Relationship Between Sonographic Characteristics and Afirma Gene Expression Classifier Results in Thyroid Nodules With Indeterminate Fine-Needle Aspiration Cytopathology

Qing-Li Zhu1,2
William C. Faquin2
Anthony E. Samir2

OBJECTIVE. The purpose of this article is to investigate whether specific clinical and sonographic characteristics are predictive of a benign Afirma test result.

MATERIALS AND METHODS. We conducted a retrospective study of Afirma gene expression classifier analysis performed in 44 patients with 45 indeterminate thyroid fine-needle aspiration (FNA) cytologic results between March 2013 and April 2014. Of these, 33 of 45 nodules (73.3%) were repeat atypia of undetermined significance (AUS) and follicular lesions of undetermined significance (FLUS), or follicular neoplasm (FN) and suspicious for a follicular neoplasm (SFN) before Afirma testing.

RESULTS. Of the 45 nodules, 21 (46.7%) were cytologically diagnosed as FLUS, 16 (35.6%) were diagnosed as AUS, and eight (17.8%) were diagnosed as FN or SFN. By Afirma testing, 23 of the 45 nodules (51.1%) were benign, 21 (46.7%) were suspicious, and one (2.2%) had nondiagnostic results. The mean (± SD) nodule size was smaller in the Afirma-benign group than in the Afirma-suspicious group (1.8 ± 0.8 cm [95% CI, 1.4–2.1] vs 2.2 ± 0.8 cm [95% CI, 1.8–2.6]; p < 0.035). No sonographic feature was statistically significantly different between the Afirma-benign and -suspicious groups, including nodule solidity (p = 0.225), echogenicity (p = 0.543), calcification (p = 0.542), and hypervascularity (p = 0.976). All nodules were ovoid shaped and had circumscribed margins in both Afirma groups.

CONCLUSION. Smaller nodule size was the only characteristic associated with a benign diagnosis on Afirma testing. Sonographic characteristics are not helpful in cases that had a repeat indeterminate FNA finding before Afirma testing.
strategy might be to perform confirmatory Afirma testing in patients whose indeterminate nodules on FNA are considered more likely to be benign on the basis of nodule sonographic characteristics and patient clinical characteristics. Clinical and nodule sonographic characteristics are known to be predictive of malignancy risk. For example, in one study of 165 AUS-FLUS nodules, the malignancy rate of nodules was much higher (79.3%) in a selected patient population with suspicious sonographic findings than in a population with indeterminate features on ultrasound (24.7%) [12].

The goal of this study was to investigate whether specific clinical and sonographic characteristics are predictive of a benign Afirma test result.

Materials and Methods

Patients

The institutional review board at Massachusetts General Hospital waived the requirement to obtain informed consent for the retrospective study. Our hospital is a quaternary referral center providing comprehensive care for patients with thyroid cancer, including dedicated thyroid endocrinology, surgery, and nuclear medicine facilities. Afirma gene expression classifier analysis has been performed since March 2013, which was chosen as the starting point of our study. From March 2013 through April 2014, the electronic medical record was reviewed for patients who underwent thyroid FNA biopsy, had received a diagnosis of AUS-FLUS or FN-SFN, according to the Bethesda System for Reporting Thyroid Cytopathology [4], and then were sent for Afirma testing. A total of 44 patients (27 women and 17 men) with 45 thyroid nodules were analyzed in the current study. Of these, 33 of 45 nodules (73.3%) were repeat AUS-FLUS or FN-SFN before Afirma testing. For each subject, we collected demographic, sonographic, cytologic, Afirma gene expression classifier, histopathologic, and subsequent management and outcome data.

Thyroid Ultrasound

One radiologist with 10 years of experience in thyroid ultrasound reviewed each patient’s images. The reviewer took the lesion size measured on the original images as the nodule size and determined the ultrasound features by image review. The original reports were not used to determine image features. The reviewer was blinded to FNA or Afirma results when classifying ultrasound results. All patients had their electronic medical records or charts reviewed for the following sonographic variables: nodule solidity, echogenicity, shape, margin, calcifications, and vascularity. Nodule solidity was classified as solid, predominantly solid (> 50% solid for a mixed nodule), and predominantly cystic (> 50% cyst for a mixed nodule). Nodule shape was classified as ovoid, irregular, or taller than wide. Nodule margin was classified as circumscribed or noncircumscribed (i.e., ill defined). Echogenicity was classified into markedly hypoechoic (i.e., more hypoechoic than the adjacent strap muscle), hypoechoic, isoechoic to thyroid, or hyperechoic. Calcification was classified as no calcifications, microcalcifications (< 1 mm in diameter and visualized as tiny punctuate hyperechoic foci without acoustic shadows), or macrocalcifications (hyperechoic foci > 1 mm in size).
Nodule vascularity was assessed by color Doppler imaging and was classified as hypervascular (i.e., greater vascularity in the nodule than in the normal thyroid parenchyma), hypovascular, or none.

**Fine-Needle Aspiration Procedure and Afirma Gene-Expression Classifier Analysis**

Under ultrasound guidance, multiple FNA specimens were obtained using 25-gauge needles. Samples were preserved on alcohol-fixed slides in Cytolyt (BD Diagnostics-TriPath) and in Afirma testing medium. FNA cytologic findings were read by board-certified subspecialty cytopathologists and were classified according to the Bethesda system [4]. When classified as cytologically indeterminate, either on initial or repeat FNA biopsy, depending on the preference of the referring physician, aspirates were sent for Afirma analysis.

**Statistical Analysis**

SPSS statistics software (version 13, IBM) was used for the data analysis. Univariate analysis was done using the Fisher exact test for categoric variables (patient sex), and the independent samples t test was used to compare means of the continuous normal data (patient age). Nodule size was not normally distributed and, thus, the Mann-Whitney U test was used. A p < 0.05 was considered to indicate a statistically significant difference.

**Results**

Forty-five Afirma gene expression classifier analyses were performed for 44 patients with 45 indeterminate thyroid nodules between March 2013 and April 2014. The 44 patients in our cohort consisted of 27 (61.4%) women and 17 (38.6%) men, with a mean age of 56.8 ± 13.1 years. None of the patients had a history of ionizing radiation exposure in childhood, thyroid carcinoma, or family history of thyroid carcinoma.

Of the 45 nodules, 51.1% (23/45) were reported as Afirma benign, 46.7% (21/45) as Afirma suspicious, and the remaining 2.2% (1/45) as nondiagnostic. The mean nodule size in the Afirma-suspicious group was larger than that in the Afirma-benign group (2.2 ± 0.8 cm [95% CI, 1.8–2.6 cm] vs 1.8 ± 0.8 cm [95% CI, 1.4–2.1 cm]; p = 0.035). No statistically significant differences were observed for patient age (58.3 ± 13.7 years vs 56.6 ± 11.8 years; p = 0.648), sex (p = 0.537), or thyroid function test results (thyroid-stimulating hormone level, 1.7 ± 0.9 mIU/L vs 2.7 ± 1.9 mIU/L; p = 0.097).

Of the 45 nodules, 21 (46.7%) were cytologically diagnosed as FLUS, 16 (35.6%) were diagnosed as AUS, and eight (17.8%) were FN-SFN. For the 21 FLUS nodules, Afirma analyses were benign in 12 cases and suspicious in nine cases. For the 16 AUS nodules, Afirma analyses were benign in seven cases (Fig. 1), suspicious in eight cases (Fig. 2), and nondiagnostic in one case. This nondiagnostic case was diagnosed as malignant on a subsequent FNA biopsy and then confirmed as papillary carcinoma on thyroidectomy specimen histologic analysis. In eight FN-SFN, Afirma analyses were benign in four cases and suspicious in four cases.

Patients with indeterminate cytologic findings and suspicious Afirma results were recommended for surgery, and 10 of 21 (47.6%) completed the surgery. Of these, three of 21 (14.3%) nodules proved to be follicular variant papillary thyroid carcinoma, one (4.8%) was a classic papillary thyroid carcinoma, one (4.8%) was oncocyctic follicular variant papillary thyroid carcinoma, one (4.8%) was a follicular carcinoma, two (9.5%) were Hurthle cell adenomas, one (4.8%) was nodular hyperplasia, and one (4.8%) was adenomatous nodule with oncocyctic features.

Eleven patients did not complete surgery. Of these, five had committed to future surgery that has not yet been performed, four were lost to follow-up, and two declined the recommendation. Patients with a benign Afirma result were all recommended for follow-up, and none of them had undergone thyroid surgery at the end of the data analysis period.

Ultrasound features and Afirma Results are shown in Table 1. On preoperative sonography, 73% (32/44) of nodules were solid. Of these, 59% (19/32) were Afirma benign and 41% (13/32) were Afirma suspicious. Twenty-seven percent (12/44) of the nodules were complex nodules. Of these, 33% (4/12) were Afirma benign and 67% (8/12) were Afirma suspicious. This difference was not statistically significant (p = 0.179). Seven of eight complex cystic nodules with an Afirma suspicious diagnosis were more solid than cystic. Surgery revealed three cancers and two benign entities in these seven nodules. One nodule with an Afirma-suspicious diagnosis was predominantly cystic. This nodule was lost.

**TABLE 1: Ultrasound Features and Afirma (Veracyte) Results in 44 Thyroid Nodules With Indeterminate Fine-Needle Aspiration Cytologic Findings**

<table>
<thead>
<tr>
<th>Ultrasound Features</th>
<th>Total (n = 44)</th>
<th>Afirma Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign (n = 23)</td>
<td>Suspicious (n = 21)*</td>
</tr>
<tr>
<td>Nodule size (cm), mean ± SD (95% CI)</td>
<td>1.8 ± 0.8 (1.4–2.1)</td>
<td>2.2 ± 0.8 (1.8–2.6)</td>
</tr>
<tr>
<td>Component</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solid</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>Complex cystic &lt; 50%</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Complex cystic ≥ 50%</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Vascularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypervascular</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Hypovascular</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Calcium</td>
<td>33</td>
<td>21</td>
</tr>
<tr>
<td>Microcalcification</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Mixed calcification</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Vascularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypervascular</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>Hypovascular</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Calcium</td>
<td>38</td>
<td>21</td>
</tr>
<tr>
<td>Microcalcification</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>None</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Unknown</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Note—Except for nodule size, data are number of thyroid nodules. One nondiagnostic Afirma case was not included.

*Ten of 21 patients with suspicious Afirma results completed the surgery, and the final pathologic outcomes of the nodules are included in parentheses.

Ultrasound Versus Afirma Testing of FNA-Indeterminate Thyroid Nodules

Downloaded from www.ajronline.org by John Hanna on 11/09/15 from IP address 71.6.50.242. Copyright ARRS. For personal use only; all rights reserved
to follow-up. All nodules had an ovoid shape and circumscribed margin. No sonographic feature was associated with Afirma-benign or -suspicious status. Logistic regression with forward stepwise selection showed no predictive power for any combination of factors.

Discussion

According to the Bethesda system, the malignancy risk of nodules diagnosed as AUS-FLUS on FNA cytology is estimated as 5–15%, and repeat FNA is recommended [4]. However, marked variability in incidence and malignancy in resection specimens has been reported [13]. In a retrospective multicenter study, Alexander et al. [14] reported a malignancy rate of 44% in patients with cytologically indeterminate and suspicious Afirma nodules. Our 70.0% malignant rate in this subgroup was higher than that reported by Alexander et al. This can be explained by the approach used in our institution for selecting patients for Afirma testing. Most (33/45) of our patients underwent repeat FNA biopsy before Afirma testing. Because a repeat FNA biopsy can accurately reclassify more than 50% of nodules in the AUS-FLUS category as benign without the use of ancillary molecular tests [15–17], it is likely that our subject sample contained a higher proportion of malignant nodules than the subjects in the article by Alexander et al.

Although clinical and ultrasound features have been shown to be predictive of malignancy in thyroid nodules in general, the literature concerning their utility for the prediction of malignancy in patients with an indeterminate FNA diagnosis is mixed. Some authors have found that ultrasound appearance is helpful to further stratify indeterminate FNA cases into high and low risk for malignancy [12, 18–20]. In a study of 165 AUS-FLUS nodules, Jeong et al. [12] found that the malignancy rate was 79.3% for nodules with suspicious ultrasound features and 24.7% for nodules with indeterminate ultrasound features. Others have shown that sonographic features alone may be insufficient to predict malignancy in these patients [7, 9, 21, 22]. Mehta et al. [21] found that one or more suspicious ultrasound features was identified only in 33% of nodules and occurred regardless of histologic findings. Of the 21 nodules with suspicious Afirma results in our series, none showed suspicious sonographic findings such as taller-than-wide dimension or ill-defined margins, and only two of these nodules showed microcalcification. The reason for the lack of sonographic characteristics associated with malignancy in this group is uncertain. It is possible that nodules with indeterminate cytologic findings are less likely to show sonographic characteristics considered suspicious for malignancy; however, this is speculative, and our study is not powered to make this determination. Nonetheless, this finding highlights the lack of utility of sonographic features for malignancy risk stratification in thyroid nodules indeterminate on FNA biopsy.

Limitations

Our study has several limitations, including small sample size, limited pathologic follow-up, and retrospective study design. In addition, it is possible that there is an unrecognized selection bias in the subgroup of patients referred for Afirma analysis. At our institution, radiologists, endocrinologists, and endocrine surgeons perform approximately 2500 thyroid biopsies per year. The decision to proceed to Afirma testing is made by the referring physician and varies considerably among referring physicians. Typically, endocrinologists perform two FNA procedures before Afirma testing, whereas primary care physicians commonly perform Afirma testing after a single nondiagnostic aspirate. Most (33/45) of our subjects underwent two FNA biopsies of the indeterminate nodule before Afirma testing. We did not review all cytology reports to isolate the number of subjects with indeterminate FNA results, and we did not compare the characteristics of nodules in patients with indeterminate biopsy samples who were referred for Afirma testing with those who were not referred for Afirma testing.

Conclusion

Except for nodule size, sonographic characteristics are not useful in cases that had an indeterminate FNA finding before Afirma testing and cannot be used as a substitute for the Afirma gene expression classifier in such patients.

Acknowledgments

We thank Ronald N. Arpin III and Beverly Pizzi for their help in data collection.

References

1. Cooper DS, Doherty GM, Haugen BR, et al.; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid...
Ultrasound Versus Afirma Testing of FNA-Indeterminate Thyroid Nodules

cancer. Thyroid 2009; 19:1167–1214
12. Jeong SH, Hong HS, Lee EH, Cha JG, Park JS, Kwak JJ. Outcome of thyroid nodules characterized as atypia of undetermined significance or follicular lesion of undetermined significance and correlation with ultrasound features and BRAFV600E mutation analysis. AJR 2013; 201:[web]W854–W860
17. 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
19. Rosario PW. Thyroid nodules with atypia or follicular lesions of undetermined significance (Bethesda category III): importance of ultrasonography and cytological subcategory. Thyroid 2014; 24:1115–1120
20. Yoo WS, Choi HS, Cho SW, et al. The role of ultrasound findings in the management of thyroid nodules with atypia or follicular lesions of undetermined significance. Clin Endocrinol (Oxf) 2014; 80:735–742