Journal of Pediatric Surgery 56 (2021) 1120–1125

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.elsevier.com/locate/jpedsurg.org

Pediatric adaptions are needed to improve the diagnostic accuracy of thyroid ultrasound using TI-RADS

Hira Ahmad[a,1], Ameer Al-Hadidi[a,1], Adam Bobbey[b], Summit Shah[b], Joseph Stanek[c], Kathleen Nicol[d], Robert P. Hoffman[e], Jennifer H. Aldrink[a,*]

[a] Department of Surgery, Division of Pediatric Surgery, The Ohio State University College of Medicine, Nationwide Children's Hospital, FB Suite 6B.1, 700 Children's Drive, Columbus, OH, United States
[b] Department of Radiology, Nationwide Children's Hospital, Columbus, OH, United States
[c] Division of Hematology, Oncology, and Bone Marrow Transplantation, Nationwide Children's Hospital, Columbus, OH, United States
[d] Department of Pathology, Nationwide Children's Hospital, Columbus, OH, United States
[e] Department of Pediatrics, Division of Endocrinology, Nationwide Children's Hospital, Columbus, OH, United States

ARTICLE INFO

Article history:
Received 18 January 2021
Accepted 5 February 2021

Keywords:
Thyroid nodule
Ultrasound
Thyroid imaging and reporting data system
Thyroid carcinoma
Pediatric

ABSTRACT

Background/Purpose: Thyroid Imaging Reporting and Data System (TI-RADS) is validated in adults but not yet in children. The purpose of this study was to determine the sensitivity, specificity, and accuracy of TI-RADS in predicting thyroid malignancy for pediatric nodules, and to compare the diagnostic accuracy to the current American Thyroid Association (ATA) guidelines.
Methods: A single institution retrospective review was performed of patients younger than 21 years who underwent thyroid nodule fine needle aspiration biopsy (FNAB). Two radiologists were blinded to the pathology and independently classified all biopsied thyroid nodules based on TI-RADS. ATA and TI-RADS guidelines were analyzed to determine the diagnostic sensitivity and specificity of both scoring systems.
Results: 115 patients (median age 15.5 years, 90 females) with 138 nodules were scored using TI-RADS. There was moderate inter-rater agreement between radiologists (Kappa = 0.51; $p < 0.0001$). Evaluating several potential TI-RADS criteria, 23.2%-68.1% of nodules were recommended for FNAB, compared to 82.6% of nodules using ATA guidelines. Using TI-RADS ≥ 3 (without size cutoff) as an indication for FNAB had 100% sensitivity with no missed suspicious or malignant nodules on cytology or pathology.
Conclusions: Using TI-RADS for diagnostic management of pediatric thyroid nodules improves accuracy in predicting malignancy.

1. Introduction

Pediatric thyroid cancers represent approximately 2% of all malignancies in patients under 15 years of age, with an annual incidence of 2.4 per 100,000 [1–3]. The incidence varies with age, gender, and race. While differentiated thyroid cancer represents only 1% of malignancies in prepubertal children, the incidence increases to 7% of malignancies in adolescents with a female preponderance (5:1 female-male ratio) and is the second most common malignancy in females between the ages of 15 and 19 [2–5].

Identification of a thyroid nodule in a child warrants further evaluation, as solitary nodules may carry up to a 6-fold increased risk of malignancy compared to adults (25–30% versus 5%) [3,6]. Pediatric thyroid malignancies are also more likely to present with cervical lymph node (40–80%) and distant (20–30%, most commonly pulmonary) metastases than adult thyroid malignancies [3,5]. Following physical examination, evaluation of thyroid nodules is initiated with ultrasound imaging. Ultrasound alone cannot always distinguish between benign and malignant lesions; therefore, nodules of concern are then further evaluated with a fine needle aspiration biopsy (FNAB). FNAB for thyroid nodules is the most accurate and cost-effective means of tissue sampling with an accuracy rate of 80–95% in pediatric patients [2,7].

The American Thyroid Association (ATA) has developed criteria with recommendations for biopsy of nodules ≥ 1 cm in size [8]. Thyroid Imaging Reporting and Data System (TI-RADS), another classification system that has been increasingly utilized in adults, similarly provides recommendations based on nodule size and ultrasound characteristics including composition, echogenicity, shape, margins, and echogenic foci (Table 1) [9].

Abbreviations: ACR, American College of Radiology; PED, Pediatric; TI-RADS, Thyroid Imaging Reporting and Data System; FNAB, Fine Needle Aspiration Biopsy; ATA, American Thyroid Association; CI, Confidence Interval; IQR, Interquartile range.

Type of Study: Retrospective Comparative Study.

Level of Evidence: Level III.

* Corresponding author.
E-mail address: jennifer.aldrink@nationwidechildrens.org (J.H. Aldrink).
[1] Co-first authors.

https://doi.org/10.1016/j.jpedsurg.2021.02.034
0022-3468/

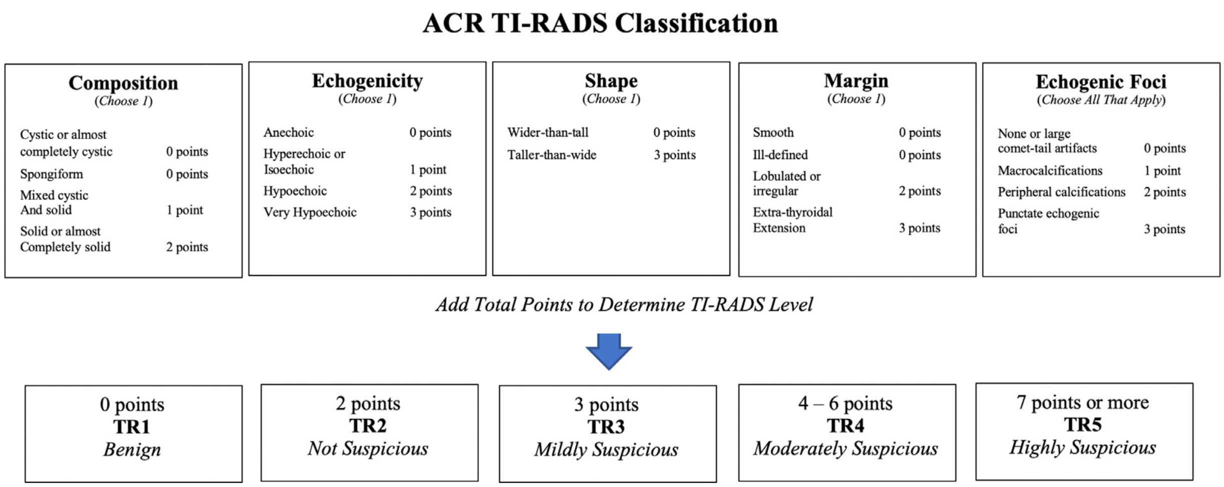

Fig. 1. ACR TI-RADS Classification [9]. ACR = American College of Radiology. TI-RADS = Thyroid Imaging Reporting and Data System.

Table 1
Comparison of the ACR and PED TI-RADS Recommendations for FNA per TI-RADS Level [9].

Recommendations for FNA per TI-RADS Level		ACR TI-RADS	PED TI-RADS
TR1	Benign	*No FNA*	*No FNA*
TR2	Not Suspicious	*No FNA*	*No FNA*
TR3	Mildly Suspicious	*FNA if ≥ 2.5 cm* *Follow if ≥ 1.5 cm*	*FNA if ≥ 1.5 cm*
TR4	Moderately Suspicious	*FNA if ≥ 1.5 cm* *Follow if ≥ 1.0 cm*	*FNA if ≥ 1.0 cm*
TR5	Highly Suspicious	*FNA if ≥ 1.0 cm* *Follow if ≥ 0.5 cm*	*FNA for any technically feasible size*

ACR = American College of Radiology.
PED = Pediatric.
TI-RADS = Thyroid Imaging Reporting and Data System.
FNA = Fine Needle Aspiration.

Until recently, recommendations for the management of thyroid nodules in children and adults have followed the same guidelines. This is problematic for two reasons: 1) validation of criteria is based solely on data from the adult literature and 2) the ATA guidelines depend significantly on nodule size, where thyroid volume changes with increasing age and the size of the nodules are not predictive of malignancy in children [10,11]. In 2015, the ATA commissioned a task force to create guidelines specific for pediatric thyroid nodules and thyroid malignancies. In these guidelines, the preferential use of ultrasound appearance and clinical context was emphasized over nodule size alone to identify lesions that warrant a FNAB. The American College of Radiology (ACR) TI-RADS classification provides an easy, reproducible, and objective method of grading thyroid nodules based on ultrasound findings (Fig. 1) [9]. The purpose of this study was to determine the sensitivity, specificity, and accuracy of predicting thyroid malignancy in nodules using the TI-RADS classification system, and to compare the diagnostic accuracy to the traditional pediatric ATA guidelines. We hypothesized that applying TI-RADS to guide FNAB would provide enhanced diagnostic accuracy in identifying malignant nodules in pediatric patients.

2. Methods

2.1. Patient population

A single institution retrospective review was performed following Institutional Review Board Approval (#00000545). All patients 21 years or younger who underwent a thyroid nodule FNAB from January 2015 to March 2019 and had ultrasound imaging available for review were included in the study. A chart review was conducted to obtain data on patient demographics, age at the time of diagnosis of the nodule, associated symptoms, the number, location, and size of thyroid nodules present, nodule cytology, and surgical procedure performed with the resulting pathology. All patients over the age of 21 years and those without complete imaging or pathology data available were excluded from the study.

2.2. Image analysis

All ultrasound images of the nodules that underwent a FNAB were independently reviewed by two board-certified pediatric radiologists who were blinded to the FNAB cytology and surgical pathology results. Nodules were assessed using a combination of gray-scale and color doppler multi-planar ultrasound images and cine clips on a variety of systems. Thyroid nodule characteristics included: composition, echogenicity, shape, margins, and echogenic foci. Nodule size was measured in centimeters (cm) of maximal dimensions and graded according to TI-RADS guidelines. Retrospective recommendations for FNAB were offered based upon the ultrasound interpretation and TI-RADS scoring completed by the radiologists. When scores assigned by each radiologist resulted in disparate recommendations (whether to perform a FNAB or not), the images were re-reviewed by both radiologists and a score was subsequently agreed upon. When more than one nodule was present, each nodule was evaluated independently.

2.3. Statistical analysis

Data was summarized using standard descriptive statistical methods: frequency and percentage for qualitative variables and median and range or interquartile range (IQR) for quantitative variables. The agreement of radiologists' initial retrospective ratings was assessed using weighted Kappa statistics and the correlation between scores was computed with a Spearman's Rho correlation coefficient (r_s). Using suspicious or malignant FNAB cytology (Bethesda IV, V, VI) and malignant surgical pathology as true positives, the sensitivity, specificity, accuracy, and positive/negative predictive value of ATA and TI-RADS criteria at predicting suspicious or malignant cytopathology were measured and presented with 95% confidence intervals (CI) calculated using either Clopper-Pearson (sensitivity/specificity/accuracy) or standard logit methods (predictive values). Additionally, novel proposed pediatric-specific criteria that either lowered or removed the size threshold for recommending FNAB were also applied using the TI-RADS classifica-

Table 2
Bethesda System for Reporting Thyroid Cytopathology.

Bethesda System for Reporting Thyroid Cytopathology		# of Nodules	% of Nodules
I	*Cyst fluid/Non-diagnostic*	13	9.42%
II	*Benign follicular nodule*	88	78.98%
	Chronic lymphocytic thyroiditis	21	
III	*Atypia/Follicular lesion of undetermined significance*	6	4.35%
IV	*Suspicious for follicular neoplasm*	2	1.45%
V	*Suspicious for malignancy*	2	1.45%
VI	*Malignant*	6	4.35%

tion system and were assessed in a similar fashion. One of the proposed criteria, PED TI-RADS, utilized the TI-RADS classification system and recommended biopsy on smaller nodules compared to the traditional criteria, influenced by Richman et al. [12] (Table 1). Additional proposed novel criteria recommended FNAB of nodules TI-RADS $\geq$ 3 and TI-RADS $\geq$ 4, irrespective of nodule size. All statistical analyses were conducted using SAS 9.4 (SAS Institute Inc., Cary, NC, USA). All p-values were two-sided and those < 0.05 were considered statistically significant.

2.4. Fine needle aspiration biopsy

The decision to perform a FNAB was made by the patient's physician based on clinical consideration in conjunction with ATA guidelines. The samples from FNAB were classified according to the Bethesda System for Reporting Thyroid Cytopathology. The classifications are detailed in Table 2. Each biopsy sample must contain a minimum of 6 groupings of well-preserved thyroid epithelia cells, consisting of at least 10 cells per group [13].

3. Results

3.1. Patient characteristics

A total of 115 patients were included in this study, with 138 thyroid nodules undergoing FNAB. The median age of the cohort was 15.5 years (range 5.0–20.2) (IQR 13–16.9). The majority of patients were female (n = 90, 78.3%). 12.2% of patients had a family history of thyroid cancer, and 4.3% had a genetic predisposition to thyroid cancer. A history of radiation therapy to the neck and/or chest was noted in 2.6% of patients. Most patients (n = 47, 40.9%) were asymptomatic at presentation. When symptoms were present, the most common included: neck fullness (n = 39, 33.9%), fatigue (n = 27, 23.5%), weight/appetite changes (n = 17, 14.8%), and dysphagia (n = 13, 11.3%).

3.2. Nodule cytology and surgical pathology

FNAB revealed nodule cytology was overwhelmingly benign, with 79% of nodules being either benign follicular nodules or chronic lymphocytic thyroiditis (Table 2). 7.3% of nodules biopsied were suspicious or malignant, 4.4% of nodules were lesions of undetermined significance, and 9.4% were non-diagnostic. Thirty-five patients proceeded to surgical resection for a total of 36 operations. Thirteen of these 35 underwent surgical resection after a variable period of observation following FNAB demonstrating Bethesda I or II cytology (median observation time 11 months (range 4–29)). Sixteen patients (45.7%) underwent a lobectomy (1 with a concomitant lymph node dissection) and 19 patients (54.3%) underwent a total thyroidectomy (7 with a concomitant lymph node dissection). One patient required a completion thyroidectomy following the identification of a follicular malignancy on final surgical pathology (identified as Bethesda IV, suspicious for follicular neoplasm on FNAB). In total, there were 13 malignant nodules (11 papillary carcinoma, 2 follicular carcinoma) and 21 benign nodules on surgical pathology. Pathology report and any subsequent follow-up were absent for one patient.

The concordance between FNAB cytology and surgical pathology results was 88.6%. One patient had malignant (Bethesda VI) cytology with no malignancy identified on surgical pathology. Three patients had benign (Bethesda II) cytology that underwent resection due to the size of the nodule (2.6 cm, 3.1 cm, and 4.5 cm), with surgical pathology revealing a malignancy (2 of which were microscopic or minimally invasive).

3.3. Outcomes and comparative analysis

The TI-RADS scores comprised in our cohort included: 13 TR1, 31 TR2, 51 TR3, 35 TR4, and 8 TR5 nodules (Fig. 2). There was no correlation between TR score or Bethesda group to patient age ($r_s = -0.01$, $p= 0.87$; $r_s = 0.06$, $p= 0.46$, respectively). Malignant surgical pathology was only observed in nodules graded TR3 or higher. After initial independent grading and interpretation of the nodules' appearances and characteristics by two radiologists, 21 nodules had TR scores that recommended discrepant management strategies based upon assigned TI-RADS criteria by the radiologists (Table 1). Overall, the Spearman correlation and Kappa statistic were both 0.51 ($p < 0.0001$), indicating moderate inter-rater agreement between the radiologists.

ATA criteria recommended the performance of a FNAB on 82.16% (114 of 138) of the nodules. In comparison, TI-RADS criteria recommended a FNAB on 23.19% (32/138) of the nodules. ATA and TI-RADS had different recommendations on 59.42% (82 of 138) of the nodules, with the majority of cytology (93.90%) being non-diagnostic or benign (Bethesda I and II) when ATA recommended FNAB and TI-RADS recommended against. When using proposed criteria for recommending FNAB in pediatrics, PED TI-RADS criteria recommended a FNAB on 46.38% (64 of 138) of nodules, TI-RADS $\geq$ 3 criteria recommended FNAB on 68.12% (94 of 138) of nodules, and TI-RADS $\geq$ 4 criteria recommended FNAB on 31.16% (43 of 138) of nodules (Table 3).

Evaluation of nodules based upon size less than 1 cm was then performed. Twenty-four nodules less than 1 cm in size underwent FNAB, the smallest of which were two nodules measuring 0.4 cm. Both nodules were scored as TR2 and both had non-diagnostic cytology (Bethesda I). Sixteen nodules less than 1 cm in size had a TR score $\geq$ 3, with the smallest nodule in this group measuring 0.6 cm. One nodule had non-diagnostic cytology and 13 nodules were benign (Bethesda II). However, two of the nodules measured 0.9 cm with a TR4 score and had atypia of undetermined significance (Bethesda III) and suspicious for malignancy (Bethesda V) cytology. Both nodules were consistent with papillary thyroid carcinoma on surgical pathology.

4. Discussion

One-hundred and fifteen patients were included in this study, and 138 thyroid nodules underwent FNAB. Two radiologists (blinded to pathology) independently reviewed the imaging with

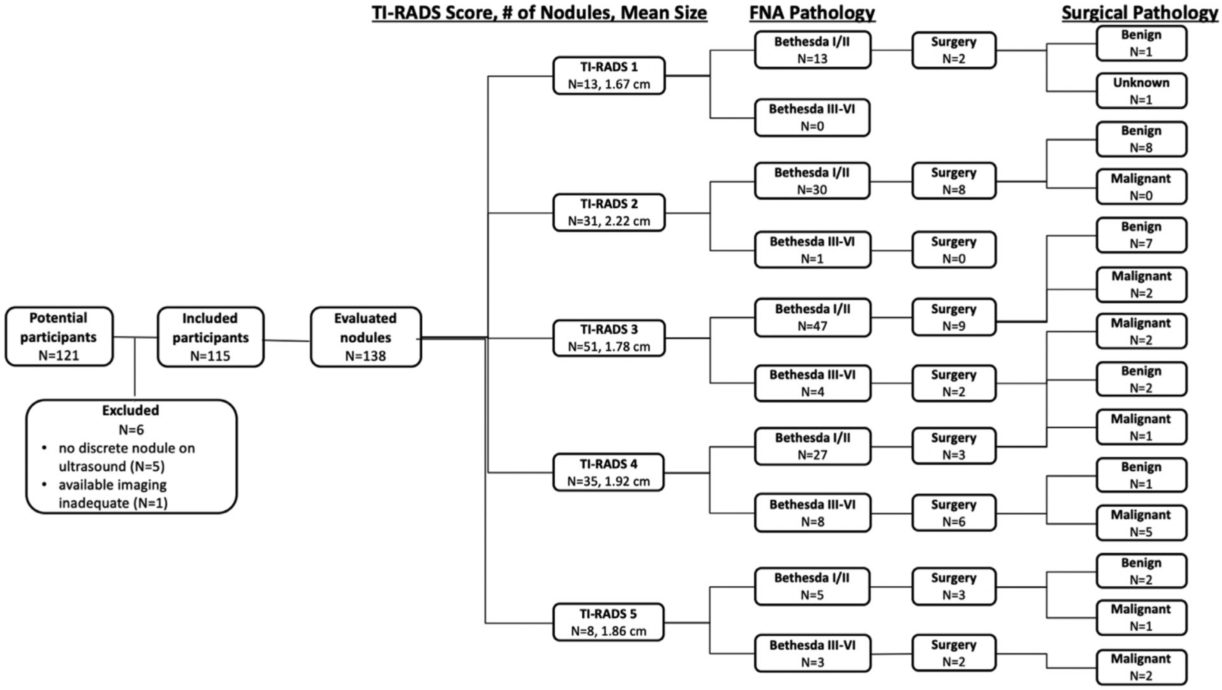

Fig. 2. Flowchart of nodules based upon TI-RADS score, FNA pathology, and surgical pathology.

Table 3
Comparison of criteria in identifying malignant surgical pathology. 95% Confidence Interval in parenthesis. (all $p > 0.05$).

	ATA	TI-RADS	PEDTI-RADS	TI-RADS $\geq$ 3	TI-RADS $\geq$ 4
Sensitivity	84.62% (54.55–98.08)	76.92% (46.19–94.96)	84.62% (54.55–98.08)	100.00% (75.29–100.00)	69.23% (38.57–90.91)
Specificity	9.52% (1.17–30.38)	71.43% (47.82–88.72)	52.38% (29.78–74.29)	42.86% (21.82–65.98)	76.19% (52.83–91.78)
Accuracy	38.24% (22.17–56.44)	73.53% (55.64–87.12)	64.71% (46.49–80.25)	64.71% (46.49–80.25)	73.53% (55.64–87.12)
Positive Predictive Value	36.67% (30.65–43.13)	62.50% (44.32–77.73)	52.38% (39.90–64.57)	52.00% (42.79–61.07)	64.29% (43.57–80.76)
Negative Predictive Value	50.00% (13.78–86.22)	83.33% (64.12–93.33)	84.62% (59.06–95.45)	100.00%	83">80.00% (63.10–90.34)
# FNA recommended ($n = 138$)	114	32	64	94	43
Missed Suspicious or Malignant Cytology ($n = 10$)	1	3	1	0	1
Missed Malignant Pathology ($n = 13$)	2	3	2	0	4

TI-RADS = Thyroid Imaging Reporting and Data System.
FNA = Fine Needle Aspiration.
PED = Pediatric.

moderate correlation. We aimed to determine the sensitivity, specificity, and accuracy of TI-RADS in predicting thyroid malignancy for nodules with available cytopathology, and to compare the diagnostic accuracy to the current American Thyroid Association (ATA) guidelines (Table 3). When using malignant surgical pathology as the true positive, ATA had a sensitivity of 84.6%. However, it had the lowest specificity (9.5%) and accuracy (38.2%) of all groups while recommending 114 (highest number of recommendations) FNAB. Following ATA criteria would have missed two nodules with confirmed malignancy on surgical pathology. TI-RADS criteria had higher specificity, accuracy, positive and negative predictive values while only recommended 32 FNAB (lowest number of recommendations). However, following these criteria would have resulted in three missed nodules with confirmed malignancy. From the additional proposed criteria, TI-RADS $\geq$ 3 recommended FNAB on 94 nodules and demonstrated 100% sensitivity, 42.9% specificity, and 64.7% accuracy without missing any nodules with malignant pathology. There was no statistically significant difference between any of the outcomes of these criteria ($p > 0.05$).

Since only 35 patients underwent surgical resection, we further evaluated the impact of each of the criteria using the Bethesda classification of all 115 nodules. When suspicious or malignant cytology (Bethesda IV, V, VI) was used as the true positive to compare the sensitivity, specificity, and accuracy of the different criteria, they demonstrated comparable results (Table 4). ATA had a sensitivity of 90.0%; however, it again displayed the lowest specificity (17.9%) and accuracy (23.2%) of all groups. Following ATA criteria would have missed one nodule with suspicious or malignant cytology. TI-RADS criteria had a substantially higher specificity, accuracy, positive and negative predictive values; however, following these criteria would have resulted in three missed nodules with suspicious or malignant cytology. From the additional proposed criteria, TI-RADS $\geq$ 3 demonstrated 100% sensitivity, 34.4% specificity, and 39.1% accuracy without missing any nodules with suspicious or malignant cytology. There was no statistically sig-

Table 4
Comparison of criteria in identifying suspicious or malignant cytology (Bethesda IV,V,VI). 95% Confidence Interval in parenthesis. (all $p > 0.05$).

	ATA	TI-RADS	PEDTI-RADS	TI-RADS $\geq$ 3	TI-RADS $\geq$ 4
Sensitivity	90.00% (55.50–99.75)	70.00% (34.75–93.33)	90.00% (55.50–99.75)	100.00% (69.15–100.00)	80.00% (44.39–97.48)
Specificity	17.97% (11.74–25.73)	80.47% (72.53–86.94)	57.03% (47.99–65.74)	34.38% (26.21–43.28)	72.66% (64.08–80.16)
Accuracy	23.19% (16.43–31.13)	79.71% (72.03–86.07)	59.42% (50.74–67.69)	39.13% (30.94–47.80)	73.19% (64.99–80.37)
Positive Predictive Value	7.89% (6.42–9.67)	21.87% (14.07–32.39)	14.06% (10.94–17.90)	10.64% (9.50–11.89)	18.60% (13.07–25.80)
Negative Predictive Value	95.83% (77.55–99.35)	97.17% (92.99–98.89)	98.65% (91.87–99.79)	100.00%	97.89% (93.06–99.38)

ATA = American Thyroid Association.
TI-RADS = Thyroid Imaging Reporting and Data System.
FNA = Fine Needle Aspiration.
PED = Pediatric.

nificant difference between any of the outcomes of these criteria ($p > 0.05$).

Using TI-RADS criteria alone in our retrospective cohort had better sensitivity, specificity, positive predictive value and negative predictive value compared to ATA criteria, while potentially avoiding 82 biopsies. However, 6 suspicious or malignant nodules on cytology and pathology would have been missed using TI-RADS criteria, compared to 3 missed using ATA criteria. When a modified PED TI-RADS model was employed, potentially missed ($n = 3$) suspicious or malignant cytology and pathology was comparable to ATA, while 50 biopsies could have been avoided. In addition, using TI-RADS $\geq$ 3 as an indication for FNAB, with no defined size cutoff, had 100% sensitivity with no missed suspicious or malignant nodules on cytology or pathology in our patient cohort. Furthermore, 20 biopsies could have been avoided by using these criteria.

Thyroid nodules are much more common in adults than in pediatric patients. In adults, most thyroid nodules are benign [3,5,14]. The radiographic classification of thyroid nodules in adults is based on ATA risk stratification method and TI-RADS [15]. ATA classifies thyroid nodules by composite constellation of ultrasound patterns associated with suspicion for malignancy, ranging from benign to highly suspicious. However up to 4% of nodules cannot be classified using ATA [8]. TI-RADS ranges from TR1 (benign) to TR5 (high suspicion for malignancy), and nodules are scored according to the number of suspicious ultrasound features of malignancy within the nodule [16,17]. TI-RADS, an ultrasound-based risk stratification system to identify nodules, was implemented in adults to identify nodules most likely to represent malignant nodules and appropriate candidates for FNAB, while decreasing the number of biopsies performed on benign nodules [9]. In TI-RADS, recommendations for thyroid nodule management, biopsy, or sonographic follow-up are based on both TI-RADS level and its maximum diameter. The goal of TI-RADS in adults is not to diagnose every thyroid malignancy, but to balance the benefit of identifying clinically important cancers against the risk and cost of subjecting patients with benign nodules or indolent cancers to biopsy and treatment [9]. In children, however, a more aggressive approach is required, as thyroid nodules in children are more likely to be malignant than in adults [3].

Thyroid nodules are not common in pediatric patients, especially before 10 years of age. However, when a thyroid lesion is identified in a child, it carries a risk of malignancy 2–6 times higher (25–30%) than in adults (5%), with a higher risk of recurrence (39%). Adult-based systems, such as ATA and TI-RADS, have been attempted to be validated in pediatric patients [6,18]. Martinez-Rios et al. demonstrated that both ATA and TI-RADS are similar in their characteristics when classifying benign and malignant nodules in pediatric patients [6]. In this study, the ATA method was unable to categorize 5.8% of malignant and 2.8% of benign nodules. Based upon more objective components, TI-RADS has also been shown to have greater inter-rater reproducibility compared to ATA criteria [6]. ATA guidelines for pediatric patients differ from guidelines for adult patients, in that recommendations for FNAB are based on ultrasound characteristics (composition, echogenicity, shape, margins, echogenic foci and vascularity of the nodule) and not solely dependent on nodule size [11,19]. The features found to be most important in predicting the odds of a malignant diagnosis per ATA guidelines were 1) predominantly solid components (10.5x more than cystic), 2) irregular margins (53.2x more than smooth margins), and echogenic foci (3.5x more than no echogenic foci) with good intra- and interobserver agreements [14,18]. Furthermore, it has been shown that combining the ultrasound parameters increases diagnostic accuracy rather than using individual parameters alone [8,13–16].

It is important to distinguish between the Kwak and American College of Radiology (ACR) TI-RADS. Both scoring systems allocate points to suspicious ultrasound characteristics of the nodules, however Kwak version gives equal weight to all the features, whereas ACR version gives different weight to different characteristics [9,15,16]. We implemented ACR TI-RADS at our institution in May 2019. Even though ATA guidelines show promise, a recent retrospective study of ACR TI-RADS in 74 nodules (62 children) demonstrated very good discrimination between malignant and benign thyroid nodules in pediatric patients, making it a useful decision-making tool in the management of pediatric thyroid nodules [19]. In another retrospective study, eight radiologists reviewed 100 thyroid nodules that were histopathologically proven to be either benign (85%) or malignant (15%) [20]. After applying ACR TI-RADS criteria, overall sensitivity, specificity, and accuracy were 92% (95% CI: 68%, 98%), 44% (95% CI: 33%, 56%), and 52% (95% CI: 40%, 63%), respectively, whereas without ACR TI-RADS, the overall sensitivity, specificity and accuracy were 95%, 20%, and 28% respectively. Although fewer malignancies were recommended to undergo biopsy using the ACR TI-RADS criteria, the majority met the criteria for follow-up ultrasound, with only three of 100 (2.5%) malignancy encounters requiring no follow-up or biopsy. This study showed a significant improvement in the accuracy of recommendations for thyroid nodule management in pediatric patients.

We propose using a modified algorithm for the diagnostic management of pediatric thyroid nodules, which would take into consideration both the need to weigh each ultrasound characteristic and also the need to marginalize size criteria given thyroid volume changes with age [10,21,22]. When implemented, one such proposed pediatric specific criteria, TI-RADS $\geq$ 3, had 100% sensitivity with no missed suspicious or malignant nodules on cytology or pathology. Further efforts to incorporate concerning lymph node features into a pediatric specific algorithm should also be considered since children are more likely to present with metastatic disease to cervical lymph nodes at the time of diagnosis compared to adults [3,14,23].

The design of this study has certain limitations. Since ultrasound is a dynamic modality, a retrospective review of static images or limited sweeps results in inherent selection bias by the reviewer. Additionally, ultrasound is highly user-dependent and subjective to images captured, with human experts not perfectly consistent in how findings are interpreted, making a retrospective analysis of such images somewhat constrained. While the radiologists in this study demonstrated moderate inter-rater reliability, this could certainly be improved with live imaging and real-time interpretation, as well as further refinement of reporting tools including TIRADS. Other imaging modalities such as neck computed tomography or magnetic resonance imaging were not utilized for the purpose of this retrospective analysis. In addition, radiographic appearance of lymph nodes was not analyzed in this patient cohort, focusing on thyroid nodules only for evaluation and criteria for biopsy. It is possible that additional patients who were initially observed with benign cytology will undergo surgical resection in the future, altering the current results of the study over time. This is also a single-center study from a large tertiary referral center with perhaps bias towards capturing larger or more aggressive appearing thyroid nodules with enhanced oncological potential, and therefore the results may not be generalizable to other institutions without further prospective analysis and validation. Furthermore, due to the limited number of malignancies and since the TI-RADS criteria were applied retrospectively to nodules obtaining an FNAB using the ATA criteria (the analysis would be bias towards the ATA criteria), we were limited in our statistical power to compare the different criteria ability to diagnose malignant nodules. Finally, while the total TIRADS score was determined by the radiologists, the individual components of the scores were not recorded in this study, limiting the ability to compare specific sonographic features as more or less predicitive of malignancy. This is certainly data that would have been valuable to compare to ATA predictive features, and will be done in subsequent investigations.

Future area of prospective study will include validation of TI-RADS scoring in pediatric and adolescent patients, and a comparison of TI-RADS to ATA scoring system to accurately identify which patients should undergo FNAB. Confirmation bias, using two different types of gold standard for tissue diagnosis (cytology or surgical pathology) is likely unavoidable even with a prospective study since patients with benign FNAB will not be subjected to a surgery. In addition, a quality improvement analysis is ongoing with our institution's pediatric radiologists in efforts to improve objectivity and decrease variability in the interpretation of thyroid imaging.

5. Conclusion

Although both ATA and TI-RADS are comparable, using TI-RADS for diagnostic management of pediatric thyroid nodules improves accuracy in predicting malignancy, thereby decreasing unnecessary biopsies. We recommend using a modified pediatric algorithm with a scaled size criterion for improved selection of FNAB in pediatric and adolescent patients. Since the thyroid volume changes with age, nodule size should be considered less in the prediction of malignancy in pediatric patients. Further efforts to validate TI-RADS for pediatric thyroid nodules are required. Although our study is small and further prospective and larger studies are needed, we show an important trend in supporting TI-RADS guidelines in management of pediatric patients with thyroid nodules.

References

[1] Tomà P, Rossi UG, Magnaguagno F, Granata C, Laron Z. Imaging of the normal and affected thyroid in children with emphasis on sonography. Pediatr Endocrinol Rev 2003.

[2] Radhakrishnan R, Kline-Fath B. The pediatric head and neck. Diagnostic ultrasound. Elsevier; 2018. p. 1628–71.

[3] Qaisi M, Eid I. Pediatric head and neck malignancies. Oral Maxillofac Surg Clin N Am 2016. doi:10.1016/j.coms.2015.07.008.

[4] Ries L, Melbert D, Krapcho M, Stinchcomb D, Howlader N, Horner M, et al. SEER cancer statistics review, 1975–2005, National Cancer Institute. Bethesda, MD. Seer cancer stat rev 1975–2005. MD: Natl Cancer Institute Bethesda; 2008.

[5] Dinauer CA, Breuer C, Rivkees SA. Differentiated thyroid cancer in children: diagnosis and management. Curr Opin Oncol 2008. doi:10.1097/CCO.0b013e3282f30220.

[6] 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. doi:10.1007/s00247-017-3974-y.

[7] Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2006. doi:10.1089/thy.2006.16.109.

[8] Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, 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. doi:10.1089/thy.2015.0020.

[9] 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. doi:10.1016/j.jacr.2017.01.046.

[10] Ueda D. Normal volume of the thyroid gland in children. J Clin Ultrasound 1990. doi:10.1002/jcu.1870180602.

[11] Francis GL, Waguespack SG, Bauer AJ, Angelos P, Benvenga S, Cerutti JM, et al. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. Thyroid 2015. doi:10.1089/thy.2014.0460.

[12] Richman DM, Benson CB, Doubilet PM, Wassner AJ, Asch E, Cherella CE, et al. Assessment of American college of radiology thyroid imaging reporting and data system (TI-RADS) for pediatric thyroid nodules. Radiology 2020. doi:10.1148/radiol.2019191326.

[13] Gharib H, Papini E, Paschke R, Duick D, Valcavi R, Hegedus L. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. J Endocrinol Investig 2010;33:51–6.

[14] Moon W, Jung S, Lee J, Na D, Baek J, Lee Y. Benign and malignant thyroid nodules: US differentiation-multicenter retrospective study. Radiology 2008;247:762–70.

[15] Frates M, Benson C, Charboneau J, Cibas E, Clark O, Coleman B. Management of thyroid nodules detected at US: society of Radiologists in Ultrasound consensus conference statement. Ultrasound Q 2006;22:231–8.

[16] Grant E, Tessler F, Hoang J, Langer J, Beland M, Berland L. Thyroid ultrasound reporting Lexicon: white paper of the ACR Thyroid Imaging, Reporting and Data System (TIRADS) committee. J Am Coll Radiol 2015;12:1272–9.

[17] Shapira-Zaltsberg G, Miller E, Martinez-Rios C, Bass J, Goldbloom EB, Tang K, et al. Comparison of the diagnostic performance of the 2017 ACR TI-RADS guideline to the Kwak guideline in children with thyroid nodules. Pediatr Radiol 2019. doi:10.1007/s00247-019-04385-6.

[18] Lim-Dunham J, Toslak IE, Alsabban K, Aziz A, Martin B, Okur G. 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;47:429–36.

[19] Lim-Dunham JE, Toslak IE, Reiter MP, Martin B. Assessment of the American college of radiology thyroid imaging reporting and data system for thyroid nodule malignancy risk stratification in a pediatric population. Am J Roentgenol 2019. doi:10.2214/AJR.18.20099.

[20] Hoang JK, Middleton WD, Farjat AE, Langer JE, Reading CC, Teefey SA, et al. Reduction in thyroid nodule biopsies and improved accuracy with American college of radiology thyroid imaging reporting and data system. Radiology 2018. doi:10.1148/radiol.2018172572.

[21] Suzuki S, Midorikawa S, Fukushima T, Shimura H, Ohira T, Ohtsuru A. Systematic determination of thyroid volume by ultrasound examination from infancy to adolescence in Japan: the Fukushima Health Management Survey. Endocr J 2015;62:261–8.

[22] Ueda D. Sonographic measurement of the volume of the thyroid gland in healthy children. Acta Paediatr Jpn Overseas Ed 1989;31:352–4.

[23] Rivkees S, Mazzaferri E, Verburg F, Reiners C, Luster M, Breuer C. The treatment of differentiated thyroid cancer in children: emphasis on surgical approach and radioactive iodine therapy. Endocr Rev 2011;32:798–826.