

OPEN

# Short-term Follow-up US Leads to Higher False-positive Results Without Detection of Structural Recurrences in PTMC

Jung Hyun Yoon, MD, PhD, Hye Sun Lee, MS, Eun-Kyung Kim, MD, PhD, Ji Hyun Youk, MD, PhD, Hyun Gi Kim, MD, PhD, Hee Jung Moon, MD, PhD, and Jin Young Kwak, MD, PhD

**Abstract:** To investigate the value of the annual follow-up neck ultrasonography (US) for postoperative surveillance in patients with papillary thyroid microcarcinoma (PTMC).

This retrospective study has been approved by our institutional review board (IRB) with waiver for informed consent. A total of 375 patients diagnosed as PTMCs, who underwent total thyroidectomy with radioiodine remnant ablation were included, to identify the recurrence rate and the false-positive rate of annual ultrasound. The number, interval, and the results of follow-up US or fine needle aspiration were obtained from electronic medical records.

Four (1.1%, 4/375) recurrences were found 3 years after the initial treatment, and only 1 patient (0.3%, 1/375) had a metastatic lymph node larger than 8 mm in the shortest diameter on US found 7.6 years after initial treatment with biochemical abnormalities. Cumulative risk of having at least 1 false-positive exam was 8.3% by the 8th US, and 8.1% by the 8–9 year follow-up. Cox multivariate regression showed shorter interval of follow-up US and presence of lymph node metastasis at initial surgery are independent predictors affecting the cumulative false-positive results (hazard ratio [HR], 0.60; 95% confidence interval [CI]: 0.49–0.73;  $P < 0.001$  and HR, 2.19; 95% CI: 1.01–4.75;  $P = 0.048$ , respectively).

Short-term follow-up US can result in higher cumulative false-positive results without detection of meaningful recurrences in patients with PTMCs who do not have biochemical abnormalities.

(*Medicine* 95(1):e2435)

**Abbreviations:** CI = confidence interval, DTC = differentiated thyroid cancer, FNA = fine-needle aspiration, HR = hazard ratio, PTMC = papillary thyroid microcarcinoma, Tg = thyroglobulin, TSH = thyroid-stimulating hormone, US = ultrasound.

Editor: Patrick Wall.

Received: October 1, 2015; revised: December 10, 2015; accepted: December 11, 2015.

From the Department of Radiology, Severance Hospital, Research Institute of Radiological Science (JHY, E-KK, HJM, JYK); Department of Research Affairs, Biostatistics Collaboration Unit (HSL); Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul (JHY); and Department of Radiology, Ajou University School of Medicine, Suwon, Korea (HGK).

Correspondence: Jin Young Kwak, Department of Radiology, Severance Hospital, Research Institute of Radiological Science, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, 120-752 Seoul, Korea (e-mail: docjin@yuhs.ac).

Supplemental Digital Content is available for this article.

The authors have no funding and conflicts of interest to disclose.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0, where it is permissible to download, share and reproduce the work in any medium, provided it is properly cited. The work cannot be changed in any way or used commercially.

ISSN: 0025-7974

DOI: 10.1097/MD.0000000000002435

## INTRODUCTION

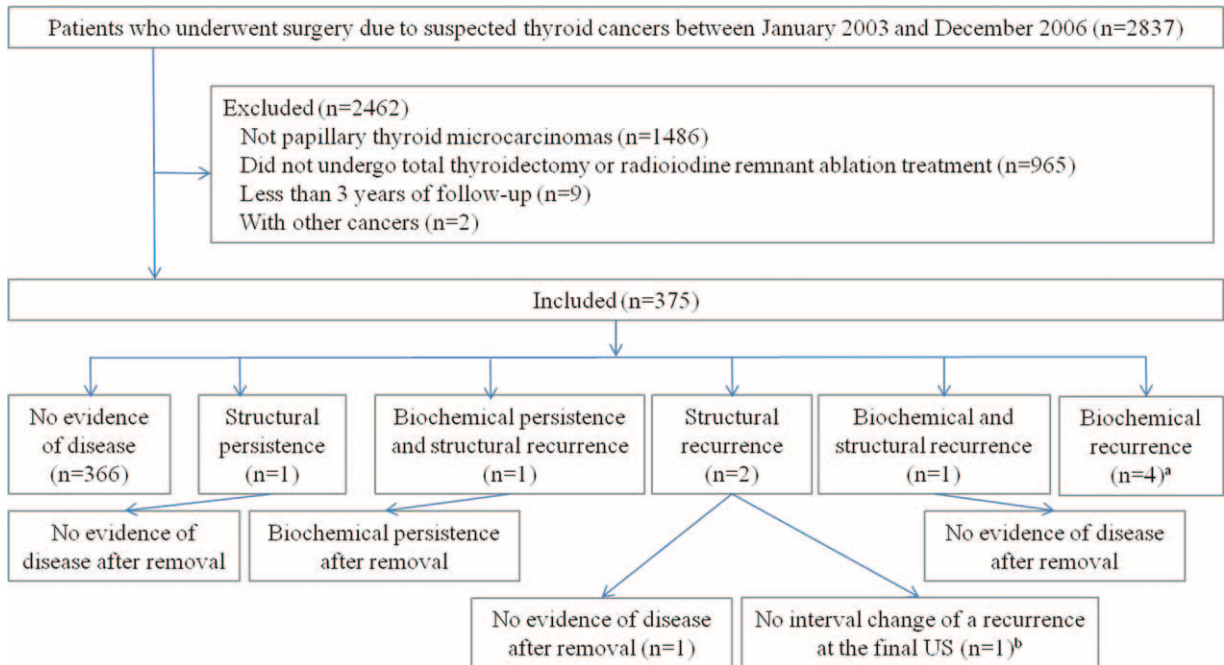
Ever since Pacini et al<sup>1</sup> suggested that the recombinant human thyroid-stimulating hormone (rhTSH)-stimulated thyroglobulin (Tg) measurement and neck ultrasonography (US) is a superior combination to rhTSH-stimulated Tg measurement and <sup>131</sup>I whole body scan in the postoperative surveillance of patients treated with differentiated thyroid cancers (DTCs), US has been widely accepted as the standard follow-up imaging modality for patients with DTCs.<sup>2</sup> The combination of stimulated Tg and neck US showed sensitivity of up to 100% for diagnosing recurrences in patients with DTCs.<sup>3</sup> US is a highly sensitive diagnostic tool that detects locoregional recurrences in the neck, which is the most common site of recurrence for thyroid cancers,<sup>1,4,5</sup> detecting recurrences even when abnormality of Tg levels have not been identified.<sup>5–7</sup> However, the clinical value of US surveillance is uncertain for patients with papillary thyroid microcarcinomas (PTMCs), because most PTMCs have an indolent clinical course, and the tumor recurrences come in small volumes that may not affect overall survival of the patient, in which treatment may cause more morbidity than its natural disease progression.<sup>8–10</sup>

The primary goal of postoperative surveillance is to detect clinically significant recurrences, not to reveal all indolent lesions. The 2015 American Thyroid Association (ATA) guidelines do not recommend fine-needle aspiration (FNA) for very small-sized lymph nodes detected on postoperative surveillance US, even when abnormal US features are observed.<sup>2</sup> Although annual US is generally recommended for 3 to 5 years as postoperative surveillance for patients with DTCs,<sup>11</sup> we came to wonder how annual US performed for postoperative surveillance may have impact on PTMC patients after surgery. Therefore, we investigated the value of annual follow-up neck US for postoperative surveillance in patients with PTMCs, evaluating the frequency of locoregional recurrences detected on US, and the false-positive rates.

## METHODS

### Study Cohort

Our institutional review board (IRB) approved of this retrospective study, and the requirement to obtain informed consent was waived. From January 2003 to December 2006, 1486 patients were confirmed to have PTMCs. Total thyroidectomy with radioiodine remnant ablation treatment was performed in 421 patients. Prophylactic central lymph node dissections were routinely performed at our institution. Of the 421 patients, 46 patients were excluded from the study; 9 for not receiving a follow-up for at least 36 months, 2 due to coexisting ovarian cancer and breast cancer, respectively, and 35 with interference of the anti-Tg antibody. Finally, 375 patients (333 women and 41 men) were included (Figure 1). The mean



**FIGURE 1.** Flow chart of the study group, <sup>a</sup>No structural recurrence, but increased serum thyroglobulin level on final follow-up in 2 patients, and follow-up loss in 2 patients after diagnosis of biochemical recurrence. <sup>b</sup>The patient had a normal serum thyroglobulin level on final follow-up.

age at surgery was 48.5 years (standard deviation [SD], 10 years; median age, 48 years; range, 20–78 years). The mean size of the 375 PTMCs was 6.3 mm (SD, 2.1 mm; median size, 6 mm; range, 2–10 mm).

**Follow-Up Protocol**

All patients underwent thyroid-stimulating hormone (TSH)-suppressive therapy with levothyroxine after surgery and were followed every 6 months for the first 3 years after initial treatment and then yearly afterward. For these patients, we recommended clinical examinations for every 6 months, neck US, chest X-ray, and measurements of serum TSH, free T4, Tg, and anti-Tg antibody for every 12 months.<sup>12</sup> Whole body scan, chest computed tomography (CT), magnetic resonance imaging (MRI), or fluorodeoxyglucose positron emission tomography (PET) was performed only in selective cases (eg, detectable serum Tg or persistent anti-Tg antibody without evident recurrence on US or whole body scan).

Before 2011, monoclonal antibody immunoradiometric assay (IRMA; CIS Bio International, Gif-sur-Yvette, France) was used for serum Tg level assays.<sup>13</sup> Radioimmunoassay was used for measuring serum anti-Tg antibodies (Brahms, Hennigsdorf/Berlin, Germany, functional sensitivity of 10 IU/mL; positive defined as value exceeding 60 U/mL), and serum TSH levels (Trinity Biotech, Co. Wicklow, Ireland, reference range of 0.4–3.1 μIU/mL, functional sensitivity of 0.0038 mIU/L). From 2011 to the present, electrochemiluminescence immunoassay using the Cobas e601 immunoassay analyzer (Roche Diagnostic, Mannheim, Switzerland) was used for serum Tg level assay (functional sensitivity, defined by the lowest analyte concentration that can be reproducibly measured with a precision CV of 20%, was 0.1 ng/mL). Serum TSH levels were measured by a

chemiluminescence immunoassay with an automated immunoassay analyzer (Architect i2000, Abbott Laboratories, Abbott Park, IL, reference range: 0.35–4.94 μIU/mL, functional sensitivity of 0.01 μIU/mL).

Neck US was performed using a 5- to 15-MHz linear array transducer (HDI 5000 or 3000, Philips-Advanced Technology Laboratories, Bothell, WA; iU22, Philips Medical Systems; Logic 9, GE Medical Systems, Milwaukee, WI; Acuson Sequoia 512, Siemens Medical Solutions USA, Inc, Mountain View, CA; EUB-7500, Hitachi Medical, Tokyo, Japan) by 77 radiologists, including 3 faculties who specialized in thyroid imaging, 15 faculties of other specialties, 4 fellows, and 55 residents of the radiologic department. When a lymph node showed loss of fatty hilum, diffuse or focal hyperechogenicity, cystic change, microcalcifications, round shape, or chaotic or peripheral vascularity on US, we regarded it as an abnormal lymph node.<sup>14</sup> Although the follow-up protocol is specifically defined at our institution, a detailed follow-up interval that differs from the institutional guidelines has been chosen at the discretion of the physician or the preference of each individual patient.

**Recurrence**

No clinical evidence of disease was defined as suppressed Tg < 1 ng/mL, stimulated Tg < 2 ng/mL, no detectable anti-Tg antibody, or no evidence of disease on US, CT, MRI, PET/CT, or whole body scan at final follow-up. Persistence was defined as suppressed Tg ≥ 1 ng/mL or stimulated Tg ≥ 2 ng/mL, or any evidence of disease on US, CT, MRI, PET/CT, or whole body scan after initial surgery and radioiodine remnant ablation therapy. Recurrence was defined as “malignancy” on cytology results, elevated serum Tg or uptake on whole body scan in

patients with no clinical evidence of disease. US-guided FNA (US-FNA) was usually performed for a suspected recurrent mass on US. Cytopathologic results were obtained by US-FNA and surgical excision.

**Data and Statistical Analysis**

Regarding follow-up US after complete treatment, we comprehensively reviewed the number, interval, and the result of follow-up US, using electronic medical records. If US-FNA was performed to diagnose focal lesions of the neck, these results were also recorded. Relevant data of the included patients were obtained for recurrences or patient mortality from our institution database.

Positive US results were considered true-positive if any atypical or malignant cells were seen on either cytology or pathology. Positive US results were considered false-positive if no atypical or malignant cells were seen on cytology or if the lesion was not visible on the final follow-up US which was performed at least 1 year after the prior follow-up showing positive results. Regarding cases with structural persistence or recurrence as true-positives, we estimated the cumulative risk of false-positive results on follow-up US based on both time-to-event and tests-to-event using the Kaplan–Meier approach.<sup>15</sup> Patients who had false-positive or true-positive results were no longer included in the Kaplan–Meier analysis despite the ongoing examinations performed afterwards.

Cox proportional hazards models were used to evaluate factors affecting the cumulative risk of false-positive results on follow-up US by univariate analysis. Multivariate analysis was performed after adjustment for baseline characteristics and known prognostic factors. Firth correction was applied because large point estimates were for the covariate “Dose of RAI” (radioactive iodine). Two-sided *P* value < 0.05 was considered to indicate statistical significance. Statistical analyses were performed using a computerized statistics program (SAS version 9.2 for Windows; SAS Institute, Cary, NC) and R package, version 3.1.1 (<http://www.R-project.org>).

**RESULTS**

Clinical characteristics are presented in Table 1. None of the patients in the study population died during the study period. Mean follow-up duration from initial treatment to the last clinical follow-up was 98.8 months (SD, 19.6 months; median, 98.5 months; range, 36.8–141.9 months). During follow-up, 1.1% (4/375) of the patients had structural recurrences, 0.3% (1/375) had structural persistence, and 1.1% (4/375) had biochemical recurrences (Table 2). All recurrences were detected 3 years after the initial treatment. Of the 4 patients with structural recurrences, 2 had elevated serum Tg levels that was normalized after surgery, and 2 had normal serum Tg levels among which 1 had surgery and was disease-free at final follow-up, and the other patient was followed without any changes observed during follow-up. Among the 4 patients with biochemical recurrence, 2 had stable serum Tg levels during follow-up, and the remaining 2 were lost for follow-up. One patient with structural persistence underwent surgery and was disease-free at final follow-up.

There were 33 patients (8.8%, 33/375) who had one or more abnormal focal lesion(s) detected on postoperative surveillance US (Appendix, <http://links.lww.com/MD/A599>). Most of them (72.7%, 24/33) showed loss of fatty hilum or other nonspecific US features. US-FNAs were performed on 28 focal lesions of 25 patients, and 5 patients (15.2%, 5/33) were

**TABLE 1.** Characteristics of the Study Cohort

Characteristic	
Age (years, mean ± SD)	48.5 ± 10
Gender	
Male	42 (11.2%)
Female	333 (88.8%)
Primary tumor	
Pathologic size (mm, mean ± SD)	6.3 ± 2.1
Multifocality in 1 lobe	36 (9.6%)
Bilaterality	96 (25.6%)
Extrathyroidal extension	180 (48%)
Lymph node metastasis	
Central LNM	107 (28.5%)
Lateral LNM	22 (5.9%)
Distant metastasis	1 (0.3%)
AJCC/UICC staging	
I	225 (60%)
III	137 (36.5%)
Iva	13 (3.5%)
RAI ablation	
30 mCi	365 (97.3%)
60 mCi	8 (2.1%)
90 mCi	1 (0.3%)
150 mCi	1 (0.3%)
ATA risk stratification	
Low-risk	141 (37.6%)
Intermediate-risk	232 (61.9%)
High-risk	2 (0.5%)
Status at final follow-up	
No evidence of disease	366 (97.6%)
Structural persistence	1 (0.3%)
Biochemical persistence and structural recurrence	1 (0.3%)
Biochemical recurrence	4 (1%)
Biochemical and structural recurrence	1 (0.3%)
Structural recurrence	2 (0.5%)

AJCC/UICC TNM staging = the American Joint Committee on Cancer (AJCC)/Union Internationale Contre le Cancer (UICC) staging (7th ed., 2010). ATA = American Thyroid Association, LNM = lymph node metastasis, RAI = radioactive iodine, SD = standard deviation.

diagnosed as metastasis. Among the focal lesions, only 1 measured larger than 8 mm at the shortest diameter.

During the study period, 2346 neck US were performed for postoperative surveillance (Table 3). For the 375 patients, the median number of neck US examinations obtained was 6 (range, 1–14). Among the 375 patients who were examined by neck US, 8.3% (28/375) had at least 1 false-positive US (Table 3). The cumulative risks of having at least 1 false-positive exams were 8.3% by the 8th US and 8.1% by the 8 to 9 year follow-up. After that, the cumulative risk of having at least 1 false-positive exam did not increase.

We evaluated the factors affecting the cumulative risk of false-positive results on follow-up US (Table 4). On univariate analysis, the shorter interval of follow-up US (hazard ratio [HR], 0.60; 95% confidence interval [CI]: 0.50–0.73; *P* < 0.001) and LNM (HR, 2.4; 95% CI: 1.15–5.06; *P* = 0.02) were statistically significant factors affecting the cumulative false-positive results on follow-up US. Cox multivariate regression analysis showed that the shorter interval of

TABLE 2. Patients With Persistences or Recurrences

	Age at Diagnosis	Initial Cancer Size, mm	AJCC/UICC Staging	ATA Risk	Time to Recurrence, year	Recurrence or Persistence	Location	Size, mm	Treatment	Final Follow-Up	Time of Final Follow-Up, year
1*	≥45	4	III	Intermediate	4	Biochemical persistence and structural recurrence	Level 4	9 × 5	Surgery	Second persistent biochemical disease <sup>†</sup>	7.3
2*	≥45	9	III	Intermediate	7.6 <sup>‡</sup>	Biochemical and structural recurrence	Level 4	16 × 9	Surgery	Disease free	8.5
3	<45	5	I	Intermediate	4.2	Structural recurrence	Op bed	4 × 2	None	Persistent structural disease	8.1
4	≥45	3	I	Low	3.2	Structural recurrence	Level 4	6 × 3	Surgery	Disease free	7.5
5	<45	5	I	Intermediate	4.3	Biochemical recurrence				Persistent biochemical disease <sup>‡</sup>	7.3
6	≥45	4	III	Intermediate	3.9	Biochemical recurrence				Persistent biochemical disease <sup>†</sup>	7.9
7	≥45	2	I	Low	7.8	Biochemical recurrence				Not available <sup>§</sup>	7.8
8	≥45	3	III	Intermediate	7.1	Biochemical recurrence				Not available <sup>§</sup>	7.1
9	≥45	9	III	High		Structural persistence	Level 4	7 × 7	Surgery	Disease free	6.8

AJCC/UICC TNM staging = the American Joint Committee on Cancer (AJCC)/Union Internationale Contre le Cancer (UICC) staging (7th ed., 2010), ATA = American Thyroid Association.

\*The patients had two suspicious lymph nodes.

†All patients had persistently high serum thyroglobulin levels without any increases.

‡Time to structural recurrence.

§No further follow-up data was available after a biochemical recurrence was discovered.



**TABLE 3.** Cumulative Probability of a First False-Positive Result by Number of Neck USs Performed and by Follow-Up Year in 375 Patients

Number of Neck USs	Number Exams, %	Cumulative Number Exams, %	Number of False-Positive Results	Number of True-Positive Results	Cumulative False-Positive Results (SE), %
1	375 (100)	375 (16.98)	5	1	1.33 (0.59)
2	368 (98.1)	743 (31.67)	5	0	2.67 (0.83)
3	360 (96)	1103 (47.02)	10	1	5.38 (1.17)
4	336 (90)	1439 (61.34)	3	1	6.22 (1.26)
5	302 (80.5)	1741 (74.21)	4	1	7.46 (1.39)
6	243 (64.8)	1984 (84.57)	0	1	7.46 (1.39)
7	161 (42.9)	2145 (91.43)	0	0	7.46 (1.39)
8	107 (28.5)	2252 (95.99)	1	0	8.33 (1.62)
9	60 (16)	2312 (98.55)	0	0	8.33 (1.62)
10	21 (5.6)	2333 (99.45)	0	0	8.33 (1.62)
11	7 (1.9)	2340 (99.74)	0	0	8.33 (1.62)
12	3 (0.8)	2343 (99.87)	0	0	8.33 (1.62)
13	2 (0.5)	2345 (99.96)	0	0	8.33 (1.62)
14	1 (0.3)	2346 (100)	0	0	8.33 (1.62)

Duration, year	Number Subjects, %	Cumulative Number Subjects, %	Number of False-Positive Results	Number of True-Positive Results	Cumulative False-Positive Results (SE), %
0–1	295 (78.67)	295 (12.57)	5	1	1.34 (0.59)
1–2	316 (84.27)	611 (26.04)	1	0	1.61 (0.65)
2–3	306 (81.6)	917 (39.09)	9	0	4.03 (1.02)
3–4	299 (79.74)	1216 (51.83)	7	2	5.94 (1.23)
4–5	283 (75.47)	1499 (63.9)	4	1	7.08 (1.34)
5–6	234 (62.4)	1733 (73.87)	1	0	7.37 (1.37)
6–7	218 (58.14)	1951 (83.16)	0	0	7.37 (1.37)
7–8	190 (50.67)	2141 (91.26)	0	1	7.37 (1.37)
8–9	112 (29.87)	2253 (96.04)	1	0	8.13 (1.55)
9–10	68 (18.14)	2321 (98.93)	0	0	8.13 (1.55)
10–11	22 (5.87)	2343 (99.87)	0	0	8.13 (1.55)
11–12	3 (0.8)	2346 (100)	0	0	8.13 (1.55)

SE = standard error.

follow-up US and lymph node metastasis are independent predictors affecting the cumulative false-positive results on follow-up US (HR, 0.60; 95% CI: 0.49–0.73;  $P < 0.001$  and HR, 2.19; 95% CI: 1.01–4.75;  $P = 0.048$ , respectively).

### DISCUSSION

In this study of patients with PTMCs who underwent total thyroidectomy and radioiodine remnant ablation treatment, locoregional structural recurrence was detected in 1.1% (4/375) of patients at a median follow-up of 98.5 months. Considering the clinical insignificance of small lymph nodes, even with cancer foci, the 2015 ATA guidelines suggest performing US-FNA for suspicious lymph nodes of 5 to 8 mm in the smallest diameter.<sup>2</sup> When considering the size cut-off for diagnostic FNA in lymph nodes, only 1 (0.3%) patient had a metastatic lymph node measuring larger than 8 mm in the shortest diameter on US in this study, who had biochemical recurrence. There were no significant structural recurrences detected among patients without biochemical recurrences.

The goal of postoperative surveillance is to properly detect recurrences which can affect morbidity and mortality from the disease, while minimizing unnecessary invasive biopsies and

patient anxiety. Although several guidelines recommend periodic US with serum Tg assessments as postoperative surveillance for patients with DTC, these guidelines do not provide an optimal time length of follow-up intervals for postoperative surveillance US,<sup>2,16</sup> and annual US has been performed empirically for postoperative surveillance.<sup>4,5,17</sup> Recently, false-positive abnormalities have been reported to result from frequent neck US of patients with intermediate risk for PTC.<sup>18</sup> In this study, shorter follow-up US intervals and lymph node metastasis at initial surgery resulted in a higher cumulative false-positive results on follow-up US. Also, all 4 structural recurrences were found 3 years after initial treatment, and 1 metastatic lymph node measuring larger than 8 mm in the shortest diameter was found on US performed 7.6 years after the initial treatment. From our results, we can conclude that the frequent use of highly sensitive US may lead to overuse of invasive procedures such as US-FNA, and cause subsequent psychological distress or economic burden on patients, despite of the lack in established benefits. On the other hand, negative results on follow-up US can lessen patient anxiety that should not be neglected, but as there are no consensus on which interval or duration that postoperative surveillance US should be performed, we expect our results to be a start of future prospective

**TABLE 4.** Factors Affecting the Cumulative Risk of False-Positive Results in Follow-up US by Cox Regression Analysis Using the Firth Correction Method

Characteristic	Univariable		Multivariable	
	HR (95% CI)	P Value	HR (95% CI)	P Value
US follow-up interval	0.60 (0.50–0.73)	<0.001	0.60 (0.49–0.73)	<0.001
Age				
≥45	1.17 (0.53–2.59)	0.692	1.20 (0.54–2.66)	0.659
<45	1 [Reference]		1 [Reference]	
Gender				
Male	2.29 (0.93–5.64)	0.073	1.64 (0.66–4.08)	0.293
Female	1 [Reference]		1 [Reference]	
Pathologic size, mm				
≥5	0.98 (0.40–2.41)	0.96	0.77 (0.30–1.94)	0.578
<5	1 [Reference]		1 [Reference]	
Extrathyroidal extension				
Positive	0.81 (0.38–1.71)	0.577	0.98 (0.46–2.09)	0.948
Negative	1 [Reference]		1 [Reference]	
LNM				
Positive	2.41 (1.15–5.06)	0.02	2.19 (1.01–4.75)	0.048
Negative	1 [Reference]		1 [Reference]	
Dose of RAI	0.64 (0.04–10.99)	0.756	0.55 (0.03–10.20)	0.688

CI = confidence interval, HR = hazard ratio, LNM = lymph node metastasis, RAI = radioactive iodine.

studies providing evidence for more cost-effective surveillance programs for PTMC patients.

Lymph node metastasis has been reported to be up to 64.1% in the central compartment and 44.5% in the lateral compartment in patients with PTMCs, when prophylactically dissected.<sup>19</sup> Another issue to contemplate is that the recurrence and persistence after surgery can be intermingled at both the central and lateral compartments of the neck, because prophylactic compartment dissection is not universally accepted as most lymph node metastases are regarded clinically insignificant.<sup>20</sup> Until now, there has been no worldwide consensus on suspicious US features that are suggestive of metastatic cervical lymph nodes.<sup>1,6,21–24</sup> In a study including patients with low risk thyroid cancers (stage I or II), all recurrences of 44 patients (8.9%) were identified in the neck, among which 30 patients (6.1%) had recurrences detected by applying sensitive US criteria. A substantial portion (70.7%) of recurrences occurred in the first 5 years after surgery.<sup>23</sup> Among the 149 aspirated lymph nodes, 80.5% (120/149) had benign results on cytology, which signifies a large number of false-positive findings on US. It is unclear how many neck recurrences would be clinically meaningful because the size of the recurrences were not available.<sup>23</sup> Nixon et al<sup>25</sup> showed a very low recurrence rate of 0.6% (5/889) after applying specific US features in diagnosing locoregional recurrences in patients who underwent thyroid lobectomy for DTC. Durante et al<sup>4</sup> also demonstrated that there were no locoregional recurrences when applying the size criterion of 5 mm on a short axis with sensitive US criteria during a median follow-up of 6.7 years in 312 PTMC patients with T1N0M0 stage. These contradictory results may be due to different primary cancer staging, different US criteria for FNA, and differences in size criteria for applying US-FNA at focal lesions in the neck.<sup>4,23</sup> The majority of false-positive results in this study were from the loss of fatty hilum within cervical LNs detected on US. The loss of fatty hilum is a highly sensitive feature but with low

specificity, which the 2015 ATA guideline does not consider a significant finding in assessing metastatic LNs;<sup>2</sup> therefore, there were no additional false-positive results at follow-up US. In addition, we can understand how different guidelines on abnormal US features of cervical lymph nodes can have impact on the recurrence rates in patients with thyroid cancers.

There are some limitations to our study. First, this study is of a retrospective design and the exclusion of some cases was inevitable. Also, the interval or duration of follow-up US was not strictly controlled, which may have affected the cumulative false-positive rates of this study. Second, we included patients with PTMCs who underwent total thyroidectomy, bilateral prophylactic central neck dissections, and radioiodine remnant ablation treatment. Future studies investigating the role of follow-up US on patients who have been treated less aggressively are anticipated to validate our results. Third, surveillance US was performed by 77 radiologists with variable experiences. Although this reflects data obtained from everyday practice, the quality of US may have been affected. Last, the radiologists were not blinded to the pathology results of the patient, such as tumor stage or LN status, which was not considered in analysis.

In spite of these limitations, results of this study suggest that short-term follow-up US can result in higher cumulative false-positive results without detection of meaningful recurrences in patients with PTMCs who do not have biochemical abnormalities.

## REFERENCES

- Pacini F, Molinaro E, Castagna MG, et al. Recombinant human thyrotropin-stimulated serum thyroglobulin combined with neck ultrasonography has the highest sensitivity in monitoring differentiated thyroid carcinoma. *J Clin Endocrinol Metab*. 2003;88:3668–3673.
- Haugen BR, Alexander EK, Bible KC, et al. 2014 American Thyroid Association Management Guidelines for Patients with Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2014.

3. Mazzaferri EL, Robbins RJ, Spencer CA, et al. A consensus report of the role of serum thyroglobulin as a monitoring method for low-risk patients with papillary thyroid carcinoma. *J Clin Endocrinol Metab.* 2003;88:1433–1441.
4. Durante C, Attard M, Torlontano M, et al. Identification and optimal postsurgical follow-up of patients with very low-risk papillary thyroid microcarcinomas. *J Clin Endocrinol Metab.* 2010;95:4882–4888.
5. Durante C, Montesano T, Torlontano M, et al. Papillary thyroid cancer: time course of recurrences during postsurgery surveillance. *J Clin Endocrinol Metab.* 2013;98:636–642.
6. Torlontano M, Attard M, Crocetti U, et al. Follow-up of low risk patients with papillary thyroid cancer: role of neck ultrasonography in detecting lymph node metastases. *J Clin Endocrinol Metab.* 2004;89:3402–3407.
7. Antonelli A, Miccoli P, Ferdeghini M, et al. Role of neck ultrasonography in the follow-up of patients operated on for thyroid cancer. *Thyroid.* 1995;5:25–28.
8. Johnson NA, Tublin ME. Postoperative surveillance of differentiated thyroid carcinoma: rationale, techniques, and controversies. *Radiology.* 2008;249:429–444.
9. Udelsman R. Treatment of persistent or recurrent papillary carcinoma of the thyroid – the good, the bad, and the unknown. *J Clin Endocrinol Metab.* 2010;95:2061–2063.
10. Sippel RS, Chen H. Controversies in the surgical management of newly diagnosed and recurrent/residual thyroid cancer. *Thyroid.* 2009;19:1373–1380.
11. Tuttle RM, Leboeuf R, Martorella AJ. Papillary thyroid cancer: monitoring and therapy. *Endocrinol Metab Clin North Am.* 2007;36:753–778vii.
12. Moon HJ, Kim EK, Chung WY, et al. Minimal extrathyroidal extension in patients with papillary thyroid microcarcinoma: is it a real prognostic factor? *Ann Surg Oncol.* 2011;18:1916–1923.
13. Kim MJ, Kim EK, Kim BM, et al. Thyroglobulin measurement in fine-needle aspirate washouts: the criteria for neck node dissection for patients with thyroid cancer. *Clin Endocrinol (Oxf).* 2009;70:145–151.
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;193:871–878.
15. Croswell JM, Kramer BS, Kreimer AR, et al. Cumulative incidence of false-positive results in repeated, multimodal cancer screening. *Ann Fam Med.* 2009;7:212–222.
16. National Comprehensive Cancer Network Clinical Practice Guidelines. Available at: <http://www.nccn.org/default.aspx>. [Accessed August 10, 2015].
17. Pacini F. Thyroid microcarcinoma. *Best Pract Res Clin Endocrinol Metab.* 2012;26:421–429.
18. Peiling Yang S, Bach AM, Tuttle RM, et al. Frequent screening with serial neck ultrasound is more likely to identify false-positive abnormalities than clinically significant disease in the surveillance of intermediate risk papillary thyroid cancer patients without suspicious findings on follow-up ultrasound evaluation. *J Clin Endocrinol Metab.* 2015;100:1561–1567.
19. Wada N, Duh QY, Sugino K, et al. Lymph node metastasis from 259 papillary thyroid microcarcinomas: frequency, pattern of occurrence and recurrence, and optimal strategy for neck dissection. *Ann Surg.* 2003;237:399–407.
20. Cooper DS, Doherty GM, Haugen BR, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2009;19:1167–1214.
21. Tuttle RM, Tala H, Shah J, et al. Estimating risk of recurrence in differentiated thyroid cancer after total thyroidectomy and radioactive iodine remnant ablation: using response to therapy variables to modify the initial risk estimates predicted by the new American Thyroid Association staging system. *Thyroid.* 2010;20:1341–1349.
22. Lebouleux 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;92:3590–3594.
23. Pelttari H, Laitinen K, Schalin-Jantti C, et al. Long-term outcome of 495 TNM stage I or II patients with differentiated thyroid carcinoma followed up with neck ultrasonography and thyroglobulin measurements on T4 treatment. *Clin Endocrinol (Oxf).* 2008;69:323–331.
24. 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;18:1734–1741.
25. Nixon IJ, Ganly I, Patel SG, et al. Thyroid lobectomy for treatment of well differentiated intrathyroid malignancy. *Surgery.* 2012;151:571–579.