Prospective Validation of ATA and ETA Sonographic Pattern Risk of Thyroid Nodules Selected for FNAC

Fabio Maino,[1] Raffaella Forleo,[1] Martina Martinelli,[1] Noemi Fralassi,[1] Filomena Barbato,[1] Tania Pilli,[1] Marco Capezzone,[1] Lucia Brilli,[1] Cristina Ciuoli,[1] Giovanni Di Cairano,[1] Laura Nigi,[1] Furio Pacini,[1] and Maria Grazia Castagna[1]

[1]Department of Medical, Surgical, and Neurological Sciences, University of Siena, 53100 Siena, Italy

Context: Recently, the American Thyroid Association (ATA) and the European Thyroid Association (ETA) have proposed that thyroid ultrasound (US) should be used to stratify the risk of malignancy in thyroid nodules and to aid decision-making about whether fine-needle aspiration cytology (FNAC) is indicated.

Objective: To validate and to compare the ATA and ETA US risk stratification systems of thyroid nodules in a prospective series of thyroid nodules submitted to FNAC.

Setting: We prospectively evaluated 432 thyroid nodules selected for FNAC from 340 patients. Cytology reports were based on the five categories according to the criteria of the British Thyroid Association.

Results: The proportion of Thy2 nodules decreased significantly, whereas the proportion of Thy4/Thy5 nodules significantly increased with increasing US risk class ($P < 0.0001$). The ability to identify benign and malignant nodules was similar between ATA and ETA systems. According to ATA and ETA US risk stratification systems, 23.7% and 56.0% nodules did not meet the criteria for FNAC, respectively. Considering only categories at lower risk of malignancy, the cumulative malignancy rate in these nodules was 1.2% for ATA and 1.7% for ETA US risk stratification systems.

Conclusions: ETA and ATA US risk stratification systems provide effective malignancy risk stratification for thyroid nodules. In clinical practice, using this approach, we should be able to reduce the number of unnecessary FNAC without losing clinically relevant thyroid cancer. (***J Clin Endocrinol Metab*** **103: 2362–2368, 2018)**

The "routine" use of neck ultrasound (US) in clinical practice has contributed to the increased detection of thyroid nodules and in particular of small nonpalpable nodules (1, 2). Among the vast amount of thyroid nodules detected in clinical practice, only 2% to 16% are malignant (3). For the differential diagnosis between benign and malignant thyroid nodules, fine-needle aspiration cytology (FNAC) is considered an accurate and cost-effective method with a high diagnostic sensitivity and specificity (4). To reduce unnecessary FNAC, US stratification systems have been proposed to help the physician in the management of thyroid nodules (5–10). According to the 2015 American Thyroid Association (ATA) US risk stratification, thyroid nodules are classified into five categories according to the combination of US features. FNAC is recommended in high- and intermediate-suspicion categories for nodules larger than 1 cm and in low-suspicion categories for nodules >1.5 cm. In very low–suspicion categories, FNAC should be considered for nodules larger than 2.0 cm, whereas it is not recommended for benign nodules (6). Similar to the 2015 ATA US risk stratification system, a well-structured reporting system based on US features has been recently proposed by the European Thyroid

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 2 February 2018. Accepted 11 April 2018.
First Published Online 16 April 2018

Abbreviations: ATA, American Thyroid Association; ETA, European Thyroid Association; EU-TIRADS, European Thyroid Imaging and Reporting Data System; FN, false negative; FNAC, fine-needle aspiration cytology; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive; US, ultrasound.

Association (ETA) (10). The European Thyroid Imaging and Reporting Data System (EU-TIRADS) classified thyroid nodules into four categories, and FNAC is recommended for EU-TIRADS 4 and 5 nodules larger than 1.5 and 1 cm, respectively. For EU-TIRADS 3 nodules, FNAC is indicated if nodules are larger than 2.0 cm, whereas it is not indicated for EU-TIRADS 2 nodules (10). ATA US risk stratification systems of thyroid nodules have been recently validated in a prospective study based upon cytological and surgical pathology results in 206 thyroid nodules (11). No prospective study has been conducted according to the ETA US stratification system.

The first aim of our study was to validate the ATA and ETA US risk stratification systems of thyroid nodules in a prospective series of thyroid nodules submitted to FNAC. The second aim was to compare the two US staging systems to verify the diagnostic accuracy and the ability to reduce the number of unnecessary FNACs without impact in the clinical diagnosis of thyroid cancer.

Patients and Methods

Study population

From November 2016 to June 2017, we prospectively collected from routine clinical care, all nodules submitted to FNAC for diagnostic purposes. All patients provided informed consent for both traditional cytology and molecular analysis at the time of FNAC. The study was approved and exempted by the institutional committee. The indication for FNAC was taken by the treating physician and was based on clinician judgment. In total, 432 thyroid nodules in 340 patients were included in the study. There were 263 (77.4%) females and 77 (22.6%) males. The mean patient age was mean age 57 ± 14.3 years (range: 16 to 86 years). The median diameter of nodules was 20 mm (range: 9 to 83 mm).

US examinations and US-guided FNA procedures

Neck US was performed by the same experienced endocrinologist of our staff using a high-resolution US color Doppler apparatus (My Laboratory 40 HD; Esaote Biomedica, Firenze, Italy) with a 7.5-MHz linear transducer. US features of each thyroid nodule were described and recorded in the database by the endocrinologist who performed the examinations, and nodules were stratified using sonographic patterns, as described and published in the 2015 ATA guidelines (6) and in the ETA US risk stratification systems (10). Based on ATA US risk assessment, thyroid nodules were classified into five groups: "benign," "very low suspicion," "low suspicion," "intermediate suspicion," and "high suspicion" (Table 1, panel A). Based on ETA US risk assessment, thyroid nodules were classified into four groups: "EU-TIRADS 2" (benign), "EU-TIRADS 3" (low risk), "EU-TIRADS 4" (intermediate risk), and "EU-TIRADS 5" (high risk) (Table 1, panel B).

At the time of the FNAC, a sample of the nodules was collected for both cytological and molecular analysis. All patients provided informed consent for both traditional cytology and molecular analysis. US-guided FNAC was performed for at least two separate passes for each thyroid nodule by using a 23/25-gauge needle. The material was air dried, stained with May-Grunwald Giemsa, and interpreted by the same experienced cytologist. The pathologist performing the cytological diagnosis was blinded to the ultrasonography risk of thyroid nodules. Cytology reports from US-guided FNAC of thyroid nodules were based on the five

Table 1. Definition of Thyroid Nodule Categories by ATA (Panel A) and ETA (Panel B) US Risk Stratification

Panel A: ATA US System (6)	
Risk categories	
Benign	Purely cystic nodules (no solid component)
Very low suspicion	Spongiform or partially cystic nodules without any of the sonographic features described in low-, intermediate-, or high-suspicion patterns
Low suspicion	Isoechoic or hyperechoic solid nodule, or partially cystic nodule with eccentric solid areas, without microcalcification, irregular margin, or extrathyroidal extension, or taller-than-wide shape
Intermediate suspicion	Hypoechoic solid nodule with smooth margins without microcalcification, extrathyroidal extension, or taller-than-wide shape
High suspicion	Solid hypoechoic nodule or solid hypoechoic component of a partially cystic nodule with one or more of the following features: irregular margins, microcalcification, taller-than-wide shape, rim calcification with small extrusive soft tissue component, and evidence of extrathyroidal extension
Panel B: EU-TIRADS (10)	
Risk categories	
EU-TIRADS 2	Pure/anechoic cysts and entirely spongiform nodules
EU-TIRADS 3	Oval shape, smooth margins, or isoechoic or hyperechoic, without any feature of high risk
EU-TIRADS 4	Oval shape, smooth margins, or mildly hypoechoic, without any feature of high risk
EU-TIRADS 5	Nodules with at least one of the following high-risk features: nonoval shape, irregular margins, microcalcifications, and marked hypoechogenicity

categories according to the criteria of the British Thyroid Association (Thy1: nondiagnostic; Thy 2: benign; Thy 3: undetermined significance; Thy 4: suspicious for malignancy; and Thy 5: malignant) (12).

Management of thyroid nodules according to cytological results

All patients with Thy4/Thy5 cytologies were sent to surgery. In the benign nodule categories (Thy2), only selected patients, those with compressive symptoms, were sent to surgery. Patients with Thy2 cytologies not submitted to surgery were observed with annual follow-up consisting of clinical examination, neck US, and serological testing (free-triiodothyronine, free-thyroxine, and thyroid-stimulating hormone measurement). Patients with indeterminate cytology (Thy3) were managed according to the additional information gained from the search for genetic mutations using a six-gene panel including point mutations of BRAF (mainly V600E), hTERT promoter, and H-K-NRAS and rearrangements of RET/PTC, TRK, and PAX8/PPARγ. In case of gene mutations, patients were submitted to surgery. In case of negative mutational status, the management was modulated on the basis of size and US features of thyroid nodules. Specifically, patients with a large nodule and/or large multinodular goiter and patients with a high-suspicion nodule at neck US were submitted to surgery.

Statistical analysis

Epidemiological data are presented as the mean ± standard deviation. The frequency distribution in the different sonographic risk categories was compared using a χ^2 test. For statistical analysis, benign to low-suspicion nodules with the ATA US risk stratification system were considered at lower risk and intermediate- to high-suspicion nodules were considered at higher risk for thyroid nodule malignancy. Similarly, EU-TIRADS 2 and EU-TIRADS 3 nodules were considered at lower risk for thyroid malignancy, and EU-TIRADS 4 and EU-TIRADS 5 nodules were considered at higher risk for thyroid malignancy. Cytopathology results from US-guided FNA were considered the standard of reference. For the statistical analysis, the Thy4 and Thy5 categories were combined, and Thy1 categories were excluded from the beginning.

Diagnostic accuracy was calculated according to Galen (13) and was based on true positive (TP), true negative (TN), false positive, and false negative (FN) results. The positive predictive value (PPV) was TP/(TP + false positive), and the negative predictive value (NPV) was TN/(FN + TN). The 95% confidence interval of all estimates was also evaluated. The agreement between different risk stratification systems was calculated using Cohen κ coefficient. A value of 1 implies perfect agreement, and values <1 imply less than perfect agreement. It was evaluated using the Landis and Koch semiquantitative scale (poor agreement $\leq$0.20, fair agreement = 0.20 to 0.40, moderate agreement = 0.40 to 0.60, good agreement = 0.60 to 0.80, and very good agreement = 0.80 to 1.0).

Negative values of κ express disagreement. Because both ATA and ETA US risk stratification systems specifically recommend diagnostic FNAC according to risk class and nodule diameter, subanalysis incorporating these size criteria for diagnostic FNAC was performed in the calculation of estimated malignancy rates. Statistical analysis was performed using the software StatView for Windows version 5.0.1 (SAS Institute, Cary, NC) and the SPSS Statistics version 22.0 (IBM SPSS statistics, Hong Kong, China). A P value <0.05 was considered statistically significant.

Results

Cytological results

According to the British Thyroid Association (12), FNAC was inadequate (Thy1) in 12% (51/432), benign (Thy2) in 77.7% (336/432), indeterminate (Thy3) in 7.1% (31/432), and suspicious or malignant (Thy4/Thy5) in 3.2% (14/432) of the nodules.

Distribution of thyroid nodules according to ATA and ETA US risk stratification and cytological results

Thy1 nodules were excluded from the statistical analysis. The final study group included 381 nodules. The distribution of nodules classified by ATA US risk pattern were as follows: benign (1.6%; n = 6), very low (53.0%; n = 202), low (20.7%; n = 79), intermediate (22.3%; n = 85), and high (2.4%; n = 9) suspicion (Table 2, panel A). The distribution of cytological classification varied significantly by ATA US risk assessment ($P < 0.0001$; Table 2, panel A). The proportion of Thy2 nodules significantly decreased with increasing of US risk class ($P < 0.0001$): 100% (n = 6/6) in benign, 99.5% (n = 193/202) in very low–suspicion, 92.4% (n = 73/79) in low-suspicion, 72.9% (n = 62/85) in intermediate-suspicion, and 22.2% (n = 2/9) in high-suspicion nodules (Table 2, panel A). On the contrary, with the increasing of US risk class, the proportion of Thy4/Thy5 nodules increased significantly ($P < 0.0001$): 0% (0/6) in benign, 0.5% (1/202) in very low–suspicion, 2.5% (2/79) in low-suspicion, 8.2% (7/85) in intermediate-suspicion, and 44.4% (4/9) in high-suspicion nodules. Similar trend was observed for the Thy3 nodules ($P < 0.0001$) (Table 2, panel A).

The distribution of nodules classified by ETA US risk pattern were: EU-TIRADS 2 (32.1%; n = 122), EU-TIRADS 3 (42.8%; n = 163), EU-TIRADS 4 (17.5%; n = 67), and EU-TIRADS 5 (7.6%; n = 29) (Table 2, panel B). The distribution of cytological classification varied significantly by ETA US risk assessment ($P < 0.0001$). The proportion of Thy2 nodules significantly decreased with increasing of US risk class ($P < 0.0001$): 95.9% (n = 117/122) in EU-TIRADS 2, 93.9% (n = 153/163) in EU-TIRADS 3, 74.6% (n = 50/67) in EU-TIRADS 4, and 55.2% (n = 16/29) in EU-TIRADS 5. (Table 2, panel B). On the contrary, with increasing of US risk class, the proportion of Thy4/Thy5 nodules increased significantly ($P < 0.0001$): 0.8% (n = 1/122) in EU-TIRADS 2, 1.2% (n = 2/163) in EU-TIRADS 3, 6.0% (n = 4/67) in EU-TIRADS 4, and 24.1% (n = 7/29) in EU-TIRADS 5. Similar trend was observed for the Thy3 nodules ($P < 0.0001$) (Table 2, panel B).

Table 2. Distribution of Thyroid Nodules According to ATA (Panel A) and ETA (Panel B) US Risk Stratification and Cytological Results

Panel A				
ATA US System (6)		British Cytology Classification (12)		
Risk Categories	n (%)	Thy2: n (%)	Thy3: n (%)	Thy4/5: n (%)
Benign	6 (1.6)	6 (100)	0 (0)	0 (0)
Very low suspicion	202 (53.0)	193 (95.5)	8 (4.0)	1 (0.5)
Low suspicion	79 (20.7)	73 (92.4)	4 (5.1)	2 (2.5)
Intermediate suspicion	85 (22.3)	62 (72.9)	16 (18.8)	7 (8.2)
High suspicion	9 (2.4)	2 (22.2)	3 (33.3)	4 (44.4)
Total	381 (100)	336 (88.2)	31 (8.1)	14 (3.7)
Panel B				
EU-TIRADS (10)		British Cytology Classification (12)		
Risk Categories	n (%)	Thy2: n (%)	Thy3: n (%)	Thy4/5: n (%)
EU-TIRADS 2	122 (32.1)	117 (95.9)	4 (3.3)	1 (0.8)
EU-TIRADS 3	163 (42.8)	153 (93.9)	8 (4.9)	2 (1.2)
EU-TIRADS 4	67 (17.5)	50 (74.6)	13 (19.4)	4 (6.0)
EU-TIRADS 5	29 (7.6)	16 (55.2)	6 (20.7)	7 (24.1)
Total	381 (100)	336 (88.2)	31 (8.1)	14 (3.7)

Diagnostic performance of the ATA and ETA US risk stratification systems

To evaluate the accuracy of the ATA and ETA US stratification systems, we excluded from the statistical analyses the Thy3 cytology and divided nodules in two categories according to US risk class. When considering benign to low-suspicion nodules at lower risk and intermediate- to high-suspicion nodules at higher risk for thyroid nodule malignancy, the diagnostic performance of the ATA US risk stratification system was as follows: sensitivity, 78.5% (11/14 nodules); specificity, 80.9% (272/336 nodules); PPV, 14.6% (11/75 nodules); and NPV, 98.9% (272/275 nodules) (Table 3). To evaluate the diagnostic accuracy of the EU-TIRADS stratification, we divided nodules in two categories according to US risk class: at lower risk (EU-TIRADS 2 and EU-TIRADS 3) and at higher risk of malignancy (EU-TIRADS 4 and EU-TIRADS 5). The diagnostic performance of the ETA US risk stratification system was as follows: sensitivity, 78.6% (11/14 nodules); specificity, 80.3% (270/336 nodules); PPV, 14.3% (11/77 nodules); and NPV, 98.9% (270/273 nodules) (Table 3). Using the Cohen κ coefficient, the agreement between ATA and ETA US risk stratification was "very good" (κ = 0.913), demonstrating that the ability to identify benign and suspicious nodules is similar between ATA and ETA systems.

Estimation of the decrease in unnecessary FNAC using ATA and ETA US risk stratification systems

After exclusion of Thy3 nodules, according to the ATA US risk stratification system, 83/350 (23.7%) nodules did not meet the criteria for the FNAC (Table 4, panel A). They were stratified as follows: 6/6 (100%) in the benign, 69/194 (35.6%) in the very low–suspicion, 7/75 (9.3%) in the low-suspicion, and 1/69 (1.4%) in the intermediate-suspicion category. The rate of malignancy (Thy4/5) was 0% (0/6) in the benign category, 1.4% (1/69) in the very low–suspicion category, 0% (0/1) in the low-suspicion category, and 100% (1/1) in the intermediate-suspicion category. Therefore, the cumulative malignancy rate for nodules that did not meet the criteria for the FNAC was 2.4% (2/83 nodules). Considering only categories at lower risk of malignancy (benign to low-suspicion nodules), cumulative malignancy rate for nodules that did not meet

Table 3. Diagnostic Performance of the ATA and ETA US Risk Stratification Systems

	Sensitivity	Specificity	PPV	NPV	Accuracy
ATA US system (6)	78.6% (49.2%–95.3%)	80.9% (76.3%–85.0%)	14.6% (10.7%–19.6%)	98.9% (97.0%–99.6%)	80.9% (76.3%–84.8%)
EU-TIRADS (10)	78.6% (49.2%–95.3%)	80.3% (75.7%–84.5%)	14.3% (10.5%–19.1%)	98.9% (97.0%–99.6%)	75.7% (75.7%–84.3%)

Table 4. Estimation of the Decrease in Unnecessary FNAC Using ATA (Panel A) and ETA (Panel B) US Risk Stratification Systems and Malignant Risk in Nodules That Did Not Meet the Criteria for FNAC

Panel A

ATA US Risk Stratification (6) (N = 350 Nodules)	Benign (n = 6)	Very Low Suspicion (n = 194)	Low Suspicion (n = 75)	Intermediate Suspicion (n = 69)	High Suspicion (n = 6)
Nodules that did not meet the criteria for the FNAC (n = 83; 23.7%)	6/6 (100%)	69/194 (35.6%)	7/75 (9.3%)	1/69 (1.4%)	0/6 (0%)
Malignant nodules (n)	0	1	0	1	0
Malignancy rate (n; %)			2/83 (2.4%)		

Panel B

EU-TIRADS (10) (N = 350 Nodules)	EU-TIRADS 2 (n = 118)	EU-TIRADS 3 (n = 155)	EU-TIRADS 4 (n = 54)	EU-TIRADS 5 (n = 23)
Nodules that did not meet the criteria for the FNAC (n = 196; 56.0%)	118/118 (100%)	58/155 (37.4%)	19/54 (35.2%)	1/23 (4.3%)
Malignant nodules (n)	1	2	3	1
Malignancy rate (n; %)		7/196 (3.6%)		

Only nodules with diagnostic cytologies (THY2 and THY4/5) were included in this analysis (n = 350).

the criteria for the FNAC was only 1.2% (1/82 nodules) (Table 4, panel A).

We performed the same analysis considering all very low–suspicion nodules as nodules that did not meet the criteria for FNAC because observation without FNAC is also a reasonable option for these categories of nodules. Using this approach, the rate of unnecessary FNAC increased from 27.3% to 59.4%. Consequently, because only two nodules were malignant at cytology, the rate of malignancy in those nodules decreased from 2.4% to 0.96%.

According to the ETA US risk stratification system, 196/350 (56.0%) did not meet the criteria for FNAC (Table 4, panel B). They were stratified as follows: 118/118 (100%) in the EU-TIRADS 2, 58/155 (37.4%) in the EU-TIRADS 3, 19/54 (35.2%) in the EU-TIRADS 4, and 1/23 (4.3%) in the EU-TIRADS 5 category. The rate of malignancy was: 0.8% (1/118) for EU-TIRADS 2, 3.4% (2/58) for EU-TIRADS 3, 15.8% (3/19) for EU-TIRADS 4, and 100% (1/1) for EU-TIRADS 5. Therefore, cumulative malignancy for nodules that did not meet the criteria for the FNAC was 3.6% (7/196 nodules). Considering only categories at lower risk of malignancy (EU-TIRADS 2 and 3), cumulative malignancy rate for nodules that did not meet the criteria for the FNAC was only 1.7% (3/176 nodules) (Table 4, panel B). The proportion of malignancy in nodules that did not meet the criteria for the FNAC was not different between ATA and ETA US risk stratification systems ($P > 0.99$).

Discussion

With the advances in US technology and the routine use of neck US in clinical practice, there has been a significant increase in the detection of thyroid nodules resulting in a proportional rise in the number of thyroid FNAC procedures. However, only 2% to 16% of thyroid nodules are malignant, whereas a "benign" result is obtained in approximately 60% to 70% of thyroid FNACs (3). To decrease these unnecessary FNACs in the majority of cytologically benign thyroid nodules, the use of US features has been proposed. Therefore, various combinations of features have been studied for that purpose. Thyroid nodules are classified into categories related to their US patterns, and the indications for FNAC is based on these categories (5–10). Recently, the ATA (6) and the ETA (10) have proposed two different US risk stratification systems to help the clinician in selecting thyroid nodules for FNAC. The goal of these US assessments of thyroid nodules proposed is to distinguish benign nodules that can be managed conservatively from those with suspicious or malignant features requiring FNAC. However, both ATA and ETA US risk stratification systems of thyroid nodules (6, 10) remain to be applied and tested in large prospective studies. In our study, we validated and compared the ATA and ETA US risk stratification systems of thyroid nodules in a prospective series of 432 thyroid nodules submitted to FNAC. Categories for US patterns correlated with cytological diagnosis. The majority of the nodules at "lower risk of malignancy" were associated with benign cytology (NPV for benign nodules was 98.9% for both ATA and ETA US systems), and the benign rates observed in these categories were within or close to the given range described in the ATA and ETA guidelines.

Also, in the categories at "higher risk of malignancy," the rates of malignant nodules were within or close to the

given range described in the two guidelines (6, 10), although it was relatively lower (67% for ATA and 30.4% for ETA) compared to the one reported in the ATA (70% to 90%) and in the ETA (26% to 87%) systems (6, 10). Our results imply that an improvement in selecting nodules for FNAC may be achieved using either ATA and ETA US risk stratification systems in the clinical practice. Specifically, both ATA and ETA US risk categories are able to predict the benignity of the nodule with a very low rate of FN results (1.1%). However, the PPV is very low, suggesting that US high-risk categories (especially intermediate risk for ATA and EU-TIRADS 4 for ETA) are not able to predict the malignancy of the nodule (PPV around 14% for both ATA and ETA US risk stratification systems). This may have been due to the low rate of the suspicious/malignant nodules observed in our series of nodules that may affect the PPV and, also, to the inclusion in the categories at "higher risk of malignancy" (intermediate to high risk for ATA and EU-TIRADS 4 and 5 for ETA) of US features with different clinical importance. Indeed, suspicious US features have a different relative risk of being associated with malignancy, ranging from 30% (for a hypoechoic nodule) to 70% (for a markedly hypoechoic nodule or nodule with microcalcifications) (14). Our results are in agreement with recent prospective (11) and retrospective studies (14). Tang *et al.* (11) prospectively compared ATA US risk stratification systems to cytological diagnosis. The authors found a high correspondence between US categories and cytological diagnosis using the Bethesda System for Reporting Thyroid Cytopathology. Differently to our study, the authors correlated US risk categories with histological results in indeterminate nodules. They also confirmed in this subgroup of nodules a significant correlation between ATA US risk categories and histological rate of thyroid cancer.

Our study group also included nodules that did not meet the criteria for FNAC. This permitted us to evaluate how many unnecessary FNACs can be avoided using the ATA and ETA US risk stratifications in clinical practice and how many malignancy nodules were lost in this subgroup of nodules. The proportion of patients in whom FNAC can be avoided was significantly higher using the ETA US risk stratification (56.0% vs 23.7% for the ATA US risk stratification system) without significantly increasing the risk of losing thyroid malignancy in thyroid nodules at lower risk of malignancy (<2.0% for both US risk stratification systems). However, if all nodules at ATA very low risk (and not only nodules ≤2 cm) were considered as nodules that did not meet the criteria for the FNAC, the percentage of patients in whom FNAC is unnecessary increased from 23.7% to 59.4%, without significant change in malignant rate.

Our study has a few limitations. First, we used cytological results as the reference standard, and FN results in some thyroid nodules with benign results could have occurred. In addition, the absence of a histological series does not allow us to ascertain the true nature of Thy3 nodules. A recent study evaluated the malignancy rate in a small subgroup of Thy3 nodules submitted to surgery (11). The authors reported a higher rate of malignant nodules in the lower-risk categories (low and very low) when compared with the whole group (17% vs 8% and 12% vs 2%, respectively). The NPV of the ATA US risk stratification system in this study was around 85%, lower than that observed in our study group (around 98%). These results suggest that additional studies, including a large series of Thy3 nodules submitted to surgery, should be performed to define the diagnostic accuracy of US risk stratification systems in Thy3 nodules. Second, although our center is an academic referral center and not a primary care setting, a low rate of suspicious/malignant nodules at cytology was observed in our study group. Our results are similar to that reported in a recent study in which the authors evaluated the diagnostic performance of FNAC of a primary care setting in Germany (15). The low rate of suspicious/malignant nodules might affect the PPV and NPV of ATA and ETA US risk stratification systems. The diagnostic accuracy is directly related to the prevalence of the disease in the population. Assuming all other factors remain constant, the PPV will increase with increasing prevalence and NPV decreases with increase in prevalence. Nevertheless, it is possible to hypothesize that the diagnostic accuracy of ATA and ETA US risk stratification systems might be different in higher-risk populations, with a higher rate of FN results and a lower NPV. In a large retrospective study including 1293 thyroid nodules with histological diagnosis and a malignant rate of 18.1%, both the TIRADS system and ATA US risk stratification system showed a very high NPV (98.1% and 97.3%, respectively) (14). These data suggest that, probably also in higher-risk thyroid nodule populations, both ATA and ETA US risk stratifications might have a good diagnostic accuracy in selecting nodules for FNAC. However, prospective studies, including a large group of nodules at high risk of malignancy, might be necessary to clarify this issue.

In conclusion, both ETA and the ATA US risk stratification systems provide effective malignancy risk stratification for thyroid nodules without a significant difference between the two scoring systems. In clinical practice, using this approach, we should be able to reduce the number of unnecessary FNACs without losing clinically significant thyroid cancer.

Acknowledgments

Financial Support: This work was supported in part by Ministero Italiano dell'Università e Ricerca Grant 2012Z3F7HE (to F.P.).

Correspondence and Reprint Requests: Maria Grazia Castagna, MD, PhD, Section of Endocrinology and Metabolism, Department of Medical, Surgical, and Neurological Sciences, University of Siena, 53100 Siena, Italy. E-mail: m.g.castagna@ao-siena.toscana.it.

Disclosure Summary: The authors have nothing to disclose.

References

1. Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest*. 2009;**39**(8):699–706.
2. Russ G, Leboulleux S, Leenhardt L, Hegedüs L. Thyroid incidentalomas: epidemiology, risk stratification with ultrasound and workup. *Eur Thyroid J*. 2014;**3**(3):154–163.
3. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: a meta-analysis. *Acta Cytol*. 2012;**56**(4):333–339.
4. Ozdemir D, Bestepe N, Faki S, Kilicarslan A, Parlak O, Ersoy R, Cakir B. Comparison of thyroid fine needle aspiration biopsy results before and after implementation of Bethesda classification. *Cytopathology*. 2017;**28**(5):400–406.
5. Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G, Gilbert J, Harrison B, Johnson SJ, Giles TE, Moss L, Lewington V, Newbold K, Taylor J, Thakker RV, Watkinson J, Williams GR; British Thyroid Association. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)*. 2014;**81**(Suppl 1):1–122.
6. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 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):1–133.
7. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, Paschke R, Valcavi R, Vitti P; AACE/ACE/AME Task Force on Thyroid Nodules. AACE/ACE/AME task force on thyroid nodules 2016 American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules—2016 update. *Endocr Pract*. 2016;**22**(5):622–639.
8. Shin JH, Baek JH, Chung J, Ha EJ, Kim JH, Lee YH, Lim HK, Moon WJ, Na DG, Park JS, Choi YJ, Hahn SY, Jeon SJ, Jung SL, Kim DW, Kim EK, Kwak JY, Lee CY, Lee HJ, Lee JH, Lee JH, Lee KH, Park SW, Sung JY; Korean Society of Thyroid Radiology (KSThR) and Korean Society of Radiology. Ultrasonography diagnosis and imaging-based management of thyroid nodules: revised Korean Society of Thyroid Radiology consensus statement and recommendations. *Korean J Radiol*. 2016;**17**(3):370–395.
9. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, Cronan JJ, Beland MD, Desser TS, Frates MC, Hammers LW, Hamper UM, Langer JE, Reading CC, Scoutt LM, Stavros AT. ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol*. 2017;**14**(5):587–595.
10. 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**(5):225–237.
11. Tang AL, Falciglia M, Yang H, Mark JR, Steward DL. Validation of American Thyroid Association ultrasound risk assessment of thyroid nodules selected for ultrasound fine-needle aspiration. *Thyroid*. 2017;**27**(8):1077–1082.
12. British Thyroid Association and Royal College of Physicians. Guidelines for the Management of Thyroid Cancer. 2nd ed. Available at: http://www.btf-thyroid.org. Accessed December 2017.
13. Galen RS. Application of the predictive value model in the analysis of test effectiveness. *Clin Lab Med*. 1982;**2**(4):685–699.
14. 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**(3):917–924.
15. Eszlinger M, Ullmann M, Ruschenburg I, Böhme K, Görke F, Franzius C, Adam S, Molwitz T, Landvogt C, Amro B, Hach A, Feldmann B, Graf D, Wefer A, Niemann R, Bullmann C, Klaushenke G, Santen R, Tönshoff G, Ivancevic V, Kögler A, Bell E, Lorenz B, Kluge G, Hartenstein C, Paschke R. Low malignancy rates in fine-needle aspiration cytologies in a primary care setting in Germany. *Thyroid*. 2017;**27**(11):1385–1392.