

Diagnostic Value of Molecular Testing in Sonographically Suspicious Thyroid Nodules

Maxwell M. Wang,¹ Katrina Beckett,² Michael Douek,² Rinat Masamed,² Maitraya Patel,² Chi-Hong Tseng,³ Michael W. Yeh,⁴ Angela M. Leung,^{5,6} and Masha J. Livhits⁴

¹David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, California 90095;

²Department of Radiology, David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, California 90095; ³Department of Medicine, David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, California 90095; ⁴Section of Endocrine Surgery, David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, California 90095; ⁵Division of Endocrinology, Diabetes, and Metabolism, David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, California 90095; and ⁶Division of Endocrinology, Diabetes, and Metabolism; VA Greater Los Angeles Healthcare System, Los Angeles, California 90073

ORCID numbers: 0000-0003-3329-3615 (M. Patel); 0000-0001-8935-9332 (A. M. Leung); 0000-0002-8436-392X (M. J. Livhits).

Objective: Molecular testing can refine the diagnosis for the 20% of thyroid fine-needle aspiration biopsies that have indeterminate cytology. We assessed the diagnostic accuracy of molecular testing based on ultrasound risk classification.

Methods: This retrospective cohort study analyzed all thyroid nodules with indeterminate cytology at an academic US medical center (2012-2016). All indeterminate nodules underwent reflexive molecular testing with the Afirma Gene Expression Classifier (GEC). Radiologists performed blinded reviews to categorize each nodule according to the American Thyroid Association (ATA) ultrasound classification and the American College of Radiology Thyroid Imaging, Reporting and Data System. GEC results and diagnostic performance were compared across ultrasound risk categories.

Results: Of 297 nodules, histopathology confirmed malignancy in 65 (22%). Nodules by ATA classification were 8% high suspicion, 44% intermediate, and 48% low/very low suspicion. A suspicious GEC result was more likely in ATA high-suspicion nodules (81%) than in nodules of all other ATA categories (57%; $P = .04$). The positive predictive value (PPV) of GEC remained consistent across ultrasound categories (ATA high suspicion, 64% vs all other ATA categories, 48%; $P = .39$). The ATA high-suspicion category had higher specificity than a suspicious GEC result (93% vs 51%; $P < .01$). A suspicious GEC result did not increase specificity for the ATA high-suspicion category.

Conclusion: The PPV of molecular testing remained consistent across ultrasound risk categories. However, a suspicious GEC result was very likely in ATA high-suspicion nodules and did not improve specificity in this sonographic category.

© Endocrine Society 2020.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Key Words: molecular testing, ultrasound, ATA, ACR TI-RADS, thyroid, nodule

Abbreviations: ACR TI-RADS, American College of Radiology Thyroid Imaging, Reporting and Data System; ATA, American Thyroid Association; FNA, fine-needle aspiration; GEC, Afirma Gene Expression Classifier; GSC, Genomic Sequencing Classifier; NPV, negative predictive value; PPV, positive predictive value.

Received 24 April 2020

Accepted 12 June 2020

First Published Online 20 June 2020

Corrected and Typeset 21 August 2020

September 2020 | Vol. 4, Iss. 9

doi: 10.1210/endo/bvaa081 | Journal of the Endocrine Society | 1–10

Thyroid nodules are present in up to two-thirds of adults [1]. Most thyroid nodules are benign; only 5% to 10% are malignant and require intervention [2]. Given the high prevalence of nodules but low risk of cancer, accurate diagnostic modalities are essential to avoid unnecessary surgery. The American Thyroid Association (ATA) and the American College of Radiology Thyroid Imaging, Reporting and Data System (ACR TI-RADS) systems are commonly used ultrasound classification schemes to stratify the risk of malignancy in thyroid nodules [3, 4]. Fine-needle aspiration (FNA) with cytopathological analysis has been the cornerstone of malignancy workup for thyroid nodules. Using the Bethesda System, cytopathology yields a benign result in 70% of biopsied nodules, a malignant result in 5%, and an indeterminate result in 25% [1, 5]. Nodules with indeterminate cytology have an overall 25% risk of malignancy and are often referred for diagnostic surgery [1, 5-7].

Recent advances in molecular testing have further refined malignancy risk stratification for indeterminate nodules and allow more patients to avoid diagnostic surgery [8, 9]. The Afirma Gene Expression Classifier (GEC) analyzes the expression of 167 cancer-associated genes using messenger RNA microarray technology [9]. While GEC has a high negative predictive value (NPV) (> 90%), the positive predictive value (PPV) is 40% to 50% [9, 10]. Thus, although molecular tests like GEC can reliably “rule out” malignancy, positive results remain inconclusive. GEC was updated in 2017 to the current version, called Genomic Sequencing Classifier (GSC), with a slight improvement in specificity (from 52% to 68%) while maintaining its high sensitivity (from 92% to 91%) [10, 11].

The aim of this study was to compare the diagnostic accuracy of molecular testing across ultrasound risk categories for cytologically indeterminate thyroid nodules. We used GEC because of the longer duration of follow-up available compared to the more recently available GSC. We hypothesized that combining GEC results with ultrasound classification would improve diagnostic accuracy.

1. Materials and Methods

A. Study Population

This retrospective cohort study included all patients who underwent FNA of a thyroid nodule within the University of Los Angeles (UCLA) Health System from September 2012 through March 2016. All nodules with an indeterminate cytology result (Bethesda 3, atypia of undetermined significance or Bethesda 4, follicular neoplasm or suspicious for a follicular neoplasm) were reflexively sent for molecular testing via Afirma GEC in South San Francisco, California. Nodules with a suspicious GEC result were recommended for surgical resection, whereas nodules with a benign GEC result were generally observed. The study was approved by our institutional review board.

B. Ultrasound Imaging Technique

Thyroid ultrasound was performed using 1 of 2 units (iU22 or Epiq 7; Philips Ultrasound) by technologists licensed in the state of California with Registered Diagnostic Medical Sonographer certification in an ultrasound department accredited by the ACR. Grayscale and color Doppler images were obtained of the thyroid gland and cervical lymph nodes using a linear 12- to 5-MHz and a curvilinear 8- to 5-MHz transducer. When a nodule was identified during real-time scanning, the sonographer would obtain images in longitudinal and transverse planes demonstrating margins, size, composition, echogenicity, shape, presence of color Doppler flow, and ancillary features such as colloid or calcifications. If multiple nodules were present, the sonographer would image the largest nodules in each lobe and those with higher risk features. Images were electronically stored and reviewed in a picture archiving and communication system (Centricity, GE Healthcare).

C. Study Design

FNA was performed according to individual clinician assessment of nodule size and presence of suspicious sonographic features. In January 2016, the ATA published new management guidelines for thyroid nodules including ultrasound risk stratification and indications for FNA [3]. Four experienced radiologists in our institution's Department of Radiology (all board certified, with fellowship training; M.D., M.P., K.B., and R.M. have 11, 10, 7, and 6 years of experience following fellowship, respectively) were recruited to perform blinded, retrospective reads on the electronically stored ultrasound images. All nodules were randomly assigned to 2 of the 4 radiologists. Without knowledge of patient demographics, nodule outcome, or original radiology interpretation, the radiologists assigned each nodule to both an ATA and ACR TI-RADS classification.

ATA classification is organized into 5 categories: benign (< 1% malignancy risk), very low suspicion (< 3% malignancy risk), low suspicion (5%-10% malignancy risk), intermediate suspicion (10%-20% malignancy risk), and high suspicion (70%-90% malignancy risk) [3]. A nodule is classified into a suspicion category based on its pattern of sonographic findings, including echogenicity, margin regularity, presence of calcifications, shape, and presence of extraglandular extension (Fig. 1). Previous studies have demonstrated that 5.0% to 14.2% of all thyroid nodules are unclassifiable using ATA [12, 13]. ACR TI-RADS also has 5 categories: TR1: benign (< 2% malignancy risk), TR2: not suspicious (< 2% malignancy risk), TR3: mildly suspicious (5% malignancy risk), TR4: moderately suspicious (5%-20% malignancy risk), and TR5: highly suspicious (< 20% malignancy risk) [4, 14]. Points are given for nodule composition, echogenicity, shape, margin, and presence of echogenic foci; the final category assigned is based on the total points.

For each nodule, the 2 radiologists' ATA classifications were assessed for concordance. If the classifications were not concordant, a third radiologist was assigned to perform a tie-breaking read. The ATA low- and very low-suspicion classifications were combined in the analyses given the small number of very low-suspicion nodules and the similarities in nodule characteristics within those ATA classifications. Similarly, the TR2 and TR1 classifications were combined in the ACR TI-RADS analyses.

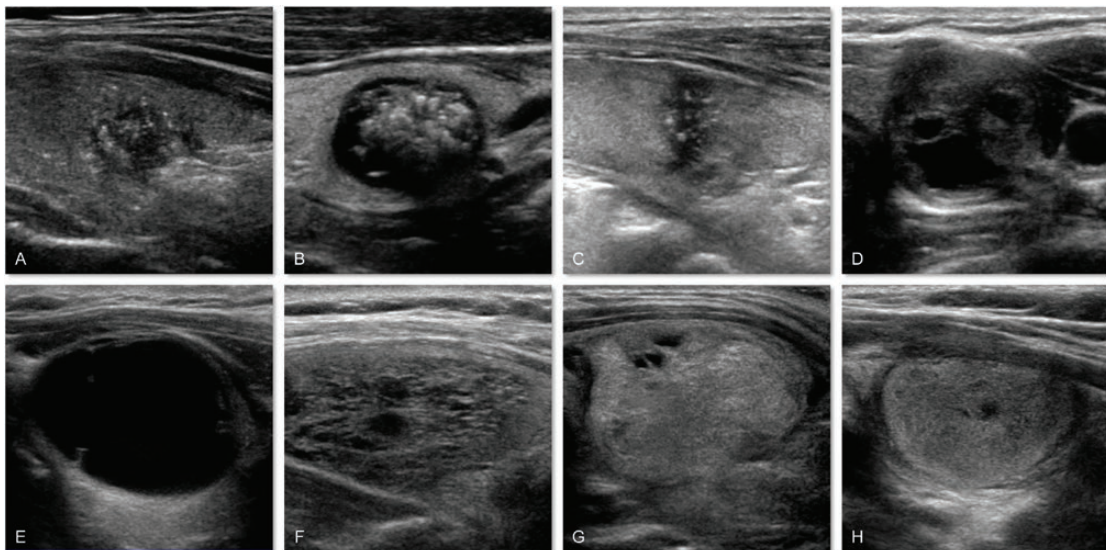


Figure 1. Representative thyroid nodule images from the study cohort exhibiting ultrasound features that are more suspicious vs less suspicious for malignancy. More suspicious features include A, irregular margins; B, microcalcifications presenting as punctate echogenic foci; C, taller-than-wide shape on transverse view; and D, extrathyroidal extension. Less suspicious features include E, cystic-appearing composition; F, spongiform composition; G, hyperechogenicity; and H, smooth margins.

“Index nodule” refers to the nodule that underwent initial FNA with a cytologically indeterminate result. We included only nodules that had a real-time ultrasound performed to guide the FNA biopsy to ensure that the retrospectively reviewed nodule was the same as the originally biopsied nodule. We rigorously analyzed surgical pathology reports for the size and location (ie, thyroid quadrant or isthmus) of nodules within the surgical specimen. Nodules with any uncertainty or discrepancy when comparing the reports were excluded. Thyroid microcarcinomas (< 1 cm) were considered malignant only if located in the same quadrant as the biopsied nodule.

Nodules with a benign GEC result that were managed nonoperatively were considered benign if a surveillance ultrasound performed more than 1 year after the initial FNA demonstrated stable nodule size and appearance. We consider this a robust threshold for benignity because of the well-established low false-negative rate of GEC (7%) [10]. Nodules with a benign GEC result that were managed nonoperatively but had no ultrasound surveillance and nodules with a suspicious GEC result that did not undergo resection were excluded from analysis of performance characteristics for the diagnostic modalities.

D. Statistical Analysis

Descriptive statistics (mean, SD, median, interquartile range, and frequency distribution) were generated for patient demographic and baseline clinical information. Sensitivity, specificity, PPV, and NPV were calculated for each diagnostic modality (GEC, ATA, and ACR TI-RADS ultrasound risk stratification) and combinations of modalities with 95% Wilson CIs (95% CIs). We used chi-square or Fisher exact tests to compare proportions between independent groups. Because multiple tests were performed on each nodule, generalized estimating equation models were used to compare the performance among tests, accounting for clustering of tests within nodules.

All analyses were carried out using R software (www.r-project.org) and *P* values less than .05 were considered statistically significant. When analyzing the diagnostic performance of ATA, the high-suspicion category was considered a positive test for malignancy; this cutoff has been used by previous studies [15]. Similarly, the TR5 category was considered a positive test for malignancy with ACR TI-RADS to maintain consistency with the ATA analysis.

2. Results

A. Patients and Afirma Gene Expression Classifier Results

At our institution from 2012 to 2016, a total of 398 biopsied thyroid nodules returned with indeterminate cytopathology (ie, Bethesda 3 or 4) (Fig. 2). However, 77 of these nodules were excluded from this study because ultrasound images were not uploaded to the patients' electronic medical records. Another 24 nodules were excluded because the sample sent for Afirma GEC was insufficient. Thus, the final study cohort comprised 297 nodules (mean size, 2.4 cm) from 281 patients (81% female; mean age, 51 years) (Fig. 3). Of these nodules, 60% (179) had a suspicious result and 40% (118) had a benign result (Fig. 4). Of the 179 GEC suspicious nodules, 146 nodules (82%) were surgically resected with a 44% malignancy rate based on histopathology: 86% papillary thyroid carcinoma, 8% Hürthle cell carcinoma, and 6% follicular carcinoma. The remaining 33 (18%) nodules were not resected because the patient declined surgery for 14 (42%) nodules, the presence of significant medical comorbidities for 12 (36%) nodules, and lost to follow-up for 7 (21%) nodules. Of these 33 unresected nodules, 16 (48%) nodules had a surveillance ultrasound more than 1 year after GEC testing demonstrating stability in size and sonographic appearance. The remaining 17 (52%) nodules were lost to follow-up.

Of the 118 GEC benign nodules, 34 nodules (29%) were surgically resected with 1 malignant nodule (follicular carcinoma) found on histopathology. The indications for resection in order of descending frequency were clinician or patient preference for definitive

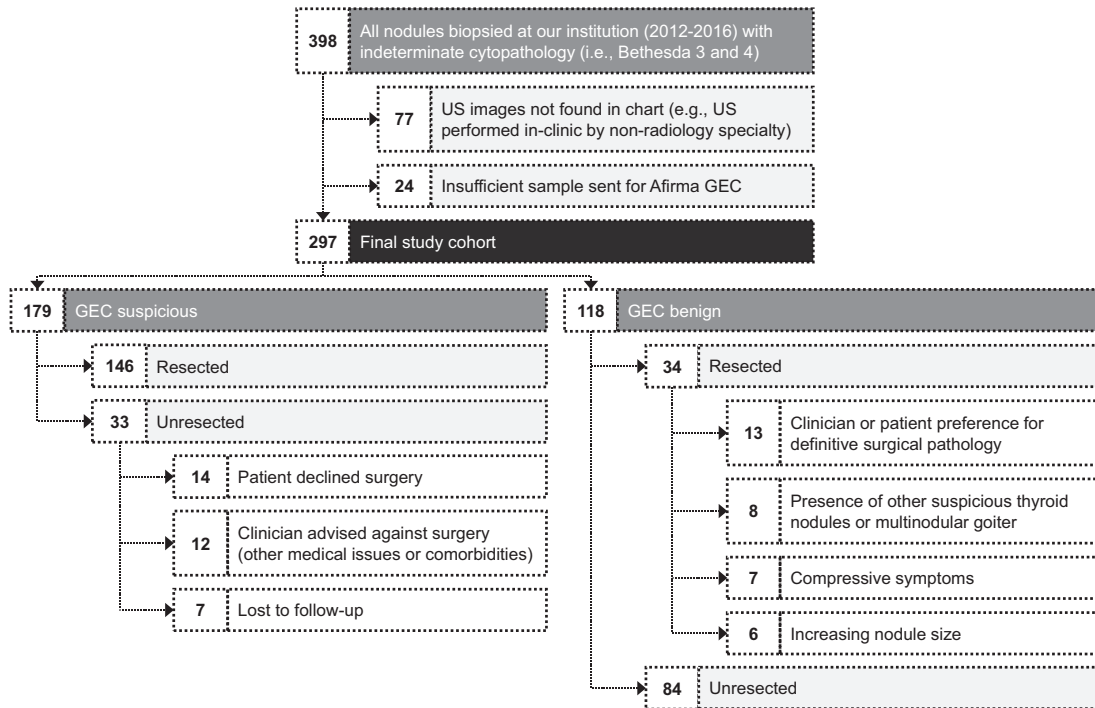


Figure 2. Flow diagram illustrating the pathway to the final study cohort (297 nodules).

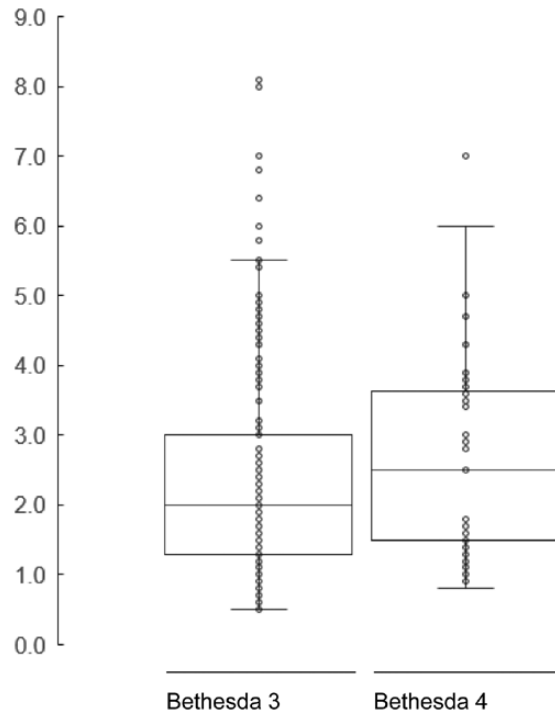


Figure 3. Nodule sizes (in centimeters), by Bethesda cytopathologic classification. Each dot represents one nodule. Box and whisker plots indicate median and quartiles.

surgical pathology for 13 (38%) nodules, presence of other suspicious thyroid nodules or multinodular goiter warranting surgery for 8 (24%) nodules, compressive symptoms for 7 (21%) nodules, and increasing size for 6 (18%) nodules.

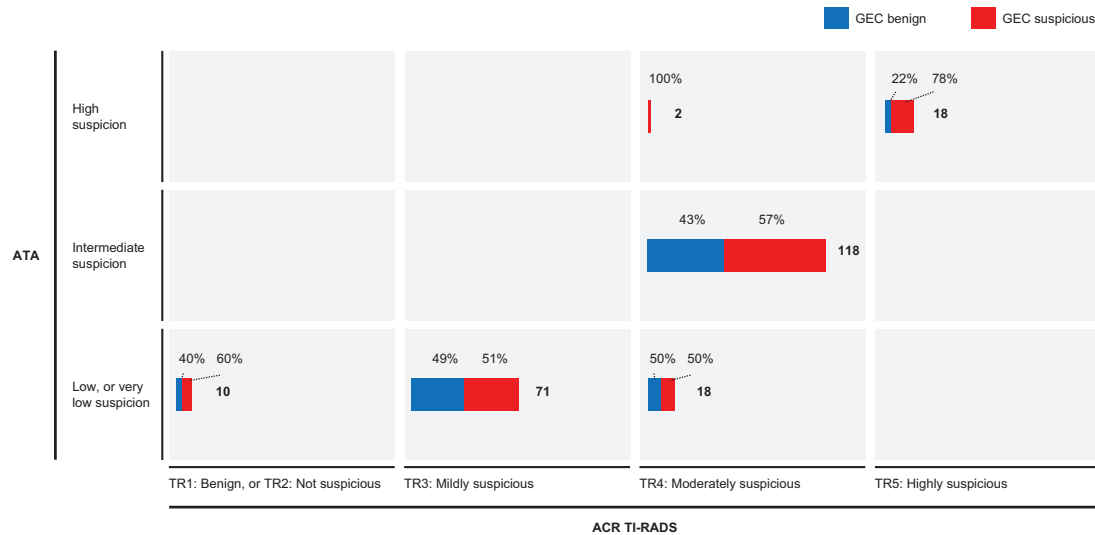


Figure 4. Intersection of nodule ultrasound classifications and Afirma GEC result. Bold numbers represent the total number of nodules in each cell. A total of 237 nodules were included in this analysis (excludes the 60 nodules lacking a concordant ATA and/ or ACR-TIRADS classification). ACR TI-RADS, American College of Radiology Thyroid Imaging, Reporting and Data System; ATA, American Thyroid Association ultrasound classification system; GEC, Afirma Gene Expression Classifier.

An additional 39 unresected GEC benign nodules were deemed benign and included in the analysis because of the stability during follow-up ultrasound surveillance. GEC had high sensitivity (98% [95% CI, 92%-100%]) and NPV (99% [95% CI, 92%-100%]) and lower specificity (51% [95% CI, 42%-59%]) and PPV (49% [95% CI, 40%-58%]).

B. American Thyroid Association and American College of Radiology Thyroid Imaging, Reporting and Data System Ultrasound Classification

The rate of concordance in ATA ultrasound risk classification among radiologists was 90%. A total of 14 nodules (4.7%) received an “ATA unclassifiable” read from one of the radiologists. Of the 268 nodules with concordant ATA classifications, 21 (8%) were ATA high suspicion, 118 (44%) were intermediate suspicion, and 129 (48%) were low or very low suspicion. Malignancy rates as stratified by ATA classification were 53% in the high-suspicion, 25% in the intermediate-suspicion, and 34% in the low- or very low-suspicion categories. The ATA classification showed a sensitivity of 16% (95% CI, 7%-27%), NPV of 70% (95% CI, 62%-77%), specificity of 93% (95% CI, 88%-97%), and PPV of 53% (95% CI, 28%-77%).

The rate of concordance in ACR TI-RADS malignancy risk classification was 87%. Of the 258 nodules with concordant ACR TI-RADS classifications, 19 (7%) were TR5, 153 (59%) TR4, 73 (28%) TR3, and 13 (5%) TR2 or TR1. Malignancy rates as stratified by ACR TI-RADS classification were 50% in the TR5, 24% in the TR4, 35% in the TR3, and 33% in TR2 or TR1 categories. The ACR TI-RADS classification showed a sensitivity of 14% (95% CI, 6%-26%), NPV of 72% (95% CI, 65%-79%), specificity of 94% (95% CI, 88%-98%), and PPV of 50% (95% CI, 23%-77%).

C. Intersection Between Afirma Gene Expression Classifier Results and Ultrasound Classification

Nodules in the ATA ultrasound high-suspicious category were more likely to have a suspicious GEC result compared to nodules in the remaining ATA categories (81% vs 57%, $P = .04$) (Fig. 5). A similar trend was observed among nodules in the ACR TI-RADS TR5 category compared to nodules in the remaining ACR TI-RADS categories (74% vs 56%, $P = .16$).

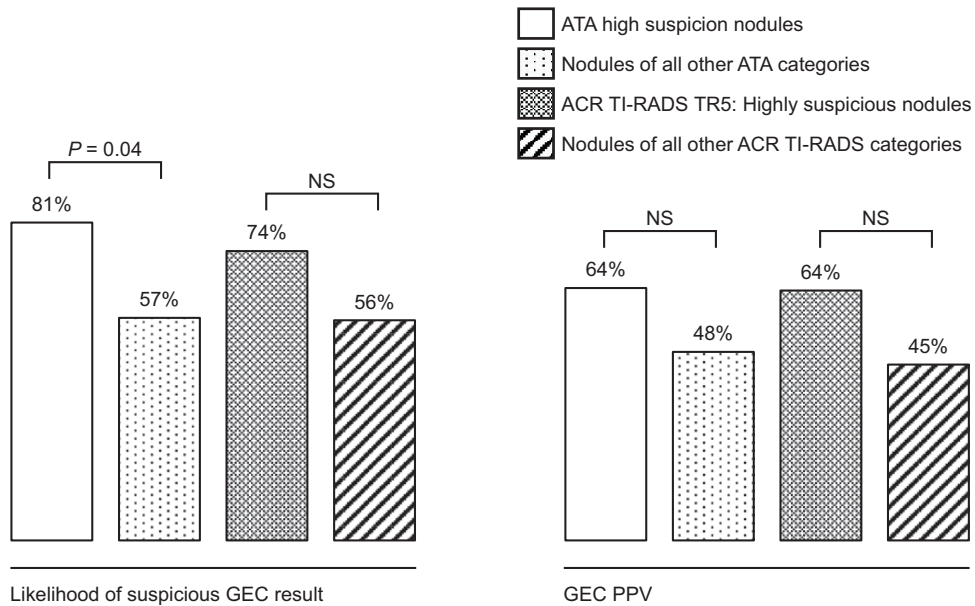


Figure 5. Intersection of ultrasound and GEC. Left: Overlap in ultrasound and GEC results, reflected as the proportion of nodules in each ultrasound group that also had a suspicious GEC result. Right: GEC's PPV stratified by ultrasound classification. ACR TI-RADS, American College of Radiology Thyroid Imaging, Reporting and Data System; ATA, American Thyroid Association ultrasound classification system; GEC, Afirma Gene Expression Classifier; NS, nonsignificant ($P > .05$); PPV, positive predictive value.

The PPV of GEC was 64% among ATA high-suspicion nodules and 48% among nodules in the remaining ATA categories ($P = .39$). The PPV of GEC was also similar comparing across ACR TI-RADS categories (TR5, 64% vs all remaining categories, 45%; $P = .34$).

GEC had higher sensitivity ($P < .01$) and NPV ($P < .01$) compared to ATA and ACR TI-RADS, whereas ATA and ACR TI-RADS had higher specificity ($P < .01$) compared to GEC. The specificity and PPV of the ATA high-suspicion category were 93% and 53%, respectively. Adding a concurrent suspicious GEC result did not significantly increase the specificity (96%, $P = .40$) or PPV (64%, $P = .53$). Similarly, the specificity and PPV of the ACR TI-RADS TR5 category was 94% and 50%, respectively. Adding a concurrent suspicious GEC result did not increase the specificity (97%, $P = .36$) or PPV (64%, $P = .50$).

Although the presence of high-suspicion ultrasound features (fully solid composition, hypoechoic echogenicity, taller-than-wide shape, irregular margin, presence of extrathyroidal extension, and presence of microcalcifications) generally increased the likelihood of a suspicious GEC result, the association was not statistically significant for any of these features on an independent basis (Fig. 6).

3. Discussion

In our analysis of the synergistic accuracy of ultrasound and molecular testing results for cytologically indeterminate thyroid nodules, the diagnostic performance of GEC remained consistent across ultrasound risk categories. However, nodules in the high-suspicion ultrasound category returned a suspicious GEC result in 81% of cases—significantly more likely than nodules in the lower-suspicion ultrasound categories. The cost-effectiveness and overall utility of molecular testing for nodules classified as high suspicion based on ultrasound may be limited.

Despite multiple modalities now available for diagnosis of thyroid nodules, they are generally used in a linear, sequential manner. Recent studies have explored combining ultrasound findings with cytological features to improve diagnostic accuracy [16–18]. Ultrasound

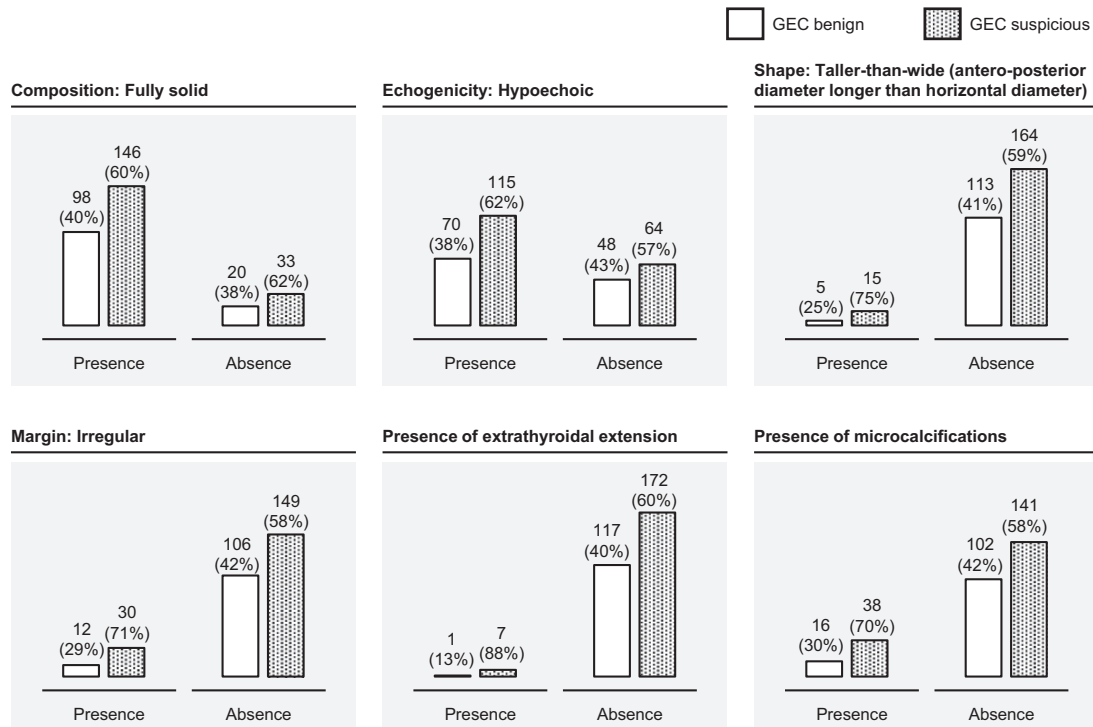


Figure 6. Likelihood of an Afirma Gene Expression Classifier (GEC) suspicious result for nodules with suspicious sonographic features. “Presence” indicates that one or more of the radiologists evaluating the nodule noted the presence of the feature. Although there are positive trends between the presence of these suspicious features and increased likelihood of a suspicious result on GEC, these associations were not statistically significant using the Fisher exact test.

characteristics have also been combined with limited mutation analysis (eg, *BRAF*^{V600E} or *NRAS*) to improve diagnostic accuracy for papillary thyroid carcinoma [16, 18]. However, no previous study has investigated the concordant and cross-disciplinary use of ultrasound findings based on 2 leading standardized classifications systems for thyroid nodules, with a clinically used, multigene molecular test panel.

Our findings indicate that the PPV of GEC remains consistent across ultrasound risk categories. However, molecular tests like GEC may have limited clinical utility for nodules with highly suspicious sonographic characteristics. A suspicious GEC result was very likely for nodules in the highest-risk ultrasound categories and did not improve the specificity or PPV over ultrasound alone. In other words, the incremental diagnostic value provided by GEC for highly sonographically suspicious nodules was minimal. We therefore posit that molecular testing may be cost-effective for nodules with low- or intermediate- but not high-risk ultrasound characteristics [19].

Although thyroid ultrasound is highly operator and radiologist dependent, studies have demonstrated robust specificity and PPV for malignancy, particularly with the highest-suspicion ultrasound categories [3, 15]. However, because only 14% to 18% of thyroid nodules are classified in the highest-suspicion ATA or ACR TI-RADS categories based on sonographic characteristics, ultrasound alone is insufficiently sensitive to rule out malignancy [14, 20]. ATA classification is highly reader dependent because the radiologist must recognize specific patterns of features that raise suspicion for malignancy. In contrast, the ACR TI-RADS methodology focuses more on inventorying and quantifying the number of suspicious features, and less on evaluating co-occurring feature patterns [4]. Thus, its methodology may be considered less reader dependent than ATA. That the presence of individual high-suspicion ultrasound features cannot predict a suspicious GEC result with statistical significance in our study further corroborates the diagnostic value of

sonographic classification systems (like ATA and ACR TI-RADS) that integrate multiple sonographic features for risk stratification. While specific features may presumably be more concerning than others in predicting malignancy (such as extrathyroidal extension or irregular margins), they may occur less frequently and be less predictive than considering the entirety of a nodule's sonographic profile.

Despite their methodological differences, the performance of ATA and ACR TI-RADS was similar in our study. Given the lower expected malignancy risk in the ACR TI-RADS TR5 compared to the ATA high-suspicion category, we would expect more nodules to be classified in the former compared to in the latter group. However, a recent study of more than 3000 nodules classified 16% and 18% of the nodules into the ACR TI-RADS TR5 and ATA high-suspicion categories, respectively [14]. The similar number of nodules classified in the 2 highest-suspicion categories suggests the diagnostic performance of the 2 systems may be more similar than previously established. In addition, our radiologists simultaneously assigned each nodule an ATA and ACR TI-RADS classification, which may increase the likelihood of convergent reads. Our study also exclusively evaluated cytologically indeterminate nodules, which may have contributed to the observed similarities between ATA and ACR TI-RADS performance characteristics. The high concordance rates for ATA and ACR TI-RADS observed in our study were likely due to the tie-breaking method we used. Before a third tie-breaking radiologist was introduced for nonconcordant reads, the ATA concordance rate was 64%. Only after tie-breaking reads were performed did the rate increase to 90%.

Our study's strengths include reflexive rather than selective molecular testing performed for all cytologically indeterminate nodules. Furthermore, our institution has a centralized head and neck cytopathology service, thus interpretations of nodule FNAs follow rigorous, uniform criteria. Our study is limited by its retrospective, single-institution design. Sonographic interpretations are radiologist specific and dependent on radiologist training and experience. This may limit generalizability to other institutions. We used GEC because of the longer duration of follow-up available compared to the currently available GSC. Although this is a limitation of our study, the results can likely be extrapolated to GSC given the relatively small differences in benign call rate and specificity between the test versions. Finally, the relatively low sample sizes in the highest-suspicion ultrasound categories reflect our radiologists' rigor in evaluating nodules for malignancy; however, this may diminish statistical power.

Cytologically indeterminate thyroid nodules continue to be a significant burden for patients and providers. However, opportunities exist to optimize the workup of these nodules by resourcefully wielding the various diagnostic modalities currently available. Our study provides insight into how ultrasound and molecular testing may be used synergistically in this process.

Acknowledgments

We thank Dalena Nguyen for providing logistical support during manuscript preparation

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

Additional Information

Correspondence: Masha J. Livhits, MD, Section of Endocrine Surgery, UCLA David Geffen School of Medicine, 10833 Le Conte Avenue, Suite 72-228 CHS, Los Angeles, CA 90095. E-mail: mlivhits@mednet.ucla.edu.

Disclosure Summary: The authors have nothing to disclose.

Data Availability: All data generated or analyzed during this study are included in this published article or in the data repositories listed in References.

References and Notes

1. Popoveniuc G, Jonklaas J. Thyroid nodules. *Med Clin North Am.* 2012;**96**(2):329-349.
2. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: a review. *JAMA.* 2018;**319**(9):914-924.
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):1-133.
4. Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): white paper of the ACR TI-RADS Committee. *J Am Coll Radiol.* 2017;**14**(5):587-595.
5. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid.* 2017;**27**(11):1341-1346.
6. Valderrabano P, Khazai L, Thompson ZJ, et al. Cancer risk stratification of indeterminate thyroid nodules: a cytological approach. *Thyroid.* 2017;**27**(10):1277-1284.
7. Wang CC, Friedman L, Kennedy GC, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid.* 2011;**21**(3):243-251.
8. Nikiforov YE. Molecular diagnostics of thyroid tumors. *Arch Pathol Lab Med.* 2011;**135**(5):569-577.
9. Zhang M, Lin O. Molecular testing of thyroid nodules: a review of current available tests for fine-needle aspiration specimens. *Arch Pathol Lab Med.* 2016;**140**(12):1338-1344.
10. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med.* 2012;**367**(8):705-715.
11. Patel KN, Angell TE, Babiarz J, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg.* 2018;**153**(9):817-824.
12. Lauria Pantano A, Maddaloni E, Briganti SI, et al. Differences between ATA, AACE/ACE/AME and ACR TI-RADS ultrasound classifications performance in identifying cytological high-risk thyroid nodules. *Eur J Endocrinol.* 2018;**178**(6):595-603.
13. Randolph G. *Surgery of the Thyroid and Parathyroid Glands. 3rd ed.* Elsevier Health Sciences; 2020.
14. Middleton WD, Teefey SA, Reading CC, et al. Multiinstitutional analysis of thyroid nodule risk stratification using the American College of Radiology Thyroid Imaging Reporting and Data System. *AJR Am J Roentgenol.* 2017;**208**(6):1331-1341.
15. Trimboli P, Fulciniti F, Zilioli V, Ceriani L, Giovanella L. Accuracy of international ultrasound risk stratification systems in thyroid lesions cytologically classified as indeterminate. *Diagn Cytopathol.* 2017;**45**(2):113-117.
16. De Napoli L, Bakkar S, Ambrosini CE, et al. Indeterminate single thyroid nodule: synergistic impact of mutational markers and sonographic features in triaging patients to appropriate surgery. *Thyroid.* 2016;**26**(3):390-394.
17. Khalil AB, Dina R, Meeran K, et al. Indeterminate thyroid nodules: a pragmatic approach. *Eur Thyroid J.* 2018;**7**(1):39-43.
18. Moon WJ, Choi N, Choi JW, Kim SK, Hwang TS. *BRAF* mutation analysis and sonography as adjuncts to fine-needle aspiration cytology of papillary thyroid carcinoma: their relationships and roles. *AJR Am J Roentgenol.* 2012;**198**(3):668-674.
19. Zanicco KA, Wang MM, Yeh MW, Livhits MJ. Selective use of molecular testing based on sonographic features of cytologically indeterminate thyroid nodules: a decision analysis. *World J Surg.* 2020;**44**(2):393-401.
20. Sahli ZT, Karipineni F, Hang JF, et al. The association between the ultrasonography TIRADS classification system and surgical pathology among indeterminate thyroid nodules. *Surgery.* 2019;**165**(1):69-74.