MDPI

Article

Inter-Reader Agreement of ATA Sonographic Risk in Thyroid Nodules with Bethesda Category III Indeterminate Cytology

Maria I. Linares Valderrama [1], Marilyn A. Arosemena [2,*], Anu Thekkumkattil [1], Russ A. Kuker [3], Rosa P. Castillo [3], Charif Sidani [3], Manuel Lora Gonzalez [4], Sabina Casula [5,6] and Atil Y. Kargi [5]

1   Division of Endocrinology, Jackson Memorial Hospital & University of Miami, Miami, FL 33136, USA;
     maria.linaresvalderrama@clinicasierravista.org (M.I.L.V.); anuthekk@gmail.com (A.T.)
2   Kovler Diabetes Center and the Department of Medicine, The University of Chicago, Chicago, IL 60637, USA
3   Department of Radiology, University of Miami, Miami, FL 33136, USA; rkuker2@med.miami.edu (R.A.K.);
     rcastillo3@med.miami.edu (R.P.C.); csidani@med.miami.edu (C.S.)
4   Department of Pathology, University of Alabama at Birmingham Hospital, Birmingham, AL 35233, USA;
     mlora@uab.edu
5   Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Miami Miller
     School of Medicine, Miami, FL 33136, USA; Sabina.Casula@va.gov (S.C.); akargi@med.miami.edu (A.Y.K.)
6   Department of Endocrinology, Miami Veterans Affairs Healthcare System, Miami, FL 33125, USA
*   Correspondence: Marilyn.Arosemena@uchospitals.edu

Abstract: Background: Substantial inter-observer variation has been documented in the recognition and description of specific sonographic features as well as for ATA sonographic risk (ASR). This raises the question if the risk stratification proposed by the ATA guidelines is reproducible and applicable for nodules with indeterminate cytology. The aim of the study was to determine the inter-reader agreement (IRR) among radiologists using the 2015 ASR stratification in indeterminate thyroid nodules. Methods: Three board certified radiologists who were blinded to clinical data and to each other, interpreted the ultrasound findings of 179 nodules that had Bethesda III cytology. The nodules were classified into high suspicion (HS), intermediate (IS), low (LS), very low (VLS). Echogenicity, composition, shape taller than wide, vascularity, type of margins, presence and type of calcifications were also described. Results: The majority consensus revealed that 28%, 27%, 39% and 5% were described as high, intermediate, low and very low ASR, respectively. The inter-reader agreement was near perfect (k 0.82 CI 95% (0.77–0.87)). Nodules were paired into a higher risk (HS + IS) and lower risk (LS + VLS) categories with substantial agreement (k 0.7) in both categories. Conclusion: A near perfect agreement among readers was observed when stratifying indeterminate cytology nodules for ASR.

Keywords: thyroid nodule; Bethesda III; indeterminate cytology; ultrasound; sonographic risk; ATA; inter-reader; agreement

1. Introduction

The monographic patterns proposed by the 2015 American Thyroid Association guidelines have been used to risk stratify thyroid nodules into five categories with corresponding estimated malignancy risks: High (70–90%), Intermediate (10–20%), Low (5–10%), Very Low (<3%) suspicion and benign (<1%) nodules. Nodules with high and intermediate suspicion patterns are recommended to undergo FNA biopsy at a lower threshold of 1 cm while those with low and very low suspicion patterns have higher thresholds of 1.5 and 2 cm, respectively [1].

Certain sonographic features have a higher association with risk of malignancy. For example, being hypoechoic provides enough risk to be an intermediate or high suspicion pattern nodule, whereas having a predominantly cystic composition corresponds to lower suspicion patterns. With different size thresholds for FNA biopsy, accurately describing each sonographic feature and categorizing each nodule into a certain ATA sonographic risk

(ASR) pattern conveys great clinical significance. Substantial inter-observer variation has been documented in the recognition and description of specific sonographic features as well as for ATA sonographic risk [2,3]

Interestingly, the use of the ATA sonographic risk patterns after obtaining a cytology result has not been outlined clearly. In the case of malignant or benign FNA cytology result, management is generally clear with resection or observation recommended correspondingly. However, the management is more variable with indeterminate cytology nodules that offer a significantly variable probability of malignancy. In the specimens diagnosed as atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS), the low interobserver concordance when interpreting the cytology specimens [4], paired with the wide range of management strategies that go from surveillance to diagnostic surgery pose a challenge for clinicians to determine how aggressive the evaluation and management of these particular nodules should be.

Recently, several molecular tests have been utilized to further risk stratify ITNs and guide management, but their use is limited by cost and availability. A study by our group has recently evaluated the association between ASR and Afirma gene expression classifier (GEC) and determined individual and combined diagnostic performances [5].

Some authors have recommended the use of sonographic classifications as adjunctive predictors of malignancy for indeterminate cytology nodules [6]. The ATA sonographic risk stratification in indeterminate cytology nodules has been found to be limited as shown by the low interobserver agreement of the distribution of the sonographic patterns as well as the finding of a large number of nodules that did not fit into any of the proposed risk patterns [7]. This result raises the question if the risk stratification proposed by the ATA guidelines is reproducible and applicable for nodules with indeterminate cytology.

Our objective is to determine the inter-reader agreement among radiologists for individual sonographic features as well as the ATA sonographic risk stratification when applied to thyroid nodules with Bethesda category III indeterminate cytology.

2. Materials and Methods

A retrospective chart review of patients with Bethesda category III nodules (Atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS)) from 1 January 2012 to 31 December 2017 was performed. Patients were 18 years or older and had one or more thyroid nodules confirmed by ultrasound. Three board certified radiologists at our academic medical center who were blinded to clinical data, pathology results as well as to each other's reports, interpreted the ultrasound (US) findings of 179 nodules with Bethesda category III cytology. Ultrasound used was Siemens Acusón S3000, Transducer 12 MHz.

Each radiologist had received the 2015 ATA guidelines for thyroid nodules risk stratification at least one month prior to the beginning of the study. One radiologist had experience with thyroid ultrasound imaging; one radiologist had general ultrasound imaging experience and the last radiologist had neuroimaging experience. All radiologists used picture archiving computerized system to review dedicated thyroid ultrasound images. Ultrasound machine

This study was approved by the Institutional Review Board of the University of Miami Miller School of Medicine. Informed consent was waived as the study was retrospective in nature.

The nodules were risk stratified and classified into high (HS), intermediate (IS), low (LS), very low (VLS) or non-ATA (NA) risk categories. Hypoechoic nodules exhibiting at least one suspicious feature (microcalcifications, irregular margins, shape taller than wide, presence of suspicious lymph nodes) were classified as HS pattern. Hypoechoic nodules without any suspicious feature were classified as IS pattern. Iso or hyperechoic solid, or partly cystic nodules with eccentric solid areas, were classified as LS pattern. Spongiform or partly cystic nodules without eccentric solid areas or other suspicious features were classified as VL pattern. Radiologists determined NA risk nodules when

describing characteristics that did not correspond to any risk pattern in the 2015 ATA classification. These nodules included iso or hyperechoic nodules with at least one high risk feature or mixed echogenicity nodules.

In addition, 7 features were described: echogenicity, composition, shape taller than wide, vascularity, margins, calcifications and presence of suspicious lymph nodes. Echogenicity was described as hypoechoic, isoechoic or hyperechoic in contrast to the normal thyroid parenchyma. Composition was classified into mixed solid-cystic with >50% solid, mixed cystic-solid with >50% cystic, mixed and spongiform. The shape of the nodule was classified as taller than wide or not taller than wide in the transverse view. Vascularity was classified as internal, peripheral, mixed or no vascularity. Nodule margins were classified into regular, irregular spiculated, irregular microlobulated and irregular infiltrative. Calcifications were classified into microcalcifications, macrocalcifications or none. Suspicious lymph nodes were classified into present or absent.

A consensus result was obtained for ATA risk stratification and each sonographic feature when at least 2 of the raters agreed on the result.

Inter-rater agreement between 3 observers for the ATA sonographic risk pattern and for each sonographic feature was calculated using Fleiss kappa statistics with a 95% confidence interval. Accepted kappa (k) values interpretation deems inter-rater agreement to be near perfect if k ranges from 0.81–1, substantial from 0.61–0.8, moderate if 0.41–0.6, fair if k 0.21–0.4 and poor if k ≤ 0.2.

3. Results

The population on our study consisted of 179 patients with 179 nodules with Bethesda III cytology. Of the patients, 81% were female and 49% of patients were Hispanic. Mean age was 57.7 years at the time of FNA biopsy. Three per cent of patients had a family history of thyroid cancer, and 12% had a history of hypothyroidism. Mean TSH was 2.26 mU/L (Table 1).

Table 1. Patient's demographics.

Patient (n = 179) and Thyroid Nodule Characteristics.	
Sex	n (%)
Female	145 (81)
Male	34 (19)
Age	
Mean (SD)	57.7 (44–71)
Race	
Asian	1 (0.6)
Black	13 (7.3)
Hispanic	88 (49.1)
White	46 (25.7)
Other	31 (3.9)
Family history of thyroid cancer	
Yes	6 (3.3)
No	166 (92.7)
Unknown	7 (3.9)
Hypothyroidism	
Yes	21 (11.7)
No	158 (88.2)
Nodules: Greatest diameter on ultrasound, cm	
Mean (range)	1.98 (0.95–3)
TSH	
≤0.5	11 (6.1)
0.51–2.5	104 (58.1)
2.51–4.5	34 (18.9)
>4.5	3 (1.7)
Not available TSH	27 (15.1)

Out of 179 nodules, 178 were able to be classified into an ASR category by majority consensus, revealing that 50 (28%), 48 (27%), 70 (39%) and 10 (5%) were described as high, intermediate, low and very low ASR, respectively (Table 2). One nodule showed absolute disparity in ASR classification between 3 readers. None of the nodules were classified in the Non ATA risk category by majority consensus. Only rater 3 classified 5 nodules as Non ATA, whereas rater 1 and 2 were able to classify every nodule into one ASR pattern (Table 3). For the 5 nodules described above, raters 1 and 2 classified them as low and intermediate ASR patterns. The overall inter-reader agreement for ASR across all three readers was near perfect (k 0.82 CI 95% (0.77–0.87)).

Table 2. Frequency of ATA sonographic risk, individual sonographic features and Inter-reader agreement calculated with Fleiss Kappa for 95% confidence interval.

	Majority Consensus n (%)	Inter-Reader Agreement 95% CI
ATA sonographic risk (n = 179)		k 0.82 CI (0.77–0.87)
High suspicious (HS)	50 (28%)	
Intermediate suspicious (IS)	48 (27%)	
Low suspicious (LS)	70 (39%)	
Very low suspicious (VLS)	10 (5%)	
Non ATA (NA)	0	
No consensus	1	
Higher risk categories (HS + IS)	98 (55%)	k 0.77 CI (0.66–0.87)
Lower risk categories (LS + VLS)	80 (45%)	k 0.7 CI (0.6–0.8)
Individual sonographic features		
Echogenicity (n = 179)		k 0.71 CI (0.65–0.77)
- Hypoechoic	78 (44%)	
- Isoechoic	52 (29%)	
- Hyperechoic	44 (25%)	
- No consensus	5 (2%)	
Composition (n = 179)		k 0.87 CI (0.81–0.93)
- Mixed >50% solid	141 (79%)	
- Mixed >50% cystic	10 (6%)	
- Mixed	20 (11%)	
- Spongiform	7 (4%)	
Shape (n = 179)		k 0.7 CI (0.6–0.74)
- Taller than wide	7 (4%)	
- Not taller than wide	172 (96%)	
Margins (n = 179)		k 0.72 CI (0.63–0.81)
- Irregular	103 (57.5%)	
- Regular	75 (42%)	
- No consensus	1 (0.5%)	
Irregular margins (n = 103)		k 0.5 CI (0.43–0.58)
- Spiculated	21 (20%)	
- Microlobulated	74 (72%)	
- Infiltrative	8 (8%)	
Calcifications (n = 179)		
- Present	29 (16%)	
- Not present	150 (84%)	
Type of calcification (n = 29)		k 0.82 CI (0.67–0.96)
- Microcalcification	16 (55%)	
Macrocalcification	8 (28%)	
- Peripheral (egg-shell)	5 (17%)	
Suspicious lymph node (n = 129)		k 0.72 CI (0.62–0.82)
- Present	4 (3%)	
- Not present	149 (97%)	
Vascularity (n = 129)		k 0.47 CI (0.41–0.53)
- Mixed	89 (69%)	
- Internal	11(9%)	
- Peripheral	50 (4%)	
- No vascularity	19 (15%)	
- No consensus	4 (3%)	

Table 3. Interpretation of ATA sonographic risk per rater.

ATA Sonographic Risk (ASR)	Rater 1	Rater 2	Rater 3
High suspicious (HS)	46	50	51
Intermediate suspicious (IS)	47	56	43
Low suspicious (LS)	76	64	70
Very low suspicious (VLS)	10	9	10
Non ATA (NA)	0	0	5

Due to similar clinical management and FNA size thresholds, we paired the HS and IS pattern nodules into a higher risk category and the LS and VLS pattern nodules into a lower risk category. The inter-reader agreement in the higher risk category (k 0.77 CI 95% (0.66–0.87)) as well as the lower risk category (k 0.7 CI 95% (0.6–0.8)) was substantial. We observed that raters disagreed when classifying 15 nodules between the higher and lower risk categories. Notably Rater 2 classified the majority of these nodules as intermediate risk whereas Rater 1 and 3 classified the majority of them as low risk. Eight of them were finally classified as low risk and 6 of them as intermediate risk with one nodule showing absolute disparity between raters (Table 4).

Table 4. Nodules with disparity between higher (HS + IS) and lower risk (L + VL) ASR categories as interpreted by each rater (n = 15).

ATA Sonographic Risk (ASR)	Rater 1	Rater 2	Rater 3
High suspicious (HS)	0	0	0
Intermediate suspicious (IS)	2	14	5
Low suspicious (LS)	13	1	9
Very low suspicious (VLS)	0	0	0
Non ATA (NA)	0	0	1
	LS ASR	IS ASR	No consensus
Consensus	8	6	1

There were 28 nodules in which majority consensus was achieved but no absolute agreement between raters when classifying their ASR. The overall agreement rates for individual features of these particular nodules were lower for the description of irregular margins (48%) and echogenicity (54%) (Table 5).

Table 5. Overall agreement rates for nodules that achieved majority consensus but no absolute consensus between raters when classifying for ASR (n = 28).

Individual Sonographic Features	Overall Agreement (%)
Composition	26/28 (92%)
Calcification	24/28 (86%)
Suspicious lymph node	22/28 (79%)
Vascularity	16/28 (57%)
Margins	16/28 (57%)
Echogenicity	15/28 (54%)
Type of irregular margins	10/21 (48%)

The majority of nodules were described as hypoechoic (78/179), >50% solid (141/179), with irregular margins (103/179), no calcifications (150/179), wide shape (172/179), mixed vascularity (89/179) and without a suspicious lymph node (149/179). The majority of nodules with irregular margins were microlobulated (74/103). For the nodules with calcifications, the majority were described with microcalcifications (16/29).

The inter-reader agreement for individual sonographic features was near perfect for composition (k 0.87), presence (k 0.85) and type of calcifications (k 0.82), substantial for

echogenicity (k 0.71), margins (k 0.72), shape (k 0.7) and presence of suspicious lymph node (k 0.72) and moderate for vascularity (k 0.47) and type of irregular margins (k 0.5) (Table 2).

4. Discussion

Ultrasonography has an essential role in assessing the malignancy risk of thyroid nodules. Multiple risk stratifying systems have been proposed in order to estimate risk based on a composite of thyroid nodule features [8–11]. The 2015 ATA guidelines are widely used to risk stratify thyroid nodules offering cutoffs for FNA biopsy depending on their estimated malignancy risk [1]. It has been observed that these systems allow a more reproducible stratification of thyroid nodules on the basis of a higher inter-reader agreement when compared to a classification based on single suspicious features [3].

However, when applied to indeterminate cytology nodules, the 2015 ATA sonographic risk stratification showed only fair to borderline acceptable inter-reader agreement [2,3,7,12] which raises the question if this system is applicable for this specific subset of thyroid nodules. Furthermore, its diagnostic performance in ITNs conveys conflicting results with some authors suggesting appropriate prediction of malignancy [13–15] and others who question its adjunctive diagnostic value, with a recent study suggesting it only be used to set the threshold for FNA [16].

In our study of a large cohort of nodules with Bethesda III cytology, we observed that the inter-reader agreement or inter-rater reliability (IRR) for the ATA sonographic risk stratification between 3 radiologists was near perfect. On a recent publication by Lam et al., the greatest source of disagreement was found to be the description of echogenicity with an IRR of k 0.35 (fair agreement) driven by a substantial number of "heteroechoic nodules" unable to fit into any ASR category, as well as different thresholds for the interpretation of echogenicity by the observers [7].

Interestingly, in our study, all three readers had the option to elect a non-ATA category; however, only one rater used this option for a total of five of the 179 nodules analyzed. In comparison, the other two raters were able to classify every single nodule with the ASR system. The observers in this study had different backgrounds in Radiology, as we wanted to mimic real-life circumstances and were only given the ATA sonographic risk guidelines an average of 30 days before initiating the study. This distinction probably indicates that US interpreters may be able to associate non-ATA thyroid nodules to one of the ATA risk patterns guided by presence or absence of suspicious features.

In that sense, we observed that in the 28 nodules in which no absolute consensus was achieved when classifying into ASR, the feature in which raters disagreed the most was echogenicity followed by the description of margins. Both features alone can determine the difference between low, intermediate or high suspicious ASR nodules.

Nonetheless, the overall inter-reader agreement for specific sonographic features of all nodules with Bethesda III cytology across the 3 observers was excellent for every feature except for vascularity and type of irregular margins that ultimately do not determine ATA sonographic risk. Echogenicity carries the most substantial significance in the ATA sonographic risk stratification as hypoechoic attenuation has been shown to be a strong predictor of malignancy [1,7,17]. In our study, the IRR for echogenicity was substantial with kappa value of 0.71. In previous cohorts of no specific subsets of thyroid nodules, echogenicity was consistently among the features with lower IRR (0.33–0.57) showing only fair to moderate agreement [3,18–26]. Conversely, in one study including only nodules with Bethesda III cytology, echogenicity showed the highest IRR among other 5 individual features, with k value of 0.94. Notably all other features had moderate to substantial agreement as well [26]. Along with the results of our study, it could be suggested that this disparity is explained by the fact that nodules with Bethesda III cytology are a subset with more homogenous US characteristics. However, the results by Lam et al. would contradict this suggestion [7].

The effect of ATA sonographic stratification in nodules with known Bethesda III cytology may not determine clinical management at this time, but some authors recommend

its use as an adjunctive predictor of malignancy. In that sense, different sonographic patterns will determine different thresholds for FNA biopsy. The ATA 2015 guidelines propose the same threshold for high and intermediate suspicious nodules and a similar threshold for low and very low suspicious nodules (1). We decided to pair higher risk categories (HS + IS) and lower risk categories (LS + VLS) and found a substantial IRR in both groups as expected from an overall near perfect IRR. However, we observed that 15 nodules showed disparity between the higher and lower risk ASR categories as interpreted by each rater. Interestingly, there is a notable predilection for rater 1 and 3 to classify these nodules as low risk ASR whereas rater 2 classified almost all of them as intermediate risk ASR. This finding suggests different thresholds for echogenicity characterization by each rater as already suggested by Lam et al. which could potentially impact critical patient management.

Our results indicate that in contrast to previous results, the inter-reader agreement for the ATA sonographic risk stratification of nodules with Bethesda III cytology was near perfect. None of the nodules were classified in the Non ATA risk category by majority consensus. All individual features proposed by the 2015 ATA guidelines showed substantial to near perfect agreement between the readers. The sonographic features in which more robust agreement was observed were composition and presence of calcifications. The most common features of disagreement were vascularity and description of irregular margins. Nevertheless, these two features do not determine ASR. Echogenicity showed overall substantial agreement; however, it was the most common feature of disagreement in the subset of nodules that did not achieve absolute consensus between all three raters. Disparity between the classification in the lower and higher risk categories was observed in 15 cases out of 179, which could potentially impact critical patient management.

References

1. Haugen, B.R.; Alexander, E.K.; Bible, K.C.; Doherty, G.M.; Mandel, S.J.; Nikiforov, Y.E.; Pacini, F.; Randolph, G.W.; Sawka, A.M.; Schlumberger, M.; et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **2016**, *26*, 1–133. [CrossRef] [PubMed]
2. Pandya, A.; Caoili, E.M.; Jawad-Makki, P.F.; Wasnik, A.P.; Shankar, P.R.; Bude, R.; Haymart, M.R.; Davenport, M.S. Limitations of the 2015 ATA Guidelines for Prediction of Thyroid Cancer: A Review of 1947 Consecutive Aspirations. *J. Clin. Endocrinol. Metab.* **2018**, *103*, 3496–3502. [CrossRef] [PubMed]
3. 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. [CrossRef]
4. Cibas, E.S.; Baloch, Z.W.; Fellegara, G.; Livolsi, V.A.; Raab, S.S.; Rosai, J.; Diggans, J.; Friedman, L.; Kennedy, G.C.; Kloos, R.T.; et al. A Prospective Assessment Defining the Limitations of Thyroid Nodule Pathologic Evaluation. *Ann. Intern. Med.* **2013**, *159*, 325–332. [CrossRef] [PubMed]

5. Arosemena, M.; Thekkumkattil, A.; Valderrama, M.I.L.; Kuker, R.A.; Castillo, R.P.; Sidani, C.; Gonzalez, M.L.; Casula, S.; Kargi, A.Y. American Thyroid Association Sonographic Risk and Afirma Gene Expression Classifier Alone and in Combination for the Diagnosis of Thyroid Nodules with Bethesda Category III Cytology. *Thyroid* **2020**, *30*, 1613–1619. [CrossRef] [PubMed]

6. Ahmadi, S.; Herbst, R.; Oyekunle, T.; Jiang, X.; Strickland, K.; Roman, S.; Sosa, J.A. Using the Ata and Acr Ti-Rads Sonographic Classifications as Adjunctive Predictors of Malignancy for Indeterminate Thyroid Nodules. *Endocr. Pract.* **2019**, *25*, 908–917. [CrossRef]

7. Lam, C.A.; Mcgettigan, M.J.; Thompson, Z.J.; Khazai, L.; Chung, C.H.; Centeno, B.A.; McIver, B.; Valderrabano, P. Ultrasound characterization for thyroid nodules with indeterminate cytology: Inter-observer agreement and impact of combining pattern-based and scoring-based classifications in risk stratification. *Endocrine* **2019**, *66*, 278–287. [CrossRef]

8. Tessler, F.N.; Middleton, W.D.; Grant, E.G.; Hoang, J.K.; Berland, L.L.; Teefey, S.A.; Cronan, J.J.; Beland, M.D.; Desser, T.S.; Frates, M.C.; 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. [CrossRef]

9. Gharib, H.; Papini, E.; Valcavi, R.; Baskin, H.J.; Crescenzi, A.; Dottorini, M.E.; Duick, D.S.; Guglielmi, R.; Hamilton, C.R.; Zeiger, M.A.; et al. American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. *Endocr. Pract.* **2006**, *12*, 63–102. [CrossRef]

10. Russ, G.; Bonnema, S.J.; Erdogan, M.F.; 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. [CrossRef]

11. Shin, J.H.; Baek, J.H.; Chung, J.; Ha, E.J.; Kim, J.-H.; Lee, Y.H.; Lim, H.K.; Moon, W.-J.; Na, D.G.; Park, J.S.; 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. [CrossRef]

12. Valderrabano, P.; Mcgettigan, M.J.; Lam, C.A.; Khazai, L.; Thompson, Z.J.; Chung, C.H.; Centeno, B.A.; McIver, B. Thyroid Nodules with Indeterminate Cytology: Utility of the American Thyroid Association Sonographic Patterns for Cancer Risk Stratification. *Thyroid* **2018**, *28*, 1004–1012. [CrossRef] [PubMed]

13. Grani, G.; Lamartina, L.; Ascoli, V.; Bosco, D.; Nardi, F.; D'Ambrosio, F.; Rubini, A.; Giacomelli, L.; Biffoni, M.; Filetti, S.; et al. Ultrasonography scoring systems can rule out malignancy in cytologically indeterminate thyroid nodules. *Endocrine* **2017**, *57*, 256–261. [CrossRef] [PubMed]

14. Barbosa, T.L.M.; Junior, C.O.M.; Graf, H.; Cavalvanti, T.; Trippia, M.A.; Ugino, R.T.D.S.; De Oliveira, G.L.; Granella, V.H.; De Carvalho, G.A. ACR TI-RADS and ATA US scores are helpful for the management of thyroid nodules with indeterminate cytology. *BMC Endocr. Disord.* **2019**, *19*, 1–11. [CrossRef]

15. Tang, A.L.; Falciglia, M.; Yang, H.; Mark, J.R.; Steward, D.L. Validation of American Thyroid Association ultrasound risk as-sess-ment of thyroid nodules selected for ultrasound fine-needle aspiration. *Thyroid* **2017**, *27*, 1077–1082. [CrossRef]

16. Giuliano, S.; Mirabelli, M.; Chiefari, E.; Vergine, M.; Gervasi, R.; Brunetti, F.S.; Innaro, N.; Donato, G.; Aversa, A.; Brunetti, A. Ma-lignancy Analyses of Thyroid Nodules in Patients Subjected to Surgery with Cytological-and Ultrasound-Based Risk Stratification Systems. *Endocrines* **2020**, *1*, 102–118. [CrossRef]

17. Na, D.G.; Baek, J.H.; Sung, J.Y.; Kim, J.-H.; Kim, J.K.; Choi, Y.J.; Seo, H. Thyroid Imaging Reporting and Data System Risk Stratification of Thyroid Nodules: Categorization Based on Solidity and Echogenicity. *Thyroid* **2016**, *26*, 562–572. [CrossRef] [PubMed]

18. Kim, S.H.; Park, C.S.; Jung, S.L.; Kang, B.J.; Kim, J.Y.; Choi, J.J.; Kim, Y.I.; Oh, J.K.; Oh, J.S.; Kim, H.; et al. Observer Variability and the Performance between Faculties and Residents: US Criteria for Benign and Malignant Thyroid Nodules. *Korean J. Radiol.* **2010**, *11*, 149–155. [CrossRef]

19. Choi, S.H.; Kim, E.-K.; Kwak, J.Y.; Kim, M.J.; Son, E.J. Interobserver and Intraobserver Variations in Ultrasound Assessment of Thyroid Nodules. *Thyroid* **2010**, *20*, 167–172. [CrossRef] [PubMed]

20. Lim-Dunham, J.E.; Toslak, I.E.; Alsabban, K.; Aziz, A.; Martin, B.; Okur, G.; Longo, K.C. 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. [CrossRef]

21. Park, S.H.; Kim, S.J.; Kim, E.-K.; Kim, M.J.; Son, E.J.; Kwak, J.Y. Interobserver Agreement in Assessing the Sonographic and Elastographic Features of Malignant Thyroid Nodules. *Am. J. Roentgenol.* **2009**, *193*, W416–W423. [CrossRef] [PubMed]

22. Park, C.S.; Kim, S.H.; Jung, S.L.; Kang, B.J.; Kim, J.Y.; Choi, J.J.; Sung, M.S.; Yim, H.W.; Jeong, S.H. Observer variability in the sonographic evaluation of thyroid nodules. *J. Clin. Ultrasound* **2010**, *38*, 287–293. [CrossRef] [PubMed]

23. Park, S.; Choi, Y.J.; Kim, D.W.; Son, E.J.; Lee, H.S.; Yoon, J.H.; Kim, E.-K.; Moon, H.J.; Kwak, J.Y. Interobserver Variability and Diagnostic Performance in US Assessment of Thyroid Nodule According to Size. *Ultraschall Med. Eur. J. Ultrasound* **2012**, *33*, E186–E190. [CrossRef] [PubMed]

24. Wienke, J.R.; Chong, W.K.; Fielding, J.R.; Zou, K.H.; Mittelstaedt, C.A. Sonographic Features of Benign Thyroid Nodules. *J. Ultrasound Med.* **2003**, *22*, 1027–1031. [CrossRef] [PubMed]

25. Phuttharak, W.; Boonrod, A.; Klungboonkrong, V.; Witsawapaisan, T. Interrater Reliability of Various Thyroid Imaging Reporting and Data System (TIRADS) Classifications for Differentiating Benign from Malignant Thyroid Nodules. *Asian Pac. J. Cancer Prev.* **2019**, *20*, 1283–1288. [CrossRef]

26. Norlén, O.; Popadich, A.; Kruijff, S.; Gill, A.J.; Sarkis, L.M.; Delbridge, L.; Sywak, M.; Sidhu, S. Bethesda III Thyroid Nodules: The Role of Ultrasound in Clinical Decision Making. *Ann. Surg. Oncol.* **2014**, *21*, 3528–3533. [CrossRef]