

# Association of Malignancy Prevalence With Test Properties and Performance of the Gene Expression Classifier in Indeterminate Thyroid Nodules

Zaid Al-Qurayshi, MD, MPH; Ahmed Deniwar, MD; Tina Thethi, MD; Tilak Mallik, MD; Sudesh Srivastav, PhD; Fadi Murad, MD; Parisha Bhatia, MD; Krzysztof Moroz, MD; Andrew B. Sholl, MD; Emad Kandil, MD, MBA

## + Supplemental content

**IMPORTANCE** It is crucial for clinicians to know the malignancy prevalence within each indeterminate cytologic category to estimate the performance of the gene expression classifier (GEC).

**OBJECTIVE** To examine the variability in the performance of the GEC.

**DESIGN, SETTING, AND PARTICIPANTS** This retrospective cohort study of patients with Bethesda category III and IV thyroid nodules used single-institution data from January 1, 2013, through February 29, 2016. Expected negative predictive value (NPV) was calculated by adopting published sensitivity and specificity. Observed NPV was calculated based on the true-negative rate. Outcomes were compared with pooled data from 11 studies published January 1, 2010, to January 31, 2016.

**RESULTS** A total of 145 patients with 154 thyroid nodules were included in the study (mean [SD] age, 56.0 [16.2] years; 106 females [73.1%]). Malignancy prevalence was 45%. On the basis of this prevalence, the expected NPV is 85% and the observed NPV is 69%. If the prevalence is assumed to be 25%, the expected NPV would be 94%, whereas the observed NPV would be 85%. Pooled data analysis of 11 studies comprising 1303 participants revealed a malignancy prevalence of 31% (95% CI, 29%-34%) and a pooled NPV of 92% (95% CI, 87%-96%).

**CONCLUSIONS AND RELEVANCE** In this study, variability in the performance of the GEC was not solely a function of malignancy prevalence and may have been attributable to intrinsic variability of the test sensitivity and specificity. The utility of the GEC in practice is elusive because of this variability. A better definition of the GEC's intrinsic properties is needed.

*JAMA Otolaryngol Head Neck Surg.* 2017;143(4):403-408. doi:10.1001/jamaoto.2016.3526  
Published online December 15, 2016.

**Author Affiliations:** Department of Otolaryngology-Head & Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City (Al-Qurayshi); Department of Pediatrics, Tulane University School of Medicine, New Orleans, Louisiana (Deniwar); Division of Endocrinology and Metabolism, Department of Medicine, Tulane University School of Medicine, New Orleans, Louisiana (Thethi, Mallik); Department of Biostatistics and Bioinformatics, Tulane University School of Public Health and Tropical Medicine, New Orleans, Louisiana (Srivastav); Division of Endocrine and Oncological Surgery, Department of Surgery, Tulane University School of Medicine, New Orleans, Louisiana (Murad, Bhatia, Kandil); Department of Pathology, Tulane University School of Medicine, New Orleans, Louisiana (Moroz, Sholl).

**Corresponding Author:** Emad Kandil, MD, MBA, Division of Endocrine and Oncological Surgery, Department of Surgery, Tulane University School of Medicine, 1430 Tulane Ave, New Orleans, LA 70112 (ekandil@tulane.edu).

Fine-needle aspiration (FNA) biopsy is considered the criterion standard in the initial evaluation of suspicious thyroid nodules.<sup>1-3</sup> Approximately 72% of thyroid nodules are found to be benign on FNA cytologic testing, whereas 5% to 15% are malignant.<sup>4,5</sup> The remaining nodules (10%-30%) are labeled indeterminate and are categorized into atypia of undetermined significance or follicular lesion of undetermined significance, suspicious for follicular neoplasm or Hürthle cell neoplasm, and suspicious for malignant tumors.<sup>6</sup> Of these indeterminate nodules, only 15% to 35% prove to be malignant on subsequent histologic examination, most commonly papillary thyroid carcinoma or follicular carcinoma.<sup>4,6</sup> A second FNA biopsy and/or diagnostic thyroid lobectomy has been recommended in the setting of indeterminate thyroid cytologic findings.<sup>7</sup>

Various molecular markers have been found to improve the assessment of thyroid nodules preoperatively.<sup>8</sup> The Afirma gene expression classifier (GEC) was developed by Veracyte to enhance the detection of benign nodules in the setting of indeterminate cytologic findings. The GEC measures the expression of 167 genes: 142 genes in the main classifier and 25 for medullary carcinoma and nonthyroid neoplasms. The test classifies indeterminate nodules into benign or suspicious categories. Prospective multicenter validation studies<sup>9-11</sup> reported a negative predictive value (NPV) of the test at 94% to 96% for indeterminate nodules.

The American Thyroid Association's 2015 statement on surgical application of molecular profiling for thyroid nodules concluded that it is crucial for clinicians to know the malignancy prevalence within each indeterminate cytologic category to

estimate the performance of the GEC.<sup>12</sup> In this study, we aimed to examine the effect of malignancy prevalence and test properties on the performance of the GEC and compare outcomes with pooled data from multiple studies.

## Methods

### Data and Specimen Collection

After approval by the Tulane University Medical School institutional review board, we conducted a retrospective study at Tulane Medical Center, New Orleans, Louisiana. We reviewed medical records of patients who underwent FNA biopsy from January 1, 2013, through February 29, 2016, and who had an available diagnostic Afirma GEC result. Different clinicians (T.T., T.M., E.K.) performed the FNA biopsies. At the time of the biopsy, we offered access to and discussed the utility of the Afirma GEC with all adult patients with ultrasonographically confirmed thyroid nodules 1 cm or larger at maximum diameter. It was explained to the patients that if Tulane cytopathologists categorized the specimen as indeterminate, the sample would undergo Afirma GEC testing. All patients then provided written informed consent. All data were deidentified.

Cellular adequacy of FNA biopsy specimens was confirmed on site. Additional passes were collected for the Afirma GEC and washed into the collection tube with the nucleic acid preserving solution as recommended by Veracyte's protocol. Samples for the Afirma GEC were stored refrigerated at 4°C at our institution and were then shipped to Veracyte for Afirma GEC testing at the discretion of our cytopathologists in the setting of indeterminate cytologic test results. The samples were shipped in cold containers provided by the manufacturer. The FNA biopsy specimens were evaluated by 2 specialty board-certified academic cytopathologists (K.M., A.B.S.) at Tulane Medical Center. Cytologic test results were reported according to the Bethesda System for Reporting Thyroid Cytopathology.<sup>6</sup> If the cytopathologists identified Bethesda category III or IV disease, the specimen was sent for Afirma GEC testing. Tulane Medical Center is designated as an enabled center by Veracyte and therefore not required to send the specimens for cytopathologic examination to Thyroid Cytopathology Partners, the group chosen by Veracyte for concurrent cytologic test result interpretation before the Afirma GEC testing. Active surveillance vs surgical intervention was discussed with all patients. Patients with benign Afirma GEC results who did not have other indications for surgical intervention were offered conservative follow-up with outpatient visits and ultrasonographic examination with a frequency range of every 6 to 12 months. However, some of those patients developed indications for surgery during their follow-up. All operations were performed by the endocrine surgeon (E.K.); resected nodules were matched to biopsied nodules according to size and location.

### Validity Analysis

Validity of the Afirma GEC test result compared with the final histopathologic test result was assessed by calculating sensi-

## Key Points

**Questions** What is the validity of the gene expression classifier, and is there variability in test performance among different institutes, and, if so, what is the cause of this variability?

**Findings** In this study, the negative predictive value of the gene expression classifier was lower than expected. The variability in test performance was not associated with cancer prevalence alone and may have been associated with inconsistencies in the intrinsic properties of the test.

**Meaning** The variability in the test performance limits its utility in practice; large clinical trials are warranted to better define the test's intrinsic properties.

tivity, specificity, NPV, positive predictive value (PPV), and accuracy. The NPV is a function of test properties (ie, sensitivity and specificity) and population malignancy prevalence. Expected NPV was calculated according to the Bayes theorem by applying the prevalence found in the current study and adopting the sensitivity and specificity reported by Alexander et al.<sup>9</sup> Observed NPV was calculated based on the raw frequency of true-negative and false-negative results found in the current study. We hypothesized that if the test properties are stable, no difference should exist between the expected and observed NPVs.

### Pooled Analysis

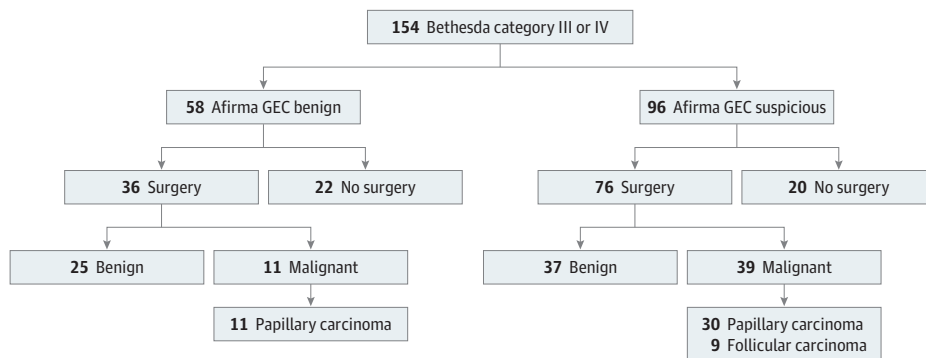
To further evaluate the variability of the Afirma GEC's performance, a pooled data analysis was performed. A PubMed search engine was used by applying the search term *afirma* to search for studies that were published from January 1, 2010, through January 31, 2016. Inclusion criteria were (1) English language, (2) original article, (3) studies that have considered Bethesda categories III and IV, and (4) studies that have published the frequencies of true-positive, true-negative, false-positive, and false-negative values. These frequencies were pooled together, and a pooled analysis of sensitivity, specificity, NPV, PPV, and malignancy prevalence was performed. Pooled data analysis allows a more precise estimation of these measures by increasing the sample size.

## Results

### Study Population

A total of 145 patients with 154 indeterminate FNA biopsy results were included in the study (mean [SD] age, 56.0 [16.2] years; 106 females [73.1%]) (Figure 1). Of those, 104 patients underwent surgery with 112 nodules excised. The characteristics of the study population are given in the Table. A total of 114 nodules (74.0%) were classified as Bethesda category III, whereas the remaining 40 nodules (26.0%) were classified as Bethesda category IV on cytopathologic evaluation. Of the 114 Bethesda category III nodules, 66 (57.9%) were categorized as suspicious on Afirma GEC test results; of the 40 Bethesda category IV nodules, 30 (75.0%) were categorized as suspicious

**Figure 1. Management of Indeterminate Thyroid Nodules, Bethesda Categories III and V, at Tulane Medical Center**



The Bethesda criteria categories are explained in Cibas and Ali.<sup>6</sup> GEC indicates gene expression classifier.

on Afirma GEC test results. Seventy-six patients (79.2%) categorized as having suspicious Afirma GEC results underwent surgery, whereas 36 patients (62.1%) categorized as having benign Afirma GEC results underwent surgery. Median time between benign Afirma GEC result and surgery was 30 days (interquartile range, 20-63 days). Patients who underwent surgery had 1 or more of the following indications: compressive symptoms, increasing size of nodule and suspicious ultrasonographic features, and/or another thyroid nodule that proved to be papillary thyroid carcinoma on FNA biopsy performed at the same time. The histopathologic test result of suspicious Afirma GEC nodules was benign in 37 nodules (48.7%) and malignant in 39 nodules (51.3%). The types of malignant tumors found on histopathologic analysis of suspicious Afirma GEC nodules were classic papillary thyroid carcinoma in 30 nodules (76.9%) and follicular thyroid carcinoma in 9 nodules (23.1%); all were stages T1a-1bNOMx, except for 1 nodule, which had stage T2N1aMx papillary thyroid carcinoma. On the other hand, the types of malignant tumors found on histopathologic analysis of benign Afirma GEC nodules were all classic papillary thyroid carcinoma, of which 8 nodules had a stage of T1a-1bNOMx, 2 had a stage of T3NOMx, and 1 had a stage of T3N1aMx.

### Validity Analysis Outcomes

The validity measures of the Afirma GEC based on the current literature data are described in eTable 1 in the Supplement. The observed sensitivity was 78%, specificity was 40%, PPV was 51%, and NPV was 69%. The malignancy prevalence in the thyroid nodules with Bethesda categories III and IV was 45%. On the basis of this prevalence and by adopting a sensitivity of 90% and a specificity of 52%, as published by Alexander et al,<sup>9</sup> the expected NPV was 85%. Figure 2 shows the difference between the observed and expected NPV by plotting the NPV curves based on our data and that of Alexander et al<sup>9</sup> in addition to curves reported by other studies.<sup>9,13-22</sup> Under the assumption of a 25% and 10% malignancy prevalence, the expected NPV was estimated to be 94% for a 25% prevalence and 98% for a 10% prevalence, whereas the observed NPV would have been 85% for a 25% prevalence and 94% for a 10% prevalence. Furthermore, if we adopted sensi-

ivities and specificities reported by other studies,<sup>9,13-22</sup> the expected NPV for a prevalence of 45% would have ranged from 74% to 100% (Figure 2).

### Pooled Analysis Outcomes

A PubMed search revealed 29 pertinent studies, 11 of which met the described inclusion criteria (eTable 2 in the Supplement).<sup>9,13-22</sup> The included studies have a range of sensitivity of 83% to 100%, specificity of 7% to 60%, PPV of 14% to 44%, and NPV of 75% to 100%. The prevalence range was 13% to 51%. Pooling these studies together with our own data generated a sample size of 1303 nodules. Pooled data analysis revealed a sensitivity of 93% (95% CI, 91%-96%), a specificity of 36% (95% CI, 33%-40%), and a malignancy prevalence of 31% (95% CI, 29%-34%). Figure 3 shows the plot of the pooled NPV curve and pooled prevalence with their respective 95% CIs. Figure 4 is a magnification of the intersection between the pooled NPV curve and the pooled prevalence described in Figure 3. On the basis of the pooled data analysis, the point of intersection between the prevalence and NPV curve occurs at an NPV of 92% (95% CI, 87%-96%).

### Discussion

In this study, we found a wide variability in the performance of the Afirma GEC. This variability cannot be accounted for solely by the differences in malignancy prevalence. Many factors could be responsible for this variability, including the intrinsic test properties represented by sensitivity and specificity. A comparison of multiple studies<sup>9,13-22</sup> revealed interinstitutional unpredictability in test characteristics. Marti et al<sup>19</sup> identified a wide interinstitutional variability as well by comparing data from 2 facilities. It is difficult to determine the underlying reasons behind this variability among the studies because the use of the test in practice is not definitively established and health care professionals have conflicting views regarding its role in directing management, which could affect the selection of patients to undergo the test and the interpretation of test outcomes. On the other hand, most studies are from a single institution, which limits their generalizability.

Table. Characteristics of the Study Population<sup>a</sup>

Characteristic	Finding
Age, mean (SD), y	56.0 (16.2)
Sex, No. (%) of patients	
Male	39 (26.9)
Female	106 (73.1)
Nodule size, mean (SD), cm	2.4 (1.6)
Bethesda category, No. (%) of nodules <sup>b</sup>	
III	114 (74.0)
IV	40 (26.0)
Afirma GEC result, No. (%) of nodules	
Benign	58 (37.7)
Suspicious	96 (62.3)
Underwent thyroid surgery, No. (%) of nodules	
No	42 (27.3)
Yes	112 (72.7)
Surgical pathologic test result, No. (%) of nodules	
Benign	62 (55.4)
Malignant	50 (44.6)

Abbreviation: GEC, gene expression classifier.

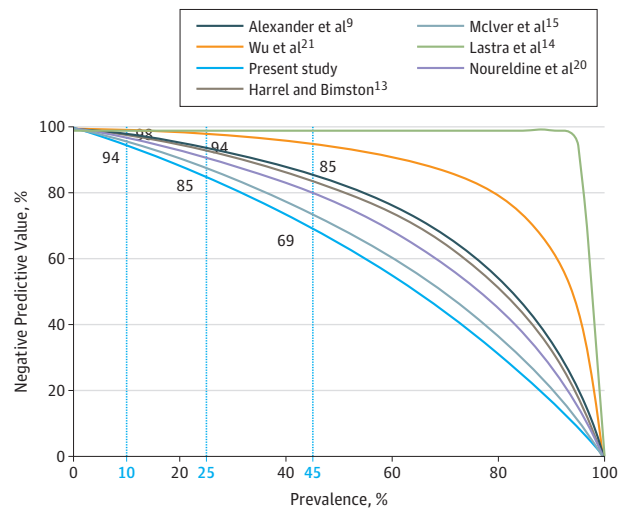
<sup>a</sup> The study population included 145 patients and 154 nodules.

<sup>b</sup> Bethesda criteria categories are explained in Cibas and Ali.<sup>6</sup>

Sensitivity and specificity are more accurately defined in a large sample data set; however, most published studies<sup>9,13-22</sup> averaged a relatively small sample size (n = 108.3). The pooled analysis performed in this study had a sample size of 1303, which enabled a precise definition of the true test characteristics. Furthermore, the pooled analysis also synthesized an estimate of the true population prevalence of thyroid malignancy in indeterminate thyroid nodules. Theoretically, the prevalence should not be influenced by sample size; however, with a small sample size, as in most of the studies<sup>9,13-22</sup> that investigated the Afirma GEC, there is a higher probability of random error. Previous studies<sup>9,13-22</sup> calculated the prevalence based on patients seen in tertiary medical centers and who also elected to undergo surgery, which makes it highly influenced by selection bias. In addition, prevalence percentages are influenced by interobserver or interoperator variability. Although the effect of selection bias cannot be overcome by pooling data, we believe pooling these data from different sources in an unbiased way can best approximate the true malignancy prevalence. In a recent meta-analysis by Santhanam et al<sup>23</sup> that included 7 studies, the authors found that the Afirma GEC has a sensitivity of 95.7% (95% CI, 92%-98%) and a specificity of 30.5% (95% CI, 26%-35%). Our analysis revealed similar data points; furthermore, we used these more accurately defined sensitivity and specificity values to plot a pooled NPV curve against prevalence to assess the variability of Afirma GEC performance that takes into account test characteristics and malignancy prevalence variability.

Genetic and molecular profiling is increasingly gaining ground in standard medical practice. Management of thyroid disorders and nodules is now largely taking into account ancillary genetic testing. Several markers and mutations are currently being investigated. For example, mutations that have

Figure 2. Negative Predictive Value Curves Based on Sensitivity and Specificity Reported by Different Studies



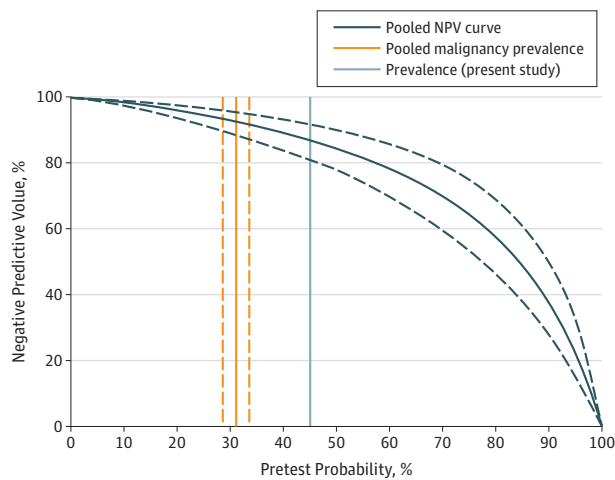
The dashed vertical lines represent chosen hypothetical malignancy prevalence from the x-axis to demonstrate the points of intersection with negative predictive curves.

been found to have a high diagnostic specificity of thyroid cancer include the V-RAF murine sarcoma viral oncogene homologue B1 at the V600E position (*BRAF*<sup>V600E</sup>) (OMIM 164757), RAS viral oncogene homologue (*RAS*) (OMIM 164790 for *NRAS*, OMIM 190070 for *KRAS*, and OMIM 190020 for *HRAS*), and gene rearrangements of paired boxed gene 8 (*PAX8*) (OMIM 167415) and peroxisome proliferator-activated receptor gamma (*PPARG*) (OMIM 601487) as well as the rearranged during transfection proto-oncogene (*RET/PTC*) (OMIM 164761). These alterations may be found in 16% of indeterminate thyroid nodules.<sup>24</sup> The presence of these mutations can direct patients to immediate total thyroidectomy. However, the mutations have low sensitivity (63.7%), missing more than a third of thyroid cancers. Thus, they are good confirmatory rather than screening tests because of high false-negative rates.<sup>25</sup> Following up indeterminate nodules by examination and ultrasonography in lieu of surgery can potentially avoid unnecessary operations and associated costs.<sup>26</sup> The Afirma GEC classifies thyroid nodules into benign and suspicious categories, depending on the expression pattern of RNA.<sup>10</sup> With a 50% benign Afirma GEC rate for indeterminate thyroid nodules, a reported 50 000 thyroid operations could be avoided every year if the test is implemented across the country because the high NPV can place these patients into a follow-up rather than surgical category.<sup>11</sup>

In an effort to improve diagnostic accuracy in cytologically indeterminate nodules, the *BRAF*<sup>V600E</sup> gene mutation is being tested along with the Afirma GEC testing. However, a study<sup>27</sup> that included 208 indeterminate thyroid nodules concluded that the addition of *BRAF* gene mutation testing did not improve Afirma GEC sensitivity or specificity.

Ancillary molecular testing, such as the Afirma GEC, can be cost-effective. Li et al<sup>26</sup> assessed the cost-effectiveness of the Afirma GEC and found that it may lower overall cost and

**Figure 3. Pooled Negative Predictive Value (NPV) Curve and Pooled Malignancy Prevalence Curve**



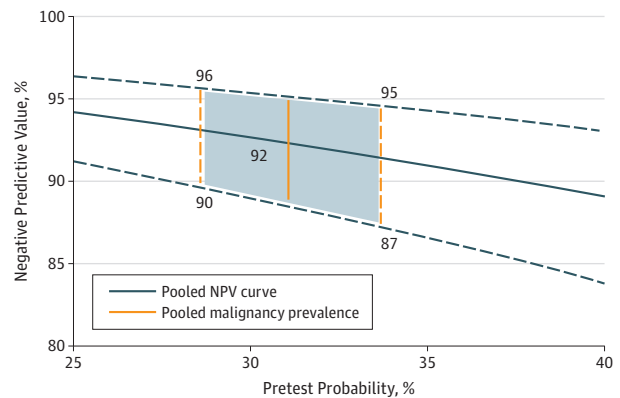
Solid light blue line indicates prevalence; solid orange line, pooled prevalence; solid dark blue line, pooled NPV; and dashed lines, 95% CIs.

moderately improve health outcome for patients with indeterminate thyroid nodules. These results were attributed to the reduction in surgery rate for patients whose nodules proved to be benign. Taking into consideration the cost of the test, an initial report<sup>11</sup> found that 1 operation was avoided for every 2 tests performed.

### Limitations

Limitations of this study include its retrospective nature and lack of long-term follow-up data. Although all patients who had benign Afirma GEC testing had clinical or sonographic indications that prompted surgical intervention, the data do not provide information regarding when these indications developed between the time of taking the test and undergoing surgery. The retrospective design lacks the control for interobserver or interoperator and intraobserver or intraoperator variability that can result from performing the FNA biopsy or reviewing the cytopathologic specimens. Cytopathologists interpreting the FNA biopsy results may be aware of the malignancy prevalence in the patient population studied and may have increased sensitivity for labeling aspirates indetermi-

**Figure 4. Expected Variability in Negative Predictive Value (NPV)**



Dashed lines indicate 95% CIs; shaded rectangle, variability.

nate. For the same reasons, the pooled data analysis lacks standardization of data sources and assessment of heterogeneity, and different methods might have been followed in different institutions that the current analysis could not adjust for. However, the pooled analysis generated a large sample size that allowed for more accurate estimations.

### Conclusions

This study has summarized the previously reported variability of the Afirma GEC's performance. Contrary to most previous investigations that focused on the malignancy prevalence as the sole culprit in assessing the unpredictability of Afirma GEC NPV, the reported NPV in this analysis revealed that variability is from test properties and malignancy prevalence.

The Afirma GEC, based on our analysis, can have an NPV with a range of 87% to 96%. Previous studies<sup>9,13-22</sup> that found an NPV above or below these limits might have overestimated or underestimated the true performance of the assay, which could have mainly been a result of the small sample size. There is a clinical need to improve the management quality of indeterminate thyroid nodules. A more precise definition of the NPV for the Afirma GEC is still required by means of large controlled clinical trials.

### ARTICLE INFORMATION

**Accepted for Publication:** September 17, 2016.

**Published Online:** December 15, 2016.  
doi:10.1001/jamaoto.2016.3526

**Author Contributions:** Dr Kandil had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Al-Qurayshi, Deniwar, Mallik, Sholl, Kandil.

**Acquisition, analysis, or interpretation of data:** Al-Qurayshi, Deniwar, Thethi, Srivastav, Murad, Bhatia, Moroz, Sholl, Kandil.

**Drafting of the manuscript:** Al-Qurayshi, Deniwar, Murad, Bhatia, Moroz.

**Critical revision of the manuscript for important intellectual content:** Al-Qurayshi, Deniwar, Thethi, Mallik, Srivastav, Sholl, Kandil.

**Statistical analysis:** Al-Qurayshi, Deniwar, Srivastav, Bhatia.

**Administrative, technical, or material support:** Thethi, Murad, Moroz, Sholl, Kandil.

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Kandil reported being a primary investigator of a multi-institutional study (Evaluation of Thyroid FNA Genomics Signature) sponsored by Veracyte. No other disclosures were reported.

**Meeting Presentation:** The study was presented at the 37th Annual Meeting of the American

Association of Endocrine Surgeons; April 11, 2016; Baltimore, Maryland.

### REFERENCES

- Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med*. 1993;328(8):553-559.
- 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 cancer. *Thyroid*. 2009;19(11):1167-1214.
- 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. *Endocr Pract*. 2010;16(suppl 1):1-43.
4. Gharib H, Goellner JR. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med*. 1993;118(4):282-289.
  5. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*. 2007;111(6):508-516.
  6. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2009;132(5):658-665.
  7. Horne MJ, Chieng DC, Theoharis C, et al. Thyroid follicular lesion of undetermined significance: evaluation of the risk of malignancy using the two-tier sub-classification. *Diagn Cytopathol*. 2012;40(5):410-415.
  8. Moses W, Weng J, Sansano I, et al. Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. *World J Surg*. 2010;34(11):2589-2594.
  9. 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.
  10. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab*. 2010;95(12):5296-5304.
  11. Duick DS, Klopper JP, Diggins JC, et al. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Thyroid*. 2012;22(10):996-1001.
  12. Ferris RL, Baloch Z, Bernet V, et al; American Thyroid Association Surgical Affairs Committee. American thyroid association statement on surgical application of molecular profiling for thyroid nodules: current impact on perioperative decision making. *Thyroid*. 2015;25(7):760-768.
  13. 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(4):364-369.
  14. 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(10):737-744.
  15. 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(11):4069-4077.
  16. Steward DL, Kloos RT. Clinical diagnostic gene expression thyroid testing. *Otolaryngol Clin North Am*. 2014;47(4):573-593.
  17. Brauner E, Holmes BJ, Krane JF, et al. Performance of the Afirma gene expression classifier in Hürthle cell thyroid nodules differs from other indeterminate thyroid nodules. *Thyroid*. 2015;25(7):789-796.
  18. Celik B, Whetsell CR, Nassar A. Afirma GEC and thyroid lesions: an institutional experience. *Diagn Cytopathol*. 2015;43(12):966-970.
  19. 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(12):3996-4001.
  20. Noureldine SI, Olson MT, Agrawal N, Prescott JD, Zeiger MA, Tufano RP. Effect of gene expression classifier molecular testing on the surgical decision-making for patients with thyroid nodules. *JAMA Otolaryngol Head Neck Surg*. 2015;141(12):1082-1088.
  21. Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW. Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. *Surgery*. 2016;159(1):118-126.
  22. 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(2):100-109.
  23. Santhanam P, Kthhir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. *Med Oncol*. 2016;33(2):14.
  24. Nikiforov YE, Ohori NP, Hodak SP, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab*. 2011;96(11):3390-3397.
  25. Ferraz C, Eszlinger M, Paschke R. Current state and future perspective of molecular diagnosis of fine-needle aspiration biopsy of thyroid nodules. *J Clin Endocrinol Metab*. 2011;96(7):2016-2026.
  26. Li H, Robinson KA, Anton B, Saldanha JJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2011;96(11):E1719-E1726.
  27. Kloos RT, Reynolds JD, Walsh PS, et al. Does addition of BRAF V600E mutation testing modify sensitivity or specificity of the Afirma gene expression classifier in cytologically indeterminate thyroid nodules? *J Clin Endocrinol Metab*. 2013;98(4):E761-E768.