

# Qualifiers of Atypia in the Cytologic Diagnosis of Thyroid Nodules Are Associated With Different Afirma Gene Expression Classifier Results and Clinical Outcomes

Sylvan C. Baca, MD, PhD<sup>1</sup>; Kristine S. Wong MD<sup>2</sup>; Kyle C. Strickland, MD, PhD<sup>2</sup>; Howard T. Heller, MD<sup>3</sup>; Matthew I. Kim, MD<sup>4</sup>; Justine A. Barletta, MD<sup>2</sup>; Edmund S. Cibas, MD<sup>2</sup>; Jeffrey F. Krane, MD, PhD<sup>2</sup>; Ellen Marqusee, MD<sup>4</sup>; and Trevor E. Angell, MD<sup>4</sup>

**BACKGROUND:** Thyroid nodules with atypia of undetermined significance (AUS) on fine-needle aspiration (FNA) have a low risk of malignancy that appears to vary based on specific features described in the AUS diagnosis. The Afirma gene expression classifier (GEC) is a molecular test designed to improve preoperative risk stratification of thyroid nodules, but its performance for different patterns of AUS has not been defined. The objective of this study was to assess GEC results and clinical outcomes in AUS nodules with architectural atypia (AUS-A), cytologic atypia (AUS-C) or both (AUS-C/A). **METHODS:** This was a retrospective review of all thyroid nodules with AUS cytopathology that underwent GEC testing at the authors' institution over a period of >4 years. **RESULTS:** In 227 nodules that had AUS cytology results and Afirma GEC testing, the rate of benign GEC results was higher in AUS-A nodules (70 of 107; 65%) than in AUS-C/A nodules (25 of 65; 38%;  $P = .0008$ ), and AUS-C nodules exhibited an intermediate rate of benign results (27 of 55 nodules; 59%). The risk of cancer among patients who had GEC-suspicious nodules, 86% of whom underwent resection, was 19% (6 of 25) for AUS-A nodules compared with 57% (21 of 37) for AUS-C/A nodules ( $P = .003$ ) and 45% (10 of 22) for AUS-C nodules ( $P = .07$ ). In nodules that had an indeterminate repeat cytology result, no difference was observed in the rate of benign GEC results or in the malignancy rate compared with nodules that had a single cytology result. **CONCLUSIONS:** The performance characteristics of Afirma GEC testing vary, depending on qualifiers of cytologic atypia. Recognition of these differences may enable clinicians to provide improved counseling and treatment recommendations to patients. *Cancer Cytopathol* 2017;125:313-22. © 2017 The Authors. *Cancer Cytopathology* published by Wiley Periodicals, Inc. on behalf of American Cancer Society. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

**KEY WORDS:** Afirma gene expression classifier; architectural atypia; atypia of undetermined significance (AUS); cytologic atypia; cytology; follicular lesion of undetermined significance (FLUS); indeterminate cytology; thyroid nodule.

## INTRODUCTION

Thyroid nodules are common in clinical practice and are most often benign.<sup>1,2</sup> For patients without a suppressed thyrotropin (TSH) concentration, ultrasound (US)-guided fine-needle aspiration (FNA) is recommended as the principle diagnostic test to assess for a potential cancer in most nodules measuring  $\geq 1.5$  cm.<sup>3,4</sup> The Bethesda

**Corresponding author:** Trevor E. Angell, MD, Thyroid Section, Division of Endocrinology, Diabetes, and Hypertension, Brigham and Women's Hospital, 77 Louis Pasteur Avenue, HIM 641, Boston, MA 02115; Fax: (617) 731-4718; tangell@bwh.harvard.edu

<sup>1</sup>Department of Medicine, The Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; <sup>2</sup>Department of Pathology, The Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; <sup>3</sup>Department of Radiology, The Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; <sup>4</sup>Thyroid Section, Division of Endocrinology, Hypertension, and Diabetes, The Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Additional supporting information may be found in the online version of this article.

**Received:** December 6, 2016; **Revised:** December 20, 2016; **Accepted:** December 22, 2016

Published online February 2, 2017 in Wiley Online Library (wileyonlinelibrary.com)

**DOI:** 10.1002/cncy.21827, wileyonlinelibrary.com

System for Reporting Thyroid Cytopathology (TBSRTC) has standardized reporting of thyroid FNA cytopathology into 6 groups stratified by risk of malignancy. Although cytopathologic evaluation is highly accurate in many cases, as many as one-third of aspirates have indeterminate findings, including those interpreted as atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS [hereafter referred to as AUS]), for which the possibility of malignancy is typically low but is not excluded.<sup>5,6</sup>

The optimal management of AUS nodules presents a clinical dilemma. Surgical resection is recommended for many such nodules because of the risk of cancer, although most prove to be benign.<sup>2,5-8</sup> For patients with benign nodules, superfluous surgery carries unnecessary risks,<sup>9,10</sup> whereas initial diagnostic surgery may be suboptimal for those with malignancy.<sup>11,12</sup>

Molecular testing has emerged as a powerful tool with the potential to improve the diagnostic assessment of indeterminate nodules preoperatively and is clinically available in various forms.<sup>13,14</sup> The Afirma gene expression classifier (GEC) from Veracyte, Inc, is a molecular diagnostic test that measures the expression of 167 genes to identify a benign gene profile with high accuracy<sup>15</sup> and is most often used for delineating the risk of cancers in nodules with cytology that is AUS or suspicious for follicular neoplasm (SFN), including suspicious for Hurthle cell neoplasm (SHCN). This test classifies nodules as “benign” or “suspicious,” which have a negative predictive value of 95% (95% confidence interval [CI], 85%-99%) and a positive predictive value (PPV) of 38% (95% CI, 27%-50%) for malignancy in AUS nodules.<sup>16</sup> The accuracy of a benign GEC result is comparable to that of benign cytopathology,<sup>17</sup> and surveillance is generally recommended in these patients,<sup>18-20</sup> whereas diagnostic thyroid surgery is most often undertaken for a suspicious GEC result.<sup>20,21</sup>

Given the rapid evolution of this field, investigations are ongoing to determine how best to utilize the Afirma test and to define its performance in specific clinical and cytologic settings. Several scenarios have been described in TBSRTC for which the diagnosis of AUS is appropriate.<sup>5</sup> TBSRTC does not require using diagnostic qualifiers of AUS; however, several studies have indicated that AUS with cytologic atypia is associated with a higher risk of malignancy than AUS with architectural atypia.<sup>22-27</sup> Recent reports have demonstrated that nodules with oncocyctic cytology (including predominantly Hurthle cell AUS)

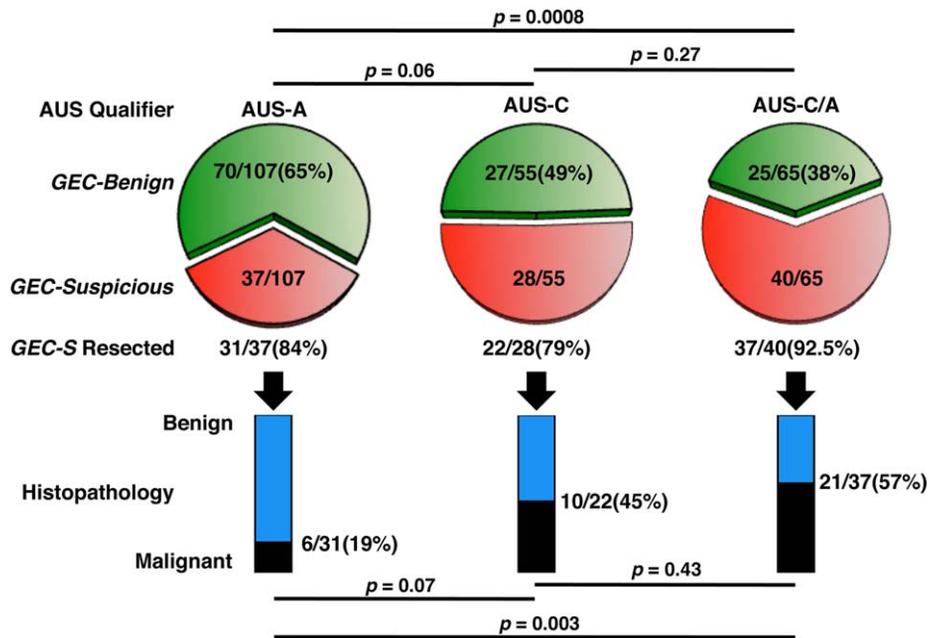
are often classified as suspicious by GEC but are unlikely to be malignant, although clinical factors, such as patient selection, may confound this observation.<sup>28-30</sup> Repeat cytologic assessment may be performed for further evaluation of a thyroid nodule with AUS cytology,<sup>2,3,31,32</sup> although the role of repeat biopsy for AUS findings in the era of molecular testing is unclear since the latter may eventually abrogate the need for additional cytologic sampling.

To further explore the performance of the Afirma GEC testing with regard to AUS qualifiers, we performed a retrospective analysis of GEC-tested nodules that had AUS cytology containing architectural atypia (AUS-A), cytologic atypia (AUS-C), or both cytologic and architectural atypia (AUS-C/A).

## MATERIALS AND METHODS

Data were retrospectively reviewed for all adult patients who visited the Thyroid Nodule Clinic at the Brigham and Women's Hospital between January 1, 2012, and July 31, 2016, and underwent FNA and Afirma GEC testing of a thyroid nodule. Patients who had an initial AUS cytology result were further investigated. To create the most clinically meaningful and accurate sample, we excluded nodules that measured <1 cm, those with a repeat cytologic result before GEC testing that was not indeterminate (benign, suspicious for malignancy [SUS], or positive for malignancy [POS]), those with an outside report of indeterminate cytology that was not repeated or reread at our institution to confirm this diagnosis, and those that had a predominantly Hurthle cell AUS cytology, because the relevance of this group has been reported elsewhere.<sup>30</sup> Nodules with a GEC result of nondiagnostic (GEC-ND) (insufficient RNA) also were excluded.

For all patients, thyroid US evaluation was performed by a radiologist with expertise in thyroid sonography using a 6-mHz to 15-mHz transducer. Nodule location, percentage of solid and cystic content, and 3-dimensional size were recorded. High-risk sonographic features, such as microcalcifications, extrathyroid extension, and abnormal lymphadenopathy, were noted when present. FNA was performed by a thyroidologist with US guidance, usually using a 25-gauge needle, and involved 2 to 4 aspirations. All aspirates were processed using a liquid-based cytology preparation (Hologic, Marlborough, Mass). Aspiration specimens were read by a cytopathologist experienced in thyroid cytopathology using TBSRTC.



**Figure 1.** Afirma gene expression classifier (GEC) results and pathologic outcomes are illustrated in thyroid nodules with different qualifiers of atypia on cytologic assessment. Pie charts comparing the proportion of Afirma GEC results that were benign (green) and suspicious (red) in thyroid nodules with a cytologic diagnosis of atypia of undetermined significance (AUS) and a qualifier of either (Left) architectural atypia (AUS-A), (Middle) cytologic atypia (AUS-C), or (Right) cytologic and architectural atypia (AUS-C/A) reveal a significant difference in the proportion of GEC results between AUS-A and AUS-C/A nodules ( $P = .0008$ ). In thyroid nodules with a suspicious GEC result that underwent surgical resection, providing a histopathologic gold-standard diagnosis, bar graphs display the proportion of benign and malignant outcomes between AUS-A, AUS-C, and AUS-C/A nodules and indicate a significant difference in the malignancy rate between AUS-A and AUS-C/A nodules ( $P = .003$ ).

Cytopathologic diagnoses of AUS with a qualifier of cytologic (AUS-C), architectural (AUS-A), or both cytologic and architectural (AUS-C/A) atypia were identified from the original cytology report descriptions provided for all AUS diagnoses. These diagnostic subcategories were defined as previously described.<sup>26,27</sup> Briefly, a diagnosis of architectural atypia was applied for a predominance of microfollicles in a sparsely cellular specimen or focal crowding/disorder of follicular cells (not attributable to preparation artifact) without associated nuclear features of papillary thyroid carcinoma (PTC) and not sufficient for a diagnosis of SFN. Cytologic atypia was regarded as the presence of nuclear features of PTC (nuclear grooves, powdery chromatin, nuclear membrane irregularity, nuclear crowding, and rare nuclear pseudoinclusions), either focally in a small number of cells in a paucicellular or normocellular specimen or similar changes that are more diffuse but mild in nature and insufficient for a diagnosis of SUS or POS. The diagnosis of AUS with cytologic and architectural atypia applied to nodules that had a combination of characteristics from both of these scenarios.

For Afirma GEC testing, 1 or 2 additional passes were performed, and the sample was sent for analysis after an initial or repeated indeterminate cytologic diagnosis (at the discretion of the treating physician). The GEC results were recorded according to the provided report as benign, suspicious, or nondiagnostic. For nodules that underwent surgical resection, histopathologic diagnosis was considered the “gold standard” in this study. For nodules in which malignancy was present, the type and extent of thyroid cancer were recorded. Cases of follicular variant of PTC (fvPTC) were reviewed to identify those that fulfilled criteria for noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP).<sup>33</sup> Because cases in the cohort were predominantly diagnosed before the introduction of NIFTP terminology and thus were diagnosed as fvPTC, NIFTPs were classified as “malignant” for the purpose of these analyses.

**Statistical Analysis**

Data are provided as numbers and percentages with mean ( $\pm$  standard deviation) or median (range) values, as

**TABLE 1.** Clinical Characteristics of the Study Cohort

Characteristic	Total	AUS-A	AUS-C	AUS-C/A
No. of nodules (%) <sup>a</sup>	227	107 (47.1)	55 (24.3)	65 (28.6)
Patient age at FNA: Mean $\pm$ SD, y		56.9 (14.6)	54.2 (15.9)	52.3 (17)
No. of women (%)		82 (76.6)	47 (85.5)	48 (73.8)
Nodule size: Median [range], cm	1.8 [1.0-6.7]	1.8 [1.0-4.6]	1.7 [1.0-6.7]	2.1 [1.0-6.4]
1-2 cm	135	68	36	31
>2-3 cm	61	24	11	26
>3 cm	31	15	8	8
Predominantly solid, no. (%) <sup>b</sup>	193 (85)	94 (87.9)	44 (80)	55 (84.6)
Multinodular goiter, no. (%) <sup>c</sup>	117 (51.5)	55 (51.4)	30 (54.5)	32 (49.2)
Repeat FNA before GEC, n. (%)	82 (36.1)	38 (35.5)	18 (32.7)	26 (40)

Abbreviations: AUS-A, atypia of undetermined significance with architectural atypia; AUS-C, with cytologic atypia, AUS-C/A with cytologic and architectural atypia; FNA, fine-needle aspiration; GEC, gene expression classifier; SD, standard deviation.

<sup>a</sup>Patients were counted in multiple categories if they had  $\geq 2$  nodules with differing cytology.

<sup>b</sup>Predominant solid nodules were those defined as solid or with <25% cystic component.

<sup>c</sup>Multinodular goiter was defined as  $\geq 2$  nodules, each measuring  $\geq 1$  cm.

indicated. Statistical testing was performed using the Fisher exact test for categorical variables and the Student *t* test for continuous variables. Comparisons were made for the frequency of GEC benign and suspicious results and for benign and malignant histopathology in nodules with a cytologic diagnosis of AUS-A, AUS-C, or AUS-C/A. Statistical significance was defined as a 2-tailed *P* value < .05 for all analyses. Analyses were performed using GraphPad v5.0d (GraphPad Software, Inc, La Jolla, Calif) and JMP (SAS, Cary, NC), and Figure 1 was created using Photoshop CS5 (Adobe Systems, Inc, San Jose, Calif).

Permission for this study was granted by the Investigational Review Board at the Brigham and Women's Hospital, which allowed a waiver of informed consent for the study. This study did not receive any financial support, approbation, or review by any commercial entity.

## RESULTS

In total, 413 eligible nodules with GEC testing were identified during this period, of which 278 nodules (67.3%) were AUS, 122 (29.5%) were SFN (including 50 SHCNs), 5 (1.2%) were SUS, and 8 (1.9%) had an outside report of indeterminate cytology that was not reconfirmed at our institution. Thirty nodules with AUS cytology (7.3%) were Hurthle cell-predominant, repeat cytology was benign in 9 nodules and SUS in 3 cases, and 9 nodules were GEC-ND, resulting in a final study population of 227 nodules from 219 patients after these exclusions. Over the study period, the overall rate of AUS diagnoses was 11.1% at our institution.

The cohort is described in Table 1 and included predominantly women (171 of 219 patients; 78.1%), ranged

in age from 21 to 87 years (median age, 56.3 years), and had a median nodule size of 1.80 cm (range, 1.0-6.7 cm). Atypia was more frequently architectural (AUS-A, 47.1%), than cytologic (AUS-C, 24.3%) or both architectural and cytologic atypia (AUS-C/A, 28.6%). A repeat FNA was performed for 36% of the included nodules at the discretion of the treating endocrinologist before obtaining an Afirma GEC test.

### GEC Result and Surgical Outcome by AUS Qualifiers

The proportion of GEC-benign results for each AUS subtype was analyzed, and the results are illustrated in Figure 1. Nodules with an initial AUS-A, AUS-C, and AUS-C/A cytology result were GEC-benign in 70 of 107 (65%), 27 of 55 (49%), and 25 of 65 (38%) cases, respectively. AUS-A nodules were significantly more likely to have a GEC-benign result than AUS-C/A nodules (*P* = .0008), and a similar trend was observed for AUS-A versus AUS-C nodules (*P* = .06). The GEC results were similar for AUS-C versus AUS-C/A nodules (*P* = .27). No significant correlations were observed between GEC results and clinical characteristics, including sex, nodule size, cystic content, and the presence of multinodularity (Table 2).

For nodules in patients who underwent surgical resection, the histopathologic diagnoses were recorded and compared by AUS subtype. Nodule resection was performed in 10 of 122 (8.2%) GEC-benign nodules with no malignancies identified on final pathology, whereas 90 of 105 (85.7%) GEC-suspicious nodules were resected. Malignancy was diagnosed in 37 of these 90 nodules (41%). Of the 15 unresected nodules, 3 are planned for surgery, 6 have not been resected because of surgical risk or

**TABLE 2.** Clinical Features and Relation to Gene Expression Classifier Results

Variable	No. of Nodules			Proportion of GEC Benign	<i>P</i> <sup>a</sup>
	Total	GEC Benign	GEC Suspicious		
Total	227	122	105	0.54	
Men	50	22	28	0.44	
Women	177	100	77	0.56	.12
Nodule size < 3 cm	194	106	88	0.55	
Nodule size > 3 cm	33	16	17	0.48	.51
Predominantly solid	193	104	89	0.54	
Complex, > 25% cystic	34	18	16	0.53	.92
Solitary	110	55	55	0.50	
Multinodular goiter <sup>b</sup>	117	66	51	0.56	.33

Abbreviation: GEC, gene expression classifier.

<sup>a</sup>Fisher exact test *P* values are listed for the indicated comparisons of GEC results.

<sup>b</sup>Multinodular goiter was defined as ≥2 nodules, each measuring ≥ 1 cm.

**TABLE 3.** Surgical Diagnoses by Subtype of Atypia

Diagnosis	No. of Nodules			Total
	AUS-A	AUS-C	AUS-C/A	
GEC-benign				
Benign histopathology				
Adenomatous/hyperplastic nodule	3	2	1	6
Follicular adenoma	2	0	2	4
GEC-suspicious				
Benign histopathology				
Adenomatous/hyperplastic nodule	15	8	14	37
Follicular adenoma	10	3	1	14
Other	0	2 <sup>a</sup>	0	2
Malignant histopathology				
Classic variant of papillary thyroid carcinoma	0	2	3	5
Follicular variant of papillary thyroid carcinoma	4	2	8	14
Noninvasive follicular neoplasm with papillary-like nuclear features	1	5	9	15
Follicular thyroid carcinoma	1	1	1	3
Total	36	25	39	100

Abbreviations: AUS-A, atypia of undetermined significance (AUS) with architectural atypia; AUS-C, with cytologic atypia, AUS-C/A with cytologic and architectural atypia; FNA, fine-needle aspiration; GEC, gene expression classifier.

<sup>a</sup>These included 1 hyalinizing trabecular tumor and 1 case of thyroid tissue with a giant cell reaction to suture in a patient with AUS-C who underwent previous thyroid surgery.

concurrent nonthyroid cancer treatment, 5 patients decided against diagnostic resection and are undergoing monitoring, and 1 patient was lost to follow-up.

Comparing diagnostic outcomes for the AUS-A, AUS-C, and AUS-C/A qualifiers, malignancy was identified in 6 of 31 (19.4%), 10 of 22 (45.5%), and 21 of 37 (56.8%) nodules, respectively (Fig. 1). AUS-A nodules were significantly less likely to be malignant than AUS-C/A nodules ( $P = .003$ ), and malignancy tended to be less common in AUS-A versus AUS-C nodules ( $P = .07$ ). The malignancy risk appeared similar between AUS-C and AUS-C/A nodules ( $P = .43$ ).

Histopathologic diagnoses for benign and malignant lesions are provided in Table 3. Among the 37 malignant lesions, there were 5 (13.5%) classic PTCs (cPTCs), 3

(8.1%) follicular thyroid carcinomas (FTCs), and 29 (78.4%) fvPTCs, of which 15 (51.7% of fvPTCs and 40.5% of all malignant cases) would be classified as NIFTP. All cPTCs were identified in aspirates with cytologic atypia (either AUS-C or AUS-C/A). FTCs and fvPTCs were distributed among the AUS qualifiers. Nodules classified as NIFTP were identified in all AUS qualifier categories but comprised 1 of 6 (16.7%) AUS-A malignancies compared with 5 of 10 (50%) AUS-C malignancies and 9 of 21 (42.6%) AUS-C/A malignancies ( $P = .37$ ). Of the NIFTPs, 1 of 15 (6.7%) were from the AUS-A group. All malignancies were classified as T1 or T2 and were either Nx or N0, but a minority were in the intermediate American Thyroid Association recurrence risk category,<sup>3</sup> including 2 angioinvasive FTCs, 3 multifocal

**TABLE 4.** Results of Gene Expression Classifier Testing and Surgical Outcomes in Nodules Diagnosed as Atypia of Undetermined Significance With or Without Repeat Cytology

Result	No. (% of Total With Repeat FNA)	GEC Benign, No. (%)	GEC Suspicious, No.	GEC Suspicious Resected, No.	Path Benign, No.	Path Malignant, No.	PPV
AUS on single FNA	145	76 (52.4)	69	59	34	25	0.42
Two indeterminate FNAs	82	46 (56.1)	36	32	20	12	0.38
AUS → AUS	54 (65.9)	27 (50)	27	24	14	10	0.42
AUS → SFN <sup>a</sup>	28 (34.1)	19 (67.9)	9	8	6	2	0.25

Abbreviations: AUS-A, atypia of undetermined significance (AUS) with architectural atypia; AUS-C, AUS with cytologic atypia, AUS-C/A AUS with cytologic and architectural atypia; FNA, fine-needle aspiration; GEC, gene expression classifier; Path, pathology; PPV, positive predictive value; SFN, suspicious for follicular neoplasm.

<sup>a</sup>These include 1 nodule that was suspicious for Hurthle cell neoplasm.

**TABLE 5.** Comparisons and *P* Values

Comparison	<i>P</i> <sup>a</sup>
[AUS on single FNA   2 indeterminate FNAs] vs [GEC benign   GEC suspicious]	.68
[AUS on single FNA   AUS → AUS] vs [GEC benign   GEC suspicious]	.87
[AUS → AUS   AUS → SFN] vs [GEC benign   GEC suspicious]	.16
[AUS on single FNA   2 indeterminate FNAs] vs [Path benign   Path malignant]	.82
[AUS on single FNA   AUS → AUS] vs [Path benign   Path malignant]	1
[AUS → AUS   AUS → SFN] vs [Path benign   Path malignant]	.68

Abbreviations: AUS, atypia of undetermined significance; FNA, fine-needle aspiration; GEC, gene expression classifier; Path, pathology; SFN, suspicious for follicular neoplasm.

<sup>a</sup>Fisher exact test *P* values are for 2 × 2 comparisons of data from Table 4.

infiltrative fvPTCs, 3 cPTCs with lymphovascular invasion, and 1 cPTC with the BRAF-V600E mutation. Despite the low number of malignancies identified in AUS-A nodules, 2 intermediate-risk cancers were from this subtype (2 of 6; 33%) compared with 3 of 8 cancers (37.5%) from AUS-C nodules and 4 of 17 cancers (23.5%) from AUS-C/A nodules.

### GEC Results and Repeat FNA

Of the 227 nodules with AUS cytology in this study, 82 had a repeat indeterminate cytologic result. This cytologic assessment was again AUS for 54 nodules (65.9%) and SFN for 28 nodules (34.1%) (Table 4). Repeat cytology was obtained in similar proportions of each AUS subtype (Supporting Table 1; see online supporting information). Nodules with AUS-C cytology had an identical repeat cytologic diagnosis in 10 of 18 (55.6%) cases compared with 15 of 38 (39.5%) AUS-A nodules. Only 1 of 17 (5.9%) AUS-C nodules had a repeat cytology result of SFN, whereas 19 of 38 (50%) AUS-A nodules had a repeat

SFN cytology result ( $P = .002$ ) (Supporting Table 1; see online supporting information).

The GEC and surgical pathology results for nodules with a single AUS cytology or repeat cytology are compared in Table 5. Nodules that had a single AUS result, compared with those that had a repeat indeterminate cytology (AUS or SFN) result, were GEC-benign in 76 of 145 (52.4%) versus 46 of 82 (56.1%) nodules, respectively ( $P = .68$ ). Nodules that had 2 AUS results were GEC-benign in 27 of 54 (50%) cases ( $P = .87$  vs a single AUS result). An evaluation of the malignancy rate within resected GEC-suspicious nodules revealed that 25 of 34 nodules (42%) with a single AUS result were malignant versus 12 of 32 nodules (38%) with repeat indeterminate cytology results ( $P = .82$ ) and 10 of 24 nodules (42%) with repeat AUS results ( $P = 1.0$ ).

To assess for selection bias that could influence which nodules underwent single versus repeat FNA, we analyzed clinical variables that may have been relevant to this decision, including patient age, sex, nodule size, cystic component, presence of multinodularity, sonographic risk, and AUS qualifier observed on first FNA (Supporting Table 2; see online supporting information). Solitary nodules comprised 58.5% of those that underwent repeat FNA compared with only 42.8% of those with a single FNA ( $P = .03$ ). No other significant differences were observed between the groups with respect to these clinical and sonographic factors.

## DISCUSSION

The management of thyroid nodules with indeterminate cytology has evolved in recent years with the adoption of molecular testing for improved preoperative risk stratification. The Afirma GEC was designed to identify benign nodules with high negative predictive value and

demonstrated both clinical validity and utility in a multicenter prospective trial and follow-up evaluation.<sup>16,20</sup> In the current study, <10% of GEC-benign nodules underwent resection; and, of those, none proved to be malignant, confirming the value of a benign GEC result in appropriately avoiding thyroid surgery.

The ideal application of the Afirma GEC is still uncertain and has been the subject of many reports.<sup>19–21,28–30,34–38</sup> A PPV of 37% to 38% for AUS and SFN nodules was initially found,<sup>16</sup> but subsequent investigations have raised concerns about the PPV of GEC testing both generally and in specific circumstances.<sup>28–30,35–38</sup> A better understanding of GEC performance in various clinical circumstances is pertinent to how clinicians advise patients. To our knowledge, this report provides the first evaluation of the influence of the AUS qualifiers of architectural (AUS-A), cytologic (AUS-C), and combined architectural and cytologic (AUS-C/A) atypia for GEC-tested nodules. The results indicate that, although GEC testing is clinically useful and delivers the expected performance across these AUS subtypes, there are significant differences in outcomes among these groups, suggesting that, by considering these AUS qualifiers, clinicians may more specifically define thyroid cancer risk and provide improved counseling to their patients.

The GEC result was benign in 54% of nodules, which is similar to many previous studies.<sup>20,29,36–40</sup> However, AUS-A nodules were most likely to be GEC-benign, and AUS-C/A nodules were least likely. In light of this, GEC testing appears to be particularly beneficial for AUS-A nodules, for which the greatest proportion of diagnostic surgeries were avoided. Even for AUS-C/A nodules, however, thyroid resection was avoided in nearly 40% based on GEC assessment. Patient characteristics, nodule size, nodule complexity, and the presence of multinodularity were not associated with GEC results.

The data also suggest that these AUS subtypes differ with respect to the risk of malignancy in nodules with a suspicious GEC result. The overall PPV for GEC-suspicious AUS nodules was 41% (for the 86% of those that had undergone resection at the time of reporting), approximating the results from several other studies.<sup>16,20,37–39</sup> In some previous reports, a lower malignancy rate was observed in GEC-suspicious nodules with Hurthle cell-predominant AUS cytology.<sup>30</sup> The current study additionally indicates a lower malignancy risk for AUS-A nodules compared with AUS-C or AUS-C/A nodules. Because

isolated architectural atypia portends a lower risk of malignancy than cytologic atypia,<sup>27,41,42</sup> our findings may reflect this lower underlying prevalence of malignancy, which decreases the PPV of any diagnostic test. It is noteworthy that, with nearly two-thirds of AUS-A nodules classified as GEC-benign, the remaining GEC-suspicious AUS-A nodules were likely enriched for malignancy and demonstrated a malignancy rate of 19%, which is greater than expected for the AUS category in general.<sup>5</sup> Others have suggested that the higher rate of suspicious GEC results and low PPV in nodules with Hurthle cell features (within either AUS or SHCN) argue against the use of the GEC in such cases.<sup>30,36</sup> In contrast, despite a lower PPV than their counterparts with cytologic atypia, AUS-A nodules have a relatively lower incidence of suspicious GEC results. Moreover, malignancies in AUS-A nodules had higher risk histopathology in 2 of 6 (33%) cases, similar to the other AUS subtypes, suggesting that resection of AUS-A nodules with a suspicious GEC result is clinically appropriate.

The association of descriptive qualifiers of AUS and an ultimate diagnosis of NIFTP was also evaluated. Although the difference was not statistically significant, NIFTP was identified more often in AUS nodules with cytologic atypia (either AUS-C or AUS-C/A). Cytologically, NIFTP is characterized by a predominantly microfollicular architecture and nuclear features suggestive of PTC but typically lacking nuclear pseudoinclusions.<sup>43,44</sup> Our series included NIFTPs that were preceded by an AUS diagnosis and thus were characterized by only mild architectural and/or cytologic atypia, which was insufficient to warrant an SFN or SUS diagnosis. Our data suggest that the presence of cytologic atypia may be more predictive of NIFTP on resection than architectural atypia.

The ideal treatment of AUS nodules is debated,<sup>3</sup> and repeat cytologic assessment is endorsed as one management strategy.<sup>2,3,45–47</sup> In the current cohort, 36% of nodules had repeat cytology performed before GEC evaluation. On repeat FNA, nodules with AUS-C were more likely to remain AUS than AUS-A nodules, which were more frequently SFN. This is reasonable, because the AUS-A subtype has cytologic features similar to those of SFN but with less abundant cellularity.<sup>48</sup> Data from others have suggested that nodules with 2 AUS aspirates have a higher risk of malignancy than those with a single AUS result, both in general and for GEC-tested nodules.<sup>29,39,45,46</sup> In a recent report, the PPV increased from approximately 66% to 91% for nodules that had 1 versus 2 AUS results.<sup>39</sup> The

data presented here demonstrate an identical malignant risk for a single versus repeated AUS cytology result in GEC-suspicious thyroid nodules and statistically similar results for a single AUS versus repeated indeterminate cytology of either kind (AUS or SFN). Thus, repeat FNA did not improve the PPV of Afirma testing of AUS nodules. The analysis of clinical factors and sonographic features did not reveal an apparent source of selection bias pertaining to which nodules underwent repeat FNA. It is likely that patient and physician preference played a substantial role in this decision, because there is no clear standard for management.

There are limitations to this study. Dividing nodules by AUS qualifiers limited the statistical power for detecting small differences between qualifier groups in this sample size. This was a single-center study, and cytopathologic evaluation of AUS qualifiers may not be reported in the same manner at other institutions. In addition, there is substantial interrater/intrater variability for indeterminate cytologic diagnoses that may limit the consistency of these findings.<sup>49</sup> In this evaluation, the predominantly Hurthle cell qualifier of AUS cytology was not included, because nodules with this cytology from our institution have recently been reported elsewhere.<sup>30</sup> The handling of tumors identified as NIFTP is particularly problematic. Cytologic evaluation cannot distinguish NIFTP from invasive/infiltrative fvPTC. A diagnosis of NIFTP can only be rendered upon histologic evaluation of the entire tumor capsule/periphery, and these lesions have recently been described with defined criteria and nomenclature, highlighting their indolent nature and suggesting that they may represent premalignant lesions.<sup>33,50</sup> Therefore, resection is still considered appropriate. This would be analogous to the removal of *RAS* mutation-positive thyroid nodules because of the possibility of malignant evolution,<sup>51</sup> although many prove benign, and some may have indolent behavior.<sup>52</sup> Here, NIFTPs were categorized as “true-positives” despite their low malignant potential.

In conclusion, AUS qualifiers of architectural, cytologic, or architectural and cytologic atypia were associated with differences in Afirma GEC test outcomes. Nodules with architectural atypia were more likely to have a benign GEC result, and nodules with both cytologic and architectural atypia were most likely to be malignant when the GEC result was suspicious. Although these data do not indicate that management decisions regarding Afirma GEC testing or results should necessarily be different based

on AUS qualifiers, patients often present with complex clinical circumstances that require individualized risk assessment. These findings provide further information for this assessment and may improve clinical decision making and counseling of patients with thyroid nodules who have an AUS cytology result. This information, along with other clinical factors, may help improve the preoperative diagnostic accuracy of molecular testing.

## FUNDING SUPPORT

No specific funding was disclosed.

## CONFLICT OF INTEREST DISCLOSURES

Trevor E. Angell receives research support from Veracyte, Inc. The remaining made no disclosures.

## AUTHOR CONTRIBUTIONS

**Sylvan C. Baca:** Data curation, formal analysis, investigation, visualization, writing—original draft, and writing—review and editing. **Kristine S. Wong:** Investigation, data curation, and writing—review and editing. **Kyle C. Strickland:** Investigation, data curation, and writing—review and editing. **Howard T. Heller:** Investigation and writing—review and editing. **Matthew I. Kim:** Writing—review and editing. **Justine A. Barletta:** Investigation and writing—review and editing. **Edmund S. Cibas:** Conceptualization, visualization, and writing—review and editing. **Jeffrey F. Krane:** Investigation and writing—review and editing. **Ellen Marqusee:** Conceptualization, methodology, data curation, visualization, and writing—review and editing. **Trevor E. Angell:** Conceptualization, formal analysis, investigation, methodology, project administration, supervision, visualization, and writing—review and editing.

## REFERENCES

1. Popoveniuc G, Jonklaas J. Thyroid nodules. *Med Clin North Am.* 2012;96:329-349.
2. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer.* 2007;111:508-516.
3. Haugen BR, Alexander EK, Bible KC, 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.
4. Gharib H, Papini E, Paschke R, et al; AACE/AME/ETA Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. *J Endocrinol Invest.* 2010;33:51-56.

5. Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*. 2009;19:1159-1165.
6. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol*. 2012;56:333-339.
7. Baloch ZW, Livolsi VA. Follicular-patterned lesions of the thyroid: the bane of the pathologist. *Am J Clin Pathol*. 2002;117:143-150.
8. Nayar R, Ivanovic M. The indeterminate thyroid fine-needle aspiration: experience from an academic center using terminology similar to that proposed in the 2007 National Cancer Institute Thyroid Fine Needle Aspiration State of the Science Conference. *Cancer*. 2009;117:195-202.
9. Sosa JA, Bowman HM, Tielsch JM, Powe NR, Gordon TA, Udelsman R. The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg*. 1998;228:320-330.
10. Bergenfelz A, Jansson S, Kristofferson A, et al. Complications to thyroid surgery: results as reported in a database from a multicenter audit comprising 3,660 patients. *Langenbecks Arch Surg*. 2008;393:667-673.
11. Esnaola NF, Cantor SB, Sherman SI, Lee JE, Evans DB. Optimal treatment strategy in patients with papillary thyroid cancer: a decision analysis. *Surgery*. 2001;130:921-930.
12. Hundahl SA, Cady B, Cunningham MP, et al. Initial results from a prospective cohort study of 5583 cases of thyroid carcinoma treated in the United States during 1996. US and German Thyroid Cancer Study Group. An American College of Surgeons Commission on Cancer Patient Care Evaluation study. *Cancer*. 2000;89:202-217.
13. Kim MI, Alexander EK. Diagnostic use of molecular markers in the evaluation of thyroid nodules. *Endocr Pract*. 2012;18:796-802.
14. Nishino M. Molecular cytopathology for thyroid nodules: a review of methodology and test performance. *Cancer Cytopathol*. 2016;124:14-27.
15. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab*. 2010;95:5296-5304.
16. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;367:705-715.
17. Wang CC, Friedman L, Kennedy GC, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid*. 2011;21:243-251.
18. Bernet V, Hupart KH, Parangi S, Woeber KA. AACE/ACE disease state commentary: molecular diagnostic testing of thyroid nodules with indeterminate cytopathology. *Endocr Pract*. 2014;20:360-363.
19. Angell TE, Frates MC, Medici M, et al. Afirma benign thyroid nodules show similar growth to cytologically benign nodules during follow-up. *J Clin Endocrinol Metab*. 2015;100:E1477-E1483.
20. Alexander EK, Schorr M, Kloppner J, et al. Multicenter clinical experience with the Afirma gene expression classifier. *J Clin Endocrinol Metab*. 2014;99:119-125.
21. McIver B, Castro MR, Morris JC, et al. An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2014;99:4069-4077.
22. Singh RS, Wang HH. Eliminating the "atypia of undetermined significance/follicular lesion of undetermined significance" category from the Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2011;136:896-902.
23. Horne MJ, Chhieng DC, Theoharis C, et al. Thyroid follicular lesion of undetermined significance: evaluation of the risk of malignancy using the 2-tier sub-classification. *Diagn Cytopathol*. 2012;40:410-415.
24. Olson MT, Clark DP, Erozan YS, Ali SZ. Spectrum of risk of malignancy in subcategories of "atypia of undetermined significance." *Acta Cytol*. 2011;55:518-525.
25. Renshaw AA. Should "atypical follicular cells" in thyroid fine-needle aspirates be subclassified? *Cancer Cytopathol*. 2010;118:186-189.
26. VanderLaan PA, Marqusee E, Krane JF. Usefulness of diagnostic qualifiers for thyroid fine-needle aspirations with atypia of undetermined significance. *Am J Clin Pathol*. 2011;136:572-577.
27. Wu HH, Inman A, Cramer HM. Subclassification of "atypia of undetermined significance" in thyroid fine-needle aspirates. *Diagn Cytopathol*. 2014;42:23-29.
28. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocyctic cell types in patients with indeterminate thyroid cytology. *Endocr Pract*. 2014;20:364-369.
29. Lastra RR, Pramick MR, Crammer CJ, LiVolsi VA, Baloch ZW. Implications of a suspicious Afirma test result in thyroid fine-needle aspiration cytology: an institutional experience. *Cancer Cytopathol*. 2014;122:737-744.
30. Brauner E, Holmes BJ, Krane JF, et al. Performance of the Afirma gene expression classifier in Hurthle cell thyroid nodules differs from other indeterminate thyroid nodules. *Thyroid*. 2015;25:789-796.
31. Dhyani M, Faquin W, Lubitz CC, Daniels GH, Samir AE. How to interpret thyroid fine-needle aspiration biopsy reports: a guide for the busy radiologist in the era of the Bethesda classification system. *AJR Am J Roentgenol*. 2013;201:1335-1339.
32. Baloch ZW, Fleisher S, LiVolsi VA, Gupta PK. Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. *Diagn Cytopathol*. 2002;26:41-44.
33. Nikiforov YE, Seethala RR, Tallini G, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol*. 2016;2:1023-1029.
34. Celik B, Whetsell CR, Nassar A. Afirma GEC and thyroid lesions: an institutional experience. *Diagn Cytopathol*. 2015;43:966-970.
35. Marti JL, Avadhani V, Donatelli LA, et al. Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules. *Ann Surg Oncol*. 2015;22:3996-4001.
36. Yang SE, Sullivan PS, Zhang J, et al. Has Afirma gene expression classifier testing refined the indeterminate thyroid category in cytology? *Cancer Cytopathol*. 2016;124:100-109.
37. Sacks WL, Bose S, Zumsteg ZS, et al. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. *Cancer Cytopathol*. 2016;124:722-728.
38. Chaudhary S, Hou Y, Shen R, Hooda S, Li Z. Impact of the Afirma gene expression classifier result on the surgical management of thyroid nodules with category III/IV cytology and its correlation with surgical outcome. *Acta Cytol*. 2016;60:205-210.
39. Villabona CV, Mohan V, Arce KM, et al. Utility of ultrasound versus gene expression classifier in thyroid nodules with atypia of undetermined significance. *Endocr Pract*. 2016;22:1199-1203.
40. Wu JX, Young S, Hung ML, et al. Clinical factors influencing the performance of gene expression classifier testing in indeterminate thyroid nodules. *Thyroid*. 2016;26:916-922.
41. Park HJ, Moon JH, Yom CK, et al. Thyroid "atypia of undetermined significance" with nuclear atypia has high rates of malignancy and BRAF mutation. *Cancer Cytopathol*. 2014;122:512-520.
42. Nishino M, Wang HH. Should the thyroid AUS/FLUS category be further stratified by malignancy risk? *Cancer Cytopathol*. 2014;122:481-483.
43. Howitt BE, Chang S, Eszlinger M, et al. Fine-needle aspiration diagnoses of noninvasive follicular variant of papillary thyroid carcinoma. *Am J Clin Pathol*. 2015;144:850-857.
44. Strickland KC, Vivero M, Jo VY, Lowe AC, et al. Preoperative cytologic diagnosis of noninvasive follicular thyroid neoplasm with

- papillary-like nuclear features: a prospective analysis. *Thyroid*. 2016;26:1466-1471.
45. Samulski TD, LiVolsi VA, Wong LQ, Baloch Z. Usage trends and performance characteristics of a “gene expression classifier” in the management of thyroid nodules: An institutional experience. *Diagn Cytopathol*. 2016;44:867-873.
  46. Faquin WC, Baloch ZW. Fine-needle aspiration of follicular patterned lesions of the thyroid: diagnosis, management, and follow-up according to National Cancer Institute (NCI) recommendations. *Diagn Cytopathol*. 2010;38:731-739.
  47. Faquin WC. Reply to can a gene-expression classifier with high negative predictive value solve the indeterminate thyroid fine-needle aspiration dilemma [letter]? *Cancer Cytopathol*. 2013;121:404.
  48. Ohori NP, Schoedel KE. Variability in the atypia of undetermined significance/follicular lesion of undetermined significance diagnosis in the Bethesda System for Reporting Thyroid Cytopathology: sources and recommendations. *Acta Cytol*. 2011;55:492-498.
  49. Cibas ES, Baloch ZW, Fellegara G, et al. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann Intern Med*. 2013;159:325-332.
  50. Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab*. 2009;94:2092-2098.
  51. Gupta N, Dasyam AK, Carty SE, et al. RAS mutations in thyroid FNA specimens are highly predictive of predominantly low-risk follicular-pattern cancers. *Clin Endocrinol Metab*. 2013;98:E914-E922.
  52. Medici M, Kwong N, Angell TE, et al. The variable phenotype and low-risk nature of RAS-positive thyroid nodules [serial online]. *BMC Med*. 2015;13:184.