Diagnostic accuracy of ultrasound, CT 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 Qiao1 2

ABSTRACT

Objectives To determine the diagnostic accuracy of ultrasound (US), 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 5 December 2021 that used US and CT to detect CLNM in patients with PTC. The primary outcomes were sensitivity, specificity and diagnostic ORs in neck-level-based (lymph nodes are analysed by neck level) or patient-based (lymph nodes are analysed by patient) analysis. Secondary outcomes were sensitivity, specificity and DORs in the central and lateral compartments.

Results Fourteen studies (1617 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% CI 0.34 to 0.37), 0.95 (95% CI 0.94 to 0.95) and 13.94 (95% CI 9.34 to 20.82) for US, were 0.46 (95% CI 0.44 to 0.47), 0.88 (95% CI 0.87 to 0.89) and 7.56 (95% CI 4.08 to 14.01) for CT, and were 0.47 (95% CI 0.45 to 0.50), 0.93 (95% CI 0.92 to 0.94) and 6.01 (95% CI 3.84 to 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 to 0.46), 0.92 (95% CI 0.89 to 0.94) and 7.36 (95% CI 4.08 to 14.01) for US, were 0.49 (0.44 to 0.54), 0.91 (0.89 to 0.94), 9.40 (5.79 to 15.27) for CT, and were 0.64 (95% CI 0.57 to 0.71), 0.83 (95% CI 0.77 to 0.88), 8.59 (95% CI 5.37 to 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.

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 dissection (ipsilateral or bilateral) is recommended by the American Thyroid Association guidelines in patients with clinically positive central nodes, especially for those with advanced primary tumours, 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. 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. 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 CT with intravenous contrast is complementary to US in patients with advanced disease. 

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, some studies suggested a superior diagnostic performance of the combination of preoperative US with CT over US alone.

Some meta-analyses examining the diagnostic accuracy of US and CT in detecting CLNM in patients with PTC have been previously conducted. 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 minimise 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 ORs 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 analysed by neck level) and patient-based (lymph nodes are analysed 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. A systematic search of Medline (via PubMed), Web of Science and Embase was conducted to identify studies published up to 5 December 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”

Figure 1 Flow chart of the literature search process.
## Table 1  Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Study design</th>
<th>Timing of imaging</th>
<th>Duration of patient recruitment</th>
<th>Sample (n) (males/ females)</th>
<th>Age (range)</th>
<th>No of lymph node</th>
<th>Diagnosis</th>
<th>Analysis methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeong et al 2006</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>July 2004-March 2005</td>
<td>26 (7/19)</td>
<td>44 (17–73)</td>
<td>312</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Kim et al 2006</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>April 2006–October 2006</td>
<td>165 (25/140)</td>
<td>48 (16–78)</td>
<td>277</td>
<td>All PTC</td>
<td>L+P</td>
</tr>
<tr>
<td>Choi et al 2009</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>February 2006–April 2007</td>
<td>299 (44/255)</td>
<td>45 (20–74)</td>
<td>352</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Choi et al 2010</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>January 2007–December 2008</td>
<td>589 (121/468)</td>
<td>46</td>
<td>589</td>
<td>All PTC</td>
<td>P</td>
</tr>
<tr>
<td>Morita et al 2010</td>
<td>Japan</td>
<td>Prospective</td>
<td>Preoperative</td>
<td>January 2007–December 2009</td>
<td>74 (12/62)</td>
<td>66 (16–84)</td>
<td>349</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Yoon et al 2011</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>February 2007–December 2007</td>
<td>113 (16/97)</td>
<td>46 (15–83)</td>
<td>122</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Seo et al 2012</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Postoperative</td>
<td>August 2008–August 2011</td>
<td>20 (4:16)</td>
<td>49.8</td>
<td>107</td>
<td>19 PTC, 1 MTC</td>
<td>L</td>
</tr>
<tr>
<td>Lee et al 2013</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>January 2007–May 2010</td>
<td>252 (45/207)</td>
<td>49 (15–82)</td>
<td>558</td>
<td>All PTC</td>
<td>L+P</td>
</tr>
<tr>
<td>Lesnik et al 2014</td>
<td>USA</td>
<td>Prospective</td>
<td>Preoperative</td>
<td>2003–2008</td>
<td>95 (NA)</td>
<td>NA</td>
<td>196</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Na et al 2015</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>March 2011–February 2012</td>
<td>176 (44/132)</td>
<td>43 (23–74)</td>
<td>352</td>
<td>All PTC</td>
<td>P</td>
</tr>
<tr>
<td>Kim et al 2017</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>January 1997–June 2015</td>
<td>3668 (NA)</td>
<td>NA</td>
<td>6557</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Eun et al 2018</td>
<td>Korea</td>
<td>Retrospective</td>
<td>Preoperative</td>
<td>January 2013–December 2015</td>
<td>302 (76:226)</td>
<td>44</td>
<td>308</td>
<td>All PTC</td>
<td>L</td>
</tr>
<tr>
<td>Lee et al 2018</td>
<td>Korea</td>
<td>Prospective</td>
<td>Preoperative</td>
<td>November 2011–December 2012</td>
<td>351 (78:273)</td>
<td>47.1</td>
<td>801</td>
<td>All PTC</td>
<td>L</td>
</tr>
</tbody>
</table>

L, neck-level-based analysis; MTC, medullary thyroid carcinoma; NA, not available; P, patient-based analysis; PTC, papillary thyroid carcinoma.
or “metastatic”) AND (“ultrasonography” or “ultrasound” or “US”) AND (computed tomography” or “CT”) (online supplemental 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 (1) 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; (2) studies with >10 patients; (3) studies with a reference standard of histopathology or cytology (the diagnostic gold standard was the pathological diagnosis of the resected lymph nodes); (4) 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; (5) studies published in English.

The exclusion criteria were (1) case reports, case series, review articles, pictorial essays, letters to editors, unpublished data, conference abstracts and proceedings on the topic of interest; (2) studies that used only US or only CT; (3) insufficient data regarding TP, FP, FN and TN; (4) duplicate publications using the same databases and studies; (5) 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; (6) 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 (Diagnostic Odds Ratios) 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. 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 OR, positive likelihood ratio (LR+), and negative LR (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 (I²). Heterogeneity in the included articles was defined as small I² <25%, moderate I² 25%–50%, and obvious I² >50%. If heterogeneity was detected (p<0.10 or I² ≥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, and forest plots were created. The pooling of sensitivity, specificity, diagnostic OR, LR+ and LR− was performed using the Meta-Disc software (V1.4, Madrid, Spain). Forest plots and summary receiver operating characteristic curves were obtained using RevMan V.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.15 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 3 were prospective studies; 13 studies were performed prospectively, and 1 study was performed postoperatively. Twelve, one and one were conducted in Korea, the USA and Japan, respectively (table 1). The studies included in this
meta-analysis were of moderate quality (online supplemental sfigures 1,2).

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

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 to 0.32), 0.97 (95% CI 0.96 to 0.98) and 14.07 (95% CI 6.66 to 29.71) from four studies. For CT, the pooled sensitivity, specificity and DOR were 0.32 (95% CI 0.28 to 0.36), 0.89 (95% CI 0.86 to 0.91) and 5.48 (95% CI 2.15 to 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 to 0.45), 0.85 (95% CI 0.82 to 0.88) and 4.32 (95% CI 2.09 to 8.92) from three studies (table 2, online supplemental sfigures 3,4).

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 to 0.78), 0.92 (95% CI 0.90 to 0.94) and 24.41 (95% CI 11.16 to 53.42) from six studies; the values for CT were 0.73 (95% CI 0.68 to 0.77), 0.89 (95% CI 0.87 to 0.91) and 15.55 (95% CI 7.98 to 30.32) from six studies; the values for the combination of US and CT were 0.88 (95% CI 0.83 to 0.91), 0.79 (95%
CI 0.73 to 0.84) and 22.59 (95% CI 11.29 to 45.19) from four studies (table 2, online supplemental sfigures 5,6).

**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 to 0.46), 0.92 (95% CI 0.89 to 0.94) and 7.56 (95% CI 4.08 to 14.01); the values for CT were 0.49 (95% CI 0.44 to 0.54), 0.91 (95% CI 0.89 to 0.94) and 9.40 (95% CI 5.79 to 15.27); the values for the combination of US and CT were 0.64 (95% CI 0.57 to 0.71), 0.83 (95% CI 0.77 to 0.88) and 8.59 (95% CI 5.37 to 13.76) (figures 4 and 5, table 2).

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 of US were 0.21 (95% CI 0.16 to 0.28), 0.95 (95% CI 0.91 to 0.97) and 4.53 (95% CI 2.34 to 8.77) for US, were 0.38 (95% CI 0.32 to 0.46), 0.90 (95% CI 0.85 to 0.93) and 5.02 (95% CI 0.46 to 54.54) for CT, and were 0.47 (95% CI 0.39 to 0.54), 0.85 (95% CI 0.80 to 0.89) and 4.88 (95% CI 2.58 to 9.23) for the combination of CT and US (table 2, online supplemental sfigures 7,8).

In contrast, the pooled estimates of sensitivity, specificity, and DOR of US were 0.87 (95% CI 0.74 to 0.95), 0.89 (95% CI 0.83 to 0.93) and 20.11 (95% CI 6.77 to 59.70); the values for CT were 0.92 (95% CI 0.81 to 0.98), 0.88 (95% CI 0.83 to 0.93) and 56.88 (95% CI 11.40 to 119.35); the values for the combination of US and CT were 0.98 (95% CI 0.89 to 0.99), 0.92 (95% CI 0.87 to 0.96) and 78.10 (95% CI 2.82 to 2160.4) (table 2, online supplemental sfigures 9,10).

**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. Differ-entiated thyroid carcinoma, particularly PTC, involves CLNM in 20%–50% of the patients, which could prevent small and intrathyroidal primary tumours. 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. 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. The preoperative US identifies lymph node or soft-tissue metastases in up to 39% of patients who had no physical examination and changed the operative management in 23% of patients.

Previous meta-analyses examined CT and US. Suh et al and Cho et al demonstrated the value of CT for CLNM but did not include US. Rajmakers et al only examined the detection of the sentinel lymph node. Wu et al and Zhao et al 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, infracavicular, 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 oesophageal 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.

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

<table>
<thead>
<tr>
<th>Studies, n</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Diagnostic OR (95% CI)</th>
<th>LR+ (95% CI)</th>
<th>LR– (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All neck level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 11</td>
<td>0.35 (0.34 to 0.37, 97.5)</td>
<td>0.95 (0.94 to 0.95, 90.8)</td>
<td>13.94 (9.34 to 20.82, 81.0)</td>
<td>6.79 (4.79 to 9.63, 84.1)</td>
<td>0.50 (0.41 to 0.60, 95.2)</td>
</tr>
<tr>
<td>CT 11</td>
<td>0.46 (0.44 to 0.47, 97.6)</td>
<td>0.88 (0.87 to 0.89, 97.9)</td>
<td>7.24 (5.46 to 9.62, 72.2)</td>
<td>3.77 (2.08 to 6.84, 98.5)</td>
<td>0.52 (0.45 to 0.61, 89.7)</td>
</tr>
<tr>
<td>US/CT 5</td>
<td>0.51 (0.49 to 0.52, 92.8)</td>
<td>0.85 (0.84 to 0.86, 97.8)</td>
<td>6.01 (3.84 to 9.40, 89.2)</td>
<td>3.04 (1.93 to 4.80, 96.3)</td>
<td>0.52 (0.45 to 0.60, 78.8)</td>
</tr>
<tr>
<td>Central neck level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 4</td>
<td>0.28 (0.24 to 0.32, 94.3)</td>
<td>0.97 (0.96 to 0.98, 53.0)</td>
<td>14.07 (6.66 to 29.71, 53.1)</td>
<td>14.07 (6.66 to 29.71, 53.1)</td>
<td>14.07 (6.66 to 29.71, 53.1)</td>
</tr>
<tr>
<td>CT 4</td>
<td>0.32 (0.28 to 0.36, 88.2)</td>
<td>0.89 (0.86 to 0.91, 84.6)</td>
<td>5.48 (2.15 to 13.98, 84.3)</td>
<td>3.71 (1.79 to 7.66, 83.0)</td>
<td>0.74 (0.62 to 0.89, 86.7)</td>
</tr>
<tr>
<td>US/CT 3</td>
<td>0.40 (0.35 to 0.45, 82.3)</td>
<td>0.85 (0.82 to 0.88, 37.0)</td>
<td>4.32 (2.09 to 8.92, 81.0)</td>
<td>2.85 (1.75 to 4.65, 77.5)</td>
<td>0.67 (0.52 to 0.86, 83.9)</td>
</tr>
<tr>
<td>Lateral neck level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 6</td>
<td>0.74 (0.69 to 0.78, 78.1)</td>
<td>0.92 (0.90 to 0.94, 89.6)</td>
<td>24.41 (11.16 to 53.42, 71.9)</td>
<td>6.67 (2.91 to 15.30, 90.5)</td>
<td>0.35 (0.28 to 0.43, 30.1)</td>
</tr>
<tr>
<td>CT 6</td>
<td>0.73 (0.68 to 0.77, 90.7)</td>
<td>0.89 (0.87 to 0.91, 91.8)</td>
<td>15.55 (7.98 to 30.32, 64.6)</td>
<td>5.54 (2.95 to 10.39, 82.9)</td>
<td>0.35 (0.21 to 0.59, 91.4)</td>
</tr>
<tr>
<td>US/CT 4</td>
<td>0.88 (0.83 to 0.91, 64.1)</td>
<td>0.79 (0.73 to 0.84, 89.6)</td>
<td>22.59 (11.29 to 45.19, 46.6)</td>
<td>3.31 (1.53 to 7.17, 91.1)</td>
<td>0.19 (0.14 to 0.25, 0)</td>
</tr>
<tr>
<td>Diagnostic accuracy of CT or US on patient level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All patient level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 4</td>
<td>0.41 (0.36 to 0.46, 95.5)</td>
<td>0.92 (0.89 to 0.94, 0)</td>
<td>7.56 (4.08 to 14.01, 51.3)</td>
<td>4.48 (3.31 to 6.05, 0)</td>
<td>0.65 (0.53 to 0.80, 75.5)</td>
</tr>
<tr>
<td>CT 4</td>
<td>0.49 (0.44 to 0.54, 95.9)</td>
<td>0.91 (0.89 to 0.94, 66.8)</td>
<td>9.40 (5.79 to 15.27, 34.4)</td>
<td>4.84 (3.66 to 6.39, 0)</td>
<td>0.53 (0.37 to 0.75, 85.6)</td>
</tr>
<tr>
<td>US/CT 2</td>
<td>0.64 (0.57 to 0.71, 0)</td>
<td>0.83 (0.77 to 0.88, 0)</td>
<td>8.59 (5.37 to 13.76, 0)</td>
<td>3.71 (2.72 to 5.08, 0)</td>
<td>0.43 (0.36 to 0.53, 0)</td>
</tr>
<tr>
<td>Central patient level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 2</td>
<td>0.21 (0.16 to 0.28, 24.8)</td>
<td>0.95 (0.91 to 0.97, 0)</td>
<td>4.53 (2.34 to 8.77, 0)</td>
<td>3.78 (2.08 to 6.86, 0)</td>
<td>0.84 (0.76 to 0.93, 36.4)</td>
</tr>
<tr>
<td>CT 2</td>
<td>0.38 (0.32 to 0.46, 92.6)</td>
<td>0.90 (0.85 to 0.93, 87.2)</td>
<td>5.02 (0.46 to 54.54, 94.4)</td>
<td>3.52 (0.52 to 23.84, 93.8)</td>
<td>0.71 (0.41 to 1.22, 95.3)</td>
</tr>
<tr>
<td>US/CT 2</td>
<td>0.47 (0.39 to 0.54, 80.8)</td>
<td>0.85 (0.80 to 0.89, 0)</td>
<td>4.88 (2.58 to 9.23, 46.4)</td>
<td>3.14 (2.23 to 4.41, 0)</td>
<td>0.64 (0.47 to 0.86, 78.4)</td>
</tr>
<tr>
<td>Lateral patient level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 2</td>
<td>0.87 (0.74 to 0.95, 88.2)</td>
<td>0.89 (0.83 to 0.93, 91.4)</td>
<td>20.11 (6.77 to 59.70, 0)</td>
<td>3.58 (0.85 to 15.16, 91.0)</td>
<td>0.22 (0.05 to 1.08, 58.0)</td>
</tr>
<tr>
<td>CT 2</td>
<td>0.92 (0.81 to 0.98, 88.6)</td>
<td>0.88 (0.83 to 0.93, 96.8)</td>
<td>36.88 (11.40 to 119.35, 0)</td>
<td>3.44 (0.29 to 40.77, 98.3)</td>
<td>0.23 (0.10 to 0.52, 0)</td>
</tr>
<tr>
<td>US/CT 2</td>
<td>0.98 (0.89 to 0.99, 58.9)</td>
<td>0.92 (0.87 to 0.96, 97.5)</td>
<td>78.10 (2.82 to 2160.4, 65.2)</td>
<td>5.30 (0.15 to 186.11, 98.6)</td>
<td>0.08 (0.02 to 0.41, 0)</td>
</tr>
</tbody>
</table>

Results are presented as n (95% CI, I² %).

LR+, positive likelihood ratio; LR–, negative likelihood ratio; US, ultrasound.
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. Besides, the lateral compartment should be carefully evaluated for skip metastases located in the upper pole or are ≤1 cm in diameter. 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. 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 images and spectral HU curve. In accordance with the findings from CT, combined gemstone spectral image parameters from DECT also demonstrated better diagnostic accuracy of CLNM in patients with PTC when compared with those that are obtained by combining the US morphological parameters especially in the lateral compartment.

Our findings revealed that compared with US or CT alone, the combination of both US and CT demonstrated higher sensitivity, that is, a meta-analytic summary sensitivity of 0.51 (95% CI 0.49 to 0.52) and 0.64 (95% CI 0.57 to 0.71), and a lower specificity, that is, a meta-analytic summary specificity of 0.85 (95% CI 0.84 to 0.86) and 0.83 (95% CI 0.77 to 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.

It should be noted that the study has strengths. First, 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. Second, 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 FN 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. First, 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, and the criteria of lymph node diameter and vascular flow for US. Second, the literature included is limited due to the study design and timing of imaging. Eleven of the 14 studies (78.6%) were retrospective, and 1 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. Third, 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. Fourth, 4 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 this 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 centres.

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% to 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.

Contributors JY is the guarantor of the integrity of the entire study. YQ designed and conceptualised the study concepts and design, JY performed the literature search, YQ analysed the literature, JY analysed the data. FZ performed the statistical analyses, YQ prepared the manuscript. FZ edited the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was waived due to the secondary data acquisition from previously published papers available in the public domain.

Provenance and peer review Not commissioned; externally peer reviewed.

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

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Ying Qiao http://orcid.org/0000-0001-9275-0113

REFERENCES
5 Mazzaferrri EL, Kloos RT. Clinical review 128: current approaches to primary therapy for papillary and follicular thyroid cancer. J Clin Endocrinol Metab 2001;86:1447–63.
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;26:1–133.
10 Laboulleux S, Girard E, Rose M, et al. Ultrasound criteria of malignancy for cervical lymph nodes in patients followed


