Update on Molecular Testing for Cytologically Indeterminate Thyroid Nodules

Michiya Nishino, MD, PhD; Marina Nikiforova, MD

● Context.—Approximately 15% to 30% of thyroid nodules that undergo fine-needle aspiration are classified as cytologically indeterminate, presenting management challenges for patients and clinicians alike. During the past several years, several molecular tests have been developed to reduce the diagnostic uncertainty of indeterminate thyroid fine-needle aspirations.

Objective.—To review the methodology, clinical validation, and recent peer-reviewed literature for 4 molecular tests that are currently marketed for cytologically indeterminate thyroid fine-needle aspiration specimens: Afirma, ThyroSeq, ThyGenX/ThyraMIR, and RosettaGX Reveal.

Data Sources.—Peer-reviewed literature retrieved from PubMed search, data provided by company websites and representatives, and authors' personal experiences.

Conclusions.—The 4 commercially available molecular tests for thyroid cytology offer unique approaches to improve the risk stratification of thyroid nodules. Familiarity with data from the validation studies as well as the emerging literature about test performance in the post-validation setting can help users to select and interpret these tests in a clinically meaningful way.

(Arch Pathol Lab Med. 2018;142:446–457; doi: 10.5858/arpa.2017-0174-RA)

F ine-needle aspiration (FNA) cytology plays an important role in the risk stratification of thyroid nodules. For patients who meet clinical, laboratory, and/or sonographic criteria for biopsy, FNA cytology can be helpful for guiding subsequent management. For example, cytologically benign nodules may be managed by a nonsurgical watchful-waiting approach in most cases, whereas cytologically malignant nodules are usually referred for surgical resection. Nevertheless, approximately 15% to 30% of thyroid FNAs are classified in one of the cytologically indeterminate categories of the Bethesda System for Reporting Thyroid Cytopathology: atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS; Bethesda III) or follicular neoplasm/suspicious for follicular neoplasm (FN/SFN; Bethesda IV).[1,2] Because of the low (5%–15% for AUS/FLUS) to moderate (15%–30% for FN/SFN) cancer risks associated with these indeterminate categories, management recommendations have generally been conservative. Repeating the FNA has been suggested as the appropriate follow-up for AUS/FLUS nodules, with the consideration of diagnostic lobectomy for nodules that remain cytologically indeterminate on repeat FNA; diagnostic lobectomy has been recommended for nodules in the FN/SFN category.

Although lobectomy may be considered necessary and adequate management for premalignant nodules or low- to intermediate-risk cancers, it is not the ideal management for all cytologically indeterminate nodules.[3] For the majority of cases in the AUS/FLUS and FN/SFN categories, lobectomy reveals a histologically benign nodule, and surgery may be regarded as overtreatment. On the other hand, lobectomy may be deemed insufficient treatment for cases in which a high-risk cancer is diagnosed on histologic examination. In the latter scenario, patients may need to undergo reoperation to complete the thyroidectomy, especially if they may require radioactive iodine ablation.

Molecular testing has thus emerged as a tool for stratifying cytologically indeterminate thyroid nodules into clinically meaningful risk categories. The goals of ancillary molecular testing for thyroid cytology include (1) avoidance of unnecessary surgery for benign nodules and (2) distinguishing high-risk cancers that merit total thyroidectomy from premalignant or low/intermediate-risk nodules for which lobectomy may be the preferred initial surgical step. Currently, 4 tests are commercially available for thyroid FNAs: Afirma (Veracyte, Inc, South San Francisco, California), ThyroSeq v2 (CBLPath, Inc, Rye Brook, New York, and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania), ThyGenX/ThyraMIR (Interpace Diagnostics, Inc, Parsippany, New Jersey), and RosettaGX Reveal (Rosetta Genomics, Inc, Philadelphia, Pennsylvania) (Figure 1, A through D, and Table 1). In this review, we discuss differences in test methodology and appraise the recent literature regarding test performance. Readers are referred

Table 1. Comparison of Commercially Available Tests and Their Clinical Validation for Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance and Follicular Neoplasm Aspirates[a]

	Afirma[12,b]	ThyroSeq[35,36,c]	ThyGenX/ThyraMIR[66,d]	Rosetta GX Reveal[69,e]	
Methodology	Expression analysis of 167 mRNAs (25 genes in screening step; 142 genes in classifier) by microarray	DNA and RNA targeted next-generation sequencing (56 genes) for mutations, gene fusions, and gene expression	Targeted next-generation sequencing for mutations (5 genes) and 3 gene fusions; expression analysis of 10 miRNA by qRT-PCR	Expression analysis of 24 miRNAs by qRT-PCR	
Sample type required for molecular testing	2 dedicated FNA passes collected in nucleic acid preservative	1–2 drops from first FNA pass or 1 dedicated pass collected into nucleic acid preservative; cell blocks and FFPE tissue can also be used	1 dedicated FNA pass collected into nucleic acid preservative	1 direct smear or ThinPrep slide with adequate cellularity for cytologic diagnosis	
Sample collection/shipping kits provided by company	Yes	Yes	Yes	Yes	
Cytology review	Centralized[f]	Local or centralized	Local or centralized	Local or centralized	
				Entire validation set[g]	Agreement set[g]
Validation study					
Indeterminate FNAs (No.)	210	239	109	150	116
Prevalence of cancer, % (No. of histologically malignant cases/No. of indeterminate FNAs)	24 (51/210)	26 (61/239)	32 (35/109)	21 (31/150)	12 (14/116)
Sensitivity,%	90	90	89	74	100
Specificity, %	52	93	85	74	80
NPV, %	94	96	94	92	100
PPV, %[h]	37	81	74	43	41
Price, $[i]	6400 (Afirma GEC + MTC) 975 (Afirma MTC alone) 475 (Afirma BRAF alone)	4056	1675 (ThyGenX) 4000 (ThyraMIR)	3700	

Abbreviations: FFPE, formalin-fixed paraffin-embedded; FNA, fine-needle aspiration; GEC, Gene Expression Classifier; miRNA, microRNA; MTC, medullary thyroid carcinoma; NPV, negative predictive value; PPV, positive predictive value; qRT-PCR, quantitative real-time reverse transcription polymerase chain reaction.

[a] Only nodules with atypia of undetermined significance/follicular lesion of undetermined significance or follicular neoplasm cytology are included in the performance characteristics summarized in this table.

[b] Afirma (Veracyte, Inc, South San Francisco, California).

[c] ThyroSeq (CBLPath, Inc, Rye Brook, New York, and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania).

[d] ThyGenX/ThyraMIR (Interpace Diagnostics, Inc, Parsippany, New Jersey).

[e] RosettaGX Reveal (Rosetta Genomics, Inc, Philadelphia, Pennsylvania).

[f] Selected academic centers have been authorized to submit specimens for Afirma testing based on interpretations made by their local cytopathologists.

[g] For comparison with the other tests, only the atypia of undetermined significance/follicular lesion of undetermined significance and follicular neoplasm cases from RosettaGX Reveal's validation cohorts are shown in this table. With the inclusion of suspicious for malignancy cases, the performance characteristics for the entire validation set (N = 189, cancer prevalence of 32%) are 85% sensitivity, 72% specificity, 91% NPV, 59% PPV, and for the agreement set (n = 150, cancer prevalence of 27%) are 98% sensitivity, 78% specificity, 99% NPV, and 62% PPV. Agreement set refers to the subset of cases for which all 3 pathologists reviewing the histology concurred on the reference diagnosis.

[h] Positive test results are defined differently for each of these tests. For the Veracyte Afirma Gene Expression Classifier and the RosettaGX Reveal, "suspicious" and "suspicious by miRNA profiling" results are considered positive test results, respectively; for ThyraMIR/ThyGenX, detection of an oncogenic mutation/fusion and/or a high-risk miRNA profile is considered positive; for ThyroSeq, detection of an oncogenic mutation/fusion is considered positive.

[i] Quoted prices as of February 2017 (personal communication with sales representatives). The quoted prices are typically different from payers' reimbursement schedules.

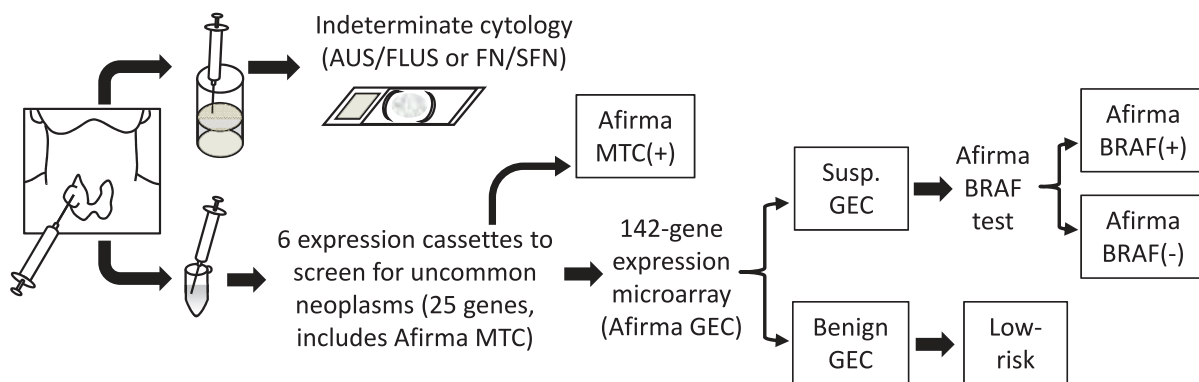to recent reviews[3–5] for additional discussion about the first 3 tests.

INFERRING TEST PERFORMANCE FROM PUBLISHED VALIDATION STUDIES

In the clinical validation studies for each of these molecular tests for thyroid FNAs, histopathology of the aspirated nodule serves as the gold standard for determining whether a nodule is malignant or benign. Clinical sensitivity and specificity refer to a test's ability to classify nodules correctly as histologically malignant or benign, respectively, and are considered fixed characteristics of a test as determined by a clinical validation study.

Estimates of cancer risk based on test results, on the other hand, are derived from positive and negative predictive values (PPVs and NPVs, respectively). In the context of molecular testing for thyroid FNAs, PPV indicates the cancer risk based on an abnormal (positive) test result, and NPV indicates the probability of benignity based on a negative test result. The complement of the NPV, or 1 − NPV, is equivalent to the cancer risk associated with a negative test result. In contrast to sensitivity and specificity, PPV and NPV

A. Afirma GEC

Indeterminate cytology (AUS/FLUS or FN/SFN)

6 expression cassettes to screen for uncommon neoplasms (25 genes, includes Afirma MTC)

Afirma MTC(+)

142-gene expression microarray (Afirma GEC)

Susp. GEC → Afirma BRAF test → Afirma BRAF(+) / Afirma BRAF(-)

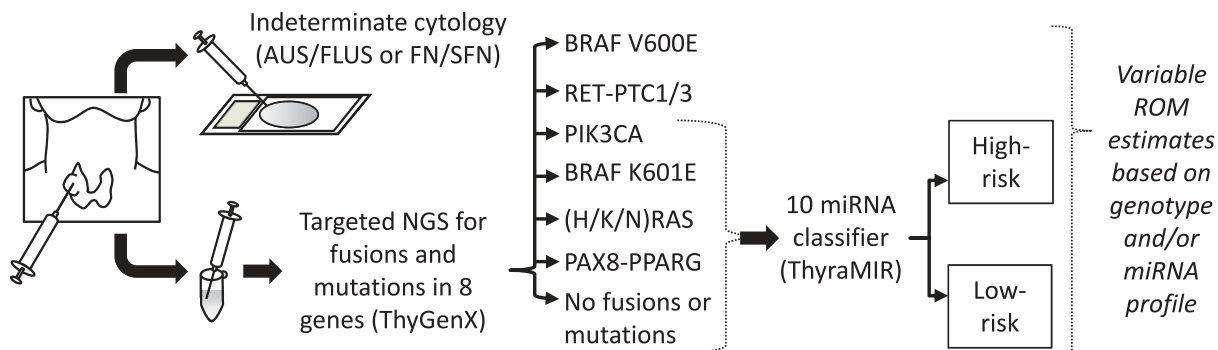Benign GEC → Low-risk

B. ThyroSeq

Indeterminate cytology (AUS/FLUS or FN/SFN)

Targeted NGS for hotspot mutations in 14 genes & 42 gene fusions (ThyroSeq)

Positive for mutations and/or gene fusions

Variable ROM estimates based on genotype

Negative for mutations or gene fusions → Low-risk

C. ThyGenX / ThyraMIR

Indeterminate cytology (AUS/FLUS or FN/SFN)

Targeted NGS for fusions and mutations in 8 genes (ThyGenX)

- BRAF V600E
- RET-PTC1/3
- PIK3CA
- BRAF K601E
- (H/K/N)RAS
- PAX8-PPARG
- No fusions or mutations

10 miRNA classifier (ThyraMIR)

High-risk / Low-risk

Variable ROM estimates based on genotype and/or miRNA profile

D. RosettaGX Reveal

Indeterminate cytology (AUS/FLUS or FN/SFN)

Nucleic acid extraction from smear

24 miRNA classifier (Reveal)

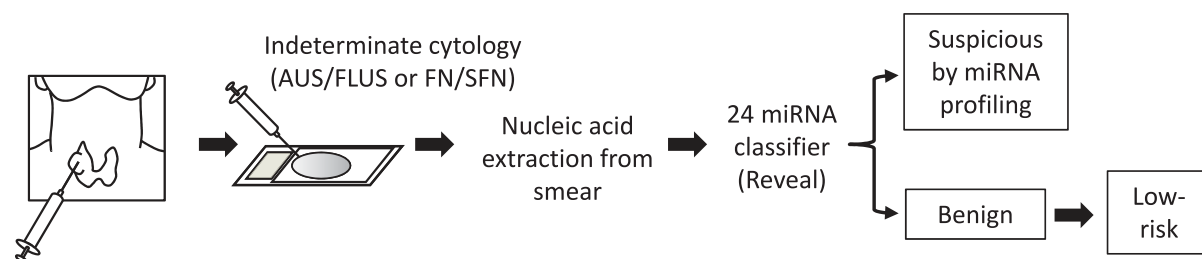Suspicious by miRNA profiling

Benign → Low-risk

Figure 1. *Overview of testing methodology for the 4 commercially available molecular tests for thyroid fine-needle aspiration (FNA) specimens. All 4 tests are performed for nodules that are interpreted as indeterminate (eg, atypia of undetermined significance/follicular lesion of undetermined significance [AUS/FLUS] or follicular neoplasm/suspicious for follicular neoplasm [FN/SFN]) by FNA cytology. A, For the Afirma Gene Expression Classifier (GEC; Veracyte, Inc, South San Francisco, California), the sample is first screened for less common lesions in the thyroid, including medullary thyroid carcinoma (MTC). If the Afirma MTC screening cassette is positive, the test result is positive and the main GEC is not performed. For samples that are negative for the screening cassettes, the expression profiles of 142 genes is analyzed by the classifier, giving*

Molecular Testing for Thyroid FNA—Nishino & Nikiforova

(or 1 − NPV) vary with the pretest probability of disease in the tested population. Thus, the PPV and 1 − NPV for each test can be plotted as functions of the pretest probability of cancer to illustrate the malignancy risk associated with positive or negative test results, respectively (Figure 2, A through H). The shapes of these predictive value curves are determined by the Bayes theorem using a test's reported sensitivity and specificity. Tests with high specificity have sharply bent PPV curves (Figure 2, C and E), and tests with high sensitivity have sharply bent 1 − NPV curves (Figure 2, B, D, F, and H). The ability of a negative molecular test result to help rule out malignancy depends on the pretest probability of cancer in the tested population (Figure 2).

The prevalence of malignancy for the cytologically indeterminate Bethesda categories (ie, AUS/FLUS and FN/SFN) is often used as an approximation for pretest probability of cancer. Cancer prevalence for cytologically indeterminate nodules may vary by institution because of dissimilarities in patient populations or the heterogeneous nature of the AUS/FLUS category, as well as differences among cytopathologists in their thresholds for using the cytologically indeterminate Bethesda categories.[6–8] Therefore, for optimal interpretation of these tests, each institution should determine the prevalence of malignancy for nodules classified in the cytologically indeterminate categories.[9] The incorporation of clinical and sonographic data may also provide a more accurate estimate of pretest disease probability.

Several caveats must be considered when comparing the performance characteristics of these tests. First, these molecular tests define a positive test result in different ways, limiting the utility of comparing PPVs as a measure of test performance. For tests with binary outcomes such as Afirma and RosettaGX Reveal, "suspicious" (or "suspicious by microRNA [miRNA] profiling" for the latter test) results are used in the calculation of true-positive and false-positive outcomes (Figure 1, A and D). In contrast, neither ThyroSeq nor ThyGenX/ThyraMIR report binary outcomes in practice. Instead, although a negative test result confers a defined low residual cancer risk, the finding of a specific genetic alteration allows estimation of a more granular cancer risk and phenotype for thyroid nodules based on the mutation/gene fusion present (Figure 1, B and C). Molecular assessment of cancer risk may be further adjusted based on the allelic frequency of the genetic change and/or the results of miRNA profiling (for samples tested by both ThyGenX and ThyraMIR). Although the detection of any oncogenic mutation/fusion or high-risk miRNA profile by these 2 tests has been considered a positive result for the purposes of statistical analysis, it should be noted that these tests provide a gradation of cancer risk that is not entirely captured by the PPV alone.[10]

Secondly, the gold standard histopathology in these validation studies generally consists of only 2 categories (benign or malignant) to facilitate statistical analysis. However, this practice runs counter to the recognition that thyroid neoplasms encompass a broad continuum of biological behavior rather than a dichotomous one.[3,11] Users of these tests should recognize that the PPVs publicized by each of these commercially available tests are therefore limited in their ability to convey a nuanced and clinically meaningful risk of disease.

AFIRMA GENE EXPRESSION CLASSIFIER

The Afirma Gene Expression Classifier (GEC) is a micro-array-based test that uses a proprietary algorithm to risk-stratify cytologically indeterminate nodules as having either benign (GEC-B) or suspicious (GEC-S) messenger RNA (mRNA) expression profiles (Table 1 and Figure 1, A). A prospective, blinded, multicenter clinical validation study characterized the performance of the test among 210 AUS/FLUS and FN/SFN nodules with a pretest cancer rate of 24% to 25%. The study had a 15% postunblinding sample exclusion rate. Based on the results of this study, Afirma has high sensitivity (90%) and modest specificity (approximately 50%) for malignancy, corresponding to posttest malignancy risks of 5% to 6% for GEC-B and 37% to 38% for GEC-S results among cytologically indeterminate nodules (Table 1 and Figure 2, A and B).[12] As a test optimized for high sensitivity and NPV, Afirma's clinical utility is based on its ability to help exclude malignancy in the setting of indeterminate cytology. The specificity of the Afirma test may be augmented by 2 "malignancy classifiers": the Afirma MTC test, which assays for gene expression profiles specific for medullary thyroid carcinoma,[13,14] and the Afirma BRAF test, which assays for the expression profile associated with the BRAF V600E mutation. Regarding the latter test, the low prevalence of BRAF V600E mutations among AUS/FLUS and FN/SFN aspirates, particularly in Western populations, may limit the utility of Afirma BRAF as a reflex test for GEC-S samples.[15,16]

The cellular composition of samples submitted for Afirma testing is determined in part by gene expression cassettes that screen for mRNA profiles associated with medullary thyroid carcinoma (the Afirma MTC test, described above), parathyroid tissue, and metastatic tumors.[12] Only samples that do not trigger these screening cassettes are analyzed by the main 142-gene classifier; samples that are flagged by these screening cassettes are resulted as having a suspicious Afirma result without subsequent analysis by the main GEC.

←

either benign or suspicious GEC results. For GEC-suspicious cases, reflex testing for the Afirma BRAF test may be performed. B, ThyroSeq (CBLPath, Inc, Rye Brook, New York, and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania) is a targeted next-generation sequencing (NGS) test for a broad panel of mutations and gene fusions that have been identified in thyroid cancer. Based on the genetic change(s) identified, the test estimates the risk of malignancy (ROM) and/or phenotype of the nodule. C, ThyGenX (Interpace Diagnostics, Inc, Parsippany, New Jersey) is a targeted NGS test for a limited panel of the most common mutations and gene fusions that have been identified in thyroid cancer. For samples with indeterminate molecular results (ie, results other than a BRAF V600E mutation or RET-PTC1/3 gene fusion), reflex testing for ThyraMIR may be performed. ThyraMIR analyzes the expression profiles of 10 microRNAs (miRNAs) to classify nodules as low or high risk. For nodules examined by both ThyGenX and ThyraMIR tests, ROM estimates are based on a combination of genotype and miRNA profiling results. D, RosettaGX Reveal (Rosetta Genomics, Inc, Philadelphia, Pennsylvania) is an miRNA classifier that classifies aspirates as benign or suspicious by miRNA profiling results. RosettaGX Reveal is distinguished as the only commercially available test to date that accepts routinely prepared cytology slides as the source for nucleic acid extraction. Abbreviation: Susp., suspicious.
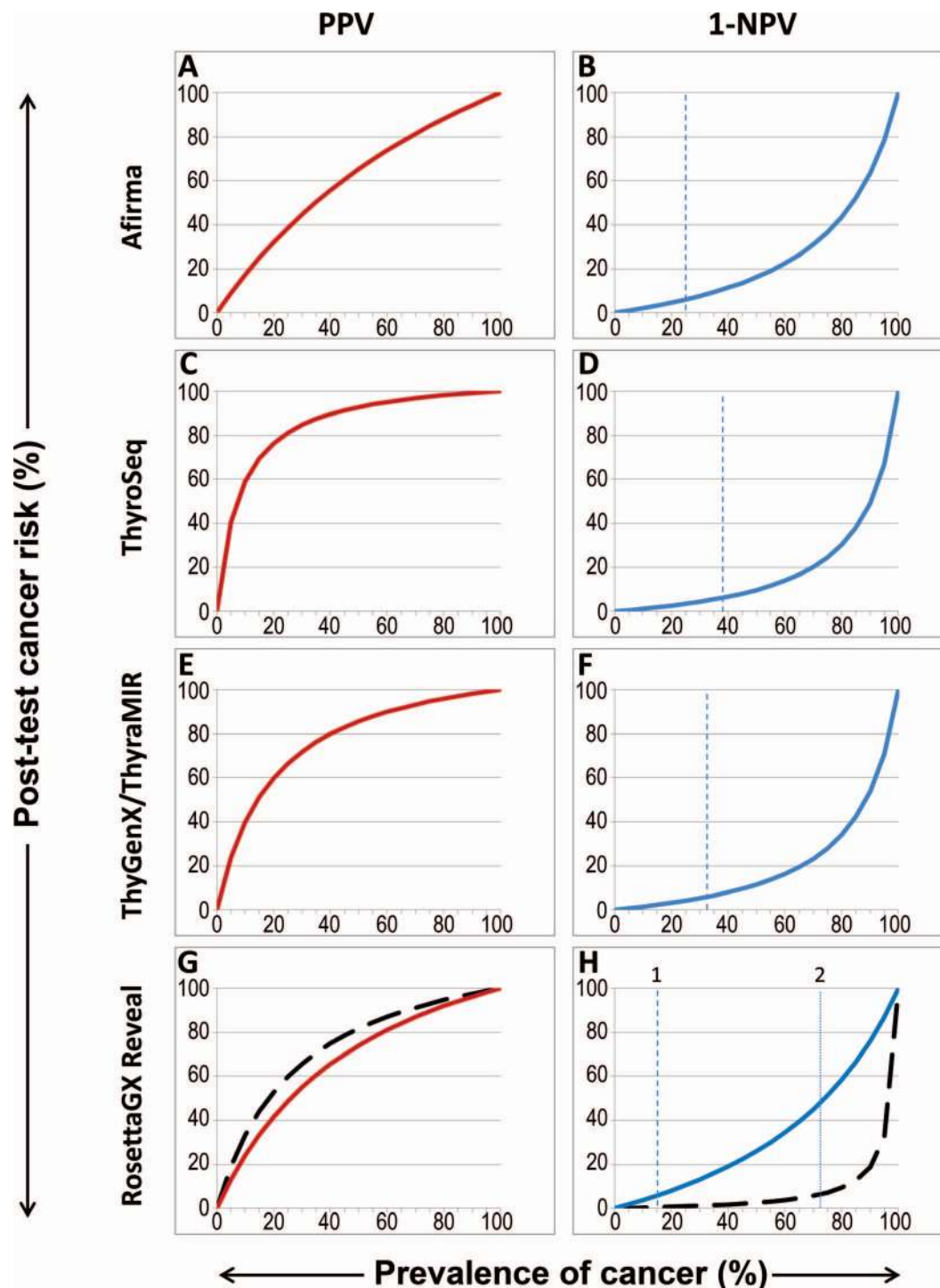
PPV — **1-NPV**

Post-test cancer risk (%)

Prevalence of cancer (%)

Figure 2. *Risk assessment of thyroid nodules based on the published sensitivity and specificity for Afirma (Veracyte, Inc, South San Francisco, California) (A and B), ThyroSeq (CBLPath, Inc, Rye Brook, New York, and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania) (C and D), ThyGenX/ThyraMIR (Interpace Diagnostics, Inc, Parsippany, New Jersey) (E and F), and RosettaGX Reveal (Rosetta Genomics, Inc, Philadelphia, Pennsylvania) (G and H). For each of these tests, the risk of thyroid cancer associated with a particular test result (y-axis) is plotted against the prevalence of thyroid cancer in the tested population (x-axis). Positive predictive value (PPV) curves (A, C, E, and G) indicate the cancer risk associated with a positive test result, and the complement to the negative predictive value (1 − NPV) curves (B, D, F, and H) show the cancer risk associated with a negative test result, both as a function of cancer prevalence. For comparison purposes, the solid PPV and 1 − NPV curves are derived from published test sensitivity and specificity data for atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) and follicular neoplasm/suspicious for follicular neoplasm (FN/SFN) cases. The additional dashed curves for RosettaGX Reveal (G and H) indicate PPV and 1 − NPV derived from analysis of a subset of cases (agreement set) in the validation cohort, with the inclusion of aspirates in the suspicious for malignancy (SFM) Bethesda category. For the 1 − NPV curves, the vertical dotted lines (B, D, F, and H) indicate the cancer prevalence (x-axis) at which a negative result is associated with a ∼6% risk of malignancy; beyond this mark, the cancer risk associated with a negative test result may be considered too high to avoid diagnostic surgery. In H, the 2 vertical lines indicate the prevalence thresholds for RosettaGX Reveal when tested among the AUS/FLUS and FN/SFN subset (line 1) versus the agreement set with inclusion of SFM cases (line 2).*

	True-Negative (GEC Benign, Histology Benign)	False-Negative (GEC Benign, Histology Malignant)	Histology of False-Negative Cases (No.)
Study			
Wu et al,[27] 2016	20	2	PTC (1), FVPTC (1)[a]
Samulski et al,[25] 2016	13	3	PTC (1), FVPTC (1), NIFTP (1)[b]
Angell et al,[17] 2015	12	1	Minimally invasive FTC
Alexander et al,[18] 2014	9	1	PTC microcarcinoma[c]
Chaudhary et al,[20] 2016	8	0	NA
Marti et al,[22] 2015	7	0	NA
Harrell and Bimston,[21] 2014	3	2	PTC (2)
Celik et al,[28] 2015	4	1	PTC
Sacks et al,[24] 2016	4	0	NA
McIver et al,[23] 2014	3	1	FTC with capsular and vascular invasion
Brauner et al,[19] 2015	3	0	NA
Villabona et al,[26] 2016	3	1	PTC
Total	**89**	**12**	

Table 2. Published Reports of Resected Nodules With Benign Afirma Gene Expression Classifier (GEC) Results

Abbreviations: FTC, follicular thyroid carcinoma; FVPTC, follicular variant of papillary thyroid carcinoma; NA, not applicable; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; PTC, papillary thyroid carcinoma.

[a] An additional case (papillary microcarcinoma in the vicinity of the 3.5-cm aspirated nodule) was classified as a false negative result by Wu et al[27]; this nodule is omitted from the current table because of the size discrepancy between the targeted nodule and the microcarcinoma.

[b] Samulski et al[25] reclassified the noninvasive subset of FVPTC as NIFTP for their analysis.

[c] The microcarcinoma was 0.6 cm by pathologic measurement but 1.0 cm as measured by preoperative ultrasound.

Can GEC-B Nodules Be Managed Like Cytologically Benign Nodules?

Based on the results of its clinical validation study, Afirma has been marketed for identifying benign nodules when the cytology is indeterminate. This statement implies that cytologically indeterminate nodules with GEC-B results can be managed with clinical observation rather than diagnostic lobectomy in most cases, similar to cytologically benign nodules. A study by Angell et al[17] endorses this approach using retrospective ultrasonographic nodule measurements as a proxy for biologic behavior; the authors found that the proportion of cytologically indeterminate GEC-B nodules exhibiting growth (greater than 20% increase in 2 or more dimensions, or greater than 50% increase in volume) was comparable to that of cytologically benign nodules during a median follow-up period of 13 to 14 months. Based on this finding, the authors suggest that GEC-B nodules may be clinically observed similarly to cytologically benign nodules. Assessment of additional sonographic parameters apart from size may be helpful in the broader confirmation of these findings.

The Afirma validation study[12] remains the benchmark for assessing its test performance. Nevertheless, the growing list of publications summarizing institutional experiences with Afirma may also shed some light on how Afirma has been functioning in the real-world clinical setting. Interpretation of these postvalidation analyses requires caution because of (1) the retrospective design of these studies and (2) the fact that in clinical practice, decisions about nodule management are not blinded to Afirma results. In particular, most GEC-B nodules are not resected, and the small fraction selected for surgery is unlikely to be representative of GEC-B nodules as a whole.

With these caveats in mind, combined analysis of published postvalidation studies reveals 101 GEC-B nodules with histologic follow-up, 12 of which were classified by authors as malignant (Table 2).[17–28] The reasons for these false-negative results are not clear, although they likely represent a combination of sampling error, the intrinsic false-negative rate of the test, and variations in thresholds for diagnosing nodules as malignant on histopathology.[7]

Taking into account the selection bias inherent in examining only GEC-B nodules that were selected for resection, the actual false-negative rate of Afirma is undoubtedly significantly lower than 12%, in keeping with the high sensitivity of the test. Furthermore, among the cases considered to be false-negative in these studies is one nodule that could presently be reclassified as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP; see discussion below).[25] Taken together, the independent reports summarizing the everyday clinical experience with Afirma appear to support its role as an effective test for identifying a subset of cytologically indeterminate nodules that may warrant a watchful waiting approach because of their low malignancy risk.

Has Afirma Reduced Unnecessary Surgeries for Cytologically Indeterminate Nodules?

Afirma has been marketed as a test that can reduce the number of unnecessary thyroid surgeries. In support of this claim, authors have pointed to the low surgical rates for GEC-B nodules compared with either (1) historical surgical rates for cytologically indeterminate nodules prior to the introduction of Afirma testing ("pre-Afirma") or (2) surgical rates for GEC-S nodules (Table 3).[18,29] Likewise, others have reported the similarity between surgical rates for GEC-B nodules and cytologically benign nodules, as well as the durability of the low surgical rates for GEC-B nodules during a 3-year follow-up period.[30,31] These comparisons demonstrate clinician and patient compliance with Afirma's recommendation for nonsurgical management for GEC-B nodules; however, these studies do not specifically address reduction in overall surgical rates due to Afirma, nor do they consider whether the surgeries were unnecessary. Indeed, for studies that provide data regarding institutional surgical rates and histologic outcomes for cytologically indeterminate nodules in the pre-Afirma versus Afirma-tested cohorts, there is no significant difference in overall surgical rates between these groups (Table 3).[20,24,27,32]

For the purposes of this analysis, if we define *unnecessary thyroid surgery* as resection of cytologically indeterminate, histologically benign nodules, the impact of Afirma appears

Table 3. Analysis of Institutional Surgical Rates Before and After Introduction of Afirma Testing

Study	Cytologically Indeterminate Cases Undergoing Thyroid Surgery, % (No. of Surgeries/No. of Indeterminate FNAs)				
	Pre-Afirma	Afirma-Tested Nodules Overall	P[a]	Afirma GEC-Benign Subset	P[b]
Sacks et al,[24] 2016	38 (83/220)	47 (60/129)	.12	8 (4/48)	<.001
Chaudhary et al,[20] 2016	61 (171/280)	54 (86/158)	.19	13 (8/63)	<.001
Yang et al,[32] 2016; Wu et al,[27] 2016	54 (272/502)	52 (128/245)	.64	21 (24/113)	<.001

Abbreviations: FNA, fine-needle aspiration; GEC, Gene Expression Classifier.

[a] Fisher exact test, 2-tailed P value comparing surgical rates pre-Afirma versus Afirma-tested nodules overall.

[b] Fisher exact test, 2-tailed P value comparing surgical rates pre-Afirma versus Afirma GEC-benign subset.

to vary among institutions (Table 4). Data from the University of California at Los Angeles experience demonstrate a significantly lower rate of unnecessary surgery in the Afirma-tested cohort (51%) compared with pre-Afirma controls (66%) (P = .004).[27,32,33] In contrast, rates of unnecessary surgery were not significantly altered by Afirma testing in the series by Chaudhary et al[20] (68% pre-Afirma, 66% post-Afirma; P = .89) or Sacks et al[24] (75% pre-Afirma, 70% post-Afirma; P = .57).

The explanation for these institutional differences is uncertain. On one hand, Sacks et al[24] speculate that in their Afirma-tested cohort, a GEC-S result may have had an unduly high effect on prompting surgery in spite of a moderate PPV of ~40%, thus offsetting the reduction in surgery afforded by GEC-B results. Different institutional criteria for selecting nodules for ancillary testing (ie, reflex testing on nodules with a single indeterminate FNA versus testing on nodules with persistently indeterminate cytology on repeat FNA) may have also contributed to the apparent differences in clinical utility of the test among various institutions. The retrospective nature of these postvalidation studies precludes definitive conclusions about the impact that Afirma has made in clinical practice. Nevertheless, the collective data suggest that the clinical utility of Afirma can be institution specific on a population-wide level; detailed examination of the differences in ordering and interpreting the test among these institutions may be helpful for refining the guidelines for optimal test use.

THYROSEQ

ThyroSeq v2 is a multigene test that is based on the targeted DNA and RNA next-generation sequencing analysis of 56 genes (ie, point mutations and small insertions/deletions in 14 genes and 42 types of gene fusions) and expression levels of 16 genes (Figure 1, B, and Table 1). Next-generation sequencing offers high sensitivity of detection and ability to quantify the proportion of cells carrying a given mutation. Primary diagnostic information provided by the test is based on the analysis of mutation hotspots and gene fusions that have been found in ~90% of

papillary thyroid carcinomas (PTCs) and other thyroid cancers.[34] In addition, the expression levels of several gene mRNAs are used to assess sample adequacy and define a proportion of thyroid follicular cells within the sample. The proportion of thyroid follicular cells is defined based on the expression of thyroglobulin, thyroid transcription factor 1, cytokeratin-7, and sodium/iodide symporter (SLC5A5) mRNAs. In addition, the expression levels of the SLC5A5 mRNAs provide information on the functional status of cells within the nodule and allow detection of hyperfunctional thyroid nodules, which are typically benign. The presence of C-cell/medullary thyroid carcinoma and parathyroid cells is defined by the expression of calcitonin mRNA and parathyroid hormone mRNA, respectively, and reported as suspicious for medullary carcinoma or parathyroid nodule. In samples with high expression of calcitonin mRNA, the findings of RET or RAS mutations allow provision of additional information related to the germline or somatic status of the disease.

Clinical validation of ThyroSeq v2 was reported in 2 single-institution studies of thyroid nodules with indeterminate cytology and known surgical outcome.[35,36] A study of 143 nodules with FN/SFN (Bethesda IV) cytology (91 retrospectively and 52 prospectively collected samples)[35] and a prospective study of 96 patients with AUS/FLUS (Bethesda III) cytology[36] reported a high sensitivity (90%–91%) and specificity (92%–93%) of ThyroSeq v2 (Table 1). In the study cohorts with a pretest cancer prevalence of 23% to 27%, ThyroSeq v2 demonstrated an NPV of 96% to 97% and a PPV of 77% to 83% (Figure 2, C and D). This general PPV of the test is further refined in each individual sample based on the type of detected genetic alteration.[35,36] Indeed, many mutations, such as BRAF V600E and RET/PTC, confer a close to 100% probability of cancer, whereas other mutations, such as RAS, are diagnostic of a clonal tumor with ~80% probability of typically low-risk cancer or NIFTP, as discussed later in this review.[11] Other mutations, such as PTEN and EIF1AX, serve as a marker of clonal neoplasm but alone are not sufficient for full cancer development.[37] Furthermore, evaluation of the expression of cell lineage

Table 4. Analysis of Institutional Rates of Unnecessary Thyroid Surgeries Before and After Introduction of Afirma Testing[a]

Study	Unnecessary Thyroid Surgeries, % (No. of Histologically Benign Cases/No. of Surgeries)		
	Pre-Afirma	Afirma-Tested Nodules Overall	P[b]
Sacks et al,[24] 2016	75 (62/83)	70 (42/60)	.57
Chaudhary et al,[20] 2016	68 (116/171)	66 (57/86)	.89
Yang et al,[32] 2016; Wu et al,[27] 2016	66 (180/272)	51 (65/128)	.004

[a] Unnecessary thyroid surgeries are defined as those on cytologically indeterminate nodules that were found to be histologically benign.

[b] Fisher exact test, 2-tailed P value.

genes established a diagnosis of parathyroid nodule in 0.6% of "thyroid" nodules with AUS/FLUS cytology, which was confirmed by subsequent parathyroid surgery or clinical and laboratory diagnosis of hyperparathyroidism.[36]

The performance and clinical utility of ThyroSeq v2 have been evaluated in recent independent studies. Valderrabano et al[38] assessed the test performance and clinical utility in 190 nodules (from 182 patients) with AUS/FLUS and FN/SFN cytology. ThyroSeq v2 performance was calculated using 102 nodules that were surgically excised. Considering NIFTP in the malignant category, the prevalence of malignancy in this cohort was 20%. ThyroSeq v2 showed 70% sensitivity, 77% specificity, 42% PPV, and 91% NPV. The performance of ThyroSeq v2 in their study was significantly better in FN/SFN than in AUS/FLUS nodules (area under the curve 0.84 versus 0.57, respectively; $P = .03$). The authors reported 5 specimens to be false-negative, with an additional case showing a low level of KRAS mutation. Among the 5 mutation-negative nodules, 2 were NIFTP, 2 were papillary carcinomas, and 1 was a minimally invasive follicular carcinoma with vascular invasion. One of the papillary carcinomas was a 0.6-cm microcarcinoma that did not match by size a 1.1-cm biopsied nodule as measured by ultrasound, making correlation between molecular testing and final pathology questionable. Of note, a previous study from the same institution reported the PPV of the Afirma test at 16%,[23] highlighting the variability in performance of molecular tests in different patient populations. The study of ThyroSeq also provided information on clinical utility of the test.[38] The authors found that in their group of AUS/FLUS nodules, a negative or positive test result did not affect significantly the rate of malignancy in the nodules that underwent surgical excision. However, in the nodules with FN/SFN cytology, a positive test result increased the risk of malignancy 2.5-fold, with PPV of 53% to 65%.[38] Overall, ThyroSeq-positive nodules in this study were more frequently resected than ThyroSeq-negative nodules (73% versus 48%, respectively; $P < .001$). Another independent study by Toraldo et al[39] evaluated ThyroSeq performance in 148 nodules with indeterminate cytology, of which 45 underwent surgery, and found 95% sensitivity, 60% specificity, 66% PPV, and 94% NPV in their patient population.

In addition to diagnostic utility, ThyroSeq and other tests based on the detection of mutations may provide information on cancer aggressiveness, which could be taken into account when considering the extent of surgery. Although prognostic association of BRAF V600E remains controversial, TERT promoter mutations have been established as independent predictors of disease recurrence and cancer-related mortality in well-differentiated thyroid cancer.[40–43] Furthermore, the co-occurrence of BRAF or RAS mutations with TERT or TP53 mutations may identify a small subset of thyroid cancers with the most unfavorable outcome.[44–46] Single studies have demonstrated that testing for mutational markers was helpful to define the optimal extent of initial surgery and reduce the proportion of 2-step surgery, that is, initial lobectomy followed by completion thyroidectomy, for many patients.[47,48] This is particularly applicable to thyroid nodules found positive for BRAF V600E, TERT mutations, and RET, ALK, or NTRK1/3 fusions, which confer a very high (close to 100%) probability of intermediate-risk or high-risk thyroid cancer. Upfront total thyroidectomy is currently a recommended surgical approach to these nodules.[3] For nodules with RAS and other RAS-like mutations, which

confer a high probability of either low-risk cancer or precancer NIFTP, diagnostic lobectomy may represent the optimal surgical approach. However, clinical utility of ThyroSeq in informing the extent of surgery is not fully defined yet and should be examined in prospective studies.

miRNA-BASED TESTS

MicroRNAs are small noncoding RNAs that regulate gene expression by influencing the stability and translation of mRNA. Compared with mRNA, miRNAs are relatively stable and can be isolated from routinely prepared formalin-fixed histopathology or alcohol-fixed cytology samples, making them a practical analyte for molecular diagnostics.[49–52] Several research studies, which used mostly surgically excised tumor tissues, reported that a number of miRNAs are differentially expressed among various benign and malignant thyroid nodules.[52–57] This provided a rationale for 2 recent commercially available tests that use miRNA expression for diagnostic use in cytologically indeterminate thyroid FNA specimens: ThyraMIR (a complementary test to ThyGenX) and RosettaGX Reveal.

ThyGenX/ThyraMIR

ThyGenX is a targeted next-generation sequencing test that assays for a small number of mutations in 5 genes (BRAF, KRAS, HRAS, NRAS, and PIK3CA) and 3 gene fusions (RET-PTC1, RET-PTC3, and PAX8-PPARG) associated with thyroid neoplasia. As described above, the detection of one of these genetic changes can help guide surgical management, depending on the mutation or gene fusion that is identified. Nodules harboring BRAF V600E mutations or clonal RET-PTC1/3 gene fusions have a virtually 100% risk of PTC and may warrant total thyroidectomy upfront.[58–63] Apart from BRAF V600E mutations and RET-PTC1/3 gene fusions, the remainder of the genetic changes in the ThyGenX panel is limited by low specificity for malignancy. RAS mutations, BRAF K601E mutations, and PAX8-PPARG gene fusions have been identified in association with both premalignant/benign and malignant neoplasms.[11,58,59,62–65] Accordingly, a diagnostic lobectomy may be the preferred initial surgical approach for nodules with these genetic changes. Similarly, a lobectomy is also recommended when ThyGenX is negative for this panel of mutations and gene fusions, because of the relatively low sensitivity and NPV of ThyGenX alone among cytologically indeterminate nodules.

The ThyraMIR miRNA expression classifier was introduced to address the challenges with the sensitivity and specificity of the ThyGenX panel. ThyraMIR measures the expression levels of 10 miRNAs (Table 5) by quantitative real-time polymerase chain reaction and uses a proprietary algorithm to classify each nodule as having either a high-risk or low-risk miRNA profile. A multi-institutional cross-sectional cohort study of ThyGenX and ThyraMIR using 109 thyroid FNA samples found that combining both tests achieved an optimal sensitivity (89%) and specificity (85%) profile among cytologically indeterminate nodules.[66,67] Based on these findings, ThyraMIR is currently offered as a reflex test when the ThyGenX test is either (1) negative or (2) positive for RAS mutations, BRAF K601E mutation, PIK3CA mutation, or PAX8-PPARG fusion (ie, any mutation/fusion other than BRAF V600E or RET-PTC1/3) (Figure 1, C).

For samples assayed by both ThyGenX and ThyraMIR, Interpace Diagnostics estimates cancer risk based on results of both tests. A negative ThyGenX result coupled with a

Table 5. Comparison of miRNA Panels Tested by ThyraMIR and RosettaGX Reveal[a]	
ThyraMIR	**RosettaGX Reveal**
hsa-miR-31-5p	hsa-miR-31-5p
hsa-miR-222-3p	hsa-miR-222-3p
hsa-miR-146b-5p	hsa-miR-146b-5p
hsa-miR-375	hsa-miR-375
hsa-miR-551b-3p	hsa-miR-551b-3p
hsa-miR-138-1-3p	hsa-miR-138-5p
hsa-miR-139-5p	hsa-miR-486-5p
hsa-miR-29b-1-5p	hsa-miR-23a-3p
hsa-miR-155	hsa-miR-574-3p
hsa-miR-204-5p	hsa-miR-152-3p
	hsa-miR-200c-3p
	hsa-miR-345-5p
	hsa-miR-5701
	hsa-miR-424-3p
	hsa-miR-3074-5p
	hsa-miR-346
	hsa-miR-342-3p
	hsa-miR-181c-5p
	hsa-miR-125b-5p
	MID-50971
	MID-20094
	MID-50976
	MID-50969
	MID-16582

[a] ThyraMIR (Interpace Diagnostics, Inc, Parsippany, New Jersey); RosettaGX Reveal (Rosetta Genomics, Inc, Philadelphia, Pennsylvania).

low-risk ThyraMIR result is associated with approximately 6% risk of malignancy among AUS/FLUS and FN/SFN nodules based on the 109 cases in the validation cohort with 32% cancer prevalence (Table 1).[66] For other permutations of ThyGenX and ThyraMIR results (eg, *RAS* mutation coupled with low-risk ThyraMIR profile), estimates of cancer risk are given based on the laboratory data. Specimen adequacy for molecular testing is based on spectrophotometric assessment of nucleic acid quantity and quality.

RosettaGX Reveal

Similar to ThyraMIR, RosettaGX Reveal analyzes miRNA expression patterns to classify cytologically indeterminate thyroid aspirates as having either a "benign" or "suspicious by miRNA profiling" result. Rosetta GX Reveal's test panel measures 24 miRNAs, 6 sequences of which closely overlap with ThyraMIR's panel of 10 miRNAs (Table 5). Markers associated with thyroid epithelial cells are included to help ensure that miRNA analysis is being performed on thyroid epithelial cells rather than blood components.[68] In contrast to the other molecular tests that are currently commercially available for thyroid aspirates, RosettaGX Reveal's miRNA classifier was analytically validated on cellular material recovered from direct cytology smears stained by Papanicolaou and Romanowsky-type stains (Figure 1, D).[68] RosettaGX Reveal is thus marketed as a test that requires only a single adequately cellular cytology slide, with both direct smears and liquid-based cytology (eg, ThinPrep) slides accepted as substrates for molecular testing.

RosettaGX Reveal was clinically validated in a retrospective multicenter study of 189 cytologically indeterminate samples with histologic follow-up.[69] Of note, this validation cohort included suspicious for malignancy (Bethesda V) cases in addition to AUS/FLUS and FN/SFN cases. In this overall validation cohort with 32% prevalence of cancer, RosettaGX Reveal was found to have 85% sensitivity, 72%

specificity, 91% NPV, and 59% PPV. The test reached superior sensitivity (98%) and NPV (99%) and modest increases in specificity (78%) and PPV (62%) when analysis was limited only to cases in which all 3 pathologists involved in the validation study (2 study pathologists and 1 referring pathologist) concurred on the reference histopathology diagnosis (n = 150; 27% prevalence of cancer). The improvement in RosettaGX Reveal's sensitivity and NPV in this agreement set was likely driven in part by the exclusion of 14 encapsulated follicular variant of PTCs from the validation set, 5 of which were misclassified as benign by RosettaGX Reveal. Importantly, the validation cohort did not include any oncocytic (Hürthle cell) carcinomas, and therefore the performance of RosettaGX Reveal in detecting these tumors in the clinical setting remains unknown. This represents a key limitation of the test, which should be taken into account particularly if the clinical intention is to avoid surgery for nodules with a negative test result.

Although the low concordance among cytopathologists for using the suspicious for malignancy Bethesda category[7] may justify its inclusion in RosettaGX Reveal's validation cohort as an indeterminate category, the validation studies for Afirma, ThyroSeq, and ThyGenX/ThyraMIR have largely limited their analyses to aspirates in the AUS/FLUS and FN/SFN categories (Table 1).[12,35,36,66,69] Such differences in the composition of the validation cohort among these studies should be considered when comparing their results; for comparison purposes, Table 1 and the solid PPV and 1 − NPV curves in Figure 2 include only data from the AUS/FLUS and FN/SFN cases from each of the validation studies. The predictive value curves for the somewhat idealized cohort (agreement set with inclusion of suspicious for malignancy cases) in the RosettaGX Reveal validation study are shown with dotted curves for comparison (Figure 2, G and H).

The ability to use cytology slides as the starting material for molecular analysis circumvents the need to collect additional FNA material into a nucleic acid preservative solution specifically for molecular testing, theoretically reducing the number of FNA passes that are required from each patient. Furthermore, the ability to analyze miRNAs from the same population of cells that have undergone visual cytologic examination may help reduce the potential for sampling error, as can be seen when separate FNA passes are obtained for cytology and molecular testing. One drawback to this approach is the need to sacrifice a diagnostic cytology slide for nucleic acid extraction. Rosetta Genomics offers slide scanning services prior to processing the slide for molecular testing.

MOLECULAR TESTING FOR THYROID FNA IN THE NIFTP ERA

The noninvasive subset of encapsulated/well-circumscribed follicular variant of PTC was recently proposed to be renamed as NIFTP, to reflect the indolent clinical behavior of these tumors in the absence of capsular or vascular invasion.[11] Retrospective analysis of archival cases has shown that the preoperative FNA cytology of NIFTPs fall predominantly into AUS/FLUS, FN/SFN, or suspicious for malignancy categories.[70–74] Cytologic and molecular features may help distinguish NIFTP from conventional-type PTCs and infiltrative (ie, not encapsulated or well-circumscribed) follicular variant of PTC, with NIFTP showing lesser degrees of nuclear atypia, increased associ-

ation with *RAS* mutations rather than *BRAF* V600E mutations, and distinct miRNA profile.[75–80] However, neither cytologic nor mutational analysis appears to reliably distinguish NIFTP from encapsulated/well-circumscribed follicular variant of PTC, and the primary distinction between these 2 entities remains the histologic identification of capsular or vascular invasion in the latter.[70,72,80]

All 4 molecular tests described in this review were developed prior to the NIFTP nomenclature change. Therefore, noninvasive, encapsulated/well-circumscribed follicular variant of PTC (tumors that would now be classified as NIFTP) in the training and validation sets for these tests were likely classified as having malignant reference histology. In keeping with this idea, several institutions have found that NIFTPs can be recognized as suspicious by Afirma[25,81,82] or harboring a mutation or gene fusion by genotyping tests such as ThyroSeq.[38,81] Although the reclassification of NIFTP out of the malignant category is expected to affect the predictive values of these tests, the overall management implications may not be significantly affected. For tests with modest PPV (Afirma GEC, Roset-taGX Reveal), an abnormal result (suspicious GEC or miRNA profile) typically prompts diagnostic lobectomy, which is currently considered appropriate treatment for NIFTP. Similarly, for tests such as ThyroSeq and ThyGenX/ThyraMIR, the granularity provided by specific genotypes (eg, isolated *RAS* mutations, *PAX8-PPARG*, or *THADA* fusions) makes it possible to suggest a diagnostic lobectomy for nodules with a high likelihood of being NIFTP. Importantly, NIFTP is considered as a premalignant tumor with a potential to invade and spread hematogenously, and as such it requires surgical excision to examine the tumor capsule to rule out invasion. Therefore, although not malignant, NIFTP is an indication for surgery, and its detection preoperatively by molecular tests may not necessarily represent a false-positive result.

CONCLUSIONS

All 4 tests described above highlight the clinical relevance of molecular diagnostics applied to thyroid cytopathology. Nevertheless, each of these tests shows differences in methodology and performance characteristics in their effort to improve risk stratification among cytologically indeterminate thyroid nodules. The emerging literature about the performance of these tests in day-to-day clinical practice can offer insights into factors that may influence the optimal use of molecular testing for thyroid cytopathology.[6,83] Prospective, blinded, multi-institutional studies that use a large set of FNA samples representing all thyroid cancer types and having low postunblinding sample exclusion rates are required for further understanding of the performance and clinical utility of these diagnostic tests. In the 2015 update to the management guidelines for thyroid nodules, the American Thyroid Association includes general recommendations about the use and interpretation of ancillary molecular testing in the context of clinical, ultrasonographic, and cytologic findings. In recognition of the relatively early point in our experience with these tests, the American Thyroid Association guidelines do not advocate a particular test or a specific algorithm for selecting nodules for molecular testing, with the exception of nodules that are cytologically suspicious for papillary carcinoma (Bethesda V), where testing for *BRAF* and other genetic markers was listed in the recommendation.[3] Studies that underpin evidence-based guidelines for test selection and interpretation will be needed to develop a standardized approach to molecular testing for thyroid cytopathology, particularly as we move towards a more personalized, risk-based approach to the diagnosis and management of thyroid nodules.

References

1. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol*. 2012;56(4):333–339.
2. 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.
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. Nishino M. Molecular cytopathology for thyroid nodules: a review of methodology and test performance. *Cancer Cytopathol*. 2016;124(1):14–27.
5. 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.
6. Baca SC, Wong KS, Strickland KC, et al. Qualifiers of atypia in the cytologic diagnosis of thyroid nodules are associated with different Afirma gene expression classifier results and clinical outcomes. *Cancer Cytopathol*. 2017; 125(5):313–322.
7. Cibas ES, Baloch ZW, Fellegara G, et al. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann Intern Med*. 2013; 159(5):325–332.
8. Nishino M, Wang HH. Should the thyroid AUS/FLUS category be further stratified by malignancy risk? *Cancer Cytopathol*. 2014;122(7):481–483.
9. Valderrabano P, Leon ME, Centeno BA, et al. Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. *Eur J Endocrinol*. 2016;174(5):621–629.
10. Najafian A, Noureldine S, Azar F, et al. RAS mutations, and RET/PTC and PAX8/PPAR-gamma chromosomal rearrangements are also prevalent in benign thyroid lesions: implications thereof and a systematic review. *Thyroid*. 2017; 27(1):39–48.
11. 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(8):1023–1029.
12. 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.
13. Kloos RT, Monroe RJ, Traweek ST, Lanman RB, Kennedy GC. A genomic alternative to identify medullary thyroid cancer preoperatively in thyroid nodules with indeterminate cytology. *Thyroid*. 2016;26(6):785–793.
14. Pankratz DG, Hu Z, Kim SY, et al. Analytical performance of a gene expression classifier for medullary thyroid carcinoma. *Thyroid*. 2016;26(11): 1573–1580.
15. 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.
16. Pusztaszeri MP, Krane JF, Faquin WC. BRAF testing and thyroid FNA. *Cancer Cytopathol*. 2015;123(12):689–695.
17. 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(11):E1477–E1483.
18. Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma gene expression classifier. *J Clin Endocrinol Metab*. 2014;99(1): 119–125.
19. 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(7):789–796.
20. 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(3):205–210.
21. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. *Endocr Pract*. 2014;20(4):364–369.
22. 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.
23. 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.
24. Sacks WL, Bose S, Zumsteg ZS, et al. Impact of Afirma gene expression classifier on cytopathology diagnosis and rate of thyroidectomy. *Cancer*. 2016; 124(10):722–728.

25. 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(11): 867–873.

26. 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(10):1199–1203.

27. 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(7):916–922.

28. Celik B, Whetsell CR, Nassar A. Afirma GEC and thyroid lesions: an institutional experience. *Diagn Cytopathol*. 2015;43(12):966–970.

29. Duick DS, Klopper JP, Diggans 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.

30. Singer J, Hanna JW, Visaria J, et al. Impact of a gene expression classifier on the long-term management of patients with cytologically indeterminate thyroid nodules. *Curr Med Res Opin*. 2016;32(7):1225–1232.

31. Sipos JA, Blevins TC, Shea HC, et al. Long-term nonoperative rate of thyroid nodules with benign results on the Afirma Gene Expression Classifier. *Endocr Pract*. 2016;22(6):666–672.

32. 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.

33. Sullivan PS, Hirschowitz SL, Fung PC, Apple SK. The impact of atypia/ follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda system. *Cancer Cytopathol*. 2014;122(12):866–872.

34. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;159(3):676–690.

35. Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. 2014;120(23):3627–3634.

36. Nikiforov YE, Carty SE, Chiosea SI, et al. Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid*. 2015;25(11):1217–1223.

37. Karunamurthy A, Panebianco F, Hsiao SJ, et al. Prevalence and phenotypic correlations of EIF1AX mutations in thyroid nodules. *Endocr Relat Cancer*. 2016; 23(4):295–301.

38. Valderrabano P, Khazai L, Leon ME, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. *Endocr Relat Cancer*. 2017;24(3):127–136.

39. Toraldo G, Godley FA, Cerda SR, et al. Large independent prospective study to evaluate the performances of ThyroSeq2 multigene next generation sequencing panel analysis on cancer diagnosis in thyroid nodules with indeterminate cytopathology [abstract]. *Thyroid*. 2016;26(S1):A148.

40. Landa I, Ganly I, Chan TA, et al. Frequent somatic TERT promoter mutations in thyroid cancer: higher prevalence in advanced forms of the disease. *J Clin Endocrinol Metab*. 2013;98(9):E1562–E1566.

41. Liu T, Wang N, Cao J, et al. The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas. *Oncogene*. 2014;33(42):4978–4984.

42. Liu X, Bishop J, Shan Y, et al. Highly prevalent TERT promoter mutations in aggressive thyroid cancers. *Endocr Relat Cancer*. 2013;20(4):603–610.

43. Melo M, da Rocha AG, Vinagre J, et al. TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. *J Clin Endocrinol Metab*. 2014;99(5):E754–E765.

44. Xing M, Liu R, Liu X, et al. BRAF V600E and TERT promoter mutations cooperatively identify the most aggressive papillary thyroid cancer with highest recurrence. *J Clin Oncol*. 2014;32(25):2718–2726.

45. Song YS, Lim JA, Choi H, et al. Prognostic effects of TERT promoter mutations are enhanced by coexistence with BRAF or RAS mutations and strengthen the risk prediction by the ATA or TNM staging system in differentiated thyroid cancer patients. *Cancer*. 2016;122(9):1370–1379.

46. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab*. 2013;98(11):E1852–E1860.

47. Yip L, Wharry LI, Armstrong MJ, et al. A clinical algorithm for fine-needle aspiration molecular testing effectively guides the appropriate extent of initial thyroidectomy. *Ann Surg*. 2014;260(1):163–168.

48. Shrestha RT, Karunamurthy A, Amin K, Nikiforov YE, Caramori ML. Multiple mutations detected preoperatively may predict aggressive behavior of papillary thyroid cancer and guide management—a case report. *Thyroid*. 2015; 25(12):1375–1378.

49. Gustafson D, Tyryshkin K, Renwick N. microRNA-guided diagnostics in clinical samples. *Best Pract Res Clin Endocrinol Metab*. 2016;30(5):563–575.

50. Jung M, Schaefer A, Steiner I, et al. Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. *Clin Chem*. 2010;56(6): 998–1006.

51. Petriella D, Galetta D, Rubini V, et al. Molecular profiling of thin-prep FNA samples in assisting clinical management of non-small-cell lung cancer. *Mol Biotechnol*. 2013;54(3):913–919.

52. Rossi ED, Bizzarro T, Martini M, et al. The evaluation of miRNAs on thyroid FNAC: the promising role of miR-375 in follicular neoplasms. *Endocrine*. 2016; 54(3):723–732.

53. Dettmer M, Perren A, Moch H, et al. Comprehensive MicroRNA expression profiling identifies novel markers in follicular variant of papillary thyroid carcinoma. *Thyroid*. 2013;23(11):1383–1389.

54. Dettmer M, Vogetseder A, Durso MB, et al. MicroRNA expression array identifies novel diagnostic markers for conventional and oncocytic follicular thyroid carcinomas. *J Clin Endocrinol Metab*. 2013;98(1):E1–E7.

55. Dettmer MS, Perren A, Moch H, et al. MicroRNA profile of poorly differentiated thyroid carcinomas: new diagnostic and prognostic insights. *J Mol Endocrinol*. 2014;52(2):181–189.

56. Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. *J Clin Endocrinol Metab*. 2008;93(5):1600–1608.

57. Yip L, Kelly L, Shuai Y, et al. MicroRNA signature distinguishes the degree of aggressiveness of papillary thyroid carcinoma. *Ann Surg Oncol*. 2011;18(7): 2035–2041.

58. Beaudenon-Huibregtse S, Alexander EK, Guttler RB, et al. Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. *Thyroid*. 2014;24(10):1479–1487.

59. Cantara S, Capezzone M, Marchisotta S, et al. Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. *J Clin Endocrinol Metab*. 2010;95(3):1365–1369.

60. Eszlinger M, Piana S, Moll A, et al. Molecular testing of thyroid fine-needle aspirations improves presurgical diagnosis and supports the histologic identification of minimally invasive follicular thyroid carcinomas. *Thyroid*. 2015;25(4): 401–409.

61. Lee ST, Kim SW, Ki CS, et al. Clinical implication of highly sensitive detection of the BRAF V600E mutation in fine-needle aspirations of thyroid nodules: a comparative analysis of three molecular assays in 4585 consecutive cases in a BRAF V600E mutation-prevalent area. *J Clin Endocrinol Metab*. 2012; 97(7):2299–2306.

62. 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.

63. 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(6):2092–2098.

64. Medici M, Kwong N, Angell TE, et al. The variable phenotype and low-risk nature of RAS-positive thyroid nodules. *BMC Med*. 2015;13:184.

65. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol*. 2011;7(10):569–580.

66. Labourier E, Shifrin A, Busseniers AE, et al. Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *J Clin Endocrinol Metab*. 2015; 100(7):2743–2750.

67. Wylie D, Beaudenon-Huibregtse S, Haynes BC, Giordano TJ, Labourier E. Molecular classification of thyroid lesions by combined testing for miRNA gene expression and somatic gene alterations. *J Pathol Clin Res*. 2016;2(2):93–103.

68. Benjamin H, Schnitzer-Perlman T, Shtabsky A, et al. Analytical validity of a microRNA-based assay for diagnosing indeterminate thyroid FNA smears from routinely prepared cytology slides. *Cancer Cytopathol*. 2016;124(10):711–721.

69. Lithwick-Yanai G, Dromi N, Shtabsky A, et al. Multicentre validation of a microRNA-based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate smears [published online ahead of print October 26, 2016]. *J Clin Pathol*. 2017;70(6):500–507. doi:10.1136/jclinpath-2016-204089.

70. Bizzarro T, Martini M, Capodimonti S, et al. Young investigator challenge: the morphologic analysis of noninvasive follicular thyroid neoplasm with papillary-like nuclear features on liquid-based cytology: some insights into their identification. *Cancer Cytopathol*. 2016;124(10):699–710.

71. Faquin WC, Wong LQ, Afrogheh AH, et al. Impact of reclassifying noninvasive follicular variant of papillary thyroid carcinoma on the risk of malignancy in the Bethesda System for Reporting Thyroid Cytopathology. *Cancer Cytopathol*. 2016;124(3):181–187.

72. Maletta F, Massa F, Torregrossa L, et al. Cytological features of "noninvasive follicular thyroid neoplasm with papillary-like nuclear features" and their correlation with tumor histology. *Hum Pathol*. 2016;54:134–142.

73. Rosario PW, Mourao GF, Nunes MB, Nunes MS, Calsolari MR. Noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Endocr Relat Cancer*. 2016;23(12):893–897.

74. Strickland KC, Howitt BE, Marqusee E, et al. The impact of noninvasive follicular variant of papillary thyroid carcinoma on rates of malignancy for fine-needle aspiration diagnostic categories. *Thyroid*. 2015;25(9):987–992.

75. Borrelli N, Denaro M, Ugolini C, et al. miRNA expression profiling of "noninvasive follicular thyroid neoplasms with papillary-like nuclear features" compared with adenomas and infiltrative follicular variants of papillary thyroid carcinomas. *Mod Pathol*. 2017;30(1):39–51.

76. Howitt BE, Jia Y, Sholl LM, Barletta JA. Molecular alterations in partially-encapsulated or well-circumscribed follicular variant of papillary thyroid carcinoma. *Thyroid*. 2013;23(10):1256–1262.

77. Rivera M, Ricarte-Filho J, Knauf J, et al. Molecular genotyping of papillary thyroid carcinoma follicular variant according to its histological subtypes

(encapsulated vs infiltrative) reveals distinct BRAF and RAS mutation patterns. *Mod Pathol*. 2010;23(9):1191–1200.

78. Strickland KC, Vivero M, Jo VY, et al. Preoperative cytologic diagnosis of noninvasive follicular thyroid neoplasm with papillary-like nuclear features: a prospective analysis. *Thyroid*. 2016;26(10):1466–1471.

79. Yoo SK, Lee S, Kim SJ, et al. Comprehensive analysis of the transcriptional and mutational landscape of follicular and papillary thyroid cancers. *PLoS Genet*. 2016;12(8):e1006239.

80. Zhao L, Dias-Santagata D, Sadow PM, Faquin WC. Cytological, molecular, and clinical features of noninvasive follicular thyroid neoplasm with papillary-like nuclear features versus invasive forms of follicular variant of papillary thyroid carcinoma. *Cancer Cytopathol*. 2017;125(5):323–331.

81. Jiang XS, Harrison GP, Datto MB. Young investigator challenge: molecular testing in noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Cancer*. 2016;124(12):893–900.

82. Wong KS, Angell TE, Strickland KC, et al. Noninvasive analysis of the transcriptional and mutational landscape of follicular and papillary thyroid carcinoma and the Afirma gene-expression classifier. *Thyroid*. 2016;26(7):911–915.

83. Noureldine SI, Olson MT, Agrawal N, et al. Effect of gene expression classifier molecular testing on the surgical decision-making process for patients with thyroid nodules. *JAMA Otolaryngol Head Neck Surg*. 2015;141(12):1082–1088.