Assessment of the American College of Radiology Thyroid Imaging Reporting and Data System for Thyroid Nodule Malignancy Risk Stratification in a Pediatric Population

Jennifer E. Lim-Dunham[1]
Iclal Erdem Toslak[1]
Michael P. Reiter[1]
Brendan Martin[2]

OBJECTIVE. The purpose of this study is to assess the diagnostic performance of the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) for malignancy risk in pediatric thyroid nodules.

MATERIALS AND METHODS. Two radiologists reviewed ultrasound images of 74 tissue-proven thyroid nodules in 62 children. Points were given for individual features and then added to determine the ACR TI-RADS category, ranging from 1 (benign) to 5 (high suspicion). Kappa coefficients were generated to assess intra- and interobserver agreement. Generalized linear mixed-effects models were used to estimate the odds of malignancy with construction of a supplementary ROC curve.

RESULTS. Fifty-four nodules were benign and 20 were malignant, with a median ACR TI-RADS category of 4 (interquartile range, 4–5). Nineteen of 20 (95.0%) malignant nodules were rated as TI-RADS category 4 or 5. There was substantial intraobserver agreement ($\kappa = 0.69$–0.77; $p < 0.001$) and moderate interobserver agreement ($\kappa = 0.37$; $p = 0.002$) for TI-RADS category. Univariable analysis showed that, with every 1-unit increase of TI-RADS category, the likelihood of malignancy increased 2.63 times (95% CI, 1.08–6.41; $p = 0.03$). After adjusting for nodule size, TI-RADS category remained marginally associated with malignancy (adjusted odds ratio, 2.27; 95% CI, 0.93–5.54; $p = 0.07$). The AUC was 0.75 (95% CI, 0.64–0.86). An optimal cut point of TI-RADS category 5 was selected, with TI-RADS category 5 nodules 10.44 times (95% CI, 2.71–40.21; $p < 0.0001$) more likely than categories 1–4 nodules to be malignant.

CONCLUSION. ACR TI-RADS discriminates well between malignant and benign nodules in a pediatric population, particularly at TI-RADS category 5.

The intent of thyroid ultrasound stratification guidelines is to identify nodules that are at high risk for malignancy on the basis of ultrasound appearance and that warrant fine-needle aspiration biopsy (FNAB). A wide array of guidelines have been proposed but are not universally accepted [1, 2]. To add to the confusion, many of these systems contain the term "Thyroid Imaging, Reporting and Data System" or "TIRADS", but they are actually distinct from each other [3–5]. One of these classification schemes is the Thyroid Imaging Reporting and Data System (TI-RADS) developed by the American College of Radiology (ACR) in 2017; this TI-RADS has the advantage of being an evidence-based multiinstitutional white paper representing consensus of expert opinion [6].

In contrast to guidelines such as the 2015 American Thyroid Association (ATA) Management Guidelines for Children with Thyroid Nodules and Differentiated Thyroid Cancer, the ACR TI-RADS was not developed specifically for the pediatric population, in whom thyroid nodules, although less common, are more frequently malignant [7–9]. Nevertheless, several aspects of the ACR TI-RADS make it attractive for use in children on a practical level. First, the system is not overly complex and is straightforward to apply with a clearly depicted graphic chart [6]. Its convenience is compounded by the advent of easily accessible online ACR TI-RADS calculators [10]. The point-based ACR TI-RADS assigns malignancy risk by allocating points corresponding to the number of suspicious ultrasound features; therefore, as opposed to some pattern-based systems, the appearance of all nodules can be accounted for in this system [6, 11, 12]. Finally, the system uses a structured lexicon that

aids in standardization of usage across practitioners of varied experience [13].

Since its introduction, the ACR TI-RADS has been validated in adults but has not yet been evaluated in children [14–17]. The purpose of this study is to apply the 2017 ACR TI-RADS to pediatric thyroid nodules and to provide an assessment of diagnostic performance.

Materials and Methods

This retrospective study was approved by the institutional review board of our medical center, and informed consent was waived. The study was HIPAA compliant.

Study Population

The patient database was extracted from our institution's electronic medical record system. Inclusion criteria were the presence of cyto- or histopathologically proven thyroid nodules in consecutive patients 18 years and younger who underwent ultrasound between 1996 and 2017 within 30 days of tissue sampling. The exclusion criterion was uncertainty in correlating the identity of the nodule seen at ultrasound with pathologic findings. Of the final study population of 62 patients with 74 nodules, 28 patients with 35 nodules underwent thyroidectomy. All but one of these patients also underwent ultrasound-guided FNAB before surgery. The remaining 34 patients with 39 nodules underwent FNAB and were followed up conservatively without further tissue sampling.

The patient population in the current study overlaps with those of two previous studies published by our group. One of these studies described 39 nodules in 33 patients who are also included in the current article, but that study evaluated these nodules in the context of the 2015 ATA Management Guidelines for Children [18]. The other study described 15 patients who are also included in the current report but it was limited to histopathologic correlation of sonographically detectable echogenic foci [19]. The current study differs in its specific focus on appraisal of the ACR TI-RADS guidelines.

Ultrasound Technique

Gray-scale sonography with color Doppler was performed using a variety of systems (Acuson Sequoia 512, XP128, and Aspen, all from Siemens Healthcare; and Logic E9, GE Healthcare) equipped with high-frequency (8-15–MHz) linear array transducers. Cine and still ultrasound images of the thyroid were electronically recorded and transferred to a PACS in DICOM format.
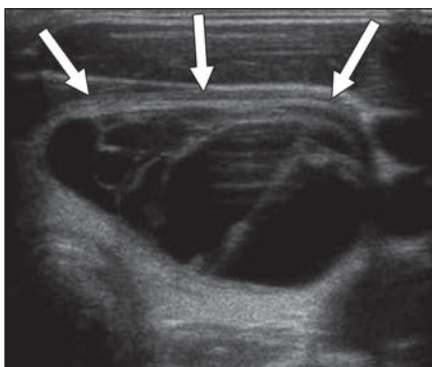
Ultrasound Image Analysis

Two board-certified pediatric radiologists blinded to tissue diagnosis independently evaluated ultrasound images on the PACS, with all 74 nodules reviewed twice, once in each of two separate sessions by each reader. Second sessions were performed at least 2 weeks after the first to minimize recall bias. A third session was performed with both radiologists to reach consensus for nodules for which there was disagreement.

The ACR TI-RADS lexicon was used to evaluate thyroid nodule characteristics. The individual parameters and possible descriptors, as well as corresponding ACR TI-RADS numeric point values, are as follows [6, 13]: for composition, cystic or almost completely cystic = 0 points, spongiform = 0 points, mixed cystic and solid = 1 point, and solid or almost completely solid = 2 points; for echogenicity, anechoic = 0 points, hyperechoic or isoechoic = 1 point, hypoechoic = 2 points, and very hypoechoic = 3 points; for shape, wider-than-tall = 0 points and taller-than-wide = 3 points; for margins, smooth = 0 points, ill-defined = 0 points, lobulated or irregular = 2 points, and extrathyroidal extension = 3 points; and for echogenic foci, none or large comet-tail artifacts = 0 points, macrocalcifications = 1 point, peripheral (rim) calcifications = 2 points, and punctate echogenic foci = 3 points.

Only one parameter each could be chosen for composition, echogenicity, shape, and margins features. On the other hand, for the echogenic foci feature, more than one parameter could be present and, therefore, all applicable parameters should be chosen [6]. An ACR TI-RADS category was then assigned to each nodule by adding the points from all of the ultrasound feature categories. The possible total point value for each nodule ranges from 0 to 17, with the exception of the value of 1, which is not a possible sum according to the criteria [6]. The five possible ACR TI-RADS categories and their corresponding point value ranges and degree of suspicion for malignancy are as follows:

TI-RADS category 1 (benign), 0 points; TI-RADS category 2 (not suspicious), 2 points; TI-RADS category 3 (mildly suspicious), 3 points; TI-RADS category 4 (moderately suspicious), 4–6 points; and TI-RADS category 5 (highly suspicious), 7 or more points [6]. Examples of each ACR TI-RADS category are shown in Figures 1–5. Nodule size was also recorded.
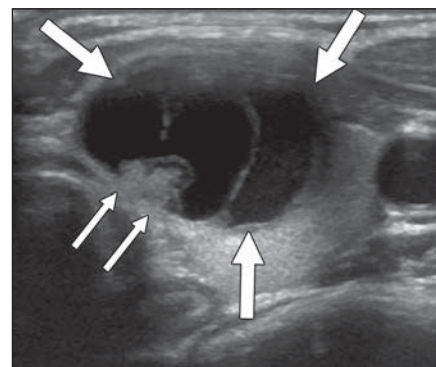
Fine-Needle Aspiration Biopsy Technique

Ultrasound-guided FNAB was performed by one of two board-certified pediatric radiologists, each with more than 10 years of experience. Care was taken to obtain samples from the solid part of the nodules. A pathologist was on site to verify the diagnostic adequacy of the sample. Technical limitations precluding ultrasound-guided FNAB included nodules that were 5 mm or smaller or were adjacent to major blood vessels.

Tissue Diagnosis

The decision to perform thyroidectomy was determined by the endocrine surgeon. Malignancy or benignity of nodules was determined with surgical pathologic results for patients who underwent thyroidectomy and with cytopathologic results for those who did not.

Samples taken from ultrasound-guided FNAB were classified on the basis of the Bethesda System for Reporting Thyroid Cytopathology, which categorizes nodules as follows: class I, nondiagnostic; class II, benign; class III, atypia or follicular lesion of undetermined significance; class IV, follicular neoplasm or suspicion for a follicular neoplasm; class V, suspicious for malignancy; and class VI, malignant [20]. For the purpose of this study, Bethesda classes II and III were considered benign, and Bethesda classes IV, V, and VI were considered malignant.

Fig. 1—18-year-old woman with American College of Radiology Thyroid Imaging Reporting and Data System category 1 nodule. Ultrasound shows 3.7-cm left almost completely cystic nodule (*arrows*) with thin internal septations. Cytologic analysis revealed Bethesda class II benign follicular nodule.

Fig. 2—5-year-old girl with American College of Radiology Thyroid Imaging Reporting and Data System category 2 nodule. Ultrasound shows 1.9-cm left mixed cystic and solid nodule (*thick arrows*) with isoechoic solid component (*thin arrows*). Cytologic analysis revealed Bethesda class II benign follicular nodule.
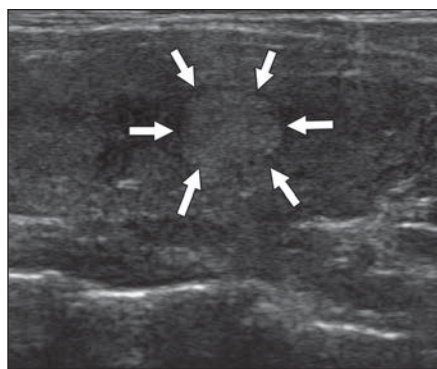
Fig. 3—17-year-old boy with American College of Radiology Thyroid Imaging Reporting and Data System category 3 nodule. Ultrasound shows 1-cm left hyperechoic solid nodule (*arrows*). Pathologic analysis revealed benign follicular nodule of colloid type with Hürthle cell metaplasia.
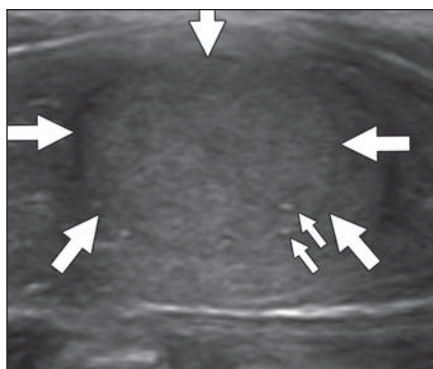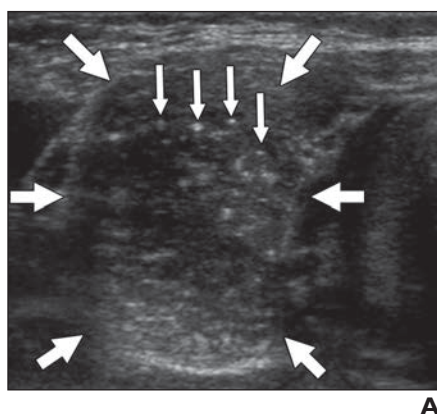

Fig. 4—12-year-old girl with American College of Radiology Thyroid Imaging Reporting and Data System category 4 nodule. Ultrasound shows 2.4-cm left solid isoechoic nodule (*thick arrows*) with smooth margins and sparse punctate echogenic foci (*thin arrows*). Pathologic analysis revealed follicular adenoma.

Fig. 5—17-year-old girl with two thyroid nodules.
A, Ultrasound shows 2.5-cm right sold hypoechoic nodule (*thick arrows*) with innumerable punctate echogenic foci (*thin arrows*) and taller-than-wide shape. Pathologic analysis revealed papillary carcinoma, assigned American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) category 5.
B, Ultrasound shows 0.8-cm left thyroid histologically proven papillary carcinoma nodule (*arrows*). Nodule was anechoic without internal vascularity and therefore was interpreted as cystic and misclassified as ACR TI-RADS category 1.
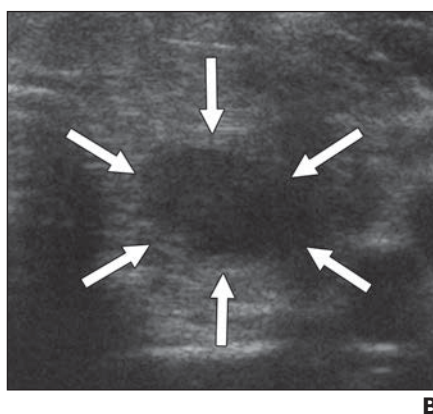
Statistical Analysis

Descriptive statistics included nodule pathologic status, patient sex and age, and tabulation of the distribution of thyroid nodule ultrasound features by pathologic status. Unweighted and weighted kappa coefficients were generated to assess intra- and interobserver reliability for all binary and ordinal ultrasound features, respectively. Observations from each radiologist's first session were compared. Kappa values and corresponding level of agreement are defined as follows: 1.00, perfect; 0.81–0.99, almost perfect; 0.61–0.80, substantial; 0.41–0.60, moderate; 0.21–0.40, fair; 0–0.20, slight; and less than 0, poor [21].

Generalized linear mixed-effects models were used to estimate the odds of malignancy as a function of univariable and multivariable nodule characteristics. To account for within-nodule correlation, random intercepts were allowed for each patient. Any associations with $p \leq 0.05$ were considered statistically significant. These models were chosen because they account for both the binary outcome of the primary dependent nodule malignancy variable and also select patients' multiple nodule observations, allowing retention of all nodules for the analysis.

To supplement this analysis and further characterize the relationship between ACR TI-RADS category and malignancy, an ROC curve was constructed to determine the optimal TI-RADS category cut point for predicting malignancy by calculating the AUC. Sensitivity, specificity, positive predictive value, and negative predictive value estimates are reported. All statistical analyses were conducted using SAS (version 9.4, SAS Institute).

Results

Demographic analysis of the 62 patients revealed 56 female patients (median age, 16.5 years; interquartile range [IQR], 15–18 years) and six male patients (median age, 12.5 years; IQR, 9–17 years). Of the total 74 nodules, 54 (73.0%) were benign and 20 (27.0%) were malignant. The median nodule size was 1.90 cm (IQR, 1.30–2.80 cm).

Of the 54 benign nodules, tissue diagnosis was confirmed by FNAB cytologic analysis in 39 (38 Bethesda class II and one Bethesda class III) and by surgical histopathologic analysis in 15 (two follicular adenomas, two benign thyroid nodules of colloid type, two lymphocytic thyroiditis, and nine multinodular goiters). Tissue diagnosis of all 20 malignant nodules was confirmed on the basis of surgical histopathologic analysis (14 papillary thyroid carcinoma, three follicular variant–papillary carcinoma, one papillary microcarcinoma, one follicular carcinoma, and one follicular neoplasm with Hürthle cell features).

For each radiologist, intraobserver agreement was substantial for TI-RADS category ($\kappa = 0.69$–0.77; $p < 0.001$). Intraobserver agreement was substantial to almost perfect for all individual parameters ($\kappa = 0.66$–0.94; $p < 0.001$), except for echogenicity, which was moderate to substantial ($\kappa = 0.46$–0.70; $p < 0.001$). In contrast, interobserver agreement, although substantial for composition and shape ($\kappa = 0.68$–0.79; $p < 0.001$), was only moderate for TI-RADS category ($\kappa = 0.37$; $p = 0.002$) (Table 1). However, a cross-tabulation calculation performed for TI-RADS category between the two observers showed that 19 of 41 (46.0%) disagreements were between adjacent categories (e.g., ACR TI-RADS category 3 vs 4, or category 4 vs 5).

The median ACR TI-RADS category was 4 (IQR, 4–5). ACR TI-RADS category categorized by pathologic status is shown in Table 2 with distribution as follows: four nodules (5.4%) with TI-RADS category 1 (three benign and one malignant); four (5.4%) with TI-RADS category 2 (four benign and zero malignant), six (8.1%) with TI-RADS category 3 (six benign and zero malignant), 24 (32.4%) with TI-RADS category 4 (22 benign and two malignant), and 36 (48.7%) with TI-RADS category 5 (19 benign and 17 malignant). All malignant nodules (19/20; 95.0%) were rated as TI-RADS category 4 or 5 except for one false-negative of a malignant 0.8-cm left thyroid nodule in a 17-year-old girl that was erroneously classified as TI-RADS category 1 (benign) (Fig. 5B); this

TABLE 1: Intra- and Interobserver Agreements for Thyroid Nodule Ultrasound Features

Nodule Characteristic	Intraobserver Agreement, Radiologist 1		Intraobserver Agreement, Radiologist 2		Interobserver Agreement	
	κ (95% CI)	p	κ (95% CI)	p	κ (95% CI)	p
Composition	0.77 (0.63–0.90)	< 0.001	0.85 (0.75–0.95)	< 0.001	0.68 (0.52–0.83)	< 0.001
Echogenicity	0.46 (0.26–0.65)	< 0.001	0.70 (0.58–0.82)	< 0.001	0.35 (0.21–0.49)	< 0.001
Shape[a]	0.94 (0.89–1.00)	< 0.001	0.83 (0.65–1.00)	< 0.001	0.79 (0.67–0.91)	< 0.001
Margin	0.76 (0.65–0.88)	< 0.001	0.66 (0.49–0.83)	< 0.001	0.25 (0.09–0.40)	0.02
Echogenic foci	0.91 (0.83–0.98)	< 0.001	0.81 (0.69–0.92)	< 0.001	0.41 (0.27–0.54)	< 0.001
ACR TI-RADS category	0.77 (0.66–0.87)	< 0.001	0.69 (0.57–0.81)	< 0.001	0.37 (0.24–0.51)	0.002

Note—All estimates are reported as weighted kappa coefficients unless otherwise specified. All *p* values were considered statistically significant (*p* < 0.05). ACR = American College of Radiology, TI-RADS = Thyroid Imaging Reporting and Data System.
[a]Unweighted kappa values are shown for shape.

TABLE 2: Distribution of Thyroid Nodule Ultrasound Features by Pathologic Status

Feature	Benign (*n* = 54)	Malignant (*n* = 20)	Total (*n* = 74)
American College of Radiology Thyroid Imaging Reporting and Data System category			
1	3 (5.6)	1 (5.0)	4 (5.4)
2	4 (7.4)	0 (0.0)	4 (5.4)
3	6 (11.1)	0 (0.0)	6 (8.1)
4	22 (40.7)	2 (10.0)	24 (32.4)
5	19 (35.2)	17 (85.0)	36 (48.7)
Composition			
Cystic or almost completely cystic or spongiform	4 (7.4)	1 (20.0)	5 (6.8)
Mixed cystic and solid	22 (40.7)	0 (0.0)	22 (29.7)
Solid or almost completely solid	28 (51.9)	19 (95.0)	47 (63.5)
Echogenicity			
Anechoic	3 (5.6)	1 (5.0)	4 (5.4)
Hyperechoic or isoechoic	33 (61.1)	1 (5.0)	34 (46.0)
Hypoechoic	15 (27.8)	15 (75.0)	30 (40.5)
Very hypoechoic	3 (5.6)	3 (15.0)	6 (8.1)
Shape			
Wider than tall	49 (90.7)	19 (95.0)	68 (91.9)
Taller than wide	5 (9.3)	1 (5.0)	6 (8.1)
Margin			
Smooth or ill-defined	45 (83.3)	4 (20.0)	49 (66.2)
Lobulated or irregular	9 (16.7)	15 (75.0)	24 (32.4)
Extrathyroidal extension	0 (0.0)	1 (5.0)	1 (1.4)
Echogenic foci			
None or large comet-tail artifacts	18 (33.3)	3 (15.0)	21 (28.4)
Macrocalcifications	0 (0.0)	50 (0.0)	0 (0.0)
Peripheral (rim) calcifications	1 (1.9)	0 (0.0)	1 (1.4)
Punctate echogenic foci alone	35 (64.8)	12 (60.0)	47 (63.5)
Punctate echogenic foci in combination with macrocalcifications	0 (0.0)	5 (25.0)	5 (6.8)

Note—Not all percentages total 100% because of rounding.

patient also had a concomitant 2.5-cm right malignant papillary carcinoma nodule that was assigned a TI-RADS category 5 rating (Fig. 5A).

On univariable analysis, for every 1-unit increase on the ordinal TI-RADS scale, the likelihood of malignancy increased 2.63 times (95% CI, 1.08–6.41; *p* = 0.03). For every 1-cm increase in size, nodules were also 69% more likely to be malignant (odds ratio, 1.69; 95% CI, 1.04–2.73; *p* = 0.03). On multivariable analysis, after adjusting for nodule size, ACR TI-RADS category, although not statistically significant, remained marginally associated with malignancy (adjusted odds ratio, 2.27; 95% CI, 0.93–5.54; *p* = 0.07), but size was no longer associated. Using the ROC curve generated from ACR TI-RADS category, the AUC estimate was 0.75 (95% CI, 0.64–0.86), which confirmed the accuracy of the predictive modeling. An optimal cut point of ACR TI-RADS category 5 was selected because it maximized both sensitivity (85%) and specificity (65%). The corresponding false-positive rate of 53% at this cut point did result in suppression of the positive predictive value, which was 47%. The negative predictive value was 92% (Table 3). Overall, nodules with an ACR TI-RADS category of 5 were 10.44 times (95% CI, 2.71–40.21; *p* < 0.001) more likely to be malignant.

The distribution of individual ultrasound features is also shown in Table 2. Some features were well differentiated by pathologic status. For the echogenicity feature, most (90.0%) malignant nodules were hypoechoic or very hypoechoic, whereas most (61.1%) benign nodules were hyper- or isoechoic. For the margin feature, most (83.3%) benign nodules were smooth or ill-defined, whereas most (75.0%) malignant nodules were lobulated or irregularly marginated. Other features, however, were not well differentiated by pathologic analysis. These included punctate echogenic foci, either alone or in combination with macrocalcifications, which were found in the majority of both malignant and benign nodules (85.0% and 64.8%, respectively). Finally, taller-than-wide shape, a suspicious feature, was found in a small percentage of both malignant (5.0%) and benign (9.3%) nodules, including in a 14-year-old boy with six thyroid nodules, all showing benign multinodular goiter on surgical pathologic analysis (Fig. 6).

Discussion

The results of the current study reveal that the ACR TI-RADS discriminates well between malignant and benign thyroid nod-

TABLE 3: Predicted Versus Actual Malignancy for American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) Category 5

Predicted[a]	Actual		
	Benign	Malignant	Total
Benign	35	3	38
Malignant	19	17	36
Total	54	20	74

Note— An optimal cut point of ACR TI-RADS category 5 was selected because it maximized both sensitivity (85%) and specificity (65%), with a positive predictive value of 47% and a negative predictive value of 92%.
[a]Predicted malignant refers to ACR TI-RADS category 5. Predicted benign refers to ACR TI-RADS category 1, 2, 3, or 4.

Fig. 6—13-year-old boy with multiple thyroid nodules. Ultrasound shows 2.8-cm isthmus nodule (*thick arrows*), assigned American College of Radiology Thyroid Imaging Reporting and Data System category 4, showing mixed solid and cystic composition and taller-than-wide shape. Isthmus nodule is squeezed between two abutting solid nodules (*thin arrows*). Pathologic analysis revealed benign multinodular goiter in all nodules.

ules in a pediatric population and is particularly effective in identifying positive cases. The TI-RADS category assigned to a nodule was proportional to the risk of malignancy when we used a univariable regression analysis. This strong association between ACR TI-RADS category and malignancy was further enhanced by a more rigorous multivariable regression analysis and ultimately was confirmed by the AUC estimate of 0.75. These results corroborate recent reports that have validated the ACR TI-RADS in large adult populations [14–17, 22]. We therefore suggest that this classification can be helpful as a decision-making tool in the management of pediatric thyroid nodules.
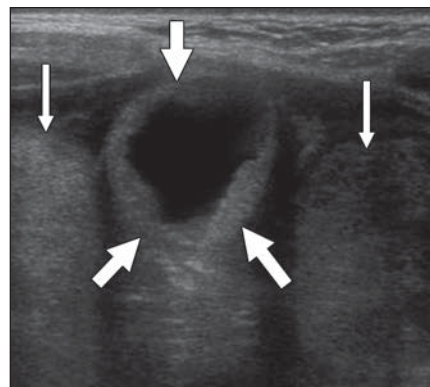
Apart from the ACR TI-RADS, other thyroid nodule ultrasound risk stratification guidelines have also recently been found to perform well for children. One study assessed the 2015 ATA Management Guidelines for Children and found that the composite pattern of ultrasound features accurately predicted malignancy, with the odds of malignancy significantly higher in nodules determined to have a high level of suspicion by ultrasound [18]. Other studies found overall good diagnostic performance in correctly classifying pediatric nodules using the ATA system and a different TIRADS developed at a single institution [3, 12]. One reason we sought in the current study to specifically appraise the TI-RADS developed by the ACR was the advantage it may have over other TIRADS classifications with regard to accessibility, applicability, and general reputation and acceptance in practice settings in the United States.

Because of the differing biologic behavior of thyroid cancer in the pediatric versus the adult population, adjustments in the way protocols are applied, particularly with regard to nodule size, may prove prudent. ACR TI-

RADS does not consider size in assignment of TI-RADS level in adults but does use size to suggest thresholds for when a nodule should undergo FNAB [6]. In children, on the other hand, it is recommended that sonographically suspicious nodules undergo FNAB regardless of size because malignancy is better predicted by ultrasound features and clinical factors rather than size alone [7, 9]. The results of the current study support this strategy, because, although there was a significant correlation between size and malignancy on univariable analysis, higher ACR TI-RADS categories remained associated with increased likelihood of malignancy on multivariable analysis, whereas nodule size did not.

Although the ACR TI-RADS was useful as a predictor of malignancy in children, its imperfections should be considered. Although intraobserver reproducibility for ACR TI-RADS category was high, interobserver reliability was only moderate, which was not as robust as hoped. Additional analysis of the interobserver data indicated that many of the disagreements were between adjacent TI-RADS categories (e.g., category 3 vs 4, rather than category 3 vs 5). Therefore, although the results were divergent, the extent of the discord was often not substantial.

Of greater concern was the large number of higher suspicion ratings, which is congruent with prior studies in children using different risk stratification guidelines [12, 18]. In our study, we found that the majority of both benign and malignant nodules were rated as ACR TI-RADS category 4 or 5. Furthermore, although ROC analysis confirmed that TI-RADS 5 was the category at which discrimination between malignant and benign nodules was most effective, a higher than desired number of false-positives were observed. Specifically, 19 of 36 (53%) benign nodules were erroneously classified

as category 5, meaning these nodules may have undergone unnecessary FNAB. This was reflected in suppressed positive predictive value and specificity at TI-RADS category 5, findings that are supported by a recent study of ACR TI-RADS in adults that reported similarly low specificities, ranging from 44% to 51% [17]. Despite these shortcomings, the correct classification of 35 of 38 (92%) of benign nodules at ACR TI-RADS categories 1, 2, 3, or 4 suggests that, in these cases, FNAB could have been avoided and ultrasound surveillance would have been an alternative management option.

Although this study did not analyze which individual ultrasound features were most responsible for the high false-positive rate of ACR TI-RADS category, some general relevant trends were observed. Punctate echogenic foci, a feature favoring malignancy, were found not only in most of our malignant nodules, but also in most of our benign nodules a finding supported in other studies [5, 23]. This result could be explained by prior sonographic and histopathologic correlation in children showing that punctate echogenic foci may represent not only exclusively psammoma bodies that are associated with papillary thyroid carcinoma, but also other entities, such as stromal calcifications and sticky colloid [19]. Nevertheless, radiologists using the ACR TI-RADS should be aware that, with the current point allocation system, the presence of punctate echogenic foci is accorded 3 points, resulting in a dispropor-

tionate influence on elevating the TI-RADS score and therefore establishing the level of suspicion. The category of punctate echogenic foci was found in a recent study in adults to be one of the ACR TI-RADS sonographic features for which interobserver agreement was lowest, suggesting that radiologists could benefit from education regarding this important distinguishing feature [16].

Taller-than-wide shape, a suspicious feature, has been reported as rare in pediatric nodules [7, 12]. However, we found this feature in 9.3% of our benign nodules, which is an unexpectedly high frequency and which we hypothesize could also contribute to the high false-positive rate of ACR TI-RADS category. All of the benign nodules with this configuration occurred in the same 14-year-old boy with multinodular goiter (Fig. 3). Inspection of his images revealed that the taller-than-wide shape could be attributed to distortion and mass effect from the sheer size and number of neighboring nodules. This case serves as a reminder that the significance of the taller-than-wide shape should be interpreted in the context of surrounding thyroid tissue because it may represent a manifestation of secondary, rather than intrinsic, contour deformity. Refinements to the current ACR TI-RADS format may be necessary to account for this specific situation.

Although the large number of false-positives poses a challenge to use of the ACR TI-RADS, of equal concern is the false-negative we encountered in which a malignant nodule was assigned a benign rating of TI-RADS category 1. The individual ultrasound features that caused this 0.8-cm malignant nodule in a 17-year-old girl to be assigned ACR TI-RADS category 1 were the selection of cystic for composition and anechoic for echogenicity, both of which correspond to 0 points (Fig. 5B). Although it is possible that the nodule was genuinely cystic, an alternative consideration is that this was a solid nodule that should have received a very hypoechoic rating. This possible misinterpretation illustrates the ease with which very hypoechoic, a feature favoring malignancy and defined as echogenicity less than that of the neck musculature, can be misconstrued as anechoic, a feature favoring benignity, and vice versa, despite their diverging clinical implications [13]. This may be an issue particularly for subcentimeter nodules. In quantitative

terms, the difference in the ACR TI-RADS between the two features is 3 points, which could be enough to erroneously shift management recommendations to deferral of FNAB in a nodule that ought to undergo tissue sampling, as well as the reverse.

This case emphasizes the point that subcentimeter nodules may be difficult to characterize accurately with sonography, mirroring prior studies in children showing that six of 52 malignant nodules were incorrectly classified as very low and low suspicion with ATA criteria [12]. Furthermore, because this patient also had a separate 2.5-cm malignant right nodule (Fig. 5A), management of pediatric nodules should be based on not only its sonographic appearance and ACR TI-RADS rating but also on the broader clinical context, including the presence of additional suspicious nodules.

Limitations of our study include its retrospective design, which resulted in the use of different ultrasound vendors and potential variation in image quality. Another restriction was the small study population and even smaller number of malignant nodules ($n = 20$), possibly limiting our ability to extrapolate the true population effect. The high rate of false-positives observed at ACR TI-RADS category 5 also requires further research to determine whether these positive but preliminary results can be replicated while simultaneously increasing the positive predictive value.

This study shows that the ACR TI-RADS discriminates well between malignant and benign nodules in a pediatric population, particularly at TI-RADS category 5, and is useful in malignancy risk stratification. However, practitioners should be aware of the high rate of false-positives at higher TI-RADS categories, as well as the possibility of false-negatives, especially in subcentimeter nodules.

References

1. Mandel SJ. The application of sonographic patterns to risk stratification of thyroid nodules. In: Milas M, Mandel S, Langer J, eds. *Advanced thyroid and parathyroid ultrasound*. Basel, Switzerland: Springer International Publishing, 2017:131–142

2. Shin JH, Baek JH, Chung J, et al. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology consensus statement and recommendations. *Korean J Radiol* 2016; 17:370–395

3. Kwak JY, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. *Radiology* 2011; 260:892–899

4. Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *Eur Thyroid J* 2017; 6:225–237

5. Zayadeen AR, Abu-Yousef M, Berbaum K. Retrospective evaluation of ultrasound features of thyroid nodules to assess malignancy risk: a step toward TIRADS. *AJR* 2016; 207:460–469

6. Tessler FN, Middleton WD, Grant EG, 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

7. Essenmacher AC, Peter H, Joyce J, et al. Sonographic evaluation of pediatric thyroid nodules. *RadioGraphics* 2017; 37:1731–1752

8. Vergamini LB, Frazier AL, Abrantes FL, Ribeiro KB, Rodriguez-Galindo C. Increase in the incidence of differentiated thyroid carcinoma in children, adolescents, and young adults: a population-based study. *J Pediatr* 2014; 164:1481–1485

9. Francis GL, Waguespack SG, Bauer AJ, et al. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2015; 25:716–759

10. Cheng P. TIRADS calculator for thyroid nodules. Keck School of Medicine of USC website. www-hsc.usc.edu/~phillimc/calc/tirads.html. Published 2018. Accessed August 22, 2018

11. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. Malignancy risk stratification of thyroid nodules: comparison between the Thyroid Imaging Reporting and Data System and the 2014 American Thyroid Association management guidelines. *Radiology* 2016; 278:917–924

12. Martinez-Rios C, Daneman A, Bajno L, van der Kaay DCM, Moineddin R, Wasserman JD. Utility of adult-based ultrasound malignancy risk stratifications in pediatric thyroid nodules. *Pediatr Radiol* 2018; 48:74–84

13. Grant EG, Tessler FN, Hoang JK, 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(12 Pt A):1272–1279

14. Grani G, Lamartina L, Cantisani V, Maranghi M, Lucia P, Durante C. Interobserver agreement of various thyroid imaging reporting and data systems. *Endocr Connect* 2018; 7:1–7

15. Middleton WD, Teefey SA, Reading CC, et al. Multiinstitutional analysis of thyroid nodule risk stratification using the American College of Radiology Thyroid Imaging Reporting and Data Sys-

tem. *AJR* 2017; 208:1331–1341

16. Hoang JK, Middleton WD, Farjat AE, et al. Interobserver variability of sonographic features used in the American College of Radiology Thyroid Imaging Reporting and Data System. *AJR* 2018; 211:162–167

17. Hoang JK, Middleton WD, Farjat AE, et al. Reduction in thyroid nodule biopsies and improved accuracy with American College of Radiology Thyroid Imaging Reporting and Data System. *Radiology* 2018; 287:185–193

18. Lim-Dunham JE, Erdem Toslak I, Alsabban K, et al. Ultrasound risk stratification for malignancy using the 2015 American Thyroid Association management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Pediatr Radiol* 2017; 47:429–436

19. Erden Toslak I, Martin B, Barkan GA, Kılıç AI, Lim-Dunham JE. Patterns of sonographically detectable echogenic foci in pediatric thyroid carcinoma with corresponding histopathology: an observational study. *AJNR* 2018; 39:156–161

20. Cibas ES, Ali SZ. The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol* 2009; 132:658–665

21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174

22. Ha SM, Ahn HS, Baek JH, et al. Validation of three scoring risk-stratification models for thyroid nodules. *Thyroid* 2017; 27:1550–1557

23. Al Nofal A, Gionfriddo MR, Javed A, et al. Accuracy of thyroid nodule sonography for the detection of thyroid cancer in children: systematic review and meta-analysis. *Clin Endocrinol (Oxf)* 2016; 84:423–430