AJNR

Clinical Validation of a Predictive Model for the Presence of Cervical Lymph Node Metastasis in Papillary Thyroid Cancer

N.U. Patel, K.E. Lind, K. McKinney, T.J. Clark, S.S. Pokharel, J.M. Meier, E.R. Stamm, K. Garg and B. Haugen

Clinical Validation of a Predictive Model for the Presence of Cervical Lymph Node Metastasis in Papillary Thyroid Cancer

[ID] N.U. Patel, [ID] K.E. Lind, [ID] K. McKinney, [ID] T.J. Clark, [ID] S.S. Pokharel, [ID] J.M. Meier, [ID] E.R. Stamm, [ID] K. Garg, and [ID] B. Haugen

ABSTRACT

BACKGROUND AND PURPOSE: Ultrasound is a standard technique to detect lymph node metastasis in papillary thyroid cancer. Cystic changes and microcalcifications are the most specific features of metastasis, but with low sensitivity. This prospective study compared the diagnostic accuracy of a predictive model for sonographic evaluation of lymph nodes relative to the radiologist's standard assessment in detecting papillary thyroid cancer metastasis in patients after thyroidectomy.

MATERIALS AND METHODS: Cervical lymph node sonographic images were reported by a radiologist (R method) per standard practice. The same images were independently evaluated by another radiologist using a sonographic predictive model (M method). A test was considered positive for metastasis if the R or M method suggested lymph node biopsy. The result of lymph node biopsy or surgical pathology was used as the reference standard. We estimated relative true-positive fraction and relative false-positive fraction using log-linear models for correlated binary data for the M method compared with the R method.

RESULTS: A total of 237 lymph nodes in 103 patients were evaluated. Our analysis of relative true-positive fraction and relative false-positive fraction included 54 nodes with pathologic results in which at least 1 method (R or M) was positive. The M method had a higher relative true-positive fraction of 1.46 (95% CI, 1.12–1.91; $P = .006$) and a lower relative false-positive fraction of 0.58 (95% CI, 0.36–0.92; $P = .02$) compared with the R method.

CONCLUSIONS: The sonographic predictive model outperformed the standard assessment to detect lymph node metastasis in patients with papillary thyroid cancer and may reduce unnecessary biopsies.

ABBREVIATIONS: LN = lymph node; FNAB = fine-needle aspiration biopsy; M method = sonographic evaluation of lymph nodes using the predictive model; PTC = papillary thyroid carcinoma; R method = sonographic evaluation of the lymph node per standard clinical practice; rFPF = relative false-positive fraction; rTPF = relative true-positive fraction; Tg = thyroglobulin; US = ultrasound

Thyroid cancer represents 3.4% of all new cancer cases in the United States, with an incidence of 14.2 per 100,000 individuals per year.[1] The National Cancer Institute estimates 56,870 new cases of thyroid cancer and 2010 deaths due to thyroid cancer in 2017.[1] Its incidence has significantly increased in recent years,

attributed to increases in papillary thyroid carcinoma (PTC).[2,3] With a common cancer, accurate assessment for recurrence is paramount. The risk of recurrence spans from <1% in very low-risk patients to >50% in high-risk patients,[4] with a recurrence rate of approximately 27% for regional lymph node (LN) metastases in patients with PTC.

Evaluation of postoperative disease status can be performed with serum thyroglobulin levels (Tg), cervical ultrasound (US), iodine radioisotope scanning, contrast-enhanced neck CT, or MR imaging. Compared with diagnostic radioiodine scan or neck CT, LN evaluation by comprehensive neck ultrasound is less expensive and exposes patients to no ionizing radiation or intravenous iodinated contrast agents.

The 2015 American Thyroid Association guidelines recommend that following an operation, cervical US should be performed at 6–12 months to evaluate the thyroid bed as well as the central and lateral cervical LN compartments, with the frequency of subsequent follow-up imaging depending on the patient's risk

```
IRB approved, HIPAA-compliant prospective paired study in post-thyroidectomy
patients with known diagnosis of PTC.

Independent review of sonograhically identified LN in surveillance neck ultrasound
by R and M methods. N=103 participants, 237 lymph nodes

R method- Standard sonographic          M method- Independent evaluation of the
evaluation of LN per clinical practice.      LN by sonographic predictive model.

Positive R method -    Negative    Negative    Positive M method -
FNAB indicated N=87 LN    test         test       FNAB indicated N=47 LN

FNAB of LN or surgery   Follow up US if Negative test or    FNAB of LN or surgery
                        FNAB is not performed
                        N=164 LN, 67 participants

Presence or absence of PTC metastasis in a LN is determined by FNAB (cytology or Tg
wash) or by surgical pathology results. N=73 LN from 39 participants.

rTPF and rFPF analysis includes LN where at least one method (M or R) was positive and
pathology result was ascertained. N=54 LN, 34 participants.
```
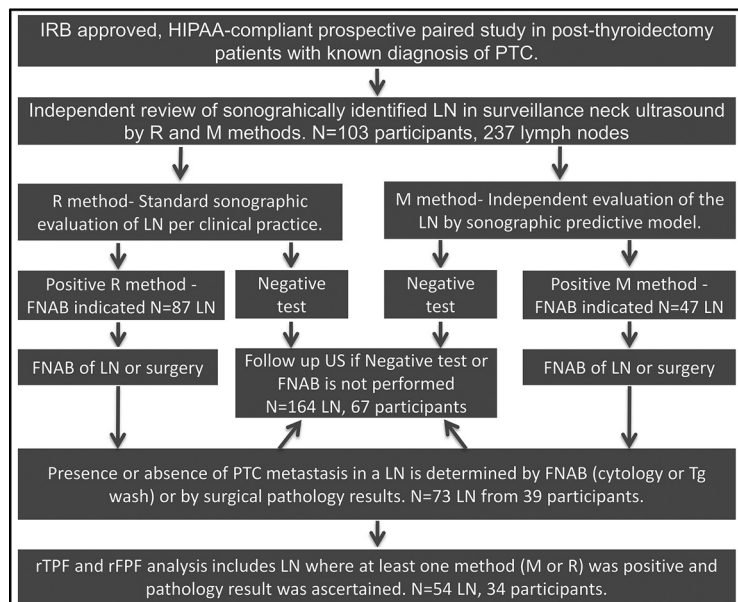
FIGURE. Diagram depicting the overall study design.

for disease recurrence and Tg status.[4] These guidelines recommend fine-needle aspiration biopsy (FNAB) of sonographically suspicious LNs of ≥8–10 mm in the smallest diameter for evaluation of cytology, with Tg measurement in the needle washout fluid if positive results would change patient management.

Many studies report sonographic features of LNs that are associated with thyroid cancer metastasis in patients with PTC; however, some of these features are not highly sensitive. Cystic changes and microcalcifications are the most specific sonographic features of PTC LN metastasis, but with low sensitivity.[5-8] In the absence of these most specific sonographic features, LN selection for FNAB can be challenging. One group's pilot data on 71 lymph nodes showed that only 32% of the abnormal or suspicious LNs by sonographic features had PTC metastasis.[9] A predictive model was developed in that retrospective pilot study based on sonographic markers (nonhomogeneous echo texture, microcalcification, and nodal volume). This model had a sensitivity of 65% (95% CI, 50%–78%) and a specificity of 85% (95% CI, 73%–94%) at a probability cut-point of 0.38.[9]

The goal of the current study was to prospectively validate this sonographic predictive model (M method) and estimate the diagnostic accuracy of the M method relative to radiologists' standard assessments (R method) in identifying LN metastasis in a patient population with PTC after thyroidectomy. We hypothesized that the M method would have higher relative true-positive fraction (rTPF) and lower relative false-positive fraction (rFPF) compared with the R method.

MATERIALS AND METHODS
Study Design
The flow diagram in the Figure depicts the overall study design. We conducted a prospective study with a sample of patients at a large academic hospital (University of Colorado Hospital at Anschutz Medical Campus), recruiting patients after thyroidectomy with a known diagnosis of PTC who had a comprehensive neck sonographic examination performed at our institution from

June 2015 to December 2015. All eligible patients during this period were screened for inclusion and exclusion criteria outlined below and recruited consecutively. This study was institutional review board–approved and Health Insurance Portability and Accountability Act–compliant. Informed consent was obtained from each participant who was willing to enroll in this study.

To be eligible for the study, participants were required to have a diagnosis of PTC, be 18 years of age or older, provide informed consent to participate, and have a sonographically identified LN of at least 5 mm in the short axis for zones 3, 4, 5, 6, and 7 or at least 10 mm in the short axis for zones 1 and 2. We excluded anyone who was younger than 18 years of age, had a cancer other than papillary thyroid carcinoma, had distant metastasis, or was unable or unwilling to provide informed consent. Each patient's eligibility per study criteria was determined at the time of the screening US. If a patient was found eligible, then study-related information was given to the patient at the time of the US. Patients who were interested in participating in the study gave consent after the screening US or were contacted later (24–48 hours) to obtain informed consent.

Age, sex, type of PTC, and laboratory, radiology, and pathology data were abstracted from electronic medical records. Participants were followed for up to 12 months for biopsy or surgery results.

Image Acquisition

Neck Ultrasound Protocol. Neck ultrasound for LN evaluation was performed using L12–5-, L17–5-, and C8–5-MHz transducers and iU22 US machines (Philips Healthcare, Best, the Netherlands). The standard neck ultrasound protocol was designed to detect, map, and characterize the lymph nodes in zones 1, 2, 3, 4, and 6. Zonal mapping of the cervical LNs was performed using anatomic landmarks.[10] Zone 5 was not routinely evaluated unless there was a palpable abnormality or clinical concern in that area. The C8–5 transducer was used to evaluate inferior zone 6 and zone 7 LNs.

The gray-scale ultrasound images included 3 axis dimensions of a nodule/LN if it met the size criteria: ≥10 mm in the short axis for zones 1 and 2, and ≥5 mm in the short axis in the other zones. These static gray-scale images of the lymph nodes were supplemented with a superior-to-inferior cine clip of the central compartment and right and left lateral compartments of the neck for the radiologist's review.

Color Doppler evaluation was performed with a L12–5 or L17–5 transducer (iU22 US; Philips Healthcare) in a lymph node plane that best assessed the hilar and nonhilar flow (longitudinal or axial image). The color scale was decreased to the Nyquist limit; then, the color gain was increased to the point of color speckle and

Table 1: Lymph node evaluation by the sonographic predictive model (features of a predicted abnormal LN)[a]

Nonhomogeneous Echo Pattern	Microcalcifications	Volume (mm³)
Present	Present	Any
Absent or indeterminate	Present	≥853
Present	Absent or indeterminate	≥690
Absent or indeterminate	Absent or indeterminate	≥2538

[a] LN FNAB recommendation was based on the combination of features at a given volume threshold.

then decreased slightly. The following parameters were selected for color Doppler evaluation: color scale around 5 (500–800 Hz), color gain in the range of 70%–85%, and low wall filter.

Image Analysis

Index Test. Neck US images were reviewed and analyzed for both M and R methods on a PACS.

Image Analysis by the R Method. Images from neck US were evaluated per standard clinical practice by 6 fellowship-trained radiologists with 3–20 years of experience. All radiologists involved in the standard reporting during the clinical practice underwent formal training regarding how to review and report neck US images for LN assessment. Reading radiologists reviewed all static and cine clips, evaluating multiple sonographic features described in the literature to determine whether LNs had normal or abnormal findings and if LNs required FNAB. We assessed the following features: size, shape, hilum, echo pattern, echogenicity, calcification, and color Doppler flow. Besides previous neck US for comparison, the reading radiologists also had access to other clinically relevant information, including pathology reports, risk of recurrence, serum Tg level, and abnormal findings on other imaging modalities (eg, CT or nuclear medicine radioiodine scan). The results of the R method were considered positive for metastasis if the interpreting radiologist recommended FNAB of a LN.

Image Analysis by the M Method. Each neck US examination was also independently evaluated by a radiologist with 10 years of experience who was not involved with and was blinded to the results of the US clinical read (R method). All static and compartment-based cine clips were reviewed for LN evaluation. LN FNAB was recommended only if the LN met sonographic feature combinations as described in the predictive model summarized in Table 1.[9] Node volume was calculated by an ellipsoid formula using 3 axis node dimensions. The results of the M method were considered positive for metastasis if they indicated FNAB of a LN.

Fine-Needle Aspiration Biopsy. At our institution, it is standard to perform compartment-based LN FNAB to assist surgeons in planning for a compartmental LN dissection operation.

FNAB of the LN positive by the R or M method was subsequently performed under US guidance per standard technique after obtaining informed consent. If the patient had multiple LNs positive by the R or M method, then the most suspicious LN in each separate compartment (central, right lateral, left lateral) was selected for FNAB. Two-to-four samples were obtained from the selected LN using a 25-ga needle for cytology and Tg wash. PTC metastasis was confirmed by positive cytology. LN samples were examined by cytopathologists with 3–15 years of experience. If LN cytology was negative for PTC metastasis, fine-needle aspiration Tg wash was performed to determine the presence or absence of PTC metastasis.[11] Fine-needle aspiration Tg wash of <1 ng/mL was considered negative for PTC metastasis.[12,13] Results of FNAB (cytology and fine-needle aspiration Tg wash if cytology was inconclusive or inadequate) of the LNs were used as the reference standard to determine the presence or absence of PTC metastasis. Surgical pathology of the LN from the same level compartment-based surgical dissection was used as a secondary reference standard if positive LNs by the R or M method did not undergo FNAB.

Follow-Up Neck US Examinations. Follow-up neck ultrasound examinations were reviewed for LNs that were identified in the index US test but did not undergo FNAB or surgical resection for definite diagnosis (inclusive of positive and negative LNs by the R and M methods). The US images were evaluated to determine whether the LN was stable (unchanged in size and appearance), not suspicious (not seen on follow-up US, became normal in appearance, or smaller), or suspicious (became abnormal in appearance or increased in size).

Statistical Design

We calculated descriptive statistics for all variables of interest, including means, medians, and SDs for continuous variables and frequencies and percentages for categoric variables.

A common problem when validating screening tests is missing data, which creates verification bias. This bias occurs when only screen-positives for cancer are referred for the reference standard test, which often happens when the reference standard is invasive. Because performing biopsies on patients who were negative with the R and M methods would be unethical, reference standard data are missing for these patients. With these data missing, classic measures of test accuracy, including sensitivity and specificity, cannot be correctly calculated. However, there are other meaningful accuracy measures that can be calculated in studies facing verification bias. Because patients who had a positive M method or R method test were referred for biopsy or an operation, we could calculate the proportion of true-positives and false-positives for each method. Furthermore, we could compare the performance of the M and R methods on these measures.

To compare the M method with the R method, we estimated the relative true-positive fraction and relative false-positive fraction using log-linear models for correlated binary data.[14] This approach is specifically designed to handle incomplete outcome data in instances in which screen-negatives for cancer never get the reference standard test. We fit separate models for rTPF and rFPF, with the binary outcomes of metastatic disease (determined from LN FNAB or surgical pathology results) positive and negative, respectively. The primary predictor in each model was an indicator variable equal to 1 for records corresponding the M method and equal to zero for records corresponding to the R method. We used a backward stepwise model-building approach. We first evaluated potential tests by sex interactions, to determine whether the performance of the tests varied by patient sex; then, we evaluated sex main effects in the absence of interaction, with a plan to drop sex main effects if nonsignificant. Covariates were retained at an α level of .20.[15]

Table 2: Patient characteristics (N = 103)

Variable	Mean (SD)	Median	Min, Max
Age (yr)	51.17 (13.49)	51.0	23, 78
Time since PTC diagnosis (yr)	6.07 (5.35)	4.5	0.23, 30.48

Note:—Min indicates minimum; Max, maximum.

As a secondary analysis, we evaluated the association between follow-up US examinations and M and R method results on the index US for the patients (and nodes) that did not undergo biopsy or surgery but had US follow-up performed later. We categorized the US follow-up examination results for LN into 3 groups: suspicious, not suspicious, or stable relative to the index US. To evaluate the association between the follow-up US results and results of R and M methods, we used a Fisher exact test for R and M methods separately (ie, we tested for an association between the 3-level category of follow-up US results and the binary result of the R or M method).

We used a type I error rate of 0.05 for all statistical tests. All analyses were conducted using SAS 9.4 (SAS Institute, Cary, North Carolina).

RESULTS

A total of 105 participants after thyroidectomy were enrolled in the study from June 2015 to December 2015; two participants were excluded because one patient had medullary and papillary thyroid carcinoma diagnoses and the other participant had distant metastasis in the lungs, yielding a final study sample of 103 patients after thyroidectomy with diagnosis of papillary thyroid carcinoma. Staging information could be retrieved from electronic medical records in 66 participants ($n = 103$): 46 with stage I; five with stage II; eight with stage III; and 7 with stage IV thyroid cancer. These 103 participants had 237 LNs evaluated by both the R and M methods. Forty-seven LNs were positive by the M method and 87 LNs were positive by R method.

Thirty-nine of the 103 participants had LN biopsy or an operation (37 had LN FNAB, and 2 had a compartment-based LN dissection operation). The median time between the index US and LN biopsy was 28 days for the subset of the participants (28 of 37) who had biopsy of the LNs (positive by the R or M method). Nine of 37 participants had LN biopsy recommended by the R or M method but chose not to undergo repeat biopsy because they already had the same LN biopsied before the index US. Of the 164 LNs without pathology results, 105 LNs from 67 patients had follow-up US examinations available for comparison (the time interval from the index US test ranged from 90 to 560 days; 7/67 [10.45%] had the follow-up US at <6 months; 12/67 [17.91%], at 6–9 months; 19/67 [28.36%], at 9–12 months; and 29/67 [43.28%], at >12 months).

Participant characteristics are presented in Tables 2 and 3. Most participants (73.79%) were women, and most patients had 1–3 nodes identified by US. Twenty-one of the 39 participants who underwent LN biopsy or surgical resection had 33 lymph nodes positive for metastasis. LN metastasis was commonly seen in zone 6 (17/33) followed by zone 4 (8/33) and zone 3 (5/33). Serum Tg information within 1 month of the index US was found in electronic medical records in 28 of 39 participants. Of 18 participants with LN metastasis, only 8 had elevated serum Tg levels (defined as >1.0 ng/mL).

Table 3: Patient characteristics (N = 103)

	Frequency (%)
Women	76 (73.79%)
Biopsy or operation performed	39 (37.86%)
Positive for LN metastasis	21 (53.85%)
No. of LNs evaluated on US	
1	30 (29.13%)
2	36 (34.95%)
3	21 (20.39%)
4	11 (10.68%)
5	4 (3.88%)
8	1 (0.97%)

Table 4: Frequencies of patient level (N = 103 participants) and lymph node level (N = 237 LNs) M and R method results[a]

	Biopsy Positive for Metastasis	Biopsy Negative for Metastasis	Not Biopsied	Total
Person level				
Total participants	21	18	64	103
M				
Positive	17	6	13	36
Negative	4	12	51	67
R				
Positive	19	13	17	49
Negative	2	5	47	54
Lymph node level				
Total lymph lodes	33	40	164	237
M				
Positive	21	9	17	47
Negative	12	31	147	190
R				
Positive	24	26	37	87
Negative	9	14	127	150

[a] For person level results, biopsy positive for metastasis indicates that the patient had at least 1 LN positive for metastasis, and positive M or R method indicates that the patient had at least 1 LN called positive by the respective method.

Table 5: LN pathology results by R and M method

	Method	
Results	+R Method	–R Method
Positive for metastasis (33 LNs)		
+M method	17	4
−M method	7	5
Negative for metastasis (40 LNs)		
+M method	6	3
−M method	20	11

Table 6: Relative true-positive and relative false-positive rates for the M method compared with the R method (n = 54 LNs with pathology results in which at least 1 method, R or M, had positive findings)

	Estimate (SE)	95% CI	P Value
rTPF M method	1.46 (0.20)	1.12–1.91	.01
rFPF M method	0.58 (0.14)	0.36–0.92	.02

Note:—SE, indicates standard error.

Lymph node level and patient level results are presented in Table 4. Table 5 presents the frequencies of R and M method positivity for the subsets of lymph nodes positive and negative for metastasis, respectively. The analysis of rTPF and rFPF includes 54 LNs with pathology results from 34 patients in whom at least 1 method (R or M) was positive for LN metastasis (Table 6). LNs with pathology results but negative by both methods (R and M) were excluded in the analysis of rTPF and rFPF. Model-based

Table 7: M and R method results compared with follow-up neck ultrasound examination for interval change in lymph nodes without pathology results (n = 105)

M and R Methods	Suspicious	Not Suspicious	Stable	P Value[a]
+M method	0	5	5	.66
−M method	4	33	58	
+R method	2	11	18	.74
−R method	2	27	45	

[a] Fisher exact test.

estimates of rTPF and rFPF are presented in Table 6. Compared with the R method, the M method had a higher rTPF of 1.46 (95% CI, 1.12–1.91; P = .01) and a lower rFPF of 0.58 (95% CI, 0.36–0.92; P = .02). There were not significant interactions between method and sex or the main effects of sex for either outcome (P > .05 for each); thus, these parameters were dropped from the final models for both outcomes. Table 7 presents the M and R method results by follow-up US comparison and reports the results of the Fisher exact tests. The index R and M method results were not associated with suspicious findings at follow-up US examination (P = .74 and P = .66, respectively).

DISCUSSION

This study demonstrated that the sonographic predictive model (M method) outperforms the current standard of care (R method) in identifying PTC nodal metastasis in patients after thyroidectomy. The M method has higher true-positive rates and lower false-positive rates compared with the R method, providing evidence that using this model in clinical practice to determine the need for LN FNAB may more accurately identify patients who should undergo FNAB.

Also in our study, we followed LNs identified in the index US that did not undergo FNAB or surgical resection. We did not find an association between R or M method results and whether follow-up US revealed suspicious change in the LN, but this finding is not surprising given the typical indolent course of the disease and low incidence of abnormal nodes during the follow-up period. In keeping with this, of 95 nodes negative by the M method on the index US, only 4.21% (4/95 nodes) changed to the suspicious category on follow-up US examination, suggesting that 95.79% (91/95) of M method–negative nodes were either benign or had less aggressive PTC metastasis. These patients might be managed nonsurgically with active surveillance because surgical complications are typically higher with repeat surgery in the same compartment.

The predictive model (M method) can help achieve the ultimate goal of screening cervical LNs with neck sonography in the intermediate-to-high risk patient population with PTC, by improving detection of LN recurrence. Also, the M method may improve the role of surveillance neck sonography in a patient population with low-risk PTC that can be monitored for local recurrence with less aggressive strategies.

This study had several strengths, including the use of reference standard pathology results, prospective design, and analytic methods unbiased when reference standard results were incomplete. Although most patients did not have R or M method-positive nodes, thereby precluding valid estimates of sensitivity and specificity, the study design and analytic methods used in this study are a valid and ethical alternative to complete ascertainment of reference standard results.

There are several limitations of our study. The predictive model (Table 1) is based on sonographic features only and does not take into account the risk of recurrence and serum Tg level. In low- and intermediate-risk patients, the risk of lymph node recurrence is low (<2%) in patients with undetectable serum Tg levels and is much higher in those with detectable/elevated serum Tg levels.[4] One of the reading radiologists for the R method was also 1 of the 2 radiologists performing LN FNAB. Our study is prospective but is of a small cohort from a single institution, limiting generalizability of the results. Also, the diagnostic performance of the predictive model in prethyroidectomy assessment of cervical lymph node metastasis in PTC is uncertain because our study included only patients after thyroidectomy; another study is needed to determine whether our results are generalizable to other patient groups.

Several studies in the literature have discussed the predictors of metastatic disease in differentiated thyroid cancer, though there were no other studies applying a sonographic predictive model for the population of patients with PTC. Other studies have noted the size, central location, echo pattern, and Doppler flow or abnormal enhancement on CT as predictors of metastatic disease: Alzahrani et al[16] found the size (7.5 mm) and central location of cervical LNs as the most important predictors of the presence of metastatic disease. In the study by Aribaş et al,[17] central location and hypoechogenicity with loss of hilum in the lateral neck were predictors of malignancy. Chammas et al[18] found that an altered vascularization (resistive index of 0.77 as a cutoff value), a short axis of ≥0.9 cm, an abnormal hilum, and a heterogeneous echotexture were the most accurate sonographic predictors of LN malignancy, with a diagnostic accuracy near 80%. However, pulse Doppler evaluation of a LN to obtain the resistive index can be time-consuming and requires additional technical skill. Liu et al[19] developed a scoring system and mathematic model using CT to diagnose metastatic central compartment nodes in PTC.[19] Using 4 risk factors of LN metastasis on CT, including cystic or necrotic change, abnormal enhancement, nodal grouping ≥2, and nodal area ≥30.00 mm^2, this method had a sensitivity and specificity of 68.8% and 85.9%, respectively, but it has the downside of exposing patients to ionizing radiation and an iodinated contrast agent.

CONCLUSIONS

The sonographic predictive model demonstrated higher true-positive rates and lower false-positive rates compared with radiologists' standard assessment of LNs to detect PTC metastasis in patients after thyroidectomy. Incorporation of this sonographic predictive model in clinical practice may improve the diagnostic accuracy in detecting PTC nodal metastasis and thereby reduce the number of unnecessary LN FNABs, especially in low-risk patients with PTC. A large multi-institutional study is needed to further validate this sonographic predictive model.

methodology workshop faculty to help design the research method for this clinical trial. The authors also acknowledge the support and assistance provided by the research coordinators of the Radiology Department at the University of Colorado.

REFERENCES

1. National Cancer Institute. Cancer Stat Facts: Thyroid Cancer. http://seer.cancer.gov/statfacts/html/thyro.html. Accessed April 20, 2017
2. Davies L, Welch HG. **Current thyroid cancer trends in the United States.** *JAMA Otolaryngol Head Neck Surg* 2014;140:317–22 CrossRef Medline
3. Enewold L, Zhu K, Ron E, et al. **Rising thyroid cancer incidence in the United States by demographic and tumor characteristics, 1980–2005.** *Cancer Epidemiol Biomarkers Prev* 2009;18:784–91 CrossRef Medline
4. 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 CrossRef Medline
5. Rosário P, Tavares W, Borges M, et al. **Ultrasonographic differentiation of cervical lymph nodes in patients with papillary thyroid carcinoma after thyroidectomy and radioiodine ablation: a prospective study.** *Endocr Pract* 2014;20:293–98 CrossRef Medline
6. 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;92:3590–94 CrossRef Medline
7. Landry CS, Grubbs EG, Busaidy NL, et al. **Cystic lymph nodes in the lateral neck as indicators of metastatic papillary thyroid cancer.** *Endocr Pract* 2011;17:240–44 CrossRef Medline
8. Sohn YM, Kwak JY, Kim EK, et al. **Diagnostic approach for evaluation of lymph node metastasis from thyroid cancer using ultrasound and fine-needle aspiration biopsy.** *AJR Am J Roentgenol* 2010; 194:38–43 CrossRef Medline
9. Patel NU, McKinney K, Kreidler SM, et al. **Ultrasound-based clinical prediction rule model for detecting papillary thyroid cancer in cervical lymph nodes: a pilot study.** *J Clin Ultrasound* 2016;44:143–51 CrossRef Medline
10. Som PM, Curtin HD, Mancuso AA. **Imaging-based nodal classification for evaluation of neck metastatic adenopathy.** *AJR Am J Roentgenol* 2000;174:837–44 CrossRef Medline
11. Torres MR, Nóbrega Neto SH, Rosas RJ, et al. **Thyroglobulin in the washout fluid of lymph-node biopsy: what is its role in the follow-up of differentiated thyroid carcinoma?** *Thyroid* 2014;24:7–18 CrossRef Medline
12. Snozek CL, Chambers EP, Reading CC, et al. **Serum thyroglobulin, high-resolution ultrasound, and lymph node thyroglobulin in diagnosis of differentiated thyroid carcinoma nodal metastases.** *J Clin Endocrinol Metab* 2007;92:4278–81 CrossRef Medline
13. Moon JH, Kim YI, Lim JA, et al. **Thyroglobulin in washout fluid from lymph node fine-needle aspiration biopsy in papillary thyroid cancer: large-scale validation of the cutoff value to determine malignancy and evaluation of discrepant results.** *J Clin Endocrinol Metab* 2013;98:1061–68 CrossRef Medline
14. Pepe MS, Alonzo TA. **Comparing disease screening tests when true disease status is ascertained only for screen positives.** *Biostatistics* 2001;2:249–60 CrossRef Medline
15. Mickey RM, Greenland S. **The impact of confounder selection criteria on effect estimation.** *Am J Epidemiol* 1989;129:125–37 CrossRef Medline
16. Alzahrani AS, Alsuhaibani H, Salam SA, et al. **Diagnostic accuracy of high-resolution neck ultrasonography in the follow-up of differentiated thyroid cancer: a prospective study.** *Endocr Pract* 2005;11: 165–71 CrossRef Medline
17. Aribaş BK, Arda K, Çiledağ N, et al. **Predictive factors for detecting malignancy in central and lateral cervical lymph nodes in papillary carcinoma of the thyroid.** *Asia Pac J Clin Oncol* 2011;7:307–14 CrossRef Medline
18. Chammas MC, Macedo TA, Lo VW, et al. **Predicting malignant neck lymphadenopathy using color duplex sonography based on multivariate analysis.** *J Clin Ultrasound* 2016;44:587–94 CrossRef Medline
19. Liu T, Su X, Chen W, et al. **A mathematical model using computed tomography for the diagnosis of metastatic central compartment lymph nodes in papillary thyroid carcinoma.** *Eur Radiol* 2014;24: 2827–34 CrossRef Medline