-Negative Papillary Thyroid Carcinoma: Is It Really Significant? Ann Surg Oncol 2017;**24**(2):442-49 doi: 10.1245/s10434-016-5552-1[published Online First: Epub Date]|.

35. Eun NL, Son EJ, Kim JA, et al. Comparison of the diagnostic performances of ultrasonography, CT and fine needle aspiration cytology for the prediction of lymph node metastasis in patients with lymph node dissection of papillary thyroid carcinoma: A retrospective cohort study. International journal of surgery 2018;**51**:145-50 doi: 10.1016/j.ijsu.2017.12.036[published Online First: Epub Date]|.

36. Scheumann GF, Gimm O, Wegener G, et al. Prognostic significance and surgical management of locoregional lymph node metastases in papillary thyroid cancer. World J Surg 1994;**18**(4):559-67; discussion 67-8 doi: 10.1007/BF00353765[published Online First: Epub Date]|.

37. Grebe SK, Hay ID. Thyroid cancer nodal metastases: biologic significance and therapeutic considerations. Surg Oncol Clin N Am 1996;**5**(1):43-63

38. Nam-Goong IS, Kim HY, Gong G, et al. Ultrasonography-guided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. Clin Endocrinol (Oxf) 2004;**60**(1):21-8 doi: 10.1046/j.1365-2265.2003.01912.x[published Online First: Epub Date]|.

39. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 2006;**6**:9 doi: 10.1186/1471-2288-6-9[published Online First: Epub Date]|.

40. Hay ID, Grant CS, van Heerden JA, et al. Papillary thyroid microcarcinoma: a study of 535 cases observed in a 50-year period. Surgery 1992;**112**(6):1139-46; discussion 46-7

41. Qubain SW, Nakano S, Baba M, et al. Distribution of lymph node micrometastasis in pN0 well-differentiated thyroid carcinoma. Surgery 2002;**131**(3):249-56 doi: 10.1067/msy.2002.120657[published Online First: Epub Date]|.

42. Arturi F, Russo D, Giuffrida D, et al. Early diagnosis by genetic analysis of differentiated thyroid cancer metastases in small lymph nodes. J Clin Endocrinol Metab 1997;**82**(5):1638-41 doi: 10.1210/jcem.82.5.4062[published Online First: Epub Date]|.

43. Kouvaraki MA, Shapiro SE, Fornage BD, et al. Role of preoperative ultrasonography in the surgical management of patients with thyroid cancer. Surgery 2003;**134**(6):946-54; discussion 54-5 doi: 10.1016/s0039-6060(03)00424-0[published Online First: Epub Date]|.

44. O'Connell K, Yen TW, Quiroz F, et al. The utility of routine preoperative cervical ultrasonography in patients undergoing thyroidectomy for differentiated thyroid cancer. Surgery 2013;**154**(4):697-701; discussion 01-3 doi: 10.1016/j.surg.2013.06.040[published Online First: Epub Date]|.

45. Yeh MW, Bauer AJ, Bernet VA, et al. American Thyroid Association statement on preoperative imaging for thyroid cancer surgery. Thyroid 2015;**25**(1):3-14 doi: 10.1089/thy.2014.0096[published Online First: Epub Date]|.

46. Kuna SK, Bracic I, Tesic V, et al. Ultrasonographic differentiation of benign from malignant neck lymphadenopathy in thyroid cancer. J Ultrasound Med

2006;**25**(12):1531-7; quiz 38-40 doi: 10.7863/jum.2006.25.12.1531[published Online First: Epub Date]|.

47. Park JH, Lee YS, Kim BW, et al. Skip lateral neck node metastases in papillary thyroid carcinoma. World J Surg 2012;**36**(4):743-7 doi: 10.1007/s00268-012-1476-5[published Online First: Epub Date]|.

48. Zhao Y, Li X, Li L, et al. Preliminary study on the diagnostic value of single-source dual-energy CT in diagnosing cervical lymph node metastasis of thyroid carcinoma. J Thorac Dis 2017;**9**(11):4758-66 doi: 10.21037/jtd.2017.09.151[published Online First: Epub Date]|.

49. He M, Lin C, Yin L, et al. Value of Dual-Energy Computed Tomography for Diagnosing Cervical Lymph Node Metastasis in Patients With Papillary Thyroid Cancer. J Comput Assist Tomogr 2019;**43**(6):970-75 doi: 10.1097/RCT.0000000000000927[published Online First: Epub Date]|.

50. Li L, Cheng SN, Zhao YF, et al. Diagnostic accuracy of single-source dual-energy computed tomography and ultrasonography for detection of lateral cervical lymph node metastases of papillary thyroid carcinoma. J Thorac Dis 2019;**11**(12):5032-41 doi: 10.21037/jtd.2019.12.45[published Online First: Epub Date]|.

Figure legend

Figure 1. Flowchart of the literature search process

Figure 2. Forest plots for the sensitivities and specificities of US, CT, and combination in neck level-based analysis

Figure 3. SROC of US, CT, and combination in detecting CLNM in patients with PTC in neck level-based analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer.

Figure 4. Forest plots for the sensitivities and specificities of US, CT, and combination in patient-based analysis

Figure 5. SROC of US, CT, and combination in detecting CLNM in patients with PTC in patient-based analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer

Table 1. Characteristics of included studies

Study	Country	Study design	Timing of imaging	Duration of patient recruitment	Sample (n) (males/ females)	Age (range)	No of lymph	Diagnosis	Analysis methods
Jeong HS 2006[28]	Korea	Retrospective	Preoperative	July 2004-March 2005	26 (7/19)	44 (17-73)		All PTC	L
Kim E 2008[16]	Korea	Retrospective	Preoperative	April 2006–October 2006	165 (25/140)	48 (16-78)		All PTC	L+P
Ahn JE 2008[13]	Korea	Retrospective	Preoperative	January 2005–December 2005	37 (7/30)	47(20-68)	481	All PTC	L
Choi JS 2009[14]	Korea	Retrospective	Preoperative	February 2006–April 2007	299 (44/255)	45 (20-74)	52	All PTC	L
Choi YJ 2010[29]	Korea	Retrospective	Preoperative	January 2007–December 2008	589 (121/468)	46	389	All PTC	P
Morita S 2010[30]	Japan	Prospective	Preoperative	January 2007–December 2009	74 (12/62)	66 (16-84)	349	All PTC	L

Yoon 2011[31]	JH	Korea	Retrospective	Preoperative	February 2007-December 2007	113 (16/97)	46	(15-83)	122	All PTC	L
Seo 2012[32]	YL	Korea	Retrospective	Postoperative	August 2008–August 2011	20 (4:16)	49.8		107	19 PTC, 1 MTC	L
Lee 2013[17]	DW	Korea	Retrospective	Preoperative	January 2007–May 2010	252 (45/207)	49	(15-82)		All PTC	L+P
Lesnik 2014[33]	D	USA	Prospective	Preoperative	2003–2008	95 (NA)	NA			All PTC	L
Na 2015[19]	DK	Korea	Retrospective	Preoperative	March 2011–February 2012	176 (44/132)	43	(23-74)	252	All PTC	P
Kim 2017[34]	SK	Korea	Retrospective	Preoperative	January 1997–June 2015	3668 (NA)	NA		455	All PTC	L
Eun 2018[34]	NL	Korea	Retrospective	Preoperative	January 2013–December 2015	302 (76:226)	44		308	All PTC	L
Lee 2018[18]	Y	Korea	Prospective	Preoperative	November 2011–December	351 (78:273)	47.1		801	All PTC	L

2012

NA not available; PTC papillary thyroid carcinoma; MTC medullary thyroid carcinoma; L, neck level-based analysis; P, Patient-based analysis

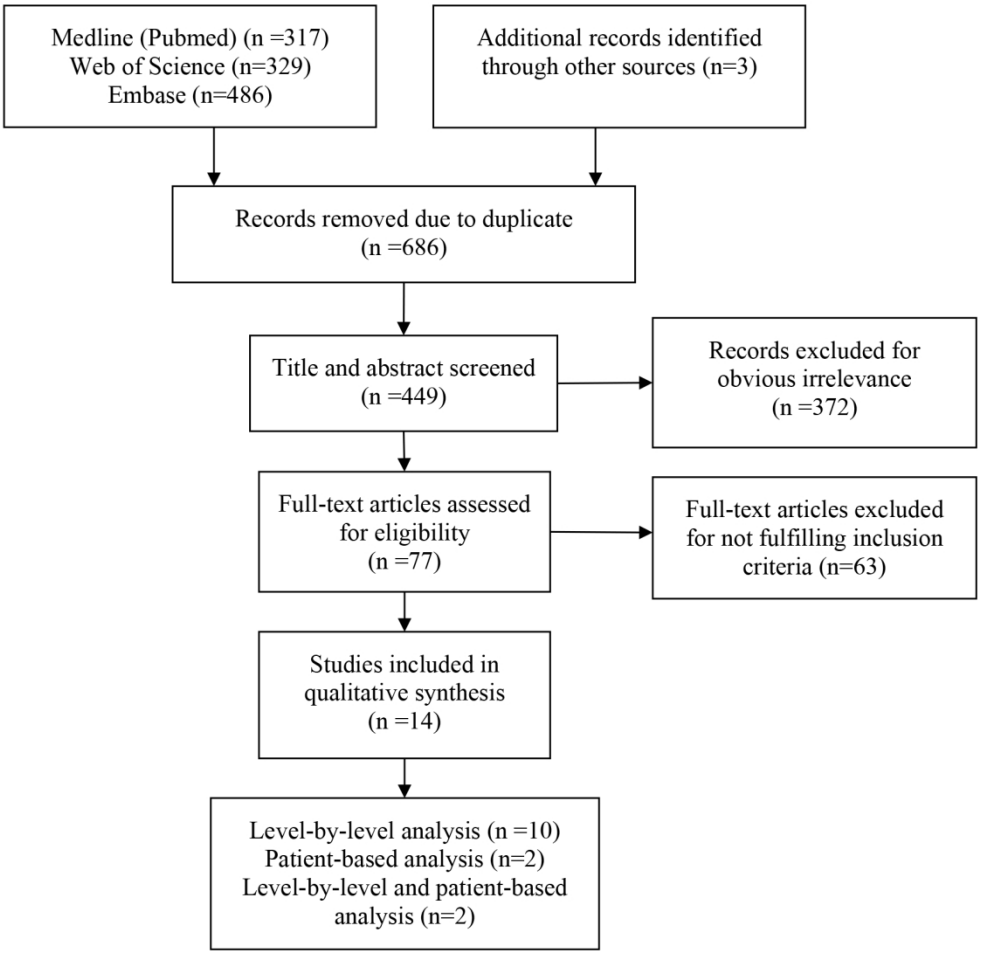
Table 2. Pooled estimates of sensitivity, specificity, diagnostic OR, LR+, LR-

	Studies,n	Sensitivity (95% CI)	Specificity (95% CI)	Diagnostic OR (95% CI)	LR+ (95% CI)	LR- (95% CI)
Diagnostic accuracy of CT or US on neck level						
All neck level						
US	11	0.35 (0.34-0.37, 97.5)	0.95 (0.94-0.95, 90.8)	13.94 (9.34-20.82, 81.0)	6.79 (4.79-9.63, 82.1)	0.50 (0.41-0.60, 95.2)
CT	11	0.46 (0.44-0.47, 97.6)	0.88 (0.87-0.89, 97.9)	7.24 (5.46-9.62, 72.2)	3.77 (2.08-6.84, 88.8)	0.52 (0.45-0.61,89.7)
US/CT	5	0.51 (0.49-0.52, 92.8)	0.85 (0.84-0.86, 97.8)	6.01 (3.84-9.40, 89.2)	3.04 (1.93-4.80, 93.3)	0.52 (0.45-0.60, 78.8)
Central neck level						
US	4	0.28 (0.24-0.32, 94.3)	0.97 (0.96-0.98, 53.0)	14.07 (6.66-29.71, 53.1)	14.07 (6.66-29.71, 53.1)	14.07 (6.66-29.71, 53.1)
CT	4	0.32 (0.28-0.36, 88.2)	0.89 (0.86-0.91, 84.6)	5.48 (2.15-13.98, 84.3)	3.71 (1.79-7.66, 93.0)	0.74 (0.62-0.89, 86.7)
US/CT	3	0.40 (0.35-0.45, 82.3)	0.85 (0.82-0.88, 37.0)	4.32 (2.09-8.92, 81.0)	2.85 (1.75-4.65, 87.5)	0.67 (0.52-0.86, 83.9)
Lateral neck level						
US	6	0.74 (0.69-0.78, 78.1)	0.92 (0.90-0.94, 89.6)	24.41 (11.16 -53.42, 71.9)	6.67 (2.91-15.30, 90.5)	0.35 (0.28-0.43, 30.1)
CT	6	0.73 (0.68-0.77, 90.7)	0.89 (0.87-0.91, 91.8)	15.55 (7.98 -30.32, 64.6)	5.54 (2.95-10.39, 82.2)	0.35 (0.21 -0.59, 91.4)
US/CT	4	0.88 (0.83-0.91, 64.1)	0.79 (0.73-0.84, 89.6)	22.59 (11.29 -45.19, 46.6)	3.31 (1.53-7.17, 91.1)	0.19 (0.14-0.25, 0)
Diagnostic accuracy of CT or US on patient level						
All patient level						

US	4	0.41 (0.36-0.46, 95.5)	0.92 (0.89-0.94, 0)	7.56 (4.08-14.01, 51.3)	4.48 (3.31 -6.05, 0)	0.65 (0.53-0.80, 75.5)
CT	4	0.49 (0.44-0.54, 95.9)	0.91 (0.89-0.94, 66.8)	9.40 (5.79-15.27, 34.4)	4.84 (3.66-6.39, 0)	0.53 (0.37-0.75, 85.6)
US/CT	2	0.64 (0.57-0.71, 0)	0.83 (0.77-0.88, 0)	8.59 (5.37-13.76, 0)	3.71 (2.72 -5.08, 0)	0.43 (0.36-0.53, 0)
Central patient level						
US	2	0.21 (0.16-0.28, 24.8)	0.95 (0.91-0.97, 0)	4.53 (2.34-8.77, 0)	3.78 (2.08-6.86, 0)	0.84 (0.76 -0.93, 36.4)
CT	2	0.38 (0.32-0.46, 92.6)	0.90 (0.85-0.93, 87.2)	5.02 (0.46-54.54, 94.4)	3.52 (0.52-23.84, 0)	0.71 (0.41-1.22, 95.3)
US/CT	2	0.47 (0.39-0.54, 80.8)	0.85 (0.80-0.89, 0)	4.88 (2.58-9.23, 46.4)	3.14 (2.23-4.41, 0)	0.64 (0.47-0.86, 78.4)
Lateral patient level						
US	2	0.87 (0.74-0.95, 88.2)	0.89 (0.83-0.93, 91.4)	20.11 (6.77-59.70, 0)	3.58 (0.85-15.16, 0)	0.22 (0.05-1.08, 58.0)
CT	2	0.92 (0.81-0.98, 88.6)	0.88 (0.83-0.93, 96.8)	36.88 (11.40 -119.35, 0)	3.44 (0.29-40.77, 98)	0.23 (0.10-0.52, 0)
US/CT	2	0.98 (0.89-0.99, 58.9)	0.92 (0.87-0.96, 97.5)	78.10 (2.82 -2160.4, 65.2)	5.30 (0.15-186.1, 96)	0.08 (0.02-0.41, 0)

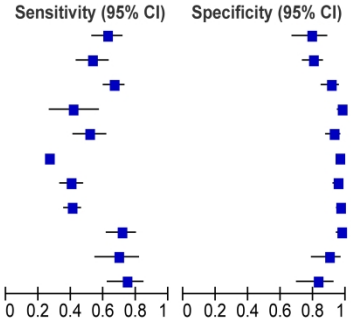
Results are presented as n (95% CI; I2, %); OR, odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio

Supplementary materials



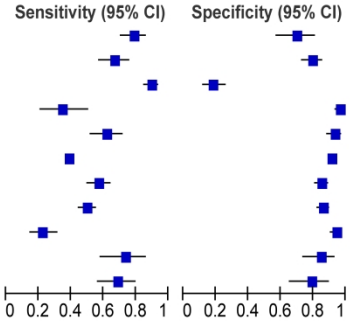
All neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	73	14	44	52	0.62 [0.53, 0.71]	0.79 [0.67, 0.88]
Choi 2009	59	38	52	150	0.53 [0.43, 0.63]	0.80 [0.73, 0.85]
Eun 2018	156	13	79	128	0.66 [0.60, 0.72]	0.91 [0.85, 0.95]
Jeong 2006	19	7	27	259	0.41 [0.27, 0.57]	0.97 [0.95, 0.99]
Kim 2008	54	13	51	159	0.51 [0.41, 0.61]	0.92 [0.87, 0.96]
Kim 2017	841	140	2296	3295	0.27 [0.25, 0.28]	0.96 [0.95, 0.97]
Lee 2013	83	18	124	333	0.40 [0.33, 0.47]	0.95 [0.92, 0.97]
Lee 2018	156	15	227	403	0.41 [0.36, 0.46]	0.96 [0.94, 0.98]
Morita 2010	79	7	32	231	0.71 [0.62, 0.79]	0.97 [0.94, 0.99]
Seo 2012	36	6	16	52	0.69 [0.55, 0.81]	0.90 [0.79, 0.96]
Yoon 2011	52	9	18	43	0.74 [0.62, 0.84]	0.83 [0.70, 0.92]



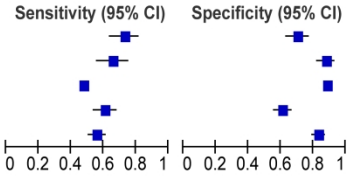
All Neck level:CT

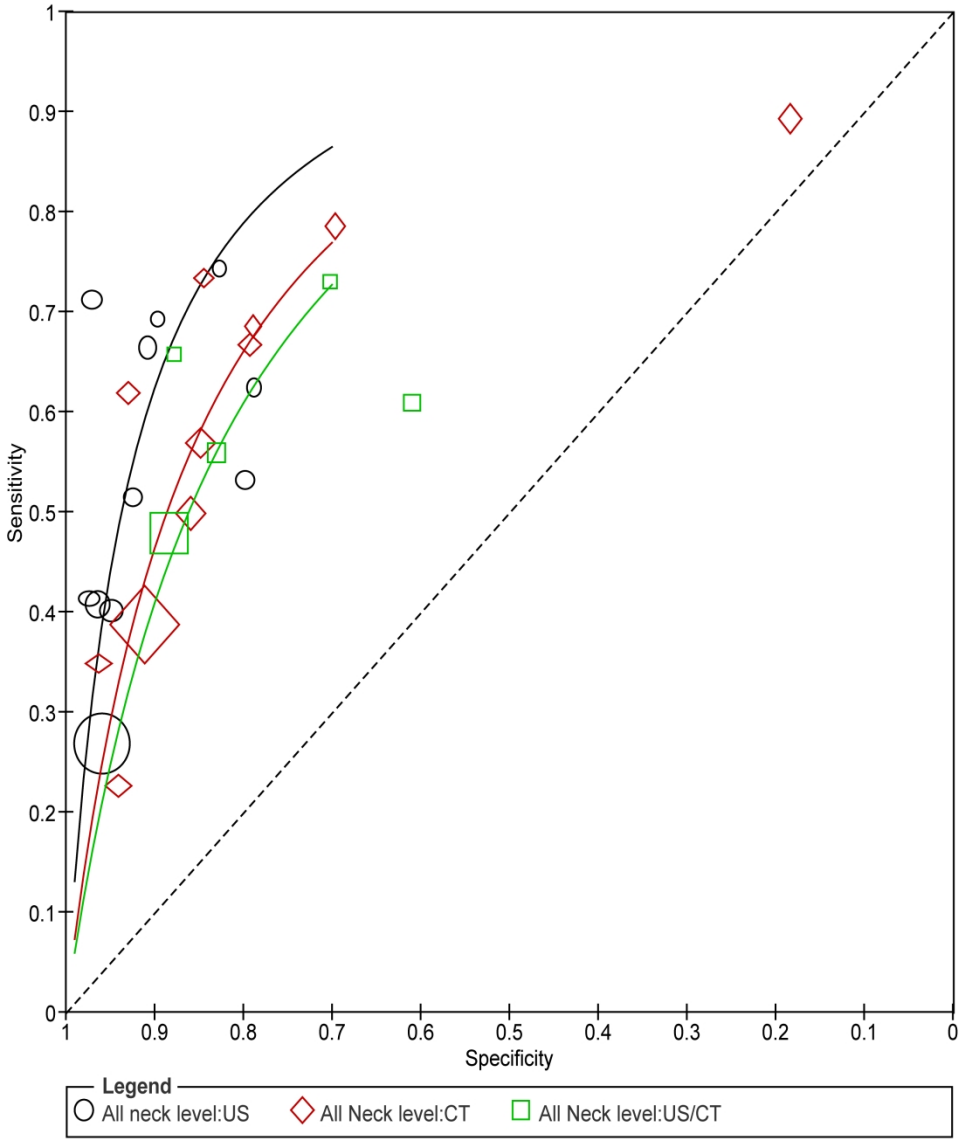
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	99	20	27	46	0.79 [0.70, 0.85]	0.70 [0.57, 0.80]
Choi 2009	74	39	37	149	0.67 [0.57, 0.75]	0.79 [0.73, 0.85]
Eun 2018	210	115	25	26	0.89 [0.85, 0.93]	0.18 [0.12, 0.26]
Jeong 2006	16	10	30	256	0.35 [0.21, 0.50]	0.96 [0.93, 0.98]
Kim 2008	65	12	40	160	0.62 [0.52, 0.71]	0.93 [0.88, 0.96]
Kim 2017	1221	306	1921	3129	0.39 [0.37, 0.41]	0.91 [0.90, 0.92]
Lee 2013	118	53	89	298	0.57 [0.50, 0.64]	0.85 [0.81, 0.88]
Lee 2018	191	59	192	359	0.50 [0.45, 0.55]	0.86 [0.82, 0.89]
Morita 2010	25	14	86	224	0.23 [0.15, 0.31]	0.94 [0.90, 0.97]
Seo 2012	33	10	12	55	0.73 [0.58, 0.85]	0.85 [0.74, 0.92]
Yoon 2011	48	11	22	41	0.69 [0.56, 0.79]	0.79 [0.65, 0.89]



All Neck level:US/CT

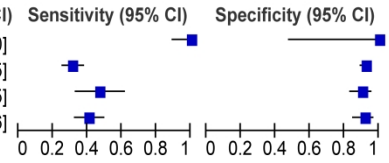
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	81	56	30	132	0.73 [0.64, 0.81]	0.70 [0.63, 0.77]
Kim 2008	69	21	36	151	0.66 [0.56, 0.75]	0.88 [0.82, 0.92]
Kim 2017	1503	400	1639	3035	0.48 [0.46, 0.50]	0.88 [0.87, 0.89]
Lee 2013	126	137	81	214	0.61 [0.54, 0.68]	0.61 [0.56, 0.66]
Lee 2018	214	71	169	347	0.56 [0.51, 0.61]	0.83 [0.79, 0.86]





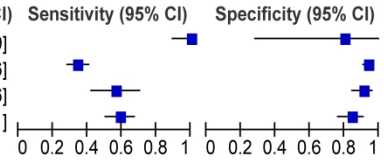
All patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	0	0	5	1.00 [0.89, 1.00]	1.00 [0.48, 1.00]
Choi 2010	75	27	163	324	0.32 [0.26, 0.38]	0.92 [0.89, 0.95]
Kim 2008	25	11	28	101	0.47 [0.33, 0.61]	0.90 [0.83, 0.95]
Lee 2013	59	9	85	99	0.41 [0.33, 0.49]	0.92 [0.85, 0.96]



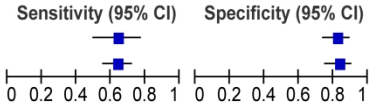
All patient level:CT

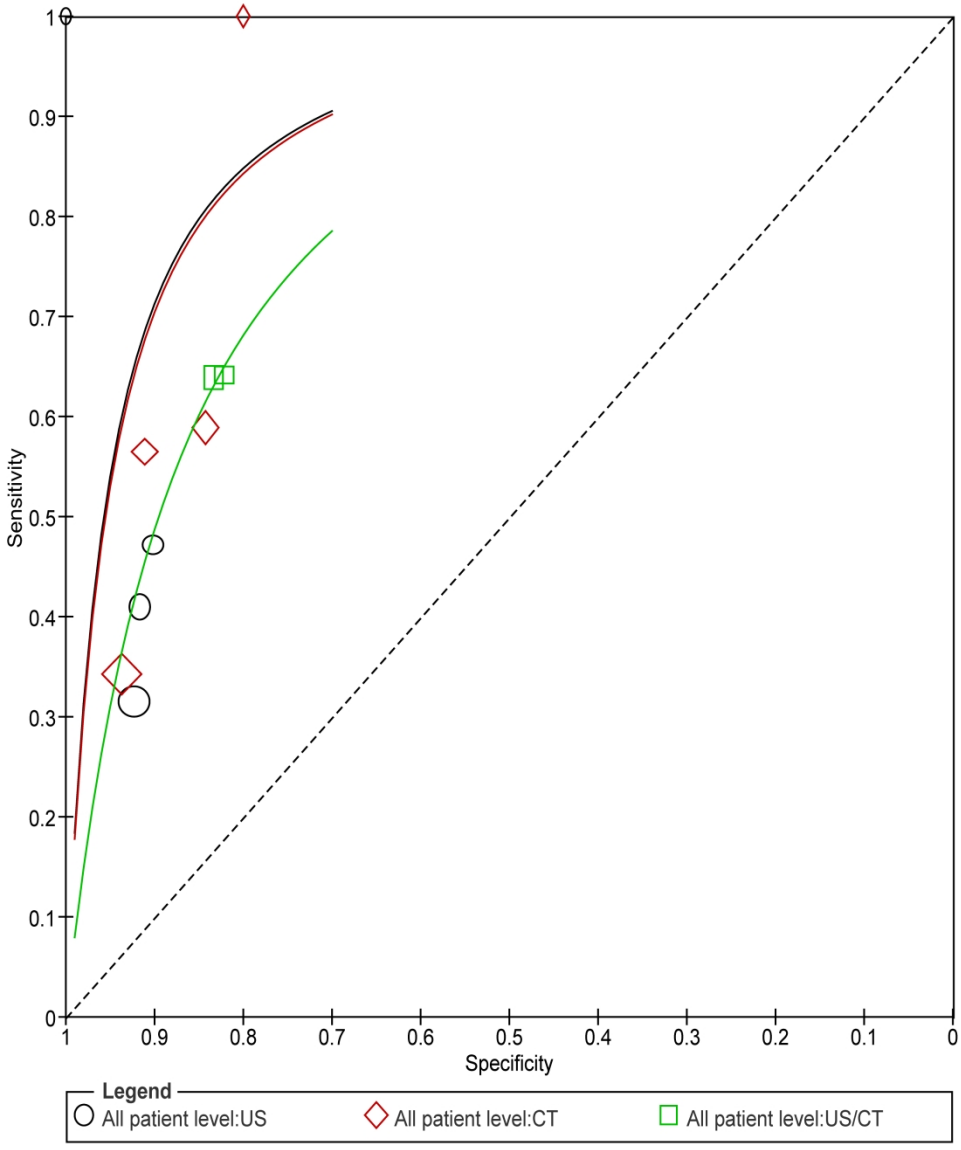
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	1	0	4	1.00 [0.89, 1.00]	0.80 [0.28, 0.99]
Choi 2010	82	22	156	329	0.34 [0.28, 0.41]	0.94 [0.91, 0.96]
Kim 2008	30	10	23	102	0.57 [0.42, 0.70]	0.91 [0.84, 0.96]
Lee 2013	85	17	59	91	0.59 [0.51, 0.67]	0.84 [0.76, 0.91]

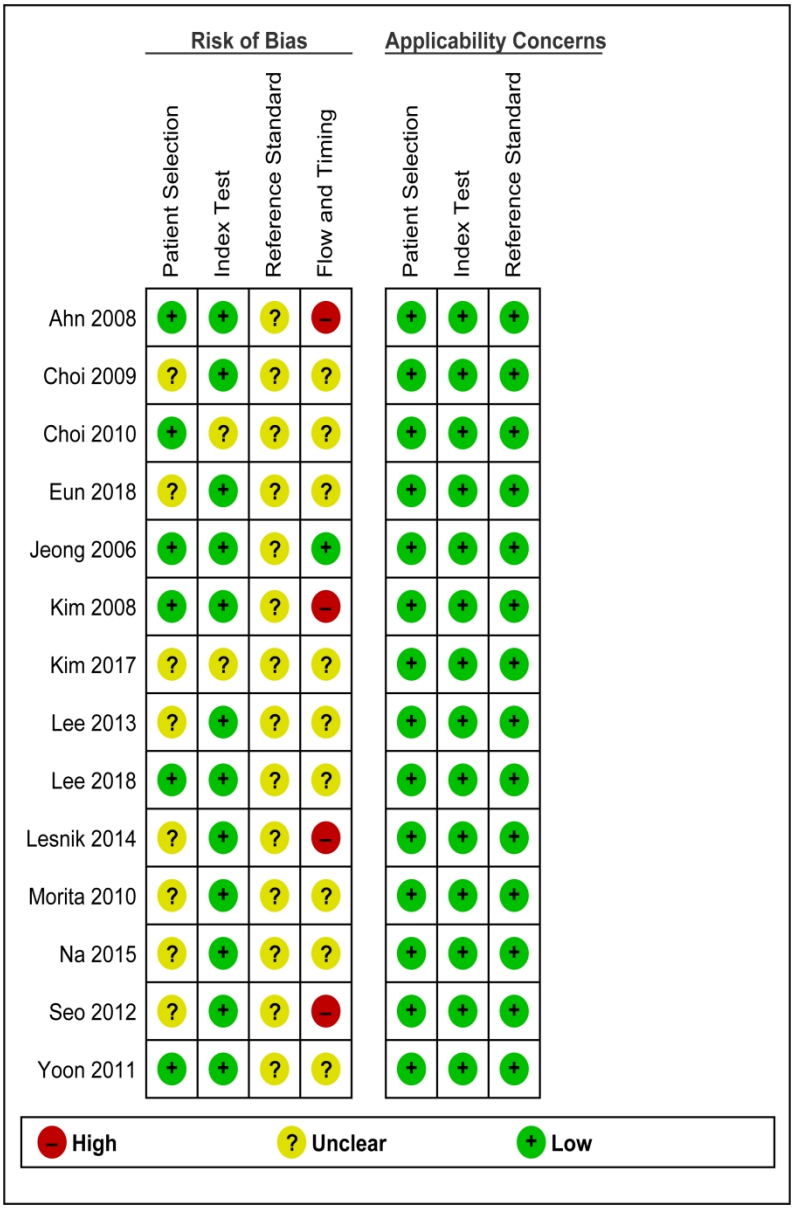


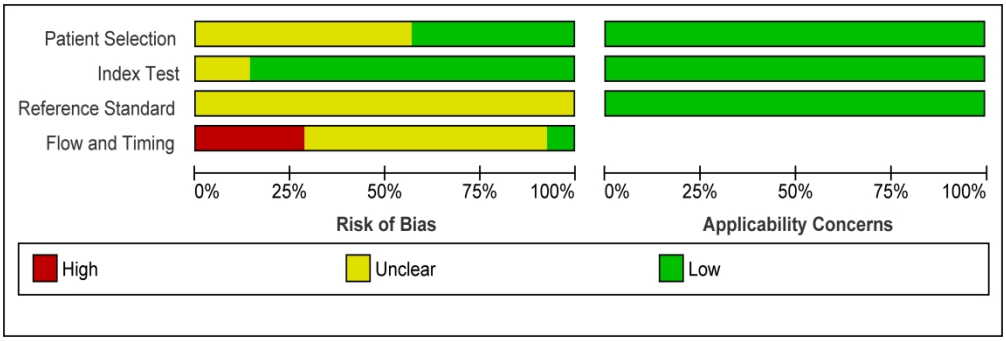
All patient level:US/CT

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	34	20	19	92	0.64 [0.50, 0.77]	0.82 [0.74, 0.89]
Lee 2013	92	18	52	90	0.64 [0.55, 0.72]	0.83 [0.75, 0.90]



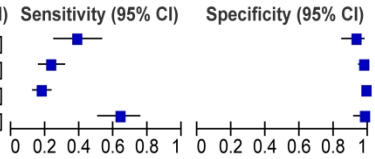






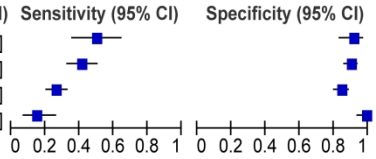
Central neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	20	6	32	75	0.38 [0.25, 0.53]	0.93 [0.85, 0.97]
Lee 2013	31	8	102	269	0.23 [0.16, 0.31]	0.97 [0.94, 0.99]
Lee 2018	39	5	180	307	0.18 [0.13, 0.24]	0.98 [0.96, 0.99]
Morita 2010	42	2	24	80	0.64 [0.51, 0.75]	0.98 [0.91, 1.00]



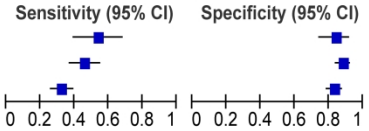
Central neck level:CT

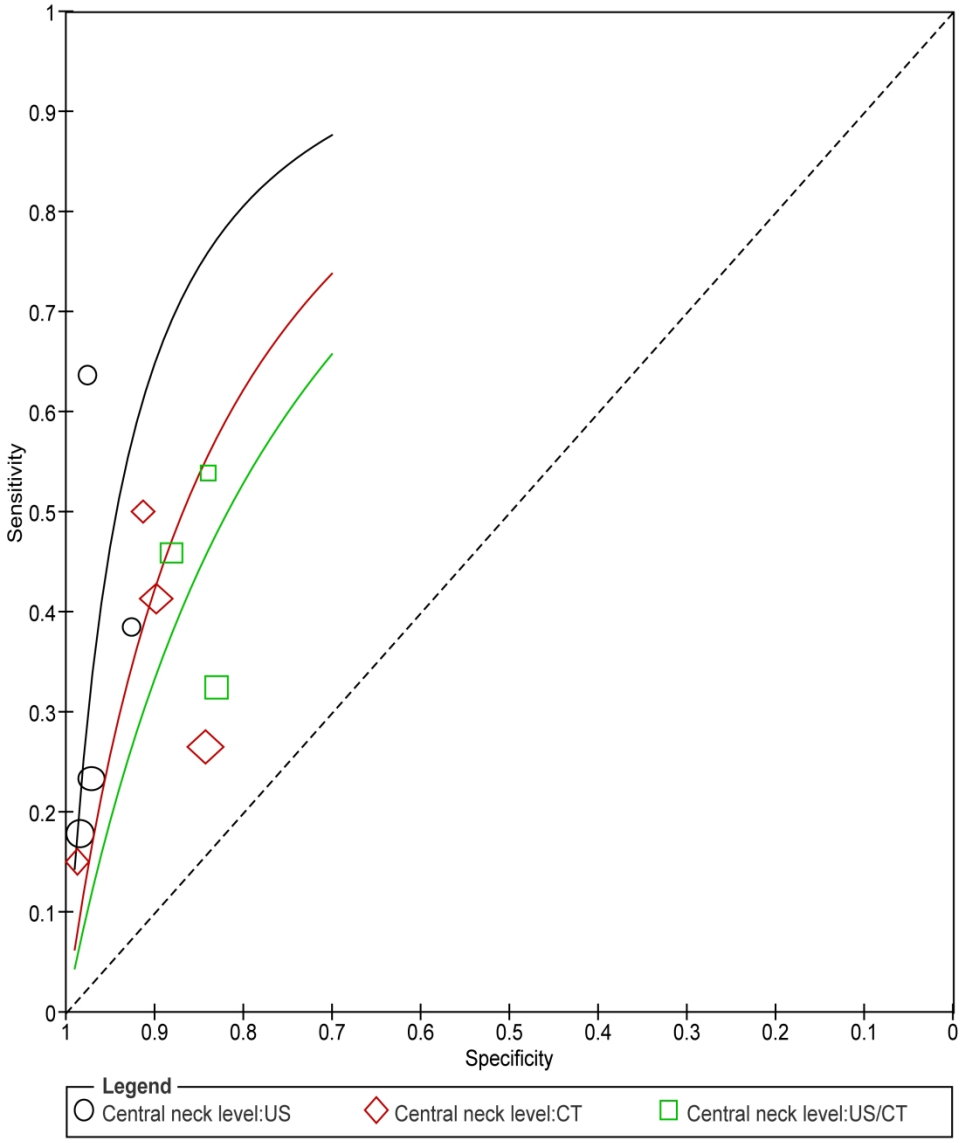
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	26	7	26	74	0.50 [0.36, 0.64]	0.91 [0.83, 0.96]
Lee 2013	55	28	78	249	0.41 [0.33, 0.50]	0.90 [0.86, 0.93]
Lee 2018	58	49	161	263	0.26 [0.21, 0.33]	0.84 [0.80, 0.88]
Morita 2010	10	1	56	81	0.15 [0.08, 0.26]	0.99 [0.93, 1.00]



Central neck level:US/CT

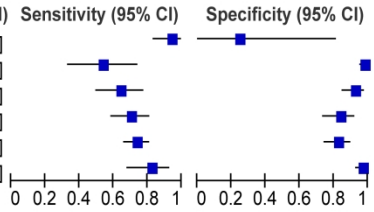
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	28	13	24	68	0.54 [0.39, 0.68]	0.84 [0.74, 0.91]
Lee 2013	61	33	72	244	0.46 [0.37, 0.55]	0.88 [0.84, 0.92]
Lee 2018	71	53	148	259	0.32 [0.26, 0.39]	0.83 [0.78, 0.87]





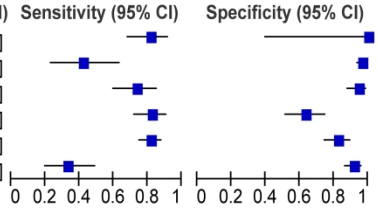
Lateral neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	46	3	3	1	0.94 [0.83, 0.99]	0.25 [0.01, 0.81]
Jeong 2006	14	5	12	229	0.54 [0.33, 0.73]	0.98 [0.95, 0.99]
Kim 2008	34	7	19	84	0.64 [0.50, 0.77]	0.92 [0.85, 0.97]
Lee 2013	52	12	22	62	0.70 [0.59, 0.80]	0.84 [0.73, 0.91]
Lee 2018	117	20	42	94	0.74 [0.66, 0.80]	0.82 [0.74, 0.89]
Morita 2010	37	5	8	151	0.82 [0.68, 0.92]	0.97 [0.93, 0.99]



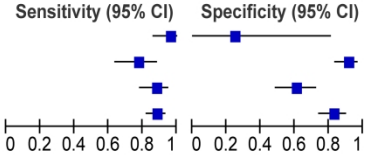
Lateral neck level:CT

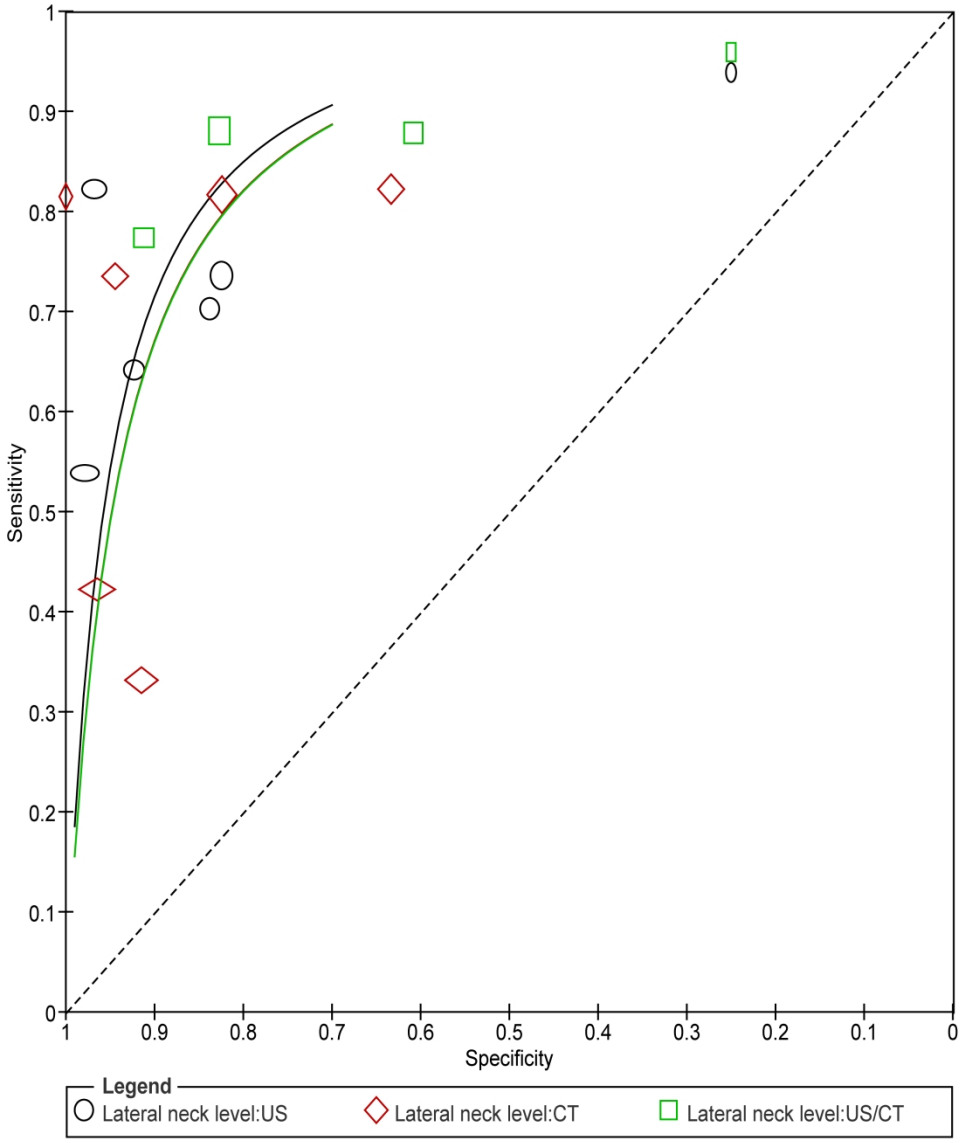
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	40	0	9	4	0.82 [0.68, 0.91]	1.00 [0.40, 1.00]
Jeong 2006	11	8	15	226	0.42 [0.23, 0.63]	0.97 [0.93, 0.99]
Kim 2008	39	5	14	86	0.74 [0.60, 0.85]	0.95 [0.88, 0.98]
Lee 2013	61	27	13	47	0.82 [0.72, 0.90]	0.64 [0.52, 0.74]
Lee 2018	130	20	29	94	0.82 [0.75, 0.87]	0.82 [0.74, 0.89]
Morita 2010	15	13	30	143	0.33 [0.20, 0.49]	0.92 [0.86, 0.95]



Lateral neck level:US/CT

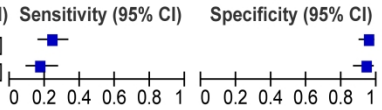
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	47	3	2	1	0.96 [0.86, 1.00]	0.25 [0.01, 0.81]
Kim 2008	41	8	12	83	0.77 [0.64, 0.88]	0.91 [0.83, 0.96]
Lee 2013	65	29	9	45	0.88 [0.78, 0.94]	0.61 [0.49, 0.72]
Lee 2018	140	18	19	86	0.88 [0.82, 0.93]	0.83 [0.74, 0.89]





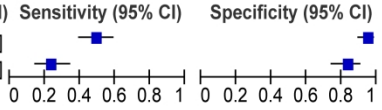
Central patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	27	7	85	133	0.24 [0.17, 0.33]	0.95 [0.90, 0.98]
Na 2015	14	6	67	89	0.17 [0.10, 0.27]	0.94 [0.87, 0.98]



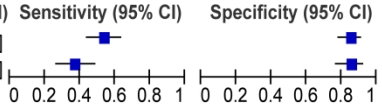
Central patient level:CT

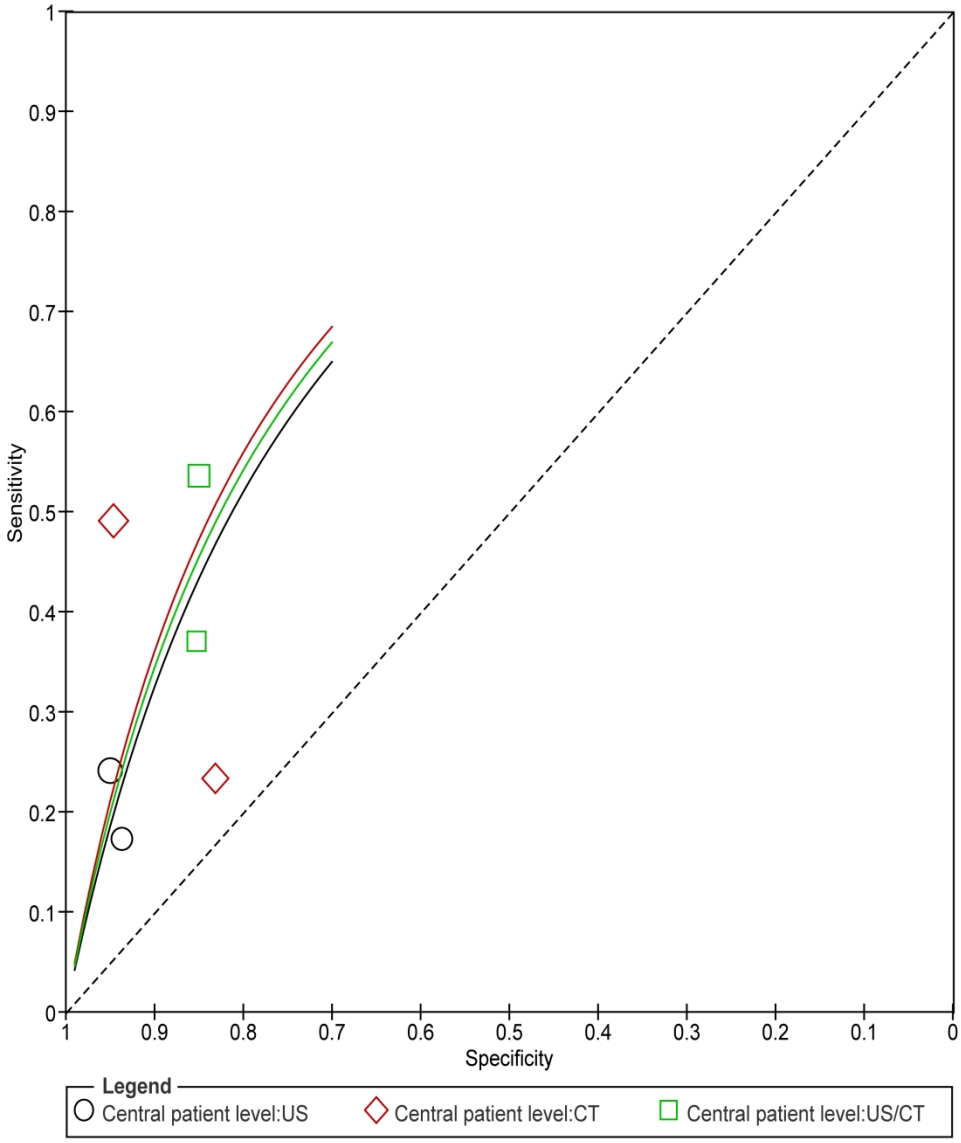
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	55	7	57	123	0.49 [0.40, 0.59]	0.95 [0.89, 0.98]
Na 2015	19	16	62	79	0.23 [0.15, 0.34]	0.83 [0.74, 0.90]



Central patient level:US/CT

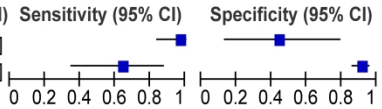
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	60	21	52	119	0.54 [0.44, 0.63]	0.85 [0.78, 0.90]
Na 2015	30	14	51	81	0.37 [0.27, 0.48]	0.85 [0.77, 0.92]





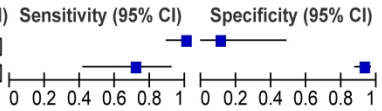
Lateral patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	31	5	1	4	0.97 [0.84, 1.00]	0.44 [0.14, 0.79]
Na 2015	9	14	5	148	0.64 [0.35, 0.87]	0.91 [0.86, 0.95]



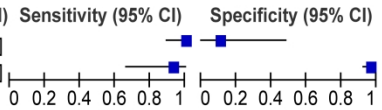
Lateral patient level:CT

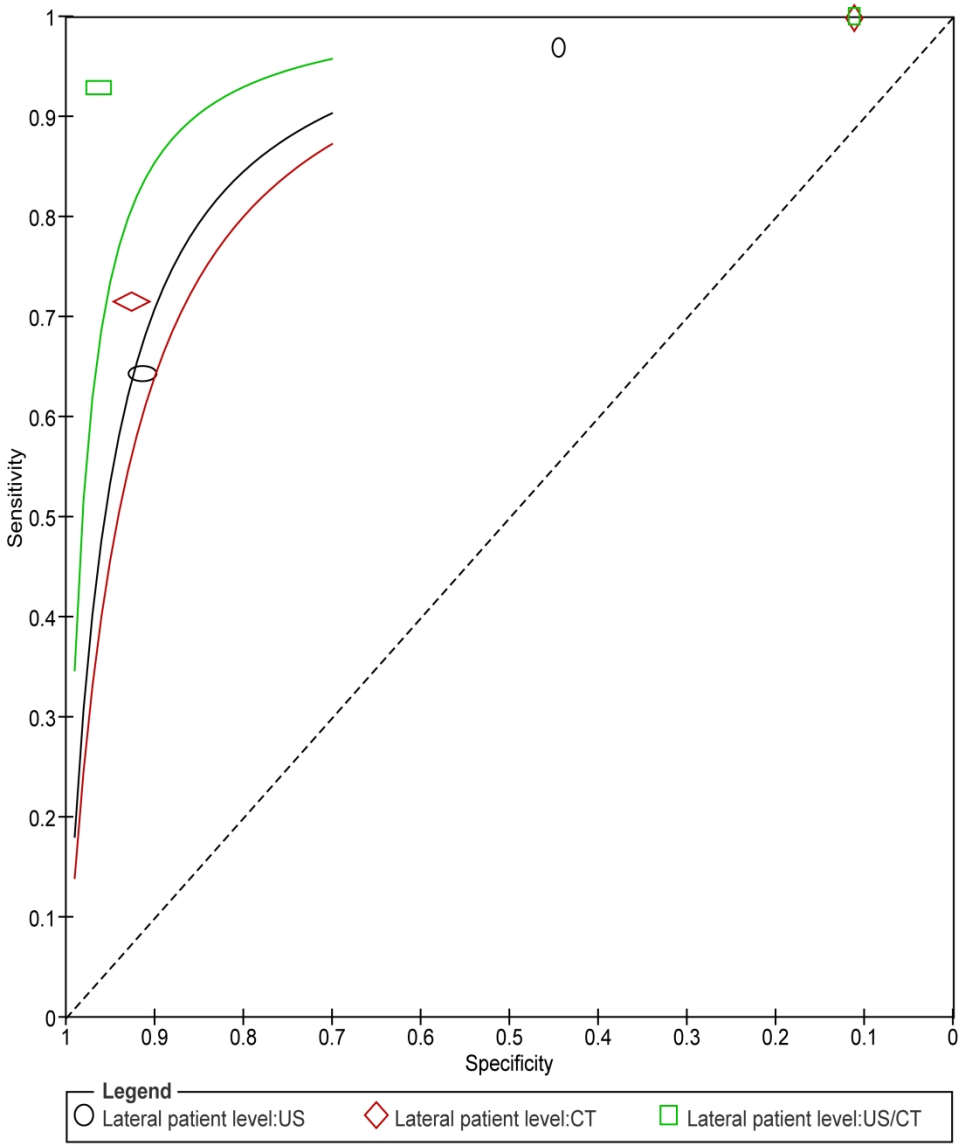
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	32	8	0	1	1.00 [0.89, 1.00]	0.11 [0.00, 0.48]
Na 2015	10	12	4	150	0.71 [0.42, 0.92]	0.93 [0.87, 0.96]



Lateral patient level:US/CT

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	32	8	0	1	1.00 [0.89, 1.00]	0.11 [0.00, 0.48]
Na 2015	13	6	1	156	0.93 [0.66, 1.00]	0.96 [0.92, 0.99]





Supplementary Table S1. Search strategy

PubMed		Search strategy	Numbers
P	#1	((((thyroid cancer) OR (thyroid carcinoma)) OR (thyroid tumor)) OR (papillary thyroid cancer)) OR (thyroid neoplasm)	94,066
	#2	(cervical lymph node) OR (neck lymph node)	38,156
	#3	#1 AND #2	5436
	#4	(metastasis) OR (metastatic)	1,436,116
	#5	#3 AND #4	4005
Intervention	#6	((ultrasonography) OR (ultrasound)) OR (US)	2,49,7,831
	#7	(computed tomography) OR (CT)	876,390
	#8	#6 AND #7	449,815
P+I	#9	#5 AND #8	317

Embase		Search strategy	Numbers
P	#1	'thyroid cancer' OR 'thyroid carcinoma' OR 'thyroid tumor' OR 'papillary thyroid cancer' OR 'thyroid neoplasm'	88,150

	#2	'cervical lymph node' OR 'neck lymph node'	18,642
	#3	#1 AND #2	3260
	#4	'metastasis' OR 'metastatic'	954,603
	#5	#3 AND #4	2772
I	#6	'ultrasonography' OR 'ultrasound' OR 'us'	1,451,034
	#7	'computed tomography' OR 'ct'	1,235,000
	#8	#6 AND #7	133,587
P+I	#9	#5 AND #8	486

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5-8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5-8
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5-8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5-8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5-8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5-8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5-8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8-11
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8-11
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	8-11
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	None

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

Diagnostic accuracy of ultrasound, computed tomography, and their combination in detecting cervical lymph node metastasis in patients with papillary thyroid cancer: a systematic review and meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-051568.R2
Article Type:	Original research
Date Submitted by the Author:	03-Jun-2022
Complete List of Authors:	Yang, Jian ; Xishan Coal Electricity Group Workers General Hospital, Department of Radiology Zhang, Fengyan; First Clinical Medical College, Shanxi Medical University, Department of Radiology Qiao, Ying; First Clinical Medical College, Shanxi Medical University, Department of Radiology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Radiology and imaging
Keywords:	Ultrasound < RADIOLOGY & IMAGING, Computed tomography < RADIOLOGY & IMAGING, Computed tomography < RADIOTHERAPY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Diagnostic accuracy of ultrasound, computed tomography, and their combination in detecting cervical lymph node metastasis in patients with papillary thyroid cancer: a systematic review and meta-analysis

Jian Yang[1,†], Fengyan Zhang[2,†], Ying Qiao[2,*]

[1]Department of Radiology, Xishan Coal Electricity Group Workers General Hospital, Taiyuan, 030053, Shanxi Province, China

[2]Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

[†]These authors contributed equally to this work.

[*]**Correspondence:**

Ying Qiao

Department of Radiology, First Clinical Medical College, Shanxi Medical University, Taiyuan, 030001, Shanxi Province, China

Tel: +86-15103462912

E-mail: qiaoying910723@163.com

Running Head: US and CT for CLNM for PTC

Abstract

Objectives: To determine the diagnostic accuracy of ultrasound (US), computed tomography (CT) and their combination in detecting cervical lymph node metastasis (CLNM) in patients with papillary thyroid cancer (PTC).

Methods: Medline (via PubMed), Web of Science, Embase were searched to identify studies published till December 5, 2021, that used US and CT to detect CLNM in patients with PTC. The primary outcomes were sensitivity, specificity, and diagnostic odds ratios (DORs) in neck level-based (lymph nodes are analyzed by neck level) or patient-based (lymph nodes are analyzed by patient) analysis. Secondary outcomes were sensitivity, specificity, and DORs in the central and lateral compartments.

Results: Fourteen studies (6167 patients with 11,601 neck lymph nodes) met the inclusion criteria. Based on the neck level-based analysis, the pooled sensitivity, specificity and DORs were 0.35 (95% confidence interval [(CI) 0.34-0.37], 0.95 (95% CI 0.94-0.95), and 13.94 (95% CI 9.34-20.82) for US, were 0.46 (95% CI 0.44-0.47), 0.88 (95% CI 0.87-0.89), and 7.24 (95% CI 5.46-9.62) for CT, were 0.51 (95% CI 0.49-0.52), 0.85 (95% CI 0.84-0.86), 6.01 (95% CI 3.84-9.40) for the combination of US and CT. In the patient-based analysis, the pooled estimates of sensitivity, specificity, and DOR were 0.41 (95% CI 0.36-0.46), 0.92 (95% CI 0.89-0.94), and 7.56 (95% CI 4.08-14.01) for US, were 0.49 (0.44-0.54), 0.91 (0.89-0.94), 9.40 (5.79-15.27) for CT, and were 0.64 (95% CI 0.57-0.71), 0.83 (95% CI 0.77-0.88), 8.59 (95% CI 5.37-13.76) for the combination of US and CT.

Discussion: These findings suggest US, with a DOR almost twice that of CT in the neck level-based analysis, was superior to CT in detecting CLNM in patients with PTC,

especially in the lateral compartment. The combination of US and CT increased the sensitivity from 41%-49% for the individual modalities to 64% for combined modalities in the patient-based analysis.

Keywords: ultrasound; computed tomography; cervical lymph node metastasis; papillary thyroid cancer; meta-analysis.

Strengths and limitations

- Only studies that analyzed CT and US were included.
- The analyses were performed based on the neck level and the patient level.
- Heterogeneity was observed due to study design and timing of the examinations.
- The use of CT for CLNM screening is not recognized everywhere globally.

Introduction

Papillary thyroid carcinoma (PTC) is an endocrine neoplasia with a high incidence of lymphatic metastasis and is associated with regional recurrence [1-3]. The incidence of cervical lymph node metastasis (CLNM) in patients with thyroid cancer has been reported to be 20%-90% [4]. The presence of CLNM might increase the risk of locoregional recurrence after surgery [5 6], worsening prognosis and survival [7]. Therefore, it is of great clinical importance to accurately evaluate CLNM and determine the extent of neck dissection [8]. Although prophylactic central compartment neck (groups VI and VII) dissection (ipsilateral or bilateral) is recommended by the American Thyroid Association (ATA) guidelines in patients with clinically positive central nodes, especially for those with advanced primary tumors, the information regarding prophylactic lateral compartment (groups I-V) neck dissection has not been clearly stated [8]. Thus, the indications for neck dissection, especially the lateral compartment, should be carefully assessed as it might lead to severe postoperative complications [9].

Preoperative staging with ultrasound (US) for cervical lymph nodes, including both central and lateral neck compartments, is the most widely accepted first imaging technique for patients with thyroid or suspicious malignancies cytologic or molecular findings. It can observe node enlargement, loss of fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, peripheral vascularity, and calcifications, which are all indicators of malignant invasion [10]. In addition, US is inexpensive, widely available, can be carried out bedside, and provide real-time imaging. Still, US is operator-dependent, and the images will vary depending on the angle and pressure of the probe on

the neck. Computed tomography (CT) provides three-dimensional neck images that avoid operator-dependency issues. On the other hand, the analysis of each layer takes time, and the use of contrast carries a risk of kidney injury. The preoperative use of computed tomography (CT) with intravenous (IV) contrast is complementary to US in patients with advanced disease [8 11 12]. A suspicious node on US can be confirmed by CT, and CT can detect nodes that were not visible because they were behind solid or air-containing structures or were not considered suspicious for various reasons. Although several studies have failed to prove the benefit of CT over US in detecting lateral lymph node metastasis [13-15], some studies suggested a superior diagnostic performance of the combination of preoperative US with CT over US alone [14 16-19].

Some meta-analyses examining the diagnostic accuracy of US and CT in detecting CLNM in patients with PTC have been previously conducted [20-24]. However, these meta-analyses studies integrated the findings of US and CT from different studies and populations.

To the best of our knowledge, no previous meta-analysis included studies that evaluated CLNM in patients with PTC using both US and CT, which could minimize the confounding effect of an operator in interpreting the diagnostic accuracy of preoperative imaging. This meta-analysis aimed to evaluate the sensitivity, specificity, and diagnostic odds ratios (DORs) of US, CT, and their combination in detecting positive CLNM in patients with PTC based on the central and lateral neck levels and by using neck level-based (lymph nodes are analyzed by neck level) and patient-based (lymph nodes are analyzed by patient, irrespective of the level) analyses.

Methods

Systematic literature research

This meta-analysis was reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines [25]. Ethical approval was waived due to the secondary data acquisition from previously published papers available in the public domain. A systematic search of Medline (via PubMed), Web of Science, and Embase was conducted to identify studies published up to December 5, 2021, that assessed the accuracy of US and CT in detecting CLNM in patients with PTC. The search strategy was developed in collaboration with a hospital librarian and included subject headings and text words: ("thyroid cancer" OR "thyroid carcinoma" OR "thyroid tumor" OR "papillary thyroid cancer" OR "thyroid neoplasm") AND ("cervical lymph node" OR "neck lymph node") AND ("metastasis" or "metastatic") AND ("ultrasonography" or "ultrasound" or "US") AND (computed tomography" or "CT") (Supplementary Table S1). The studies were initially screened by examining their titles and abstracts, and the full texts of potentially eligible studies were retrieved for further review. No language restriction was applied. A manual search of additional records and reference lists was also performed to include more relevant studies.

Study selection

The inclusion criteria of the studies were (a) prospective or retrospective studies that evaluated the diagnostic accuracy of both US and CT for detecting CLNM in patients with PTC, using neck level-based or patient-based analysis; (b) studies with >10 patients; (c) studies with a reference standard of histopathology or cytology (the diagnostic gold

standard was the pathological diagnosis of the resected lymph nodes); (d) studies that reported the absolute numbers of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) results directly or derived from the reported data or communicated by the authors in response to our request; (e) studies published in English.

The exclusion criteria were (i) case reports, case series, review articles, pictorial essays, letters to editors, unpublished data, conference abstracts, and proceedings on the topic of interest; (ii) studies that used only US or only CT; (iii) insufficient data regarding TP, FP, FN, and TN; (iv) duplicate publications using the same databases and studies; (v) if the patient population of one article is overlapping with the patient population of other or multiple articles, then the article with the largest sample size was included; (vi) studies with less than 10 cases confirmed by the reference standard. One reader reviewed the full texts of the candidate articles and selected those that met the inclusion criteria. A second reader reviewed the process of the inclusion of articles in the meta-analysis. No inter-reader disagreements were observed.

Primary and secondary outcomes

The primary outcomes were sensitivity, specificity, and DORs in a neck level-based or patient-based analysis. Secondary outcomes were sensitivity, specificity, and DORs in central and lateral compartments in neck level-based or patient-based analysis.

Data extraction and quality assessment

Two reviewers independently performed the data extraction. Data such as study characteristics, clinical and patient characteristics, reference standard or standards, cervical lymph node compartment, technical characteristics of CT and US and contrast enhancement, the definition of CLNM according to CT and US image findings, and the

diagnostic performance of CT and US, such as TP, FP, FN, TN were obtained from each study.

Two reviewers who were not blinded to the journal names, author names, and year of publication assessed the methodologic and reporting quality of each study by using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) [26]. Each study was independently assessed by two reviewers after a tutorial meeting on the guidelines for interpreting the items. Any disagreements were resolved by discussion with an experienced third reviewer.

Statistical analysis

The pooled sensitivity, specificity, diagnostic odds ratio (OR), positive likelihood ratio (LR+), and negative likelihood ratio (LR-) were calculated for US and CT in a neck level-based analysis (at neck level, central neck level, and lateral neck level) and a patient-based analysis (patient level, central patient level, and lateral level). The heterogeneity of pooled sensitivity, specificity, diagnostic OR, LR+, and LR- was measured by the inconsistency (I^2). Heterogeneity in the included articles was defined as small I^2 <25%, moderate I^2 25%-50%, and obvious I^2 >50%. If heterogeneity was detected (P-value <0.10 or $I^2 \geq 50\%$), a random-effects model was applied; otherwise, a fixed-effects model was used. A bivariate logistic regression model was used for meta-analysis of diagnostic test accuracy [27], and forest plots were created. The pooling of sensitivity, specificity, diagnostic OR, LR+, and LR- was performed using the Meta-Disc software (version 1.4, Madrid, Spain). Forest plots and summary receiver operating characteristic (SROC) curves were obtained using RevMan 5.3. A P-value of <0.05 was considered statistically significant.

Patient and public involvement

The patients and the public were not involved in this study.

Results

Characteristics of included studies

The initial literature search yielded 1135 potential studies for this meta-analysis. A total of 449 articles were screened after removing the duplicates. Of these, 372 studies were excluded after reviewing the titles and abstracts, and 63 articles were excluded after reviewing the full texts (Figure 1). Fourteen studies were ultimately selected for inclusion [13 14 16-19 28-35]: 10 studies used a neck level-based analysis, two studies used a patient-based analysis, and two studies used both. Five studies reported the diagnostic performance by combining both US and CT [14 16-19]. A total of 6167 patients with 11,601 neck lymph nodes were included, and all patients were diagnosed with PTC except one who was diagnosed with medullary thyroid cancer. The earliest study was started in 1997, whereas the latest one was started in 2012. The median number of patients per study was 171 (range 20-3668), while the median number of lymph nodes per study was 331 (range 107-6557). Eleven were retrospective studies, and three were prospective studies; 13 studies were performed preoperatively, and 1 study was performed postoperatively. Twelve, one, and one were conducted in Korea, the United States, and Japan, respectively (Table 1). The studies included in this meta-analysis were of moderate quality (supplementary Figure 1[sFigure 1], sFigure 2).

Neck level-based diagnostic accuracy of US and CT

Eleven studies used both CT and US for detecting CLNM in patients with PTC, and five of them assessed the diagnostic accuracy of the combination of CT and US. The pooled sensitivity, specificity, diagnostic OR, LR+, and LR- were 0.35 (95% CI 0.34-0.37), 0.95 (95% CI 0.94-0.95), 13.94 (95% CI 9.34-20.82), 6.79 (95% CI 4.79-9.63), and 0.50 (95% CI 0.41-0.60) for US, were 0.46 (95% CI 0.44-0.47), 0.88 (95% CI 0.87-0.89), 7.24 (95% CI 5.46-9.62), 3.77 (95% CI 2.08-6.84), and 0.52 (95% CI 0.45-0.61) for CT, and were 0.51 (95% CI 0.49-0.52), 0.85 (95% CI 0.84-0.86), 6.01 (95% CI 3.84-9.40), 3.04 (95% CI 1.93-4.80), and 0.52 (95% CI 0.45-0.60) for the combination of US and CT, with marked heterogeneity (Table 2, Figure 2, Figure 3).

Subgroup analyses of central and lateral neck levels were performed to investigate the effects of cervical lymph node compartment based on the diagnostic accuracy of US and CT. The subgroup analysis of the central neck level revealed that the pooled sensitivity, specificity, and DOR of US were 0.28 (95% CI 0.24-0.32), 0.97 (95% CI 0.96-0.98), and 14.07 (95% CI 6.66-29.71) from four studies. For CT, the pooled sensitivity, specificity, and DOR were 0.32 (95% CI 0.28-0.36), 0.89 (95% CI 0.86-0.91), and 5.48 (95% CI 2.15-13.98) from four studies. The pooled sensitivity, specificity, and DOR of the combination of US and CT were 0.40 (95% CI 0.35-0.45), 0.85 (95% CI 0.82-0.88), and 4.32 (95% CI 2.09-8.92) from three studies (Table 2, sFigure 3, sFigure 4).

In contrast, the subgroup analysis of the lateral neck level revealed that the pooled sensitivity, specificity, and DOR of US were 0.74 (95% CI 0.69-0.78), 0.92 (95% CI 0.90-0.94), and 24.41 (95% CI 11.16 -53.42) from six studies; the values for CT were 0.73 (95% CI 0.68-0.77), 0.89 (95% CI 0.87-0.91), and 15.55 (95% CI 7.98-30.32) from

six studies; the values for the combination of US and CT were 0.88 (95% CI 0.83-0.91), 0.79 (95% CI 0.73-0.84), and 22.59 (95% CI 11.29-45.19) from four studies (Table 2, sFigure 5, sFigure 6).

Patient-based diagnostic accuracy of US and CT

Four studies included both US and CT in detecting CLNM in patients with PTC, and two of them assessed the diagnostic accuracy by combining both CT and US. The pooled estimates of sensitivity, specificity, and DOR of US were 0.41 (95% CI 0.36-0.46), 0.92 (95% CI 0.89-0.94), and 7.56 (95% CI 4.08-14.01); the values for CT were 0.49 (0.44-0.54), 0.91 (0.89-0.94), and 9.40 (5.79-15.27); the values for the combination of US and CT were 0.64 (95% CI 0.57-0.71), 0.83 (95% CI 0.77-0.88), and 8.59 (95% CI 5.37-13.76) (Table 2, Figure 4, Figure 5).

Only two studies assessed the diagnostic accuracy of US, CT, and their combination on a patient basis. On the patient level, the pooled estimates of sensitivity, specificity, and DOR were 0.21 (95% CI 0.16-0.28), 0.95 (95% CI 0.91-0.97), and 4.53 (95% CI 2.34-8.77) for US, were 0.38 (95% CI 0.32-0.46), 0.90 (95% CI 0.85-0.93), and 5.02 (95% CI 0.46-54.54) for CT, and were 0.47 (95% CI 0.39-0.54), 0.85 (95% CI 0.80-0.89), and 4.88 (95% CI 2.58-9.23) for the combination of CT and US (Table 2, sFigure 7, sFigure 8).

In contrast, the pooled estimates of sensitivity, specificity, and DOR of US were 0.87 (95% CI 0.74-0.95), 0.89 (95% CI 0.83-0.93), and 20.11 (95% CI 6.77-59.70); the values for CT were 0.92 (95% CI 0.81-0.98), 0.88 (95% CI 0.83-0.93), and 36.88 (95% CI 11.40-119.35); the values for the combination of US and CT were 0.98 (95% CI 0.89-

0.99), 0.92 (95% CI 0.87-0.96), and 78.10 (95% CI 2.82-2160.4) (Table 2, sFigure 9, sFigure 10).

Discussion

This meta-analysis revealed that the DORs of US in the neck level-based analysis was higher than for CT or their combination on the central, lateral, and neck levels. Differentiated thyroid carcinoma, particularly PTC, involves CLNMs in 20%-50% of the patients [36-39], which could prevent small and intrathyroidal primary tumors[40]. Still, the clinical implications of macrometastases (≥2 mm) are more significant than micrometastases, in which 90% of patients might reach the criteria according to the sensitivity of the imaging methods [41 42]. The combination of US features might increase the likelihood of detecting CLNM as several US features are suggestive of metastatic lymph nodes, including enlargement, loss of fatty hilum, a rounded rather than oval shape, hyperechogenicity, cystic change, peripheral vascularity, and calcifications [10]. The preoperative US identifies lymph node or soft-tissue metastases in up to 39% of patients who had no physical examination [43] and changed the operative management in 23% of patients [44].

Previous meta-analyses examined CT and US. Suh et al. [20] and Cho et al. [21] demonstrated the value of CT for CNLM but did not include US. Raijmakers et al. [22] only examined the detection of the sentinel lymph node. Wu et al. [23] and Zhao et al. [24] examined the value of US for CLNMs but did not include CT. Therefore, these studies did not examine CT and US simultaneously. Our data found that the DORs of CT were higher than US and the combination, and the DORs of the combination remained

higher than US and CT by patient-based analysis. This was reasonable because the sensitivity of CT in the patient-based analysis was higher than that of US in the central, lateral, and patient analyses. This result might still need further investigation because of the inclusion of a small number of studies in the subgroup analysis. The operator-independent CT could be used as an adjunct in imaging deep anatomic structures, including the mediastinum, infraclavicular, retropharyngeal, and parapharyngeal regions and the structures that are acoustically shadowed by bone or air. In addition, preoperative knowledge on the extent of laryngeal, tracheal, and esophageal involvement, as well as bulky nodal disease from neck CT with contrast, significantly influences the surgical plan by indicating the need for sternotomy, tracheal or laryngeal resection, and reconstruction [45].

The results suggested that the sensitivity on the lateral compartment tended to be higher than for the central compartment regardless of the use of US, CT, or their combination in the neck level-based and patient-based analyses. The location of the lymph nodes helps in decision-making as most of the metastatic nodes are found in the lower third of the neck, and reactive enlarged lymph nodes are found in the upper part of the neck [46]. Besides, the lateral compartment should be carefully evaluated for skip metastases located in the upper pole or are ≤1 cm in diameter [47]. For patients who had preoperative CT and US and subsequently underwent total thyroidectomy and neck dissection, the sensitivity of CT was much better than US for evaluating CLNM on the neck level, but the sensitivity, specificity, and DORs for the lateral neck level tended to be higher than those of the central neck level for both CT and US[13]. Dual-energy CT (DECT) for assessing CLNM in patients with PTC was not included in this meta-analysis

as it can generate iodine-based material decomposition (MD) images and spectral HU curve [48-50]. In accordance with the findings from CT, combined gemstone spectral image (GSI) parameters from DECT also demonstrated better diagnostic accuracy of CLNM in patients with PTC when compared to those that are obtained by combining the US morphological parameters especially in the lateral compartment [50].

Our findings revealed that compared to US or CT alone, the combination of both US and CT demonstrated higher sensitivity, i.e., a meta-analytic summary sensitivity of 0.51 (0.49-0.52) and 0.64 (0.57-0.71), and a lower specificity, i.e., a meta-analytic summary specificity of 0.85 (0.84-0.86) and 0.83 (0.77-0.88) for evaluating CLNM in patients with PTC using neck level-based and patient-based analyses, respectively. In patients undergoing primary and revision surgical treatment for PTC, combined preoperative mapping with US and CT yielded significantly higher sensitivity for detecting macroscopic lymph nodes in both lateral and central neck, especially in the central neck [33].

It should be noted that the study has strengths. Firstly, Boolean operatives of "AND" rather than "OR" were used for combined datasets for all studies. Namely, only studies of direct head-to-head comparison by US, CT, and combination of both in the same patient population were included in this meta-analysis, avoiding bias due to differences in patient and institutional factors. Secondly, a meta-analysis of the included studies was performed by using neck level-based and patient-based analyses and on all, central, and lateral neck levels. Lastly, our data suggested that future follow-up studies should be performed to determine the comparative role of US and CT in identifying false-negative nodes that are not biopsied or excised.

Despite great clinical significance, there are several limitations in the current meta-analysis that are mostly associated with the available data and heterogeneity of design, interpretation of results, and reporting of data in primary studies. Firstly, the sources of heterogeneity among primary studies in meta-analyses have been reported by several previous studies, which included contrast amount, scan phase, and reconstruction slice thickness for CT [21], and the criteria of lymph node diameter and vascular flow for US [24]. Secondly, the literature included is limited due to the study design and timing of imaging. Eleven of the 14 studies (78.6%) were retrospective, and one of the 14 studies was a postoperative imaging study. A large proportion of retrospective studies might increase the sensitivity of CT and US. Twelve of the 14 studies were conducted in Korea, and so ethnic factors might affect the results of this meta-analysis. Thus, the complementary use of CT might be routine in Korea but not necessarily applicable to other parts of the world, especially in developing countries. Thirdly, modern high-resolution US transducers have a lateral resolution of 2 mm, which is not feasible for CT, allowing for the detection of small nodes and the presence of microcalcification. The included CT studies might not be comparable from one study to another, particularly over the decade, depending on the equipment, slice thickness, amount of contrast injected, etc. Fourthly, four of the 14 included studies were with patient-based results, and 12 of 14 studies were of suboptimal quality, and no definite recommendation could be drawn from the present study. Finally, MRI, US-guided FNA, and PET-CT were not included in the meta-analysis to directly compare CT and US, although they also play complementary roles in managing CLNMs in PTC.

Despite these potential drawbacks, this meta-analysis demonstrated the unique complementary value of CT secondary to US in detecting CLNMs in patients with PTC in the patient-based analysis. More importantly, the choice of a diagnostic test should be tailored to have feasible access to these imaging modalities at individual healthcare centers.

Conclusion

These findings suggest that US, with a DOR of almost twice that for CT in the neck level-based analysis, was superior to CT in detecting CLNM in patients with PTC, especially in the lateral compartment. The combination of US and CT increased the sensitivity from 41%-49% for the individual modalities to 64% for combined modalities in the patient-based analysis. CT might be valid a candidate imaging technique secondary to US in the management of CLNM in patients with PTC.

Ethics approval statement

Not applicable

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authorship

Jian Yang is the guarantor of the integrity of the entire study. Ying Qiao designed and conceptualized the study concepts and design, Jian Yang performed the literature search. Ying Qiao analyzed the literature. Jian Yang analyzed the data. Gengyan Zhang performed the statistical analyses, Ying Qiao prepared the manuscript. Fengyan Zhang edited the manuscript.

Acknowledgments

None.

Funding

None.

Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. Lancet 2016;**388**(10061):2783-95 doi: 10.1016/S0140-6736(16)30172-6[published Online First: Epub Date]|.

2. Gubbi S, Thakur S, Avadhanula S, et al. Comprehensive guidance on the diagnosis and management of primary mesenchymal tumours of the thyroid gland. Lancet Oncol 2020;**21**(11):e528-e37 doi: 10.1016/S1470-2045(20)30332-6[published Online First: Epub Date]|.

3. Durante C, Grani G, Lamartina L, et al. The Diagnosis and Management of Thyroid Nodules: A Review. JAMA 2018;**319**(9):914-24 doi: 10.1001/jama.2018.0898[published Online First: Epub Date]|.

4. Wang TS, Dubner S, Sznyter LA, et al. Incidence of metastatic well-differentiated thyroid cancer in cervical lymph nodes. Arch Otolaryngol Head Neck Surg 2004;**130**(1):110-3 doi: 10.1001/archotol.130.1.110[published Online First: Epub Date]|.

5. Mazzaferri EL, Kloos RT. Clinical review 128: Current approaches to primary therapy for papillary and follicular thyroid cancer. J Clin Endocrinol Metab 2001;**86**(4):1447-63 doi: 10.1210/jcem.86.4.7407[published Online First: Epub Date]|.

6. Lee BJ, Wang SG, Lee JC, et al. Level IIb lymph node metastasis in neck dissection for papillary thyroid carcinoma. Arch Otolaryngol Head Neck Surg

2007;**133**(10):1028-30 doi: 10.1001/archotol.133.10.1028[published Online First: Epub Date]|.

7. Lundgren CI, Hall P, Dickman PW, et al. Clinically significant prognostic factors for differentiated thyroid carcinoma: a population-based, nested case-control study. Cancer 2006;**106**(3):524-31 doi: 10.1002/cncr.21653[published Online First: Epub Date]|.

8. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;**26**(1):1-133 doi: 10.1089/thy.2015.0020[published Online First: Epub Date]|.

9. Cheah WK, Arici C, Ituarte PH, et al. Complications of neck dissection for thyroid cancer. World J Surg 2002;**26**(8):1013-6 doi: 10.1007/s00268-002-6670-4[published Online First: Epub Date]|.

10. Leboulleux S, Girard E, Rose M, et al. Ultrasound criteria of malignancy for cervical lymph nodes in patients followed up for differentiated thyroid cancer. J Clin Endocrinol Metab 2007;**92**(9):3590-4 doi: 10.1210/jc.2007-0444[published Online First: Epub Date]|.

11. American Thyroid Association Guidelines Taskforce on Thyroid N, Differentiated Thyroid C, Cooper DS, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated

thyroid cancer. Thyroid 2009;**19**(11):1167-214 doi: 10.1089/thy.2009.0110[published Online First: Epub Date]|.

12. Stack BC, Jr., Ferris RL, Goldenberg D, et al. American Thyroid Association consensus review and statement regarding the anatomy, terminology, and rationale for lateral neck dissection in differentiated thyroid cancer. Thyroid 2012;**22**(5):501-8 doi: 10.1089/thy.2011.0312[published Online First: Epub Date]|.
13. Ahn JE, Lee JH, Yi JS, et al. Diagnostic accuracy of CT and ultrasonography for evaluating metastatic cervical lymph nodes in patients with thyroid cancer. World J Surg 2008;**32**(7):1552-8 doi: 10.1007/s00268-008-9588-7[published Online First: Epub Date]|.
14. Choi JS, Kim J, Kwak JY, et al. Preoperative staging of papillary thyroid carcinoma: comparison of ultrasound imaging and CT. AJR Am J Roentgenol 2009;**193**(3):871-8 doi: 10.2214/AJR.09.2386[published Online First: Epub Date]|.
15. Liu X, Ouyang D, Li H, et al. Papillary thyroid cancer: dual-energy spectral CT quantitative parameters for preoperative diagnosis of metastasis to the cervical lymph nodes. Radiology 2015;**275**(1):167-76 doi: 10.1148/radiol.14140481[published Online First: Epub Date]|.
16. Kim E, Park JS, Son KR, et al. Preoperative diagnosis of cervical metastatic lymph nodes in papillary thyroid carcinoma: comparison of ultrasound, computed tomography, and combined ultrasound with computed tomography. Thyroid

2008;**18**(4):411-8 doi: 10.1089/thy.2007.0269[published Online First: Epub Date]|.

17. Lee DW, Ji YB, Sung ES, et al. Roles of ultrasonography and computed tomography in the surgical management of cervical lymph node metastases in papillary thyroid carcinoma. Eur J Surg Oncol 2013;**39**(2):191-6 doi: 10.1016/j.ejso.2012.07.119[published Online First: Epub Date]|.
18. Lee Y, Kim JH, Baek JH, et al. Value of CT added to ultrasonography for the diagnosis of lymph node metastasis in patients with thyroid cancer. Head Neck 2018;**40**(10):2137-48 doi: 10.1002/hed.25202[published Online First: Epub Date]|.
19. Na DK, Choi YJ, Choi SH, et al. Evaluation of cervical lymph node metastasis in thyroid cancer patients using real-time CT-navigated ultrasonography: preliminary study. Ultrasonography 2015;**34**(1):39-44 doi: 10.14366/usg.14030[published Online First: Epub Date]|.
20. Suh CH, Baek JH, Choi YJ, et al. Performance of CT in the Preoperative Diagnosis of Cervical Lymph Node Metastasis in Patients with Papillary Thyroid Cancer: A Systematic Review and Meta-Analysis. AJNR Am J Neuroradiol 2017;**38**(1):154-61 doi: 10.3174/ajnr.A4967[published Online First: Epub Date]|.
21. Cho SJ, Suh CH, Baek JH, et al. Diagnostic performance of CT in detection of metastatic cervical lymph nodes in patients with thyroid cancer: a systematic review and meta-analysis. Eur Radiol 2019;**29**(9):4635-47 doi: 10.1007/s00330-019-06036-8[published Online First: Epub Date]|.

22. Raijmakers PG, Paul MA, Lips P. Sentinel node detection in patients with thyroid carcinoma: a meta-analysis. World J Surg 2008;**32**(9):1961-7 doi: 10.1007/s00268-008-9657-y[published Online First: Epub Date]|.

23. Wu LM, Gu HY, Qu XH, et al. The accuracy of ultrasonography in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma: A meta-analysis. Eur J Radiol 2012;**81**(8):1798-805 doi: 10.1016/j.ejrad.2011.04.028[published Online First: Epub Date]|.

24. Zhao H, Li H. Meta-analysis of ultrasound for cervical lymph nodes in papillary thyroid cancer: Diagnosis of central and lateral compartment nodal metastases. Eur J Radiol 2019;**112**:14-21 doi: 10.1016/j.ejrad.2019.01.006[published Online First: Epub Date]|.

25. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;**6**(7):e1000100 doi: 10.1371/journal.pmed.1000100[published Online First: Epub Date]|.

26. 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**(8):529-36 doi: 10.7326/0003-4819-155-8-201110180-00009[published Online First: Epub Date]|.

27. 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**(10):982-90 doi: 10.1016/j.jclinepi.2005.02.022[published Online First: Epub Date]|.

28. Jeong HS, Baek CH, Son YI, et al. Integrated 18F-FDG PET/CT for the initial evaluation of cervical node level of patients with papillary thyroid carcinoma: comparison with ultrasound and contrast-enhanced CT. Clin Endocrinol (Oxf) 2006;**65**(3):402-7 doi: 10.1111/j.1365-2265.2006.02612.x[published Online First: Epub Date]|.

29. Choi YJ, Yun JS, Kook SH, et al. Clinical and imaging assessment of cervical lymph node metastasis in papillary thyroid carcinomas. World J Surg 2010;**34**(7):1494-9 doi: 10.1007/s00268-010-0541-1[published Online First: Epub Date]|.

30. Morita S, Mizoguchi K, Suzuki M, et al. The accuracy of (18)[F]-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography, ultrasonography, and enhanced computed tomography alone in the preoperative diagnosis of cervical lymph node metastasis in patients with papillary thyroid carcinoma. World J Surg 2010;**34**(11):2564-9 doi: 10.1007/s00268-010-0733-8[published Online First: Epub Date]|.

31. Yoon JH, Kim JY, Moon HJ, et al. Contribution of computed tomography to ultrasound in predicting lateral lymph node metastasis in patients with papillary thyroid carcinoma. Ann Surg Oncol 2011;**18**(6):1734-41 doi: 10.1245/s10434-010-1527-9[published Online First: Epub Date]|.

32. Seo YL, Yoon DY, Baek S, et al. Detection of neck recurrence in patients with differentiated thyroid cancer: comparison of ultrasound, contrast-enhanced CT and (18)F-FDG PET/CT using surgical pathology as a reference standard: (ultrasound vs. CT vs. (18)F-FDG PET/CT in recurrent thyroid cancer). Eur

Radiol 2012;**22**(10):2246-54 doi: 10.1007/s00330-012-2470-x[published Online First: Epub Date]|.

33. Lesnik D, Cunnane ME, Zurakowski D, et al. Papillary thyroid carcinoma nodal surgery directed by a preoperative radiographic map utilizing CT scan and ultrasound in all primary and reoperative patients. Head Neck 2014;**36**(2):191-202 doi: 10.1002/hed.23277[published Online First: Epub Date]|.
34. Kim SK, Woo JW, Park I, et al. Computed Tomography-Detected Central Lymph Node Metastasis in Ultrasonography Node-Negative Papillary Thyroid Carcinoma: Is It Really Significant? Ann Surg Oncol 2017;**24**(2):442-49 doi: 10.1245/s10434-016-5552-1[published Online First: Epub Date]|.
35. Eun NL, Son EJ, Kim JA, et al. Comparison of the diagnostic performances of ultrasonography, CT and fine needle aspiration cytology for the prediction of lymph node metastasis in patients with lymph node dissection of papillary thyroid carcinoma: A retrospective cohort study. International journal of surgery 2018;**51**:145-50 doi: 10.1016/j.ijsu.2017.12.036[published Online First: Epub Date]|.
36. Scheumann GF, Gimm O, Wegener G, et al. Prognostic significance and surgical management of locoregional lymph node metastases in papillary thyroid cancer. World J Surg 1994;**18**(4):559-67; discussion 67-8 doi: 10.1007/BF00353765[published Online First: Epub Date]|.
37. Grebe SK, Hay ID. Thyroid cancer nodal metastases: biologic significance and therapeutic considerations. Surg Oncol Clin N Am 1996;**5**(1):43-63

38. Nam-Goong IS, Kim HY, Gong G, et al. Ultrasonography-guided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. Clin Endocrinol (Oxf) 2004;**60**(1):21-8 doi: 10.1046/j.1365-2265.2003.01912.x[published Online First: Epub Date]|.

39. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 2006;**6**:9 doi: 10.1186/1471-2288-6-9[published Online First: Epub Date]|.

40. Hay ID, Grant CS, van Heerden JA, et al. Papillary thyroid microcarcinoma: a study of 535 cases observed in a 50-year period. Surgery 1992;**112**(6):1139-46; discussion 46-7

41. Qubain SW, Nakano S, Baba M, et al. Distribution of lymph node micrometastasis in pN0 well-differentiated thyroid carcinoma. Surgery 2002;**131**(3):249-56 doi: 10.1067/msy.2002.120657[published Online First: Epub Date]|.

42. Arturi F, Russo D, Giuffrida D, et al. Early diagnosis by genetic analysis of differentiated thyroid cancer metastases in small lymph nodes. J Clin Endocrinol Metab 1997;**82**(5):1638-41 doi: 10.1210/jcem.82.5.4062[published Online First: Epub Date]|.

43. Kouvaraki MA, Shapiro SE, Fornage BD, et al. Role of preoperative ultrasonography in the surgical management of patients with thyroid cancer. Surgery 2003;**134**(6):946-54; discussion 54-5 doi: 10.1016/s0039-6060(03)00424-0[published Online First: Epub Date]|.

44. O'Connell K, Yen TW, Quiroz F, et al. The utility of routine preoperative cervical ultrasonography in patients undergoing thyroidectomy for differentiated thyroid

cancer. Surgery 2013;**154**(4):697-701; discussion 01-3 doi: 10.1016/j.surg.2013.06.040[published Online First: Epub Date]|.

45. Yeh MW, Bauer AJ, Bernet VA, et al. American Thyroid Association statement on preoperative imaging for thyroid cancer surgery. Thyroid 2015;**25**(1):3-14 doi: 10.1089/thy.2014.0096[published Online First: Epub Date]|.
46. Kuna SK, Bracic I, Tesic V, et al. Ultrasonographic differentiation of benign from malignant neck lymphadenopathy in thyroid cancer. J Ultrasound Med 2006;**25**(12):1531-7; quiz 38-40 doi: 10.7863/jum.2006.25.12.1531[published Online First: Epub Date]|.
47. Park JH, Lee YS, Kim BW, et al. Skip lateral neck node metastases in papillary thyroid carcinoma. World J Surg 2012;**36**(4):743-7 doi: 10.1007/s00268-012-1476-5[published Online First: Epub Date]|.
48. Zhao Y, Li X, Li L, et al. Preliminary study on the diagnostic value of single-source dual-energy CT in diagnosing cervical lymph node metastasis of thyroid carcinoma. J Thorac Dis 2017;**9**(11):4758-66 doi: 10.21037/jtd.2017.09.151[published Online First: Epub Date]|.
49. He M, Lin C, Yin L, et al. Value of Dual-Energy Computed Tomography for Diagnosing Cervical Lymph Node Metastasis in Patients With Papillary Thyroid Cancer. J Comput Assist Tomogr 2019;**43**(6):970-75 doi: 10.1097/RCT.0000000000000927[published Online First: Epub Date]|.
50. Li L, Cheng SN, Zhao YF, et al. Diagnostic accuracy of single-source dual-energy computed tomography and ultrasonography for detection of lateral cervical lymph

node metastases of papillary thyroid carcinoma. J Thorac Dis 2019;**11**(12):5032-41 doi: 10.21037/jtd.2019.12.45[published Online First: Epub Date]|.

Figure legend

Figure 1. Flowchart of the literature search process

Figure 2. Forest plots for the sensitivities and specificities of US, CT, and combination in neck level-based analysis

Figure 3. SROC of US, CT, and combination in detecting CLNM in patients with PTC in neck level-based analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer.

Figure 4. Forest plots for the sensitivities and specificities of US, CT, and combination in patient-based analysis

Figure 5. SROC of US, CT, and combination in detecting CLNM in patients with PTC in patient-based analysis

SROC, summary receiver operating characteristic; CLNM, cervical lymph node metastasis; PTC, papillary thyroid cancer

Table 1. Characteristics of included studies

Study	Country	Study design	Timing of imaging	Duration of patient recruitment	Sample (n) (males/ females)	Age (range)	No. of lymph nodes	Diagnosis	Analysis methods
Jeong HS 2006[28]	Korea	Retrospective	Preoperative	July 2004-March 2005	26 (7/19)	44 (17-73)	31	All PTC	L
Kim E 2008[16]	Korea	Retrospective	Preoperative	April 2006–October 2006	165 (25/140)	48 (16-78)	27	All PTC	L+P
Ahn JE 2008[13]	Korea	Retrospective	Preoperative	January 2005–December 2005	37 (7/30)	47(20-68)	18	All PTC	L
Choi JS 2009[14]	Korea	Retrospective	Preoperative	February 2006–April 2007	299 (44/255)	45 (20-74)	352	All PTC	L
Choi YJ 2010[29]	Korea	Retrospective	Preoperative	January 2007–December 2008	589 (121/468)	46	589	All PTC	P
Morita S 2010[30]	Japan	Prospective	Preoperative	January 2007–December 2009	74 (12/62)	66 (16-84)	349	All PTC	L

Yoon 2011[31]	JH	Korea	Retrospective	Preoperative	February 2007-December 2007	113 (16/97)	46	(15-83)	121	All PTC	L
Seo 2012[32]	YL	Korea	Retrospective	Postoperative	August 2008–August 2011	20 (4:16)	49.8		107	19 PTC, 1 MTC	L
Lee 2013[17]	DW	Korea	Retrospective	Preoperative	January 2007–May 2010	252 (45/207)	49	(15-82)	552	All PTC	L+P
Lesnik 2014[33]	D	USA	Prospective	Preoperative	2003–2008	95 (NA)	NA		198	All PTC	L
Na 2015[19]	DK	Korea	Retrospective	Preoperative	March 2011–February 2012	176 (44/132)	43	(23-74)	352	All PTC	P
Kim 2017[34]	SK	Korea	Retrospective	Preoperative	January 1997–June 2015	3668 (NA)	NA		6557	All PTC	L
Eun 2018[34]	NL	Korea	Retrospective	Preoperative	January 2013–December 2015	302 (76:226)	44		308	All PTC	L
Lee 2018[18]	Y	Korea	Prospective	Preoperative	November 2011–December	351 (78:273)	47.1		801	All PTC	L

2012

NA not available; PTC papillary thyroid carcinoma; MTC medullary thyroid carcinoma; L, neck level-based analysis; P, Patient-based analysis

Table 2. Pooled estimates of sensitivity, specificity, diagnostic OR, LR+, LR-

	Studies,n	Sensitivity (95% CI)	Specificity (95% CI)	Diagnostic OR (95% CI)	LR+ (95% CI)	LR- (95% CI)
Diagnostic accuracy of CT or US on neck level						
All neck level						
US	11	0.35 (0.34-0.37, 97.5)	0.95 (0.94-0.95, 90.8)	13.94 (9.34-20.82, 81.0)	6.79 (4.79-9.63, 84.1)	0.50 (0.41-0.60, 95.2)
CT	11	0.46 (0.44-0.47, 97.6)	0.88 (0.87-0.89, 97.9)	7.24 (5.46-9.62, 72.2)	3.77 (2.08-6.84, 98.2)	0.52 (0.45-0.61,89.7)
US/CT	5	0.51 (0.49-0.52, 92.8)	0.85 (0.84-0.86, 97.8)	6.01 (3.84-9.40, 89.2)	3.04 (1.93-4.80, 96.2)	0.52 (0.45-0.60, 78.8)
Central neck level						
US	4	0.28 (0.24-0.32, 94.3)	0.97 (0.96-0.98, 53.0)	14.07 (6.66-29.71, 53.1)	14.07 (6.66-29.71, 53.1)	14.07 (6.66-29.71, 53.1)
CT	4	0.32 (0.28-0.36, 88.2)	0.89 (0.86-0.91, 84.6)	5.48 (2.15-13.98, 84.3)	3.71 (1.79-7.66, 83.9)	0.74 (0.62-0.89, 86.7)
US/CT	3	0.40 (0.35-0.45, 82.3)	0.85 (0.82-0.88, 37.0)	4.32 (2.09-8.92, 81.0)	2.85 (1.75-4.65, 77.5)	0.67 (0.52-0.86, 83.9)
Lateral neck level						
US	6	0.74 (0.69-0.78, 78.1)	0.92 (0.90-0.94, 89.6)	24.41 (11.16 -53.42, 71.9)	6.67 (2.91-15.30, 91.5)	0.35 (0.28-0.43, 30.1)
CT	6	0.73 (0.68-0.77, 90.7)	0.89 (0.87-0.91, 91.8)	15.55 (7.98 -30.32, 64.6)	5.54 (2.95-10.39, 85.9)	0.35 (0.21 -0.59, 91.4)
US/CT	4	0.88 (0.83-0.91, 64.1)	0.79 (0.73-0.84, 89.6)	22.59 (11.29 -45.19, 46.6)	3.31 (1.53-7.17, 91.1)	0.19 (0.14-0.25, 0)
Diagnostic accuracy of CT or US on patient level						
All patient level						

US	4	0.41 (0.36-0.46, 95.5)	0.92 (0.89-0.94, 0)	7.56 (4.08-14.01, 51.3)	4.48 (3.31 -6.05, 0)	0.65 (0.53-0.80, 75.5)
CT	4	0.49 (0.44-0.54, 95.9)	0.91 (0.89-0.94, 66.8)	9.40 (5.79-15.27, 34.4)	4.84 (3.66-6.39, 0)	0.53 (0.37-0.75, 85.6)
US/CT	2	0.64 (0.57-0.71, 0)	0.83 (0.77-0.88, 0)	8.59 (5.37-13.76, 0)	3.71 (2.72 -5.08, 0)	0.43 (0.36-0.53, 0)
Central patient level						
US	2	0.21 (0.16-0.28, 24.8)	0.95 (0.91-0.97, 0)	4.53 (2.34-8.77, 0)	3.78 (2.08-6.86, 0)	0.84 (0.76 -0.93, 36.4)
CT	2	0.38 (0.32-0.46, 92.6)	0.90 (0.85-0.93, 87.2)	5.02 (0.46-54.54, 94.4)	3.52 (0.52-23.84, 93	0.71 (0.41-1.22, 95.3)
US/CT	2	0.47 (0.39-0.54, 80.8)	0.85 (0.80-0.89, 0)	4.88 (2.58-9.23, 46.4)	3.14 (2.23-4.41, 0)	0.64 (0.47-0.86, 78.4)
Lateral patient level						
US	2	0.87 (0.74-0.95, 88.2)	0.89 (0.83-0.93, 91.4)	20.11 (6.77-59.70, 0)	3.58 (0.85-15.16, 9	0.22 (0.05-1.08, 58.0)
CT	2	0.92 (0.81-0.98, 88.6)	0.88 (0.83-0.93, 96.8)	36.88 (11.40 -119.35, 0)	3.44 (0.29-40.77, 92.3)	0.23 (0.10-0.52, 0)
US/CT	2	0.98 (0.89-0.99, 58.9)	0.92 (0.87-0.96, 97.5)	78.10 (2.82 -2160.4, 65.2)	5.30 (0.15-186.11, 88.6	0.08 (0.02-0.41, 0)

Results are presented as n (95% CI, I2 %); OR, odds ratio; LR+, positive likelihood ratio; LR-, negative likelihood ratio

Supplementary materials

sFigure 1. Summary of risk of bias and applicability concerns

sFigure 2. Risk of bias and applicability concerns graph

sFigure 3. Forest plots for the sensitivities and specificities of US, CT, and combination in central neck level analysis

sFigure 4. SROC of US, CT, and combination in detecting CLNM in patients with PTC in central neck level analysis

sFigure 5. Forest plots for the sensitivities and specificities of US, CT, and combination in lateral neck level analysis

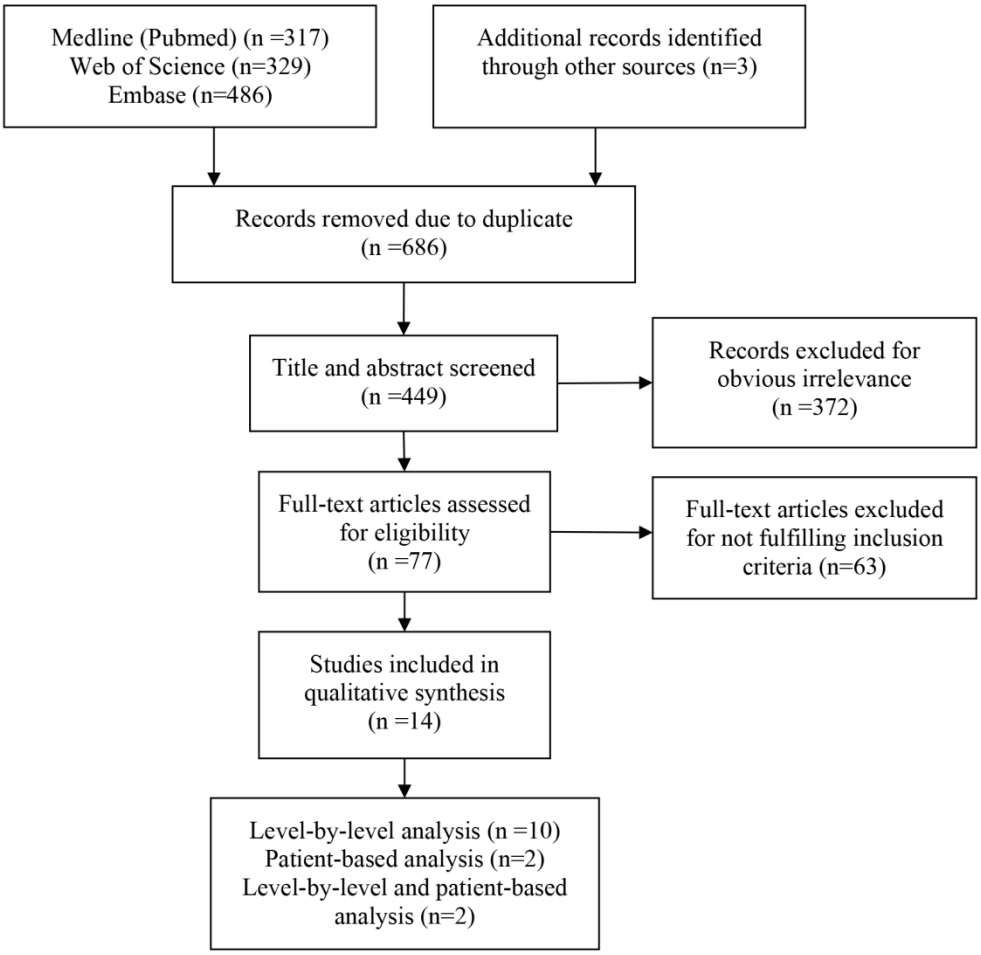
sFigure 6. SROC of US, CT, and combination in detecting CLNM in patients with PTC in lateral neck level analysis

sFigure 7. Forest plots for the sensitivities and specificities of US, CT, and combination in central patient level analysis

sFigure 8. SROC of US, CT, and combination in detecting CLNM in patients with PTC in central patient level analysis

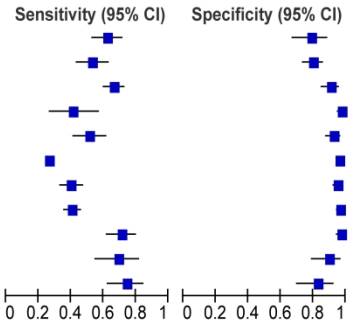
sFigure 9. Forest plots for the sensitivities and specificities of US, CT, and combination in lateral patient level analysis

sFigure 10. SROC of US, CT, and combination in detecting CLNM in patients with PTC in lateral patient level analysis



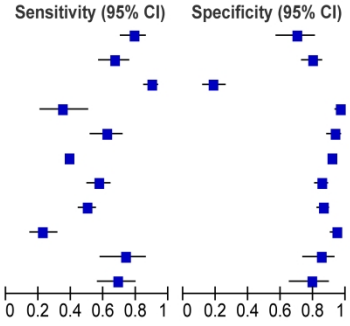
All neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	73	14	44	52	0.62 [0.53, 0.71]	0.79 [0.67, 0.88]
Choi 2009	59	38	52	150	0.53 [0.43, 0.63]	0.80 [0.73, 0.85]
Eun 2018	156	13	79	128	0.66 [0.60, 0.72]	0.91 [0.85, 0.95]
Jeong 2006	19	7	27	259	0.41 [0.27, 0.57]	0.97 [0.95, 0.99]
Kim 2008	54	13	51	159	0.51 [0.41, 0.61]	0.92 [0.87, 0.96]
Kim 2017	841	140	2296	3295	0.27 [0.25, 0.28]	0.96 [0.95, 0.97]
Lee 2013	83	18	124	333	0.40 [0.33, 0.47]	0.95 [0.92, 0.97]
Lee 2018	156	15	227	403	0.41 [0.36, 0.46]	0.96 [0.94, 0.98]
Morita 2010	79	7	32	231	0.71 [0.62, 0.79]	0.97 [0.94, 0.99]
Seo 2012	36	6	16	52	0.69 [0.55, 0.81]	0.90 [0.79, 0.96]
Yoon 2011	52	9	18	43	0.74 [0.62, 0.84]	0.83 [0.70, 0.92]



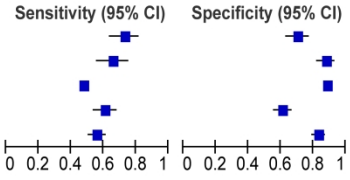
All Neck level:CT

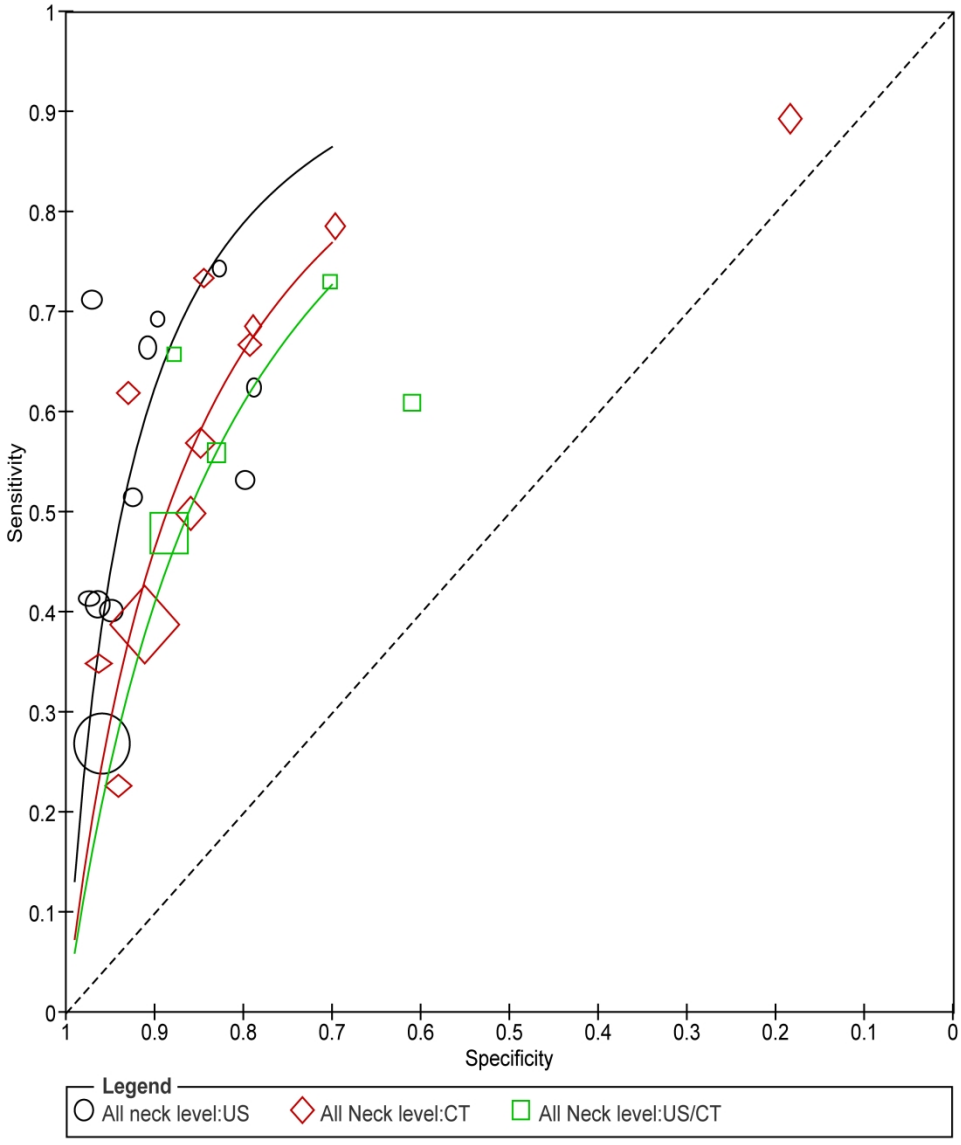
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	99	20	27	46	0.79 [0.70, 0.85]	0.70 [0.57, 0.80]
Choi 2009	74	39	37	149	0.67 [0.57, 0.75]	0.79 [0.73, 0.85]
Eun 2018	210	115	25	26	0.89 [0.85, 0.93]	0.18 [0.12, 0.26]
Jeong 2006	16	10	30	256	0.35 [0.21, 0.50]	0.96 [0.93, 0.98]
Kim 2008	65	12	40	160	0.62 [0.52, 0.71]	0.93 [0.88, 0.96]
Kim 2017	1221	306	1921	3129	0.39 [0.37, 0.41]	0.91 [0.90, 0.92]
Lee 2013	118	53	89	298	0.57 [0.50, 0.64]	0.85 [0.81, 0.88]
Lee 2018	191	59	192	359	0.50 [0.45, 0.55]	0.86 [0.82, 0.89]
Morita 2010	25	14	86	224	0.23 [0.15, 0.31]	0.94 [0.90, 0.97]
Seo 2012	33	10	12	55	0.73 [0.58, 0.85]	0.85 [0.74, 0.92]
Yoon 2011	48	11	22	41	0.69 [0.56, 0.79]	0.79 [0.65, 0.89]



All Neck level:US/CT

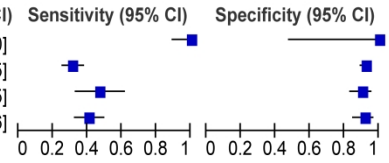
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	81	56	30	132	0.73 [0.64, 0.81]	0.70 [0.63, 0.77]
Kim 2008	69	21	36	151	0.66 [0.56, 0.75]	0.88 [0.82, 0.92]
Kim 2017	1503	400	1639	3035	0.48 [0.46, 0.50]	0.88 [0.87, 0.89]
Lee 2013	126	137	81	214	0.61 [0.54, 0.68]	0.61 [0.56, 0.66]
Lee 2018	214	71	169	347	0.56 [0.51, 0.61]	0.83 [0.79, 0.86]





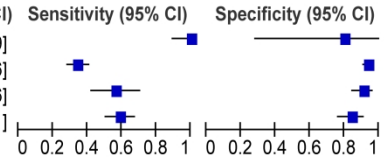
All patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	0	0	5	1.00 [0.89, 1.00]	1.00 [0.48, 1.00]
Choi 2010	75	27	163	324	0.32 [0.26, 0.38]	0.92 [0.89, 0.95]
Kim 2008	25	11	28	101	0.47 [0.33, 0.61]	0.90 [0.83, 0.95]
Lee 2013	59	9	85	99	0.41 [0.33, 0.49]	0.92 [0.85, 0.96]



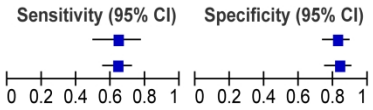
All patient level:CT

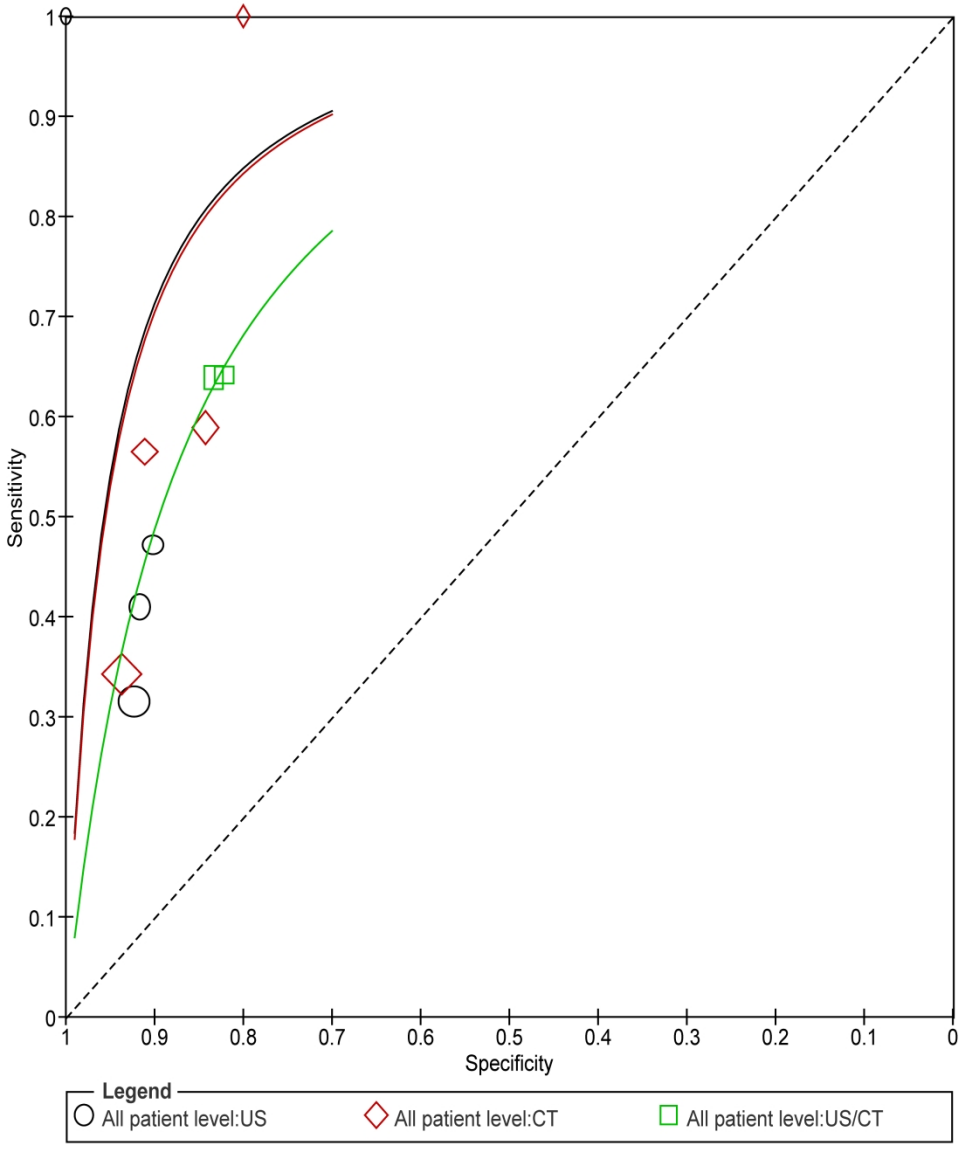
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Ahn 2008	32	1	0	4	1.00 [0.89, 1.00]	0.80 [0.28, 0.99]
Choi 2010	82	22	156	329	0.34 [0.28, 0.41]	0.94 [0.91, 0.96]
Kim 2008	30	10	23	102	0.57 [0.42, 0.70]	0.91 [0.84, 0.96]
Lee 2013	85	17	59	91	0.59 [0.51, 0.67]	0.84 [0.76, 0.91]

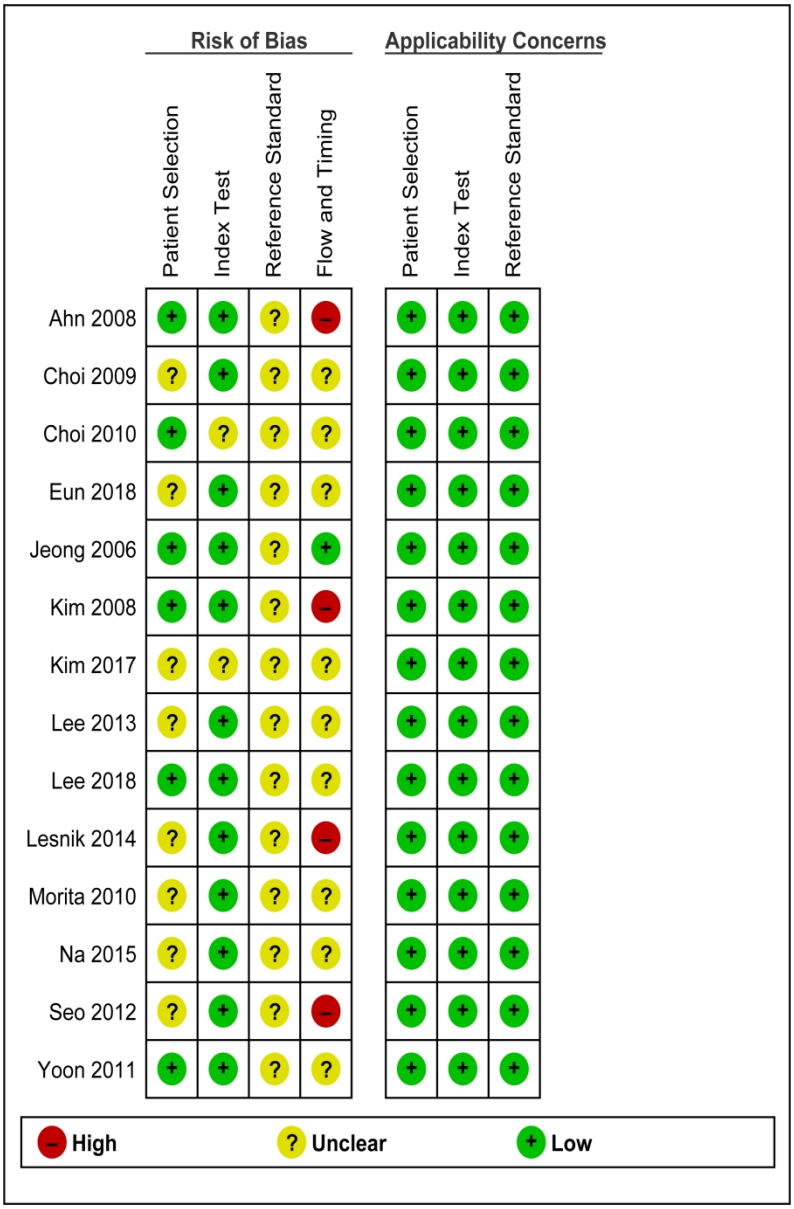


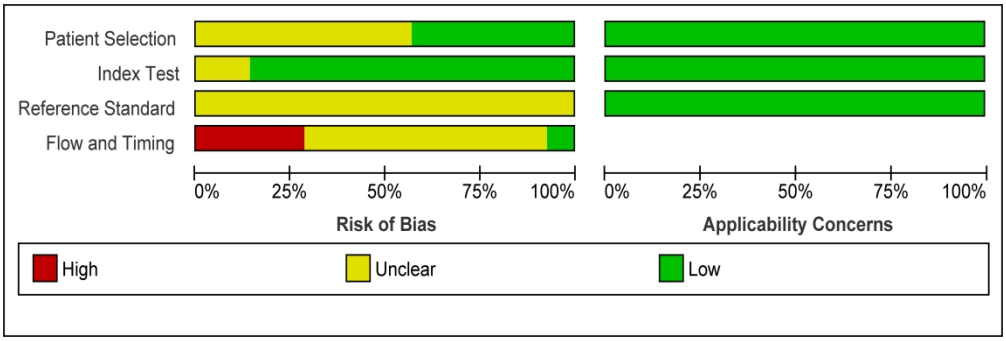
All patient level:US/CT

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	34	20	19	92	0.64 [0.50, 0.77]	0.82 [0.74, 0.89]
Lee 2013	92	18	52	90	0.64 [0.55, 0.72]	0.83 [0.75, 0.90]



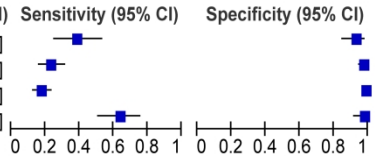






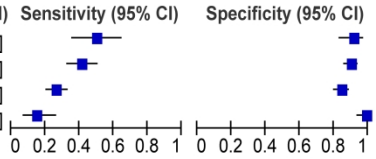
Central neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	20	6	32	75	0.38 [0.25, 0.53]	0.93 [0.85, 0.97]
Lee 2013	31	8	102	269	0.23 [0.16, 0.31]	0.97 [0.94, 0.99]
Lee 2018	39	5	180	307	0.18 [0.13, 0.24]	0.98 [0.96, 0.99]
Morita 2010	42	2	24	80	0.64 [0.51, 0.75]	0.98 [0.91, 1.00]



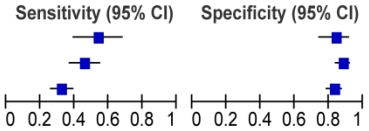
Central neck level:CT

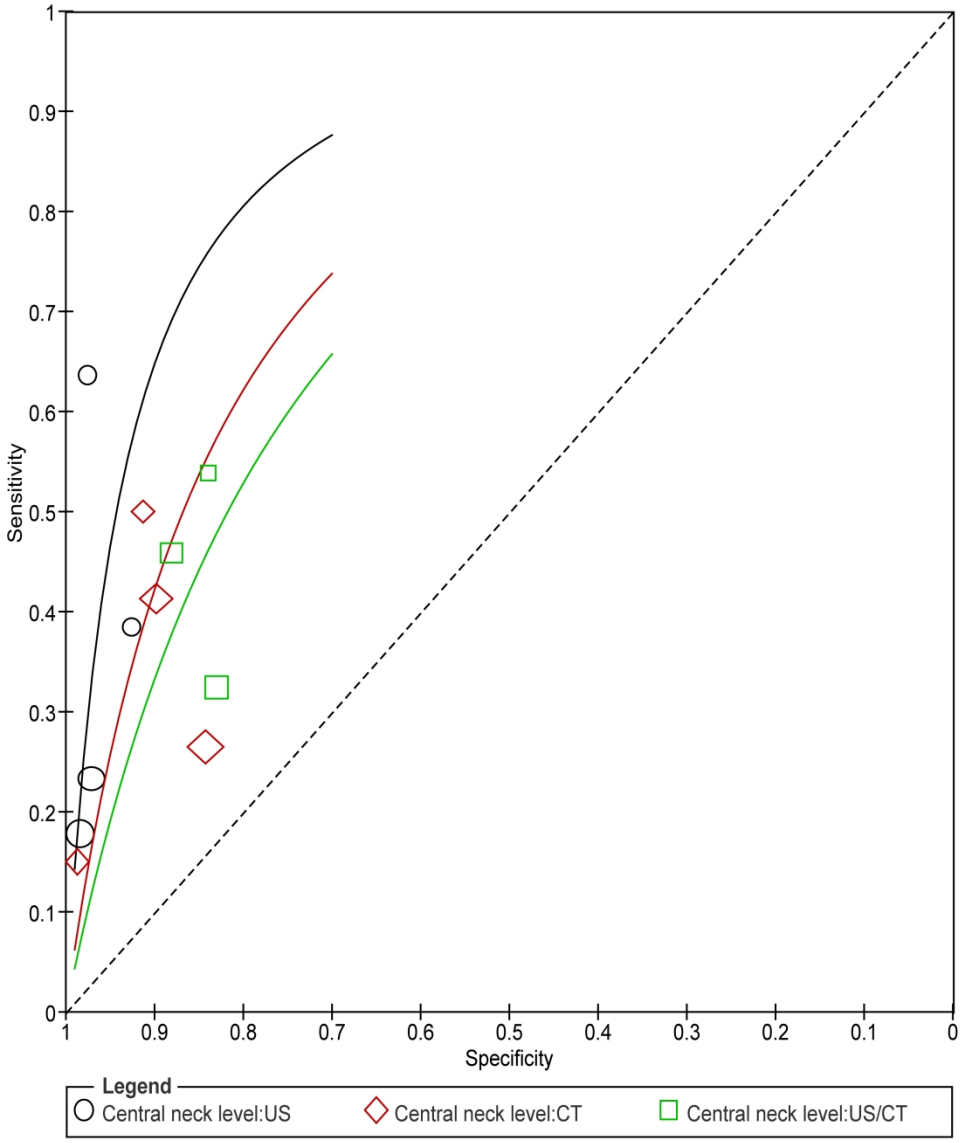
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	26	7	26	74	0.50 [0.36, 0.64]	0.91 [0.83, 0.96]
Lee 2013	55	28	78	249	0.41 [0.33, 0.50]	0.90 [0.86, 0.93]
Lee 2018	58	49	161	263	0.26 [0.21, 0.33]	0.84 [0.80, 0.88]
Morita 2010	10	1	56	81	0.15 [0.08, 0.26]	0.99 [0.93, 1.00]



Central neck level:US/CT

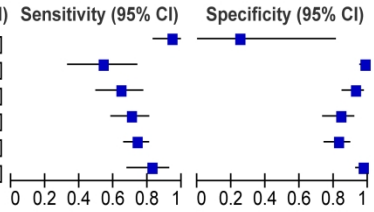
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kim 2008	28	13	24	68	0.54 [0.39, 0.68]	0.84 [0.74, 0.91]
Lee 2013	61	33	72	244	0.46 [0.37, 0.55]	0.88 [0.84, 0.92]
Lee 2018	71	53	148	259	0.32 [0.26, 0.39]	0.83 [0.78, 0.87]





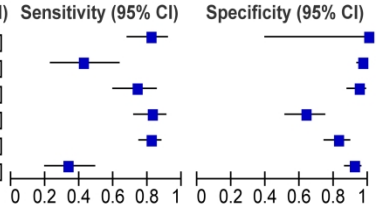
Lateral neck level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	46	3	3	1	0.94 [0.83, 0.99]	0.25 [0.01, 0.81]
Jeong 2006	14	5	12	229	0.54 [0.33, 0.73]	0.98 [0.95, 0.99]
Kim 2008	34	7	19	84	0.64 [0.50, 0.77]	0.92 [0.85, 0.97]
Lee 2013	52	12	22	62	0.70 [0.59, 0.80]	0.84 [0.73, 0.91]
Lee 2018	117	20	42	94	0.74 [0.66, 0.80]	0.82 [0.74, 0.89]
Morita 2010	37	5	8	151	0.82 [0.68, 0.92]	0.97 [0.93, 0.99]



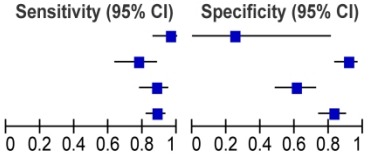
Lateral neck level:CT

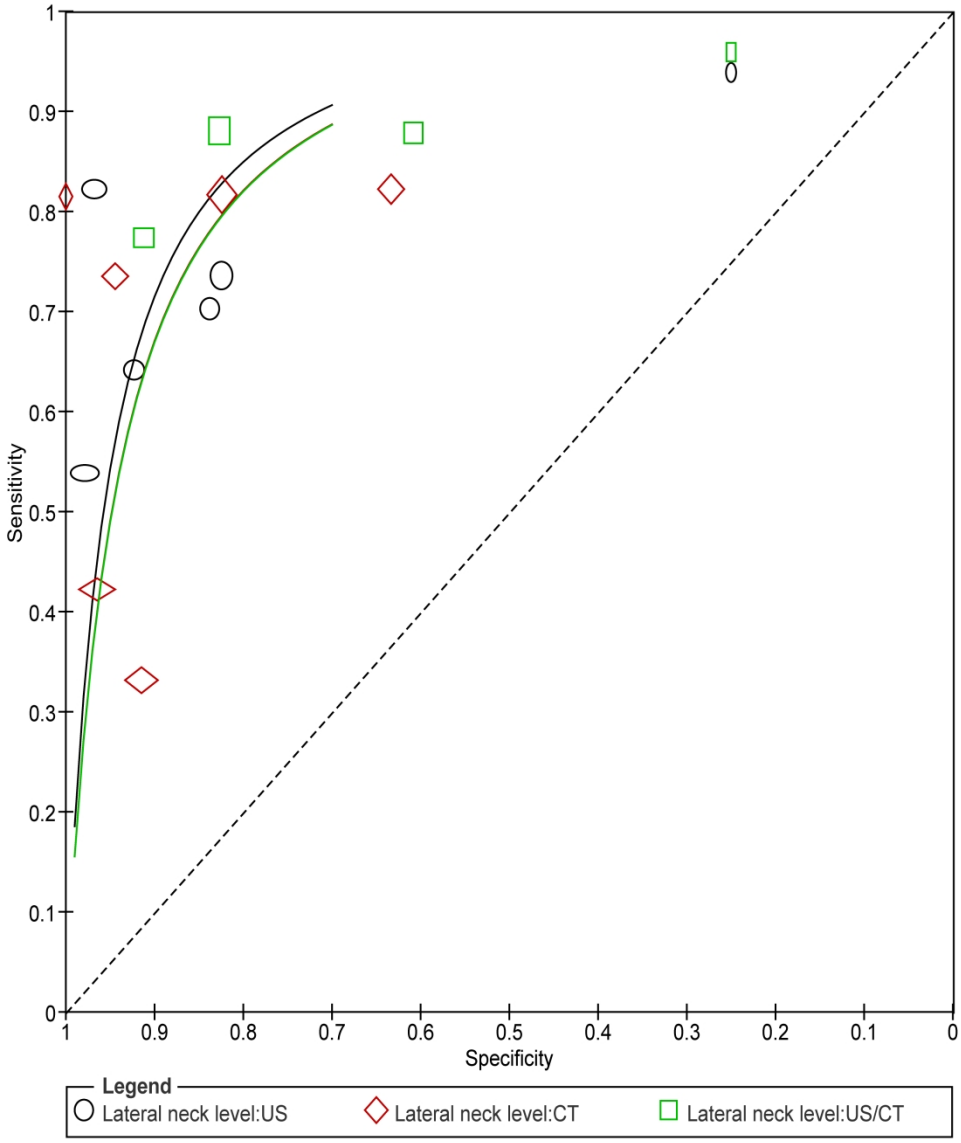
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	40	0	9	4	0.82 [0.68, 0.91]	1.00 [0.40, 1.00]
Jeong 2006	11	8	15	226	0.42 [0.23, 0.63]	0.97 [0.93, 0.99]
Kim 2008	39	5	14	86	0.74 [0.60, 0.85]	0.95 [0.88, 0.98]
Lee 2013	61	27	13	47	0.82 [0.72, 0.90]	0.64 [0.52, 0.74]
Lee 2018	130	20	29	94	0.82 [0.75, 0.87]	0.82 [0.74, 0.89]
Morita 2010	15	13	30	143	0.33 [0.20, 0.49]	0.92 [0.86, 0.95]



Lateral neck level:US/CT

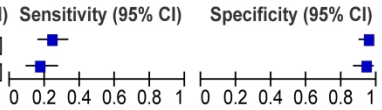
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Choi 2009	47	3	2	1	0.96 [0.86, 1.00]	0.25 [0.01, 0.81]
Kim 2008	41	8	12	83	0.77 [0.64, 0.88]	0.91 [0.83, 0.96]
Lee 2013	65	29	9	45	0.88 [0.78, 0.94]	0.61 [0.49, 0.72]
Lee 2018	140	18	19	86	0.88 [0.82, 0.93]	0.83 [0.74, 0.89]





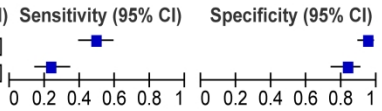
Central patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	27	7	85	133	0.24 [0.17, 0.33]	0.95 [0.90, 0.98]
Na 2015	14	6	67	89	0.17 [0.10, 0.27]	0.94 [0.87, 0.98]



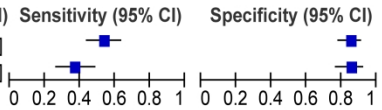
Central patient level:CT

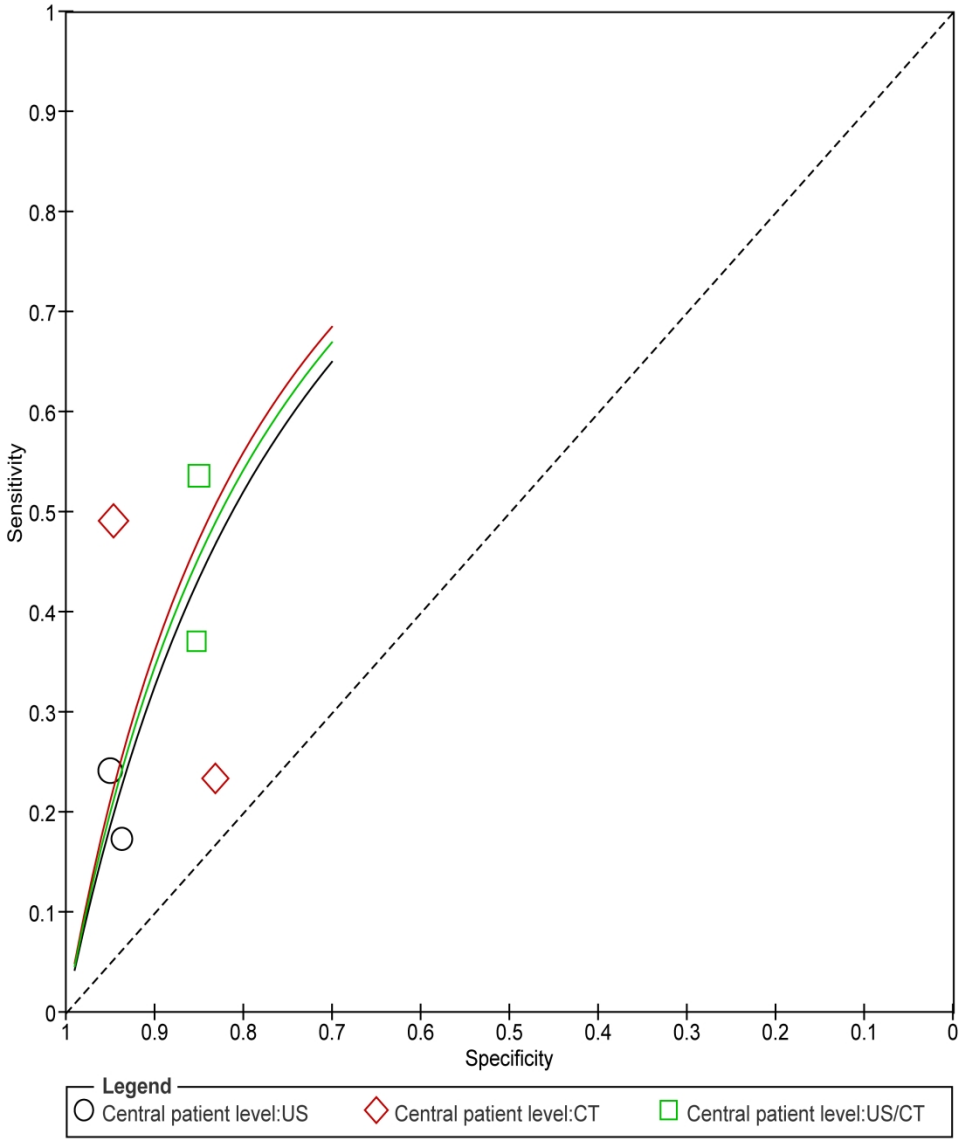
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	55	7	57	123	0.49 [0.40, 0.59]	0.95 [0.89, 0.98]
Na 2015	19	16	62	79	0.23 [0.15, 0.34]	0.83 [0.74, 0.90]



Central patient level:US/CT

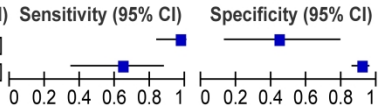
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	60	21	52	119	0.54 [0.44, 0.63]	0.85 [0.78, 0.90]
Na 2015	30	14	51	81	0.37 [0.27, 0.48]	0.85 [0.77, 0.92]





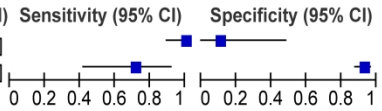
Lateral patient level:US

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	31	5	1	4	0.97 [0.84, 1.00]	0.44 [0.14, 0.79]
Na 2015	9	14	5	148	0.64 [0.35, 0.87]	0.91 [0.86, 0.95]



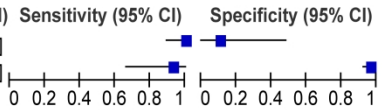
Lateral patient level:CT

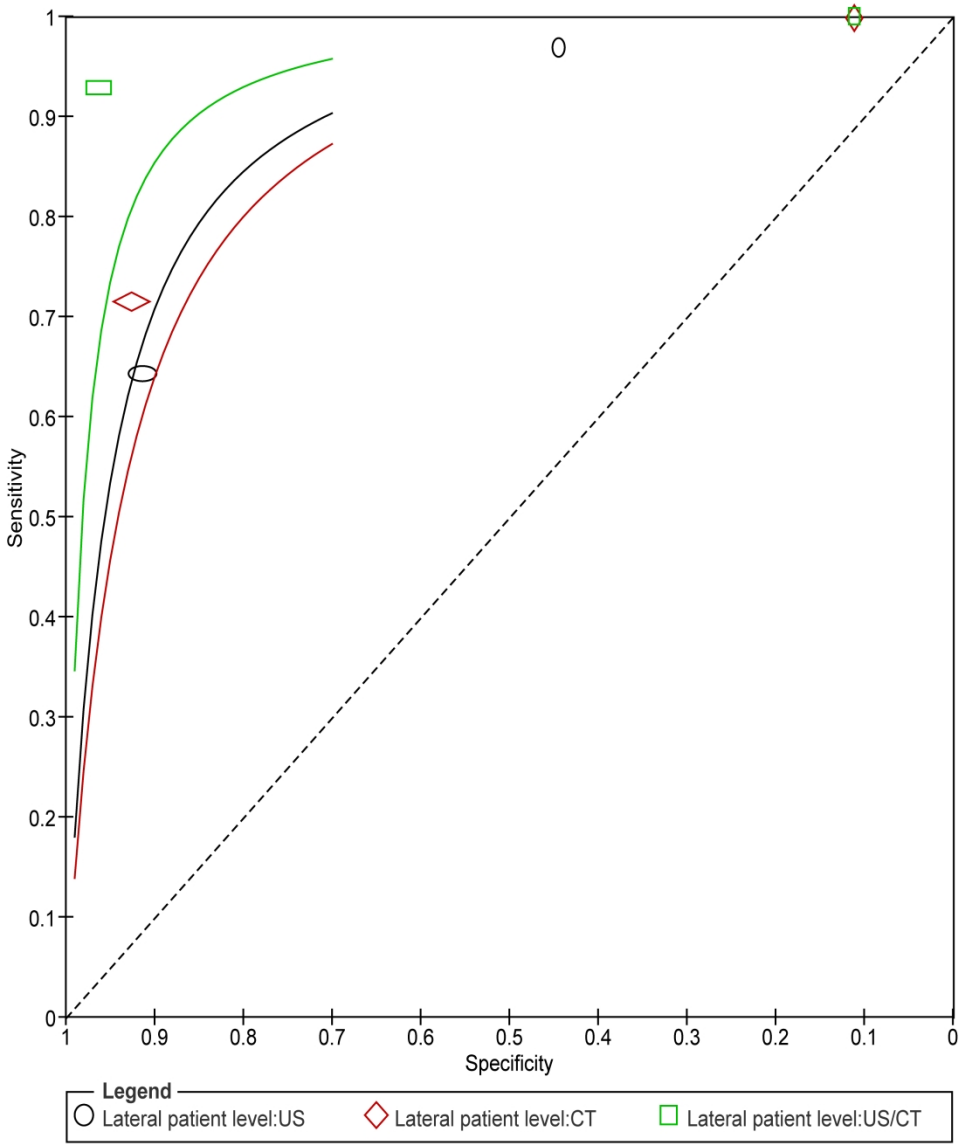
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	32	8	0	1	1.00 [0.89, 1.00]	0.11 [0.00, 0.48]
Na 2015	10	12	4	150	0.71 [0.42, 0.92]	0.93 [0.87, 0.96]



Lateral patient level:US/CT

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Lee 2013	32	8	0	1	1.00 [0.89, 1.00]	0.11 [0.00, 0.48]
Na 2015	13	6	1	156	0.93 [0.66, 1.00]	0.96 [0.92, 0.99]





Supplementary Table S1. Search strategy

PubMed		Search strategy	Numbers
P	**#1**	((((thyroid cancer) OR (thyroid carcinoma)) OR (thyroid tumor)) OR (papillary thyroid cancer)) OR (thyroid neoplasm)	94,066
	#2	(cervical lymph node) OR (neck lymph node)	38,156
	#3	#1 AND #2	5436
	#4	(metastasis) OR (metastatic)	1,436,116
	#5	#3 AND #4	4005
Intervention	**#6**	((ultrasonography) OR (ultrasound)) OR (US)	2,497,831
	#7	(computed tomography) OR (CT)	876,390
	#8	#6 AND #7	449,815
P+I	**#9**	#5 AND #8	317

Embase		Search strategy	Numbers
P	**#1**	'thyroid cancer' OR 'thyroid carcinoma' OR 'thyroid tumor' OR 'papillary thyroid cancer' OR 'thyroid neoplasm'	88,150

	#2	'cervical lymph node' OR 'neck lymph node'	18,642
	#3	#1 AND #2	3260
	#4	'metastasis' OR 'metastatic'	954,603
	#5	#3 AND #4	2772
I	**#6**	'ultrasonography' OR 'ultrasound' OR 'us'	1,451,034
	#7	'computed tomography' OR 'ct'	1,235,000
	#8	#6 AND #7	133,587
P+I	#9	#5 AND #8	486

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5-8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5-8
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5-8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5-8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	5-8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5-8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5-8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8-11
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	8-11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8-11
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	8-11
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11-13
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	None

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.