Surgery 165 (2019) 69–74

Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

The association between the ultrasonography TIRADS classification system and surgical pathology among indeterminate thyroid nodules ☆

Zeyad T Sahli, MD[a], Farah Karipineni, MD[a], Jen-Fan Hang, MD[a,e], Joseph K Canner, MHS[a], Aarti Mathur, MD, FACS[a], Jason D Prescott, MD, PhD[a], Sheila Sheth, MD[c], Syed Z Ali, MD[b,c], Martha A Zeiger, MD, FACS[d,*]

[a] Endocrine Surgery, Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD
[b] Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD
[c] Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD
[d] Department of Surgery, University of Virginia School of Medicine, Charlottesville, VA
[e] Department of Pathology and Laboratory Medicine, Taipei Veterans General Hospital, Taipei, Taiwan

ARTICLE INFO

Article history:
Accepted 3 April 2018
Available online 8 November 2018

ABSTRACT

Background: A high proportion of cytologically indeterminate, Afirma-suspicious thyroid nodules are benign. The Thyroid Imaging Reporting and Data System was proposed by the American College of Radiology in 2015 to determine appropriate management of thyroid nodules in a standardized fashion. Our aim was to determine the diagnostic value of the Thyroid Imaging Reporting and Data System in cytologically indeterminate and Afirma-suspicious nodules.

Methods: We retrospectively queried cytopathology archives for retrospectively for thyroid fine-needle aspiration specimens obtained between February 2012 and September 2016 that were associated with the following: (1) indeterminate diagnosis, (2) ultrasonographic imaging at our institution, (3) an Afirma Gene Expression Classifier–suspicious result, and (4) surgery at our institution. We then calculated the diagnostic value of the Thyroid Imaging Reporting and Data System in predicting surgical pathology.

Results: Our cohort consisted of 133 nodules among 131 patients who underwent thyroid surgery for cytologically indeterminate, Afirma-suspicious nodules. A total of 9 nodules (6.8%) were assigned TR2 "not suspicious," 25 (18.8%) TR3 "mildly suspicious," 81 (60.9%) TR4 "moderately suspicious," and 18 (13.5%) TR5 "highly suspicious." Among our cohort, the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of the Thyroid Imaging Reporting and Data System was 71.4%, 38.1%, 40.2%, 69.6%, and 50.4%, respectively.

Conclusion: Among cytologically indeterminate and Afirma-suspicious nodules, the Thyroid Imaging and Reporting and Data System was a poor predictor of final surgical pathology. Additional prospective studies are needed to validate these findings.

Introduction

Accurately predicting benign versus malignant thyroid nodules has been an ongoing clinical challenge. The widespread use of high-resolution ultrasonography for the detection of thyroid nodules has led to the increased detection of thyroid nodules, with an estimated prevalence as high as 67% in the general population.[1–3] Despite the additional diagnostic value of fine-needle aspiration (FNA), over 30% of thyroid FNAs with wide institutional variability[4,5] will be assigned to an indeterminate cytopathologic category only. These latter categories include Bethesda III or atypia of undetermined significance (AUS) and Bethesda IV or suspicious for follicular neoplasm (SFN). Over the past decade, molecular marker panels, such as Afirma Gene Expression Classifier (GEC), have been developed and utilized widely as adjuncts in the differential diagnosis of these lesions. Despite reported high sensitivity (Se; 92%) of the Afirma GEC, it has a low positive predictive value (PPV; 47%) with benign surgical pathology in more than half of those identified as Afirma-suspicious.[6] Therefore, additional and more robust means are needed to assist in identifying benign nodules and subsequently decrease the number of unnecessary operations performed.

☆ Presented at the annual meeting of the American Association of Endocrine Surgeons in Durham, North Carolina, May 6–8, 2018.

* Reprint requests: Martha A Zeiger, MD, FACS, Department of Surgery, University of Virginia School of Medicine, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22908.

E-mail address: maz4c@virginia.edu (M.A. Zeiger).

https://doi.org/10.1016/j.surg.2018.04.094
0039-6060/

Table 1
TIRADS classification.

TIRADS criteria	Nodule description
Composition	Cystic (0)
	Spongiform (0)
	Mixed cystic and solid (1)
	Solid (2)
Echogenicity	Anechoic (0)
	Hyperechoic or isoechoic (1)
	Hypoechoic (2)
	Very hypoechoic (3)
Shape	Wider-than-tall (0)
	Taller-than-wide (3)
Margin	Smooth (0)
	Ill-defined (0)
	Lobulated or irregular (2)
	Extra-thyroidal extension (3)
Echogenic foci	None or large comet-tail artifacts (0)
	Macrocalcifications (1)
	Peripheral (rim) calcification (2)
	Punctate echogenic foci (3)

TIRADS classification (points)	Description	Management
TIRADS 1 (0)	Benign	No FNA
TIRADS 2 (2)	Not suspicious	No FNA
TIRADS 3 (3)	Mildly suspicious	FNA if ≥2.5 cm; active surveillance if ≥1.5 cm
TIRADS 4 (4–6)	Moderately suspicious	FNA if ≥1.5 cm; active surveillance if ≥1 cm
TIRADS 5 (≥7)	Highly suspicious	FNA if ≥1.0 cm; active surveillance if ≥0.5 cm

In 2015, the Thyroid Imaging Reporting and Data System (TIRADS), a new ultrasonographic classification system for thyroid nodules, was proposed by the American College of Radiology to determine when to use FNA or when to recommend active surveillance of nodules.[7] In response to conflicting recommendations from several societies, the American College of Radiology proposed that TIRADS could provide clinicians with evidence-based recommendations, defined by sonographic features of a given nodule, to improve both patient management and cost-effectiveness by avoiding unnecessary FNAs and to decrease variation in the reporting of thyroid nodules.[8] TIRADS includes a standardized scoring system using several thyroid ultrasonographic characteristics: composition, echogenicity, shape, margin, and echogenic foci. TIRADS has 5 different categories (TR1–TR5), each level is associated with an increased likelihood of malignancy and therefore more aggressive recommended clinical management (Table 1).[7]

Given the high proportion of cytologically indeterminate, Afirma-suspicious nodules that ultimately are benign on final pathology, our study sought to evaluate the accuracy and predictive value of TIRADS in this subset of nodules to differentiate between benign and malignant pathology.

Methods

The current study was approved by the Institutional Review Board of the Johns Hopkins University School of Medicine. We searched retrospectively a prospectively maintained cytopathology database for consecutive thyroid FNA specimens obtained at the Johns Hopkins Hospital between February 2012 and September 2016. Thyroid nodules with an indeterminate diagnosis (AUS or SFN), available ultrasonographic imaging results, and Afirma GEC testing with a suspicious result were included. Patients who did not undergo operation or did not have ultrasonography performed at our institution were excluded. Onsite evaluation of FNA samples are performed routinely by either an experienced cytotechnologist or an attending cytopathologist for each case. Afirma GEC was performed after the final cytologic diagnosis of either AUS or SFN was made. Additional information collected and recorded included patient demographics, history of Hashimoto's disease, ultrasonography, cytology and pathology reports, and clinic and operative notes.

Table 2
Patient ($n = 131$) and thyroid nodule characteristics ($n = 133$).

Age	
Mean (SD, range)	52.2 (13.9, 17–80)
Race, no. (%)	
Asian	2 (1.5)
Black	21 (15.8)
Hispanic	1 (0.8)
White	96 (72.2)
Other	13 (9.8)
Sex, *n* (%)	
Male	38 (28.6)
Female	95 (71.4)
Hashimoto thyroiditis, *n* (%)	10 (7.5)
Nodules	
Size on ultrasonography, cm	
Mean (range)	2.4 (0.5–8.0)
Nodule size, no. of nodules (%)	
<1 cm	3 (2.2)
1.0–1.9 cm	58 (43.6)
2.0–2.9 cm	37 (27.8)
3.0–3.9 cm	21 (15.8)
≥4.0 cm	14 (10.5)
Bethesda category	
III	104 (78.2)
IV	29 (21.8)

SD, standard deviation

Our primary endpoint was to assess the diagnostic value of TIRADS, including sensitivity, specificity, PPV, and negative predictive value (NPV), in predicting surgical pathology. We also assessed whether the TIRADS ultrasonography classification would have changed the management of cytologically indeterminate, Afirma-suspicious thyroid nodules, the inference being that a benign reading would have led to no further workup (ie, whether the previously described nodules would otherwise have not undergone resection because of their corresponding TIRADS classification). All analyses were performed using STATA 14.0/MP for Linux (College Station, Texas).

Pathology slide re-review

For cases diagnosed as Follicular variant of papillary thyroid carcinoma (FVPTC) on surgical resection specimens, all histopathologic slides were retrieved and re-reviewed separately by 2 pathologists to identify tumors that would now be classified as Non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP). The diagnostic criteria for NIFTP were based on the new consensus guidelines and 20 of FVPTC met criteria.[9] For the purposes of our statistical analysis, pathology was considered either benign or malignant. Because of the indolent nature of NIFTP and current debate about whether it should be considered benign or malignant, we performed 2 analyses, with NIFTP considered as either benign or malignant.

Review of the ultrasonography imaging

For all cytologically indeterminate and Afirma-suspicious thyroid nodules, ultrasonographic imaging was retrieved and reviewed retrospectively in a blinded fashion by an experienced radiologist (S.S.) at our institution to stage the nodules according to the TIRADS classification. The nodules of interest were matched carefully according to the information regarding location and size recorded in the FNA and histopathologic reports. A search of the

Table 3
Ultrasonographic characteristics of overall cohort.

Ultrasound characteristics	All patients, *n*	Surgical pathology	
		Benign*	Malignant
Composition			
Mixed	12	11	1
Solid	121	92	29
Echogenicity			
Hyper/iso	41	36	5
Hypo	86	62	24
Very hypo	5	5	0
Unknown	1	0	1
Shape			
Wider-than-tall	126	102	24
Taller-than-wide	7	1	6
Margin			
Smooth	118	95	23
Ill-defined	3	1	2
Lobulated or irregular	12	7	5
Echogenic foci			
None or large comet-tail artifacts	109	90	19
Macrocalcifications	7	4	3
Peripheral (rim) calcifications	4	3	1
Punctate echogenic foci	13	6	7

* NIFTP considered benign

surgical pathology database was performed to identify the corresponding histopathologic diagnosis and operative procedure.

Results

Patient characteristics

Our cohort consisted of 133 nodules among 131 patients. The mean patient age was 52.2 years (range 17–80 years), and the majority of patients were white (72.2%) and female (71.4%). A total of 104 nodules (78.2%) were Bethesda III on FNA, and 29 (21.8%) were Bethesda IV. Table 2 lists the patient demographics of our study population. A total of 10 patients (7.5%) had a history of Hashimoto's thyroiditis.

Ultrasound characteristics

Sonographic features are summarized in Table 3. Tables 4 and 5 show the association between TIRADS classification and pathology; surgical pathology is summarized in Table 6. The mean thyroid nodule size was 2.4 cm, ranging from 0.5 cm to 8.0 cm.

Table 6
Thyroid nodule surgical pathology ($n = 133$).

Benign, no. of nodules; $n = 103$, (77.4%)	
Nodular hyperplasia	18 (13.5)
Adenomatoid nodule	25 (18.8)
Follicular adenoma	38 (28.6)
Hürthle cell adenoma	1 (0.8)
Lymphocytic thyroiditis	2 (1.5)
NIFTP	19 (14.3)
Malignant, no. of nodules; $n = 30$, (22.6%)	
Papillary carcinoma, conventional	13 (9.8)
Papillary carcinoma, follicular variant	8 (6.0)
Follicular carcinoma	8 (5.3)
Hürthle cell carcinoma	1 (0.8)

The vast majority of nodules were solid (121, 91.0%), hypoechoic (86, 64.7%), wider-than-tall (126, 94.7%), had a smooth margin (118, 88.7%), and had no echogenic foci or large comet-tail artifacts (109, 82.0%). The majority of nodules were assigned a TR4 or "moderately suspicious" classification (81, 60.9%), followed by TR3 or "mildly suspicious" (25, 18.8%), TR5 or "highly suspicious" (18, 13.5%), and TR2 or "not suspicious" (9, 6.8%). No thyroid nodules among our cohort were assigned a TR1 or "benign" classification.

Histopathologic diagnosis

Considering NIFTP as benign, the majority of nodules had a benign pathology (103, 77.4%) as depicted in Table 6. The most common benign pathology was follicular adenoma (38, 28.6%) followed by adenomatoid nodule (25, 18.8%), NIFTP (19, 14.3%), nodular hyperplasia (18, 13.5%), lymphocytic thyroiditis (2, 1.5%), and Hürthle cell adenoma (1, 1.5%). Malignant pathology was found in 30 nodules (22.6%). Among the malignancies, the most common pathology was classic papillary thyroid cancer (13, 9.8%) followed by follicular variant FVPTC (8, 6.0%), follicular carcinoma (8, 6.0%), and Hürthle cell carcinoma (1, 0.8%). Conversely, classifying NIFTP as malignant, 63.8% of thyroid nodules were benign and 36.8% of thyroid nodules were malignant.

Applying TIRADS

Considering NIFTP as benign (Table 7), the sensitivity, specificity, PPV, NPV, and accuracy of TIRADS was 73.3% (95% confidence

Table 4
TIRADS classification of overall cohort and risk of malignancy (ROM), NIFTP benign.

TIRADS	All patients, *n*	Cytopathology		Surgical pathology		ROM
		III	IV	Benign	Malignant	
Benign (TR1)	0	0	0	0	0	
Not suspicious (TR2)	9	8	1	8	1	11.1%
Mildly suspicious (TR3)	25	19	6	23	2	8.0%
Moderately suspicious (TR4)	81	62	19	63	18	22.2%
Highly suspicious (TR5)	18	15	3	9	9	50.0%

Table 5
TIRADS classification of overall cohort and risk of malignancy (ROM), NIFTP malignant.

TIRADS	All patients, *n*	Cytopathology		Surgical pathology		ROM
		III	IV	Benign	Malignant	
Benign (TR1)	0	0	0	0	0	0.0%
Not suspicious (TR2)	9	8	1	7	2	22.2%
Mildly suspicious (TR3)	25	19	6	18	7	28.0%
Moderately suspicious (TR4)	81	62	19	53	28	34.6%
Highly suspicious (TR5)	18	15	3	6	12	66.7%

Table 7
TIRADS' diagnostic value in predicting surgical pathology (NIFTP benign).

		Surgical pathology			
		Benign	Malignant	Diagnostic value	
TIRADS	No workup	38	8	**Se** 73.3 (95% CI: 54.1%–87.7%)	**PPV** 25.3 (95% CI: 20.7%–30.5%)
	Workup	65	22	**Sp** 36.9 (95% CI: 27.6%–47.0%)	**NPV** 82.6 (95% CI: 71.4%–90.1%)
				Accuracy 45.1% (95% CI: 36.5%–54.0%)	

Sp, specificity

Table 8
TIRADS' diagnostic value in predicting surgical pathology (NIFTP malignant).

		Surgical pathology			
		Benign	Malignant	Diagnostic value	
TIRADS	No workup	32	14	**Se** 71.4 (95% CI: 56.7%–83.4%)	**PPV** 40.2 (95% CI: 34.5%–46.2%)
	Workup	52	35	**Sp** 38.1 (95% CI: 27.7%–49.3%)	**NPV** 69.6 (95% CI: 57.6%–79.4%)
				Accuracy 50.4% (95% CI: 41.6%–59.2%)	

Sp, specificity

interval [CI]: 54.1%–87.7%), 36.9% (95% CI: 27.6%–47.0%), 25.3% (95% CI: 20.7%–30.5%), 82.6% (95% CI: 71.4%–90.1%), and 45.1% (95% CI: 36.5%–54.0%), respectively, in predicting malignancy. Applying the TIRADS criteria to our patient cohort, 46 nodules (34.6%) would not have been further evaluated because of their ultrasonographic characteristics or size, and hence would not have undergone resection (Table 3). Of those 46 nodules, 8 (17.4%) were malignant and the application of TIRADS criteria alone would have resulted in failing to diagnose them as malignant. All 8 nodules that would not have undergone resection according to TIRADS were Bethesda III on FNA and included 1 TR2, 2 TR3, 4 TR4, and 1 TR5 nodules; however, 38 benign nodules (82.6%) would have been spared further FNA evaluation and thus ultimate resection.

Considering NIFTP as malignant (Table 8), the sensitivity, specificity, PPV, NPV, and accuracy of TIRADS in predicting malignancy would have been 71.4% (95% CI: 56.7%–83.4%), 38.1% (95% CI: 27.7%–49.3%), 40.2% (95% CI: 34.5%–46.2%), 69.6% (95% CI: 57.6%–79.4%), and 50.4% (95% CI: 41.6%–59.2%), respectively. Of those 46 nodules (34.6%) that would not have undergone further workup, 14 (30.4%) were malignant. Of those malignant nodules that would not have undergone resection, 13 (92.9%) were Bethesda III on FNA and according to TIRADS were classified as 2 TR2, 4 TR3, 7 TR4, and 1 TR5. The application of TIRADS would have resulted in failing to diagnose 14 malignant nodules; however, 32 benign nodules (24.0%) would have not undergone FNA and been resected.

Discussion

To the best of the authors' knowledge, our study represents the first investigation in the literature addressing the utility of the TIRADS classification in patients with cytologically indeterminate, Afirma-suspicious thyroid nodules. Because of the novelty of the TIRADS classification, its clinical utility in relation to either cytopathology or surgical pathology is still under investigation with varying results. Our results show that, among cytologically indeterminate and Afirma-suspicious nodules, TIRADS would not have been a reliable indicator for further workup of the nodule. Endorsement of the TIRADS classification system by the American College of Radiology is intended to allow for an accurate clinical–pathologic correlation, but further evidence is needed.

The majority of published literature evaluates the diagnostic value of TIRADS and thyroid nodule cytopathology. Contrary to our findings, a prospective study by Friedrich-Rust et al including 114 patients showed a high NPV (92%–100%) for TIRADS categories 4 and 5 in excluding malignancy in the diagnostic workup of thyroid nodules, but only a "fair" degree of interobserver agreement (Cohen's kappa = 0.27, $P < .05$) was reported between 3, blinded, independent observers.[10] Moreover, a retrospective study of 100 consecutive cases comparing single-surgeon–performed ultrasonographic TIRADS findings to cytopathology, including all Bethesda categories, found a concordance rate of 83% with a sensitivity, specificity, and NPV of 70.6%, 90.4%, and 93.8%, respectively.[11] Similar to our findings, another study including 180 patients with the aim of determining the level of concordance between TIRADS and cytopathology found the greatest concordance among TIRADS 2 and 4 and Bethesda II and IV to be only 23.3% and 18.3%, respectively.[12] The differences between our reported results and those discussed in the literature may also be magnified because of the reported variable success of TIRADS prediction among AUS/Follicular Lesion of Undetermined Significance (FLUS) categories.[13] A study by Yoon et al observed that suspicious ultrasonographic features are useful in predicting malignancy among AUS subcategories but not in FLUS subcategories. For this reason, further subcategorization into separate AUS and FLUS cytology may be helpful in deciding on appropriate management of thyroid nodules when using TIRADS.[13]

There is a paucity of literature evaluating the association of TIRADS and surgical pathology. A prospective study by Horvath et al of 210 patients with 502 nodules comparing TIRADS to final pathology reported a risk of malignancy among TIRADS 2, 3, 4, and 5 to be 0%, 1.8%, 76.1%, and 98.9%, respectively. Using TIRADS 4–5 to perform FNA, the authors reported a sensitivity, specificity, PPV, and NPV of 99.6%, 74.35%, 82.1%, and 99.4%.[14] The difference in the diagnostic value of our results may be explained by our inclusion of parameters of nodule size based on the recommendations of the American College of Radiology as opposed to other variations of TIRADS that do not take into account nodule size (Table 1).[7]

Our study has several limitations. First, our report is a single-institutional study. An inherent weakness in this study also is its retrospective nature. Despite having an experienced radiologist, the use of more than 1 radiologist for interpretation of the ultrasonographic -imaging may have been helpful because image analysis can differ among radiologists and the extent of interobserver variability regarding TIRADS classification is unknown. As previously mentioned, there is active research regarding the diagnostic value of TIRADS among Bethesda category III versus category IV nodules; however, our cohort only included 29 Bethesda category IV nodules and, for this reason, a clinically meaningful comparison of the TIRADS diagnostic value between the 2 categories cannot be determined accurately.

In conclusion, our results show that among the subset of cytologically indeterminate and Afirma-suspicious nodules, TIRADS was

not a reliable indicator of malignancy. Because of its low sensitivity (71.4%–73.3%) for predicting malignancy and low specificity (36.9%–38.1%) for predicting benign pathology, our study showed no clinical benefit to using the TIRADS classification in this particular cohort of patients. Large prospective studies are needed to evaluate the diagnostic value of TIRADS in this subset of thyroid nodules. Based on our findings, we suggest that the TIRADS results be correlated with a gold standard—whether cytopathology or surgical pathology—before adoption of TIRADS becomes more widespread among clinicians.

Conflicts of interest

The authors have indicated that they have no conflicts of interest regarding the content of this article.

References

1. Dean DS, Gharib H. Epidemiology of thyroid nodules. *Bes Pract Res Clin Endocrinol Metab*. 2008;22:901–911.
2. Russ G, Leboulleux S, Leenhardt L, Hegedüs L. Thyroid incidentalomas: epidemiology, risk stratification with ultrasound and workup. *Eur Thyroid*. 2014;3:154–163.
3. Davies L, Ouellette M, Hunter M, Welch HG. The increasing incidence of small thyroid cancers: where are the cases coming from? *The Laryngoscope*. 2010;120:2446–2451.
4. Kholová I, Ludvíková M. Thyroid atypia of undetermined significance or follicular lesion of undetermined significance: an indispensable Bethesda 2010 diagnostic category or waste garbage? *Acta Cytologica*. 2014;58:319–329.
5. Straccia P, Rossi ED, Bizzarro T, Brunelli C, Cianfrini F, Damiani D, et al. A meta–analytic review of the Bethesda System for Reporting Thyroid Cytopathology: has the rate of malignancy in indeterminate lesions been underestimated? *Cancer Cytopathology*. 2015;123:713–722.
6. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;367:705–715.
7. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol*. 2017;14:587–595.
8. Grant EG, Tessler FN, Hoang JK, Langer JE, Beland MD, Berland LL, et al. Thyroid ultrasound reporting lexicon: white paper of the ACR thyroid imaging, reporting and data system (TIRADS) committee. *J Am Coll Radiol*. 2015;12:1272–1279.
9. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncology*. 2016;2:1023–1029.
10. Friedrich-Rust M, Meyer G, Dauth N, Berner C, Herrmann E, Zeuzem S, et al. Interobserver agreement of Thyroid Imaging Reporting and Data System (TIRADS) and realtime elastography for the assessment of thyroid nodules. *Ultraschall Med*. 2013;34 WS_SL11_01.
11. Singaporewalla R, Hwee J, Lang T, Desai V. Clinico-pathological correlation of thyroid nodule ultrasound and cytology using the TIRADS and Bethesda classifications. *World J Surg*. 2017;41:1807–1811.
12. Vargas-Uricoechea H, Meza-Cabrera I, Herrera-Chaparro J. Concordance between the TIRADS ultrasound criteria and the BETHESDA cytology criteria on the non-toxic thyroid nodule. *J Thyroid Res*. 2017;10:1.
13. Yoon JH, Kwon HJ, Kim EK, Moon HJ, Kwak JY. Subcategorization of atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS): a study applying Thyroid Imaging Reporting and Data System (TIRADS). *J Clin Mol Endocrinol*. 2016;85:275–282.
14. Horvath E, Silva CF, Majlis S, Rodriguez I, Skoknic V, Castro A, et al. Prospective validation of the ultrasound based TIRADS (Thyroid Imaging Reporting and Data System) classification: results in surgically resected thyroid nodules. *Eur J Radiol Open*. 2017;27:2619–2628.

Discussion

Dr Quan-Yang Duh (San Francisco, CA): The question I have for you regards the methods. I don't think you can truly calculate your sensitivity and specificity because of the patient selection that occurred. Remember, you did not analyze all of the patients that had an ultrasound. You took only patients that had ultrasound and needle biopsy. So whatever number that you come up with can only be applied to that particular group of patients.

I think you have to have a prospective study, or you have to include all of those patients that have ultrasound but never had biopsy.

Dr Zeyad T. Sahli: I agree totally with that conclusion. And the reason for studying this subset of patients is that the risk of malignancy among this group is unknown. As you know, indeterminate cytology has a variable risk of malignancy among institutions and relates to prevalence of cancer in that population.

Also, Veracyte Afirma has a variable positive predictive value. So what we wanted to do was to use this information as a third adjunct to see if it would be able to predict surgical pathology. But absolutely, you are correct.

Dr Quan-Yang Duh (San Francisco, CA): I would encourage you to do the calculation, adding in the number of patients that had ultrasound but not a biopsy, because then you have more accurate sensitivity and specificity numbers if you just want to use TIRADS.

Dr Barry Inabnet (New York, NY): Congratulations. Ultrasound is an invaluable part of what we do to evaluate patients.

In your study, these were radiologists doing the examinations. In practice, it is often endocrine surgeons that perform the ultrasound. We have 1 advantage over the radiologists as we are able to correlate what we see in the ultrasound image with what we see in the operating room. That provides a sixth dimension that is not captured in TIRADS.

Have you given that some consideration? Perhaps as you look towards a prospective study, are you going to use surgeon-based TIRADS interpretation as the next evolution of your research?

Dr Zeyad T. Sahli: I think 1 of the things that we would want to look at is the extent of interobserver variability which has been reported from fair to good, depending on specific criteria in the TIRADS scoring system.

In relation to how we could use TIRADS, I think that's a great idea and something that we would love to pursue.

Dr Shelby Holt (Dallas, TX): I wonder how the TIRADS classification relates to the sonographic patterns that are well-described in the latest iteration of the ATA guidelines.

Dr Zeyad T. Sahli: Are you asking how TIRADS compares to the ATA guidelines?

Dr Shelby Holt (Dallas, TX): The sonographic patterns that are elucidated in the ATA guidelines—are they similar to TIRADS? Do they correlate with those patterns? Or is it different? Have you looked at the ATA-described sonographic patterns alone in your cohort of patients?

Dr Zeyad T. Sahli: This study in particular did not look at the differences between the ATA and TIRADS scores, but there have been studies that show that they do have similar correlation. This wasn't in the scope of our study.

Dr Dina Elaraj (Chicago, IL): Thank you for that very interesting presentation.

Your cohort was a group of patients who had undergone FNA biopsy with indeterminate cytology and Afirma testing in whom

you tried to use the TIRADS classification to inform risk of malignancy.

Can you comment as to how the decision for molecular testing is made at your institution? And would it be possible instead to use the TIRADS classification to help inform the decision to pursue gene expression testing after that?

Dr Zeyad T. Sahli: It depends on patient and surgeon preference. Ultimately, the move will be towards following the TIRADS classification. If the nodule does not meet TIRADS criteria, there is some hesitance to perform an FNA.

But for this particular subset of patients, there were multiple surgeons involved, so it did differ by surgeon and patient preference for the decision to do so.

Dr Mahsa Javid (Charleston, SC): Having spoken to a couple of the radiologists involved in the TIRADS guideline development, there seems to be a different philosophy about what's going on here. It seems to me that they think we are doing far too many biopsies. And it's not just the fact that we are trying to pick up every malignancy. Perhaps low-risk malignancies can be followed up on future ultrasounds.

Of the malignancies that you found at pathology, did these have high-risk characteristics? Did they have lymph node metastases? Do you agree that we are doing too many biopsies and perhaps we should have a less-aggressive approach if the patient has a small microcarcinoma?

Dr Zeyad T. Sahli: That's exactly why TIRADS was initiated, because 1 of the aims was to decrease unnecessary FNAs.

In terms of the extent of malignancy and extent of metastases, that was not something that we analyzed because we were studying the nodule features in particular. But that would be something that we could go back and take a look at.