2020

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

Marilyn Arosemena
Jackson Memorial Hospital, Miami

Anu Thekkumkattil
Jackson Memorial Hospital, Miami

Maria Linares Valderrama
Jackson Memorial Hospital, Miami

Russ Kuker
University of Miami

Rosa Patricia Castillo
University of Miami

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation

Arosemena, Marilyn; Thekkumkattil, Anu; Valderrama, Maria Linares; Kuker, Russ; Castillo, Rosa Patricia; Sidani, Charif; Gonzalez, Manuel Lora; Casula, Sabina; and Kargi, Atil Yilmaz, "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*. 30, 11. 1613 - 1619. (2020). https://digitalcommons.wustl.edu/open_access_pubs/9893

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.

Authors

Marilyn Arosemena, Anu Thekkumkattil, Maria Linares Valderrama, Russ Kuker, Rosa Patricia Castillo, Charif Sidani, Manuel Lora Gonzalez, Sabina Casula, and Atil Yilmaz Kargi

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

Marilyn Arosemena,¹ Anu Thekkumkattil,² Maria Linares Valderrama,² Russ Kuker,³ Rosa Patricia Castillo,³ Charif Sidani,³ Manuel Lora Gonzalez,⁴ Sabina Casula,^{5,6} and Atil Yilmaz Kargi⁵

Background: The Afirma gene expression classifier (GEC) has been used to aid in the diagnosis and management of thyroid nodules having Bethesda category III fine-needle aspiration cytologic diagnosis (B3 nodules). The American Thyroid Association sonographic risk stratification system for thyroid nodules (ATA-US) may stratify B3 nodules and aid in the decision to order a molecular test. The aim of this study was to assess the association between ATA-US and GEC as well as to determine their individual and combined diagnostic performances when applied to B3 nodules.

Methods: A retrospective single-center study included B3 nodules that had undergone evaluation by GEC. Each ultrasound was reviewed by three radiologists, and nodules were classified using the 2015 ATA sonographic risk categories. Nodules were determined to be benign or malignant based on surgical pathology or minimum 11 months of follow-up. Positive predictive values (PPV) and negative predictive values (NPV) were calculated for GEC, ATA-US, and GEC across all ATA-US categories.

Results: One hundred twenty-six B3 nodules with GEC results were included and deemed benign or malignant based on final pathology or follow-up. Prevalence of malignancy was 32%. The rate of malignancy was similar in the ATA-US high suspicion (HS) and intermediate suspicion (IS) categories at 42% and 38%, respectively; and lower in nodules with low suspicion sonography (LS) and very low suspicion sonography (VLS) at 23% and 11%, respectively. The PPV and NPV of ATA-US was calculated by designating HS or IS sonography as a “positive” test and the lower risk categories as “negative.” ATA-US had a PPV of 40% and NPV of 79%. The GEC PPV was 40% and NPV was 83%. The PPV of GEC was 50% in nodules with HS or IS ATA-US and lower at 28% and 20%, respectively, in LS and VLS nodules. The NPV of GEC was 80% in HS, 77% in IS, 84% in LS, and 100% in VLS sonography categories.

Conclusions: In B3 nodules, ATA-US and GEC have similar diagnostic performance. The PPV of GEC varies across ATA-US categories, while the NPV remains similar. These data support the need for future prospective studies.

Keywords: GEC, thyroid nodule, Bethesda, atypia of undetermined significance, ultrasound, ATA

Introduction

THYROID NODULES ARE very common in the general population and are frequently encountered in clinical practice. While the majority of thyroid nodules are benign and not clinically significant, a minority may harbor malignancy

(1,2). Thyroid ultrasound and fine-needle aspiration cytology (FNA-C) are often performed during the diagnostic evaluation.

Despite advances in diagnostic methods, ~20–25% of thyroid nodules will be classified as indeterminate (ITNs) by FNA-C. According to the Bethesda System for Reporting Thyroid Cytopathology, thyroid nodules with indeterminate

¹Department of Medicine, Jackson Memorial Hospital, Miami, Florida, USA.

²Division of Endocrinology, Jackson Memorial Hospital, University of Miami, Miami, Florida, USA.

³Department of Radiology, University of Miami, Miami, Florida, USA.

⁴Department of Pathology, Washington University, St. Louis, Missouri, USA.

⁵Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, University of Miami Miller School of Medicine, Miami, Florida, USA.

⁶Department of Endocrinology, Miami Veterans Affairs Healthcare System, Miami, Florida, USA.

cytology are grouped as Bethesda category III (Atypia of undetermined significance or Follicular lesion of undetermined significance [AUS/FLUS]), Bethesda category IV (Follicular neoplasm [FN]), or Bethesda category V (suspicious for malignancy) (3). The rate of malignancy (ROM) among nodules with Bethesda category III cytologic diagnosis (B3) varies between 6% and 48% often necessitating surgical diagnosis (2).

In recent years, several molecular tests, including Afirma gene expression classifier (GEC) (Veracyte), have been utilized to risk stratify ITNs and to aid in the decision to perform diagnostic thyroid surgery versus surveillance (4–7). It has been proposed that the use of GEC has resulted in an ~50% decrease in diagnostic thyroid surgeries in cases of ITNs (8–10). GEC classifies ITNs as either “benign” (95% negative predictive value [NPV] for aspirates classified as AUS/FLUS and 94% for aspirates classified as FN) or “suspicious” for malignancy (positive predictive value [PPV] ~40% for malignancy) (4). While this test performance has been reported in an initial validation study, a meta-analysis of more recent studies has brought into question whether the cohort of the initial validation study was representative of the populations in which the GEC has been used due to a wide difference in the NPV results (11). While few studies have assessed the impact of individual sonographic features on GEC test performance (12,13), it is not known whether the reported NPV and PPV would be different when GEC is applied in nodules of differing American Thyroid Association (ATA) sonographic risk category.

Several studies have reported that individual ultrasound features of ITNs as well as ultrasound-based classification systems may be able to assist clinicians in predicting malignancy among ITNs and to aid in the decision to order a molecular test before triaging the patient to either clinical follow-up or surgical intervention (14–20).

The 2015 ATA management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer recommend classification of thyroid nodules into one of five sonographic patterns (ATA-US). According to these guidelines, the ATA sonographic classification scheme's primary use should be for determining the size cutoff threshold to perform FNA. The ATA sonographic risk classification system is also promoted by the guidelines authors as potentially useful in determining the optimal monitoring strategy for follow-up of thyroid nodules with benign FNA-C and for those nodules that do not meet criteria for FNA. Regarding management strategies for B3 nodules, the 2015 ATA guidelines recommend “consideration for worrisome clinical and sonographic features”; however, to date, few studies have addressed whether there could be a role of ATA-US in risk stratifying and decision-making in management of B3 nodules (14–20).

Given the considerable cost and limited availability of molecular testing, some authors have recommended to further risk stratifying ITNs based on sonographic features, triaging higher risk ultrasound patterns to surgery, lower risk ultrasound patterns to observation, and offering molecular tests, where available, on the remainder (1,14,21). To the best of our knowledge, there are currently no published studies of ITNs assessing the association of sonographic findings to GEC, comparing the performance of ultrasound-based risk prediction with that of GEC, or reporting the performance of GEC within the various ATA sonographic risk categories.

The objective of this study was to assess the correlation between ATA-US categories and GEC results as well as to assess individual diagnostic performances of ATA-US and GEC and that of ATA-US and GEC combined when applied to B3 nodules.

Materials and Methods

Following approval by the Institutional Review Board of the University of Miami Miller School of Medicine, retrospective chart review of patients who had FNA-C reported as Bethesda category III (AUS/FLUS) from January 1, 2012, to December 31, 2017, was performed. Patients were included if aged 18 years or older, presence of one or more thyroid nodules confirmed by ultrasound results, and GEC result of suspicious or benign. One hundred seventeen patients with 126 nodules met the inclusion criteria and were included in the analysis of GEC and ultrasound results. One hundred twenty-six nodules either had surgical histopathologic diagnosis ($n=85$) or had long-term follow-up after FNA-C ($n=41$) and were included in the analysis of test performance of GEC and ATA-US (Fig. 1). Informed consent was waived as the study was retrospective in nature.

Ultrasound evaluation

Ultrasound images were independently reviewed by three radiologists who were blinded to clinical data, pathology results, and to each other's assessments. Radiologists completed a checklist for each nodule including the nodule location, size, composition, shape, echogenicity, margins, calcifications, and presence of suspicious lymph nodes. Echogenicity was classified into three categories: hypoechoic, isoechoic, or hyperechoic with respect to the normal thyroid parenchyma. Composition was classified into five categories: solid, mixed cystic-solid with >50% solid, mixed cystic-solid with >50% cystic, cystic, or spongiform. Nodule margins were classified into four categories: regular, irregular spiculated, irregular microlobulated, and irregular infiltrative. Calcifications were classified into three groups: none, microcalcifications, and macrocalcifications. Suspicious lymph nodes were classified into two groups: present or absent. The shape of the nodule was classified as taller than wide or wider than tall in the transverse view.

Nodules were classified into the 2015 ATA sonographic patterns as follows. Purely cystic nodules were classified as benign (B). Spongiform or partly cystic nodules without eccentric solid areas or other suspicious features were classified as very low suspicion sonography (VLS). Isoechoic or hyperechoic solid or partly cystic nodules with eccentric solid areas were classified as low suspicion sonography (LS). Hypoechoic solid nodules without other suspicious features were classified as intermediate suspicion (IS) pattern. Solid hypoechoic nodules manifesting at least one suspicious feature (microcalcifications, irregular margins, shape taller than wide, presence of suspicious lymph nodes) were classified as high suspicion (HS) pattern. Radiologist also had the option to select “Non-ATA” as one of the group categories. Non-ATA was recommended to be chosen in the scenarios not described in the 2015 ATA classification; these scenarios included heteroechoic nodules with or without other suspicious features; and isoechoic or hyperechoic nodules with at least one suspicious feature.

FIG. 1. Bethesda III cytology nodules included in the study. Flow diagram showing nodules included in the study cohort. Afirma GEC, Afirma gene expression classifier; FNA-C, fine-needle aspiration cytology; *n*, number of nodules.

Pathology evaluation

All cytology and histopathology results were obtained from chart review and reported at our institution by board-certified pathologists experienced in thyroid cytology and histopathology. Cytology was reported using the Bethesda System for Reporting Thyroid Cytopathology (3). Bethesda III results were reported as either AUS or FLUS. Malignant histopathology results were classified as papillary thyroid cancer, follicular thyroid cancer, or “noninvasive follicular thyroid neoplasms with papillary-like nuclear features”. Afirma GEC results were categorized as “suspicious” or “benign.”

Statistical analysis

Statistical analysis was performed using software Astatsa, Statistical Analysis System version 9.3, and Excel for Mac 2016 with an alpha set to 5%. Demographic characteristics such as age, sex, race/ethnicity, and weight were recorded as

well as clinical characteristics such as family history of thyroid cancer, thyrotropin (TSH) level, history of hypothyroidism, cytology (classified as AUS, FLUS, or Hürthle cell neoplasm), GEC result, and surgical histopathology were recorded. In cases in which diagnostic surgery was not performed, data from a minimum 11 months of follow-up with ultrasound were obtained. In accordance with the procedures recommended by a systematic review of studies of Afirma GEC diagnostic performance, we decided to use a cutoff of a minimum 11 months of follow-up of those patients who did not have a surgical histologic diagnosis (22). For these nodules with available long-term follow-up data, those nodules not demonstrating enlargement greater than 3 mm in two dimensions on ultrasound, and those not being subsequently diagnosed as Bethesda category VI on repeat FNA-C were considered as true negative.

Continuous variables are reported with mean $\pm$ standard deviation, and categorical data are described with numbers

and percentages. Simple *t*-test was used for continuous data and chi-square test for categorical data. Odds ratios (OR) and associated 95% confidence intervals [CI] were calculated from the contingency tables. Association between ATA sonographic risk category and Bethesda cytology and correlation between ATA sonographic risk category and GEC results were evaluated. Diagnostic performance including NPV and PPV were calculated for GEC overall and when applied within each ATA-US category. Overall PPV and NPV for ATA classification were calculated using HS and IS nodules as a positive test and LS and VLS nodules as a negative test. We used Fleiss *k* for assessing the reliability of the inter-reader agreement.

Results

Study cohort

One hundred seventeen patients with 126 nodules with Bethesda III cytology were included. Mean age was 58 years at the time of FNA. Eighty percent of patients were female and 46.1% of patients were Hispanic. 3.4% of patients had a history of thyroid cancer, and 11.1% had a history of hypothyroidism. The mean TSH was 1.78 mU/L. Table 1 lists the patient demographics of our study population. Patient age, sex, race/ethnicity, nodule size and location, history of thyroid cancer, and TSH level were not found to be associated with risk of thyroid cancer. Overall malignancy rate was 31.7%. Only ATA-US and GEC were predictive of malignancy.

TABLE 1. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF STUDY COHORT

<i>Patient (n = 117) and thyroid nodule (n = 126) characteristics</i>	
Sex	<i>n</i> (%)
Female	94 (80)
Male	23 (20)
Age	
Mean (years, SD)	58 (43.4–72.5)
Race/ethnicity	
Asian	1 (0.9)
Black	9 (7.6)
Hispanic	54 (46.1)
White	36 (30.7)
Other	17 (14.5)
Family history of thyroid cancer	
Yes	4 (3.4)
No	112 (95.7)
Unknown	1 (0.9)
Hypothyroidism	
Yes	13 (11.1)
No	104 (88.8)
Nodules: greatest diameter on ultrasound, cm	
Mean (range)	2.03 (0.9–3.1)
TSH	
≤0.5	5 (4.2)
0.51–2.5	76 (64.9)
2.51–4.5	25 (21.4)
>4.5	1 (0.9)
Not available TSH	10 (8.5)

SD, standard deviation; TSH, thyrotropin.

Ultrasound

All 126 nodules were able to be classified into an ATA-US category by all readers, and majority consensus was achieved for all nodules as follows: 9 (7%) VL, 48 (38%) LS, 37 (29%) IS, 32 (25%) HS, and 0 (0%) Non-ATA. The mean follow-up was 20 ± 8.85 months (SD 11.2–28.9).

Inter-reader agreement for ATA-US was almost perfect, with Fleiss *k* of 0.82 [CI 0.77–0.87]. Inter-reader correlation for individual ultrasound features demonstrated substantial to near-perfect agreement: echogenicity *k* 0.71 [CI 0.65–0.77], composition *k* 0.87 [CI 0.81–0.93], margins 0.72 [CI 0.63–0.81], and calcifications *k* 0.85 [CI 0.77–0.94].

Of 126 nodules, 85 had final surgical histopathologic confirmation of diagnosis and 41 had long-term follow-up data (Fig. 1) and were included in the analysis of diagnostic performance of ATA-US and GEC. ROM of HS and IS ATA-US category was 38–44% compared with LS and VLS that had a lower ROM between 11% and 23%. The OR for malignancy for HS/IS versus LS/VLS was 6.66 [CI 1.28–11.44] (Supplementary Tables S1 and S2). Overall PPV and NPV of ATA-US was 40% [CI 0.13–0.32] and 79% [CI 0.64–0.86], respectively.

Among individual ultrasound characteristics, absence of calcifications (OR 0.24 [CI 0.07–0.79]) and irregular margins (OR 2.73 [CI 1.19–6.27]) were associated negatively and positively, respectively, to malignancy risk. All nodules with irregular margins were more likely to be malignant than those with regular margins (OR 2.73 [CI 1.19–6.27]). Predominantly solid nodules (>50% solid) were more often malignant (ROM 35%) compared with cystic and spongiform (ROM 7%), although not statistically significant (Supplementary Table S3). Of the isoechoic nodules, those with irregular margins had a ROM of 47%, while those with regular margins had a ROM of 10% (OR 8.44 [CI 1.48–48.14]) (Supplementary Table S4). Of the cases lost to follow-up (*n* = 53), 18 (34%) were HS, 13 (25%) were IS, 21 (38%) were LS, and 1 (2%) was VLS.

Afirma GEC

Among all 126 nodules, 37% had benign GEC and 63% had suspicious GEC. The rate of suspicious GEC was highest in the ATA-US high suspicion nodules (69%) and lower among the other sonographic categories (IS 65%, LS 60%, and VLS 56%) (Table 2).

TABLE 2. RATES OF SUSPICIOUS AFIRMA GENE EXPRESSION CLASSIFIER BY ATA-US CATEGORY

	<i>GEC</i>	
	<i>Suspicious</i>	<i>Benign</i>
ATA-US		
Highly suspicious	22 (69%)	10 (31%)
Intermediate suspicious	24 (65%)	13 (35%)
Low suspicious	29 (60%)	19 (40%)
Very low suspicious	5 (56%)	4 (44%)

ATA-US, American Thyroid Association sonographic risk stratification system for thyroid nodules; Afirma GEC, Afirma gene expression classifier.

Overall PPV and NPV for malignancy of GEC was 40% [CI 0.34–0.55] and 83% [CI 0.70–0.91], respectively, a test performance similar to, yet slightly better than that of ATA-US. The PPV of GEC was higher in the higher risk ATA-US categories and lower in the lower risk sonographic nodules (HS 50%, IS 50%, LS 23%, and VLS 20%). The NPV of GEC was similar across all ATA-US categories and lowest in the very low suspicion sonographic pattern nodules (HS 80%, IS 77%, LS 84%, and VLS 100%). The likelihood of malignancy in the group of nodules with both higher sonographic risk patterns (HS or IS) and suspicious Afirma was significantly greater than in the group of nodules having both low suspicion ultrasound pattern (LS or VLS) with negative Afirma (OR 6.66 [CI 1.7–25]).

Cytology

There was no difference in the ROM, rate of GEC suspicious, and distribution of ATA-US risk categories between nodules with AUS ($n=70$) and FLUS ($n=46$) cytology. Ten nodules had cytology consistent with Hürthle cell lesions, of which eight were GEC suspicious, while seven were malignant on final pathology. As a group, nodules with Hürthle cell cytology were more likely to have higher suspicion sonographic patterns compared with nodules in the other cytology categories (HS 3, IS 5, LS 1, VLS 1). Supplementary Tables S5 and S6 show the pathology results in benign and malignant nodules per each ATA-US.

Discussion

The results of this study highlight many relevant findings regarding the utility of ATA-US and GEC in the diagnostic evaluation and management of B3 nodules. First, our results demonstrate that B3 nodules were able to be categorized into an ATA-US category with a high level of inter-reader agreement. As there is no consensus definition of what the “Non-ATA” sonography pattern encompasses, it was not chosen often by the radiologists and none of the nodules received a consensus designation of Non-ATA. This is in contrast to prior reports concluding that 20–37% of ITNs could not be easily categorized into one of the five ATA sonographic risk patterns (14,19,23) and the conflicting results of the few studies of inter-reader agreement among sonographers using ATA-US in ITNs (24–26).

A second major finding of this study is the association between higher risk ATA-US patterns and suspicious GEC. Over two thirds of nodules with high suspicion sonography were found to be suspicious on GEC, and GEC suspicious nodules were more likely than GEC benign nodules to be categorized into one of the higher sonographic risk patterns.

The most important and clinically relevant findings garnered from this study regarding GEC and ATA-US as diagnostic tests in B3 nodules are (i) The diagnostic performance of ATA-US is similar to and only slightly inferior to that of GEC, (ii) Certain individual sonographic features such as border characteristics or microcalcifications may perform as well or better than ATA-US in predicting malignancy or benignity, (iii) The PPV of GEC varies across ATA-US categories, while the NPV remains fairly constant except for in the very low suspicion sonographic pattern nodules in which the NPV of a benign GEC result was 100%, and (iv) In isoechoic B3 nodules, which by definition are low suspicion

ATA-US, the presence of infiltrative margins is a predictor of malignancy.

The optimal approach to management of thyroid nodules with Bethesda category III cytology remains controversial. While molecular tests including GEC have been widely used to risk stratify such nodules these tests are relatively expensive and may not be available for all patients in all settings. It has been proposed that ATA-US can be used to risk stratify ITNs and to select which nodules should undergo molecular testing, with some authors advocating for performing molecular tests only in ITNs with LS or IS ATA-US category and others suggesting that all ITNs with high suspicion ATA-US undergo surgical management while the remainder be submitted for molecular tests (3,21). Furthermore, no studies to date have assessed the diagnostic performance of molecular tests including GEC within each ATA-US category and whether GEC test performance is different when applied to nodules of varying sonographic risk patterns. Based on the findings of our study in which the ROM in ATA-US HS pattern nodules was only 41% and in which GEC retained its NPV in this category, we suggest that GEC could be utilized in B3 nodules across all ATA-US categories, while perhaps given their low overall ROM consideration to forego molecular testing in lieu of a strategy of close surveillance could be appropriate for B3 nodules with very low sonographic risk patterns.

To the best of our knowledge, this study is first to demonstrate that the PPV of GEC varies with ATA-US category, while the NPV remains similar across the higher risk sonographic categories and low ATA risk sonographic risk B3 nodules with a 100% NPV when applied to ATA very low suspicion nodules. These findings require further validation in larger multicenter trials.

Recent reports have brought into question the applicability of the diagnostic test performance of Afirma GEC as shown in the initial validation study to real-world practice, and some have suggested that the NPV of GEC may be lower in many of the populations it is currently applied in (11). Consistent with this idea, we have found an overall NPV of GEC of 83% following a methodology outlined in a consensus article outlining a strategy for calculating test performance for Afirma GEC (22). Even if we were to assume that all GEC test negatives were true negatives, we found 8 malignancies among our total cohort of 74 negative GEC cases which would provide an NPV of 89%, lower than the 95% NPV reported in the original validation study (4). We recommend that B3 nodules with benign GEC results not having very low risk sonographic features be surveilled more closely than thyroid nodules with benign cytology. In recent years, the Afirma GEC has been replaced by Afirma Gene Sequencing Classifier with improved diagnostic performance (27). Based on findings of our study regarding GEC, we suggest that further studies be performed to assess the test performance of GSC within each ATA-US category.

Limitations of our study include that this was a single-center study and that not all patients underwent final surgical histopathologic confirmation of diagnosis or had long-term follow-up. Strengths of our study include the blinded nature of sonographic data collection and the high-inter-reader agreement among three radiologists.

Our findings suggest that ATA-US may have similar to slightly inferior diagnostic performance compared with GEC

and that the combination of information from ultrasound and GEC may alter the performance of GEC for the diagnosis of B3 nodules. We propose that in circumstances in which molecular testing is not available sonographic risk assessment may be used to risk stratify B3 nodules. Similar to prior reports, we have found that the NPV of GEC may be lower than previously reported and perhaps too low to reliably avoid the need for diagnostic surgery for most B3 nodules. Further research is needed regarding the use of ultrasound risk stratification systems to inform the performance and interpretation of molecular tests commonly utilized in the diagnosis of ITNs. Large multicenter prospective studies are needed to more definitively assess whether performance of molecular tests are altered by sonographic risk category in thyroid nodules with indeterminate FNA cytology.

Acknowledgments

We thank Dr. Jay Sosenko, Dr. Hua Li, and the University of Miami Biostatistics Collaboration and Consulting Core (BCCC).

Author Disclosure Statement

No competing financial interests exist.

Funding Information

The authors received no financial support for the research, authorship, and/or publication of this article.

Supplementary Material

Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Table S4
Supplementary Table S5
Supplementary Table S6

References

1. Seshadri KG 2017 A pragmatic approach to the indeterminate thyroid nodule. *Indian J Endocrinol Metab* **21**:751–757.
2. 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* **26**:1–133.
3. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference 2009 The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol* **132**:658–665.
4. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, Friedman L, Kloos RT, LiVolsi VA, Mandel SJ, Raab SS, Rosai J, Steward DL, Walsh PS, Wilde JJ, Zeiger MA, Lanman RB, Haugen BR 2012 Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med* **367**:705–715.
5. Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, Ghossein RA, Harrell RM, Huang J, Kennedy GC, Kim SY, Kloos RT, LiVolsi VA, Randolph GW, Sadow PM, Shanik MH, Sosa JA, Traweck ST, Walsh PS, Whitney D, Yeh MW, Ladenson PW 2018 Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg* **153**:817–824.
6. Kargi AY, Bustamante MP, Gulec S 2017 Genomic profiling of thyroid nodules: current role for Thyroseq next-generation sequencing on clinical decision making. *Mol Imaging Radionucl Ther* **26**(Suppl 1):24–35.
7. Vargas-Salas S, Martínez JR, Urra S, Domínguez JM, Mena N, Uslar T, Lagos M, Henríquez M, González HE 2018 Genetic testing for indeterminate thyroid cytology: review and meta-analysis. *Endocr Relat Cancer* **25**:R163–R177.
8. Lastra RR, Pramick MR, Crammer CJ, LiVolsi VA, Baloch ZW 2014 Implications of a suspicious afirma test result in thyroid fine-needle aspiration cytology: an institutional experience. *Cancer Cytopathol* **122**:737–744.
9. Duick DS, Klopfer JP, Diggans JC, Friedman L, Kennedy GC, Lanman RB, McIver B 2012 The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Thyroid* **22**:996–1001.
10. Kloos RT 2017 Molecular profiling of thyroid nodules: current role for the Afirma gene expression classifier on clinical decision making. *Mol Imaging Radionucl Ther* **26**(Suppl 1):36–49.
11. Valderrabano P, Hallanger-Johnson JE, Thapa R, Wang X, McIver B 2019 Comparison of postmarketing findings vs the initial clinical validation findings of a thyroid nodule gene expression classifier: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg* [Epub ahead of print]; DOI: 10.1001/jamaoto.2019.1449.
12. Villabona CV, Mohan V, Arce KM, Diacovo J, Aggarwal A, Betancourt J, Amer H, Jose T, DeSantis P, Cabral J 2016 Utility of ultrasound versus gene expression classifier in thyroid nodules with atypia of undetermined significance. *Endocr Pract* **22**:1199–1203.
13. Qing-li Z, Faquin W, Samir A 2015 Relationship between sonographic characteristics and afirma gene expression classifier results in thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Am J Roentgenol* **205**:861–865.
14. Valderrabano P, McGettigan MJ, Lam CA, Khazai L, Thompson ZJ, Chung CH, Centeno BA, McIver B 2018 Thyroid nodules with indeterminate cytology: utility of the American Thyroid Association sonographic patterns for cancer risk stratification. *Thyroid* **28**:1004–1012.
15. Kim DW, Lee EJ, Jung SJ, Ryu JH, Kim YM 2011 Role of sonographic diagnosis in managing Bethesda class III nodules. *AJNR Am J Neuroradiol* **32**:2136–2141.
16. Méndez W, Rodgers SE, Lew JI, Montano R, Solórzano CC 2008 Role of surgeon-performed ultrasound in predicting malignancy in patients with indeterminate thyroid nodules. *Ann Surg Oncol* **15**:2487–2492.
17. Gweon HM, Son EJ, Youk JH, Kim JA 2013 Thyroid nodules with Bethesda system III cytology: can ultrasonography guide the next step? *Ann Surg Oncol* **20**:3083–3088.
18. Yoo WS, Choi HS, Cho SW, Moon JH, Kim KW, Park HJ, Park SY, Choi SI, Choi SH, Lim S, Yi KH, Park DJ, Jang HC, Park YJ 2014 The role of ultrasound findings in the management of thyroid nodules with atypia or follicular lesions of undetermined significance. *Clin Endocrinol (Oxf)* **80**:735–742.

19. Ahmadi S, Herbst R, Oyekunle T, Jiang X', Strickland K, Roman S, Sosa JA **2019** Using the ATA and ACR TIRADS sonographic classifications as adjunctive predictors of malignancy for indeterminate thyroid nodules. *Endocr Pract* **25**:908–917.
20. Hong HS, Lee JY **2019** Diagnostic performance of ultrasound patterns by K-TIRADS and 2015 ATA guidelines in risk stratification of thyroid nodules and follicular lesions of undetermined significance. *Am J Roentgenol* **213**:444–450.
21. Valderrabano P, McIver B **2017** Evaluation and management of indeterminate thyroid nodules: the revolution of risk stratification beyond cytological diagnosis. *Cancer Control* **24**:1073274817729231.
22. Duh QY, Busaidy NL, Rahilly-Tierney C, Gharib H, Randolph G **2017** A systematic review of the methods of diagnostic accuracy studies of the Afirma gene expression classifier. *Thyroid* **27**:1215–1222.
23. Lam CA, McGettigan MJ, Thompson ZJ, Khazai L, Chung CH, Centeno BA, McIver B, Valderrabano P **2019** 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* **66**:278–287.
24. Pandya A, Caoili EM, Jawad-Makki F, Wasnik AP, Shankar PR, Bude R, Haymart MR, Davenport MS **2018** Limitations of the 2015 ATA guidelines for prediction of thyroid cancer: a review of 1947 consecutive aspirations. *J Clin Endocrinol Metab* **103**:3496–3502.
25. Timboli P, Deandrea M, Mormile A, Ceriani L, Garino F, Limone PP, Giovanella L **2018** American Thyroid Association ultrasound system for the initial assessment of thyroid nodules: use in stratifying the risk of malignancy of indeterminate lesions. *Head Neck* **40**:722–727.
26. Rocha TG, Rosario PW, Silva AL, Nunes MB, Silva TH, de Oliveira PHL, Calsolari MR **2018** Ultrasonography classification of the American Thyroid Association for predicting malignancy in thyroid nodules >1cm with indeterminate cytology: a prospective study. *Horm Metab Res* **50**:597–601.
27. Endo M, Nabhan F, Porter K, Roll K, Shirley LA, Azaryan I, Tonkovich D, Perlick J, Ryan LE, Khawaja R, Meng S, Phay JE, Ringel MD, Sipos JA **2019** Afirma gene sequencing classifier compared with gene expression classifier in indeterminate thyroid nodules. *Thyroid* **29**:1115–1124.

Address correspondence to:

Atil Yilmaz Kargi, MD
Division of Endocrinology, Diabetes and Metabolism
Department of Medicine
University of Miami Miller School of Medicine
1400 NW 10th Avenue Suite 807
Miami, FL 33136
USA

E-mail: akargi@miami.edu