Molecular Cytopathology for Thyroid Nodules: A Review of Methodology and Test Performance

Michiya Nishino, MD, PhD

Advances in the molecular characterization of thyroid cancers have fueled the development of genetic and gene expression-based tests for thyroid fine-needle aspirations. Collectively, these tests are designed to improve the diagnostic certainty of thyroid cytology. This review summarizes the early published experience with the commercially available versions of these tests: the Afirma Gene Expression Classifier, ThyGenX (formerly miR*Inform*)/ThyraMIR, and ThyroSeq. Key differences in testing approaches and issues regarding test performance and interpretation are also discussed. ***Cancer (Cancer Cytopathol)* 2016;124:14-27.** © *2015 American Cancer Society.*

KEY WORDS: DNA mutational analysis; fine-needle aspiration biopsy; high-throughput nucleotide sequencing; microarray analysis; molecular diagnostic techniques; thyroid neoplasms.

INTRODUCTION

Sir William Osler's description of medicine as "a science of uncertainty and an art of probability" is an apt depiction of the role fine-needle aspiration (FNA) cytology plays in the management of thyroid nodules.[1] Approximately 15% to 30% of thyroid FNAs fall in an interpretive gray zone, in which the probability of malignancy is considered too high for watchful waiting but insufficient to merit a total thyroidectomy.[2,3] Aspirates in the "follicular neoplasm/suspicious for a follicular neoplasm" (FN/SFN) category are typically associated with a 15% to 30% risk of malignancy. At this risk level, patients are generally referred for diagnostic thyroid lobectomy. In contrast, repeat FNA is the usual management for aspirates in the category of "atypia of undetermined significance/follicular lesion of undetermined significance" (AUS/FLUS) due to a 5% to 15% risk of malignancy, with diagnostic lobectomy considered for nodules with repeatedly indeterminate FNA cytology.

Management algorithms that include diagnostic lobectomy for cytologically indeterminate thyroid nodules present opportunities for improvement. The majority of nodules with AUS/FLUS or FN/SFN cytology are ultimately diagnosed as histologically benign; for these nodules, diagnostic lobectomy could be considered overtreatment. Conversely, for the subset of patients diagnosed with a malignancy in the lobectomy specimen, a return to the operating room for a completion thyroidectomy may be necessary. Coupled with these challenges is the recognition that these surgical decisions are driven in part by cytologic interpretive categories with a high propensity for interobserver variability.[4–6]

Over the past several years, molecular testing has emerged as a promising method for clarifying the gray area of indeterminate thyroid FNAs, with the aim of 1) reducing the overtreatment of benign nodules and 2) increasing the preoperative detection of malignant nodules that should be treated by a single surgery (total thyroidectomy) rather than a 2-step procedure (diagnostic lobectomy and completion thyroidectomy). This review

Corresponding author: Michiya Nishino, MD, PhD, Department of Pathology, Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215; Fax: (617) 667-7120; mnishin1@bidmc.harvard.edu

Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts

Received: July 27, 2015; **Revised:** August 13, 2015; **Accepted:** August 14, 2015

Published online September 8, 2015 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.21612, wileyonlinelibrary.com

will highlight the features of commercially available molecular tests for indeterminate thyroid FNAs, with examination of their methodology, validation data, strengths, limitations, optimal use, and interpretation.

General Comments Regarding Test Performance

The uncertainty of indeterminate thyroid FNAs can be resolved in 2 ways: toward "ruling in" or "ruling out" malignancy. The ability of a clinical test to "rule in" or "rule out" malignancy depends on its positive predictive value (PPV) and negative predictive value (NPV), respectively. Predictive values are not static properties of a clinical test but vary with the pretest probability of disease. Based on the specificity and sensitivity characteristics of a test from a validation study, Bayes' theorem can be used to extrapolate PPV and NPV for any given pretest probability of disease. For thyroid FNAs, the pretest probability of malignancy for cytologically indeterminate nodules can fluctuate with cytopathologists' thresholds for rendering AUS/FLUS or FN/SFN interpretations. For example, in a cohort with a pretest probability of malignancy of 15% to 30%, a test may have an NPV high enough to "rule out" malignancy; in a different cohort with a higher pretest probability of malignancy, the NPV of the same test may not be sufficiently high to exclude cancer. Therefore, the significance of a positive or negative test result reported by a validation study may not be universally applicable. Optimally, each end-user of these molecular tests should determine whether the pretest probability of malignancy for their patient population falls into the range for which positive and/or negative test results are clinically meaningful. The prevalence of malignancy for a particular cytologic interpretive category can be used as an approximation of the pretest probability of malignancy, although clinical and ultrasonographic parameters also can be incorporated to refine the risk level.

The Rule-Out Approach: Veracyte Afirma Test

Because a majority of nodules with indeterminate cytology are found to be benign on surgical resection, an ancillary test that can preoperatively rule out malignancy has the potential to spare a subset of these patients an unnecessary surgical procedure. The Afirma Gene Expression Classifier (GEC) from Veracyte (South San Francisco, Calif) embraces this approach by using microarray technology to assess the mRNA expression profiles of cytologically indeterminate thyroid nodules.

Test design

This test requires 2 dedicated FNA passes to be collected into a vial of RNA preservative, in addition to the FNA passes collected for routine cytomorphology. With the exception of several academic centers, Veracyte requires the concurrent cytology specimen to be interpreted at a centralized cytopathology laboratory (Thyroid Cytopathology Partners in Austin, Tex), with indeterminate aspirates (AUS/FLUS or FN/SFN) reflexed to the Afirma test. The material for Afirma testing is assayed in 2 broad steps: 1) a panel of 6 "cassettes" comprising 25 genes that screen for the expression profiles of less common entities in the thyroid, such as metastatic lesions (breast carcinoma, renal cell carcinoma, and melanoma), parathyroid tissue, medullary thyroid carcinoma, and oncocytic follicular (Hurthle cell) lesions; and 2) the main GEC, comprised of a 142-gene mRNA expression panel. Together, both parts compose the 167-gene Afirma test.[7]

The 6 screening cassettes in the first step help ensure that the samples assayed by the main GEC fall within the spectrum of the histologically benign and malignant thyroid nodules on which the GEC algorithm was trained. If the sample triggers 1 of these 6 screening cassettes, the specimen is automatically reported as having a "suspicious" gene expression profile, without further analysis by the main GEC. Samples that trigger the medullary thyroid carcinoma (MTC) cassette are flagged in the report as having a positive "Afirma MTC" test, as described in further detail below; a formal reporting system for samples that trigger the other 5 cassettes does not appear to be in place at the current time.

This main 142-gene GEC uses a proprietary algorithm to classify each aspirate as having either a "benign" or a "suspicious" GEC result. The algorithm is optimized to recognize aspirates with benign expression profiles; thus, it is considered a test with a high NPV that is useful for "ruling out" malignancy. Table 1 summarizes information regarding test methodology, cost, sample procurement, storage, and shipping methods.

Test validation

Afirma was clinically validated in a blinded prospective multicenter trial involving 265 nodules with indeterminate cytology and histologic follow-up.[7] The validation

TABLE 1. Overview of 3 Commercially Available Molecular Tests for Indeterminate Thyroid Fine-Needle Aspiration

	Afirma	ThyGenX/ThyraMIR	ThyroSeq
Company	Veracyte	Interpace Diagnostics	University of Pittsburgh Medical Center, via CBLPath
Methodology	mRNA (gene expression) microarray analysis; classification as either "benign" or "suspicious" gene expression profile by a proprietary algorithm	ThyGenX: multiplex PCR and detection of mutations (BRAF, HRAS, NRAS, and KRAS) and rearrangements (RET-PTC1, RET-PTC3, and PAX8-PPARG) by sequence-specific probes ThyraMIR: microRNA expression analysis; classification as either "negative" or "positive" by a proprietary algorithm	Next-generation sequencing to detect mutations (AKT1, BRAF, CTNNB1, GNAS, HRAS, KRAS, NRAS, PIK3CA, PTEN, RET, TP53, TSHR, TERT, and EIF1AX) and rearrangements (RET, PPARG, NTRK1, NTRK3, BRAF, and ALK)
Strengths	High NPV; validated in blinded multicenter prospective trial	High NPV and PPV (combined tests); ability to stratify risk based on the mutation; potential for prognostic and theranostic information; validated in blinded prospective multicenter study	High NPV and PPV; ability to stratify risk based on the mutation; potential for prognostic and theranostic information
Limitations	Low PPV; concern about performance for Hurthle cell lesions	ThyraMIR is a new test with limited real-world experience to date	New test with limited real-world experience to date; histology diagnoses not blinded to prior molecular testing results in validation study
Cytology interpretation	Performed by centralized cytopathology laboratory[a]	Performed by local cytopathologists	Performed by local cytopathologists or by a centralized laboratory (CBLPath)
Collection kit	Provided by company	Provided by company	Provided by company
Sample required	2 dedicated FNA passes	1 dedicated FNA pass with at least 50 ng of cellular material	1-2 drops from first pass, if sufficiently cellular
Sample stability	−20°C in NA preservative; stable up to 1 y	Room temperature in NA preservative; stable up to 6 wk	−20°C in NA preservative; stable up to 1 y; sample stability is 24 h at 4°C and 6 h at room temperature
Cost[b]	$4875 for Afirma GEC and MTC $975 for Afirma MTC alone $475 for Afirma BRAF alone	$1675 for ThyGenX alone $3300 for ThyraMIR (reflex test)	$3200
Out-of-network maximum cost[b]	$300 for Afirma GEC and MTC $80 for Afirma MTC alone $50 for Afirma BRAF alone	$500 for both tests	$300

Abbreviations: FNA, fine-needle aspiration; GEC, gene expression classifier; mRNA, messenger ribonucleic acid; MTC, medullary thyroid carcinoma; NA, nucleic acid; NPV, negative predictive value; PCR, polymerase chain reaction; PPV, positive predictive value.
[a] A few academic centers have been authorized to submit only the specimen for molecular testing, independently of the cytology specimen.
[b] Prices as of June 2015.

set was comprised of 129 AUS/FLUS, 81 FN/SFN, and 55 "suspicious for malignancy" nodules, which were associated with a cancer prevalence of 24%, 25%, and 62%, respectively. For nodules in the AUS/FLUS and FN/SFN categories, Afirma demonstrated a high NPV of 95% and 94%, respectively, corresponding to a 5% to 6% risk of malignancy. The test's NPV was suboptimal for aspirates in the "suspicious for malignancy" category, reaching only 85% (corresponding to a 15% risk of malignancy) because of to the higher pretest probability of malignancy in this interpretive category. The validation study found that a suspicious GEC result has a relatively low PPV for malignancy among cytologically indeterminate nodules (38% for AUS/FLUS and 37% for FN/SFN). Analysis of the combined AUS/FLUS and FN/SFN data from the validation study is presented in Table 2.[7–13]

Strengths and limitations of the test

For nodules with a pretest probability of malignancy <25%, the high NPV of a benign GEC result reduces the risk of malignancy to <6% , which is a level comparable to that of a cytologically benign aspirate. At this risk level, a patient can be safely triaged toward watchful waiting, with close ultrasound monitoring of the nodule and reaspiration for any nodule that demonstrates significant growth or concerning changes (Fig. 1A). In contrast, a suspicious GEC result is less informative due to its low PPV. In fact, a suspicious GEC result is best considered "still indeterminate"

TABLE 2. Comparison of the Veracyte Afirma Validation Study With Published Postvalidation Experiences on Thyroid Nodules With Indeterminate Cytology

	Veracyte Validation Study, 2012	Alexander 2014[8]	Harrell & Bimston, 2014[9]	McIver 2014[10]	Lastra 2014[11]	Marti 2015 (MSK data)[12a]	Marti 2015 (MSBI data)[12a]	Brauner 2015[13b]
No. of institutions	Multiple	Multiple	Single	Multiple	Single	Multiple	Single	Multiple
No. of cases[c]	210	309	56	60	132	94	71	71
Prevalence of malignancy based on cytology[d]	24%	39%	50%	17%	44%	55%	12%	13%
GEC result								
Benign[e]	87 (41%)	170 (55%)	20 (36%)	16 (27%)	70 (53%)	24 (26%)	37 (52%)	26 (37%)
Suspicious	123 (59%)	139 (45%)	36 (64%)	44 (73%)	62 (47%)	70 (74%)	34 (48%)	45 (63%)
Surgically resected cases	210	123	35	36	50	44	26	46
Benign GEC[f]	87 (100%)	10 (6%)	5 (25%)	4 (25%)	2 (3%)	2 (8%)	5 (14%)	3 (12%)
Suspicious GEC[f]	123 (100%)	113 (81%)	30 (83%)	32 (73%)	48 (77%)	42 (60%)	21 (62%)	43 (96%)
Test results[d]								
True-negative	82	9	4	3	2	2	5	3
False-negative	5	1	1	1	0	0	0	0
False-positive	77	66	13	27	26	18	18	37
True-positive	46	47	17	5	22	24	3	6
Test performance[d]								
Sensitivity	90%	98%	94%	83%	100%	100%	100%	100%
Specificity	52%	12%	24%	10%	7%	10%	22%	8%
NPV	94%	90%	80%	75%	100%	100%	100%	100%
PPV	37%	42%	57%	16%	46%	57%	14%	14%

Abbreviations: GEC, gene expression classifier; MSBI, Mount Sinai Beth Israel; MSK, Memorial Sloan Kettering Cancer Center; NPV, negative predictive value; PPV, positive predictive value.

[a] The analysis by Marti et al includes separate data sets from MSK and MSBI.

[b] The study by Brauner et al involved only cases of atypia of undetermined significance/follicular lesion of undetermined significance or follicular neoplasm/suspicious for a follicular neoplasm with a predominance of Hurthle cells.

[c] The table includes only nodules with atypia of undetermined significance/follicular lesion of undetermined significance or follicular neoplasm/suspicious for a follicular neoplasm cytology and satisfactory Afirma testing results; cases with nondiagnostic Afirma results were excluded.

[d] Prevalence of malignancy, test results, and test performance calculations were based only on resected nodules with reference histologic diagnoses; apart from the Veracyte validation study, true-negative and false-negative results are biased by the low rate of surgery for nodules with benign GEC results in postvalidation studies.

[e] The percentages in parentheses represent the benign call rate, which is an estimate of the percentage of patients with benign nodules who could be spared an unnecessary surgical procedure based on the results of the Afirma GEC.

[f] The percentages in parentheses represent the percentage of nodules with a benign or suspicious GEC that underwent surgical resection.

rather than the equivalent of a cytologic interpretation of "suspicious for malignancy." Therefore, total thyroidectomy should not be recommended based solely on a suspicious GEC result. Instead, Veracyte suggests that the majority of nodules with a suspicious GEC result undergo a diagnostic lobectomy, with the expectation that approximately 60% of these nodules will be found to be histologically benign. For those lobectomy specimens that are diagnosed as malignant, a second surgery (completion thyroidectomy) may be necessary. Altogether, only a benign GEC result has meaningful impact on patient management, but a suspicious GEC result merely confirms the indeterminate nature of the nodule.

Which patients would benefit from the Afirma test?

Afirma is indicated only when the clinical options for a cytologically indeterminate nodule are surgery versus active surveillance. In this regard, the usefulness of Afirma for nodules in the FN/SFN category is clear, because these nodules generally require lobectomy to examine for capsular and vascular invasion. The ability to rule out malignancy preoperatively with the Afirma test can be helpful for identifying which of these nodules can be managed safely by clinical observation, thereby sparing patients unnecessary surgical procedures for benign nodules.

The usefulness of Afirma for nodules in the AUS/FLUS category is not as straightforward because surgery is not necessarily the recommended follow-up. For nodules classified as AUS/FLUS due to suboptimal adequacy, a repeat FNA may be as effective and more economical at resolving the uncertainty of the initial aspirate.[14,15] Likewise, for aspirates placed in the AUS/FLUS category due to variable degrees of cytologic and/or architectural atypia, a consensus cytology review of the slide(s) may help to reclassify the aspirate into a more definitive category.[14]

A subset of AUS/FLUS aspirates with nuclear atypia, as well as those that are persistently categorized as AUS/FLUS on repeat FNA, may have higher malignancy rates that merit a diagnostic lobectomy compared with other types AUS/FLUS cases.[11,16,17] Although Afirma's rule-out function would be helpful in this population, it is uncertain whether the NPV will be low enough to avoid diagnostic surgery in this population, given their higher pretest probability of malignancy.

For patients who have overriding clinical indications for surgical resection of the nodule (eg, nodule size >4 cm, compressive symptoms, or personal preference), Afirma is not indicated.[18] In these patients, a benign GEC result would not alter the decision to pursue surgery, nor would the low PPV of a suspicious GEC result be helpful in guiding the decision between a diagnostic lobectomy and therapeutic total thyroidectomy.

For these reasons, reflex Afirma testing on nodules with AUS/FLUS or FN/SFN cytology (the current practice for specimens that are sent to Veracyte's centralized cytopathology laboratory) may result in overuse of the test. Given the nuances involved in the AUS/FLUS

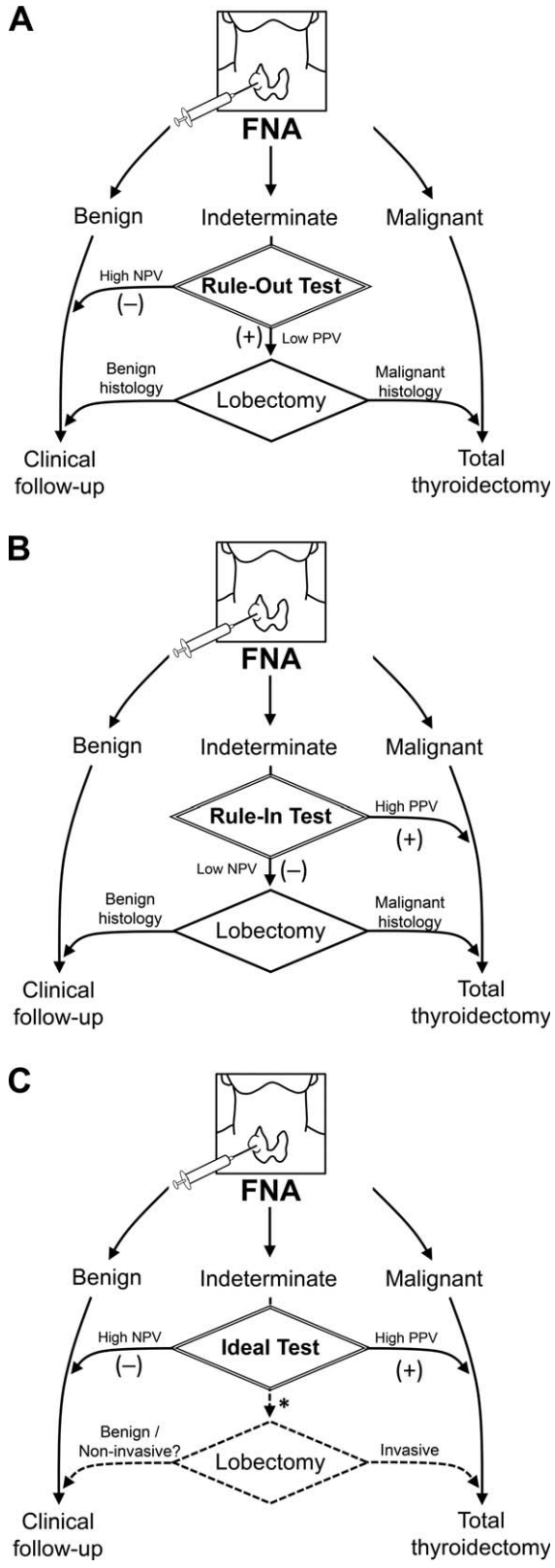

Figure 1. Differences in the usefulness of tests designed to "rule out" and/or "rule in" malignancy for indeterminate thyroid fine-needle aspirations (FNAs). A simplified management algorithm is shown for illustrative purposes, in which patients with cytologically benign nodules are followed clinically, those with cytologically malignant nodules undergo total thyroidectomy, and patients with cytologically indeterminate nodules undergo further testing. (A) Tests optimized to "rule out" malignancy such as the Afirma gene expression classifier (GEC) have a high negative predictive value (NPV); a negative ("benign" on the Afirma GEC) result indicates a low cancer risk, triaging a patient to active surveillance and reducing the rate of lobectomies performed for benign nodules. For tests with a low positive predictive value (PPV), a positive ("suspicious" on the Afirma GEC) result indicates an indeterminate risk of cancer and the consideration of diagnostic lobectomy; patients with nodules with benign histology require only clinical follow-up, whereas those with malignant histology may require a completion thyroidectomy. (B) Tests optimized to "rule in" malignancy such as the 7-gene mutation/fusion panel have a high PPV; a positive test result can direct patients to undergo total thyroidectomy upfront, sparing patients the need to return to the operating room for a completion thyroidectomy. A negative test result cannot exclude malignancy due to the low NPV, and a diagnostic lobectomy would be advised. (C) The ideal molecular test for patients with cytologically indeterminate thyroid nodules should have a sufficiently high NPV and PPV so that a negative result can safely direct patients to watchful waiting and a positive result can direct patients to a total thyroidectomy, thereby reducing the need for a diagnostic lobectomy (indicated by the dotted lines). In the future, management algorithms may consider lobectomy to be appropriate and sufficient management for patients with nodules with molecular signatures suggestive of a noninvasive follicular variant of papillary carcinoma (indicated by the asterisk).

interpretation and the diverse clinical factors that may drive the decision for surgical management, a multidisciplinary approach would be ideal to decide whether Afirma, or any ancillary test, is appropriate for a patient. The cytologic interpretation is but one of several factors that contribute to the pretest probability of malignancy for a given nodule. A personalized approach that also integrates clinical factors (eg, age, sex, nodule size, symptoms, ultrasonographic findings, family history of thyroid cancer, and patient preference for or aversion to surgery) into the equation would achieve a closer approximation of a nodule's pretest probability of malignancy and permit 1) more judicious selection of nodules for molecular testing and 2) better approximation of the NPV and PPV of a benign and suspicious GEC result, respectively, for individual patients.[14,15,18]

Postvalidation studies

Since the publication of Veracyte's clinical validation of the Afirma test in 2012,[7] several groups have published their experiences with the Afirma test outside the clinical trial setting. One of these studies was from a practice that used Veracyte's central cytopathology service,[9] whereas the others were from authorized academic centers that used their own cytopathology services to select cases for Afirma testing.[8,10–13] A summary of these postvalidation studies is shown in Table 2.[8–13]

Although the comparisons of test performance in Table 2[8–13] include only nodules with indeterminate cytology and subsequent resection, a direct comparison of postvalidation studies with the Veracyte clinical validation study is limited by fundamental differences in the populations that compose these data sets. For the thyroid nodules in the Afirma validation study, the decision to resect the nodule was made independently of the Afirma GEC results. In contrast, for the postvalidation studies, the result of the Afirma GEC was itself a factor that influenced the decision regarding whether to operate on a nodule, although the weight given to the Afirma result in this determination is uncertain. In keeping with the purpose of the Afirma test, the vast majority of nodules with benign GEC results were not resected in the postvalidation cohort. Of the published postvalidation series of 363 benign GEC cases to date, 31 nodules (8.5%) have undergone surgery; of these, 3 benign GEC cases were found to be histologically malignant (Table 2).[8–13] These false-negative cases include a 0.6-cm papillary carcinoma,[8] a 3.2-cm follicular carcinoma with "focal capsular and vascular invasion,"[10] and a 2.8-cm cystic papillary thyroid carcinoma.[9] Given the paucity of surgical follow-up for nodules with a benign GEC result, it would be misleading to compare statistical measures that require "true-negative" (eg, specificity and NPV) and "false-negative" (eg, sensitivity and NPV) values for calculation.

With these caveats in mind, the postvalidation studies offer several observations. One point of comparison is the "benign call rate" (BCR), or the percentage of cytologically indeterminate nodules with a benign GEC result. The BCR serves as an estimate of the fraction of patients who could be spared an unnecessary diagnostic lobectomy due to the Afirma test.[12] The BCR in Veracyte's validation study was 41% (Table 2).[7–13] In 2 postvalidation cohorts, the BCR was only 26% to 27%, suggesting that in these populations, the Afirma test may not be sparing as many patients from undergoing an unnecessary lobectomy as initially suggested by the Veracyte validation study.[10,12] In the remaining postvalidation cohorts, the BCR ranged from 36% to 55%, which is similar to or exceeds the BCR of 41% achieved in the Afirma validation study.[8,9,11–13] These studies indicate that the Afirma test has influenced clinical decision-making to varying degrees for patients with cytologically indeterminate thyroid nodules, with lower overall surgical rates compared with untested controls.[13,19]

Several of the postvalidation studies have highlighted other differences between their experiences and the Afirma validation study, particularly with regard to the wide range of the test's PPV among resected nodules (14%-57%) (Table 2).[8–13] Although the Afirma GEC was never intended as a rule-in test, analysis of the PPV raises some concerns regarding test performance. By culling nodules with a benign GEC result from the pool of cytologically indeterminate nodules, Afirma should enrich the malignancy rate among the remaining nodules with suspicious GEC results (Fig. 2A). This expectation would be reflected in a posttest probability of malignancy (estimated by the PPV) that is higher than the pretest probability of malignancy in a tested cohort. The Afirma validation study as well as the majority of the postvalidation studies have shown this to be the case (Table 2).[7–13] However, in 3 postvalidation populations, the PPV remained on par with the baseline prevalence of malignancy, raising concerns that Afirma may not be performing as expected in some settings (Fig. 2B).[10,12,13]

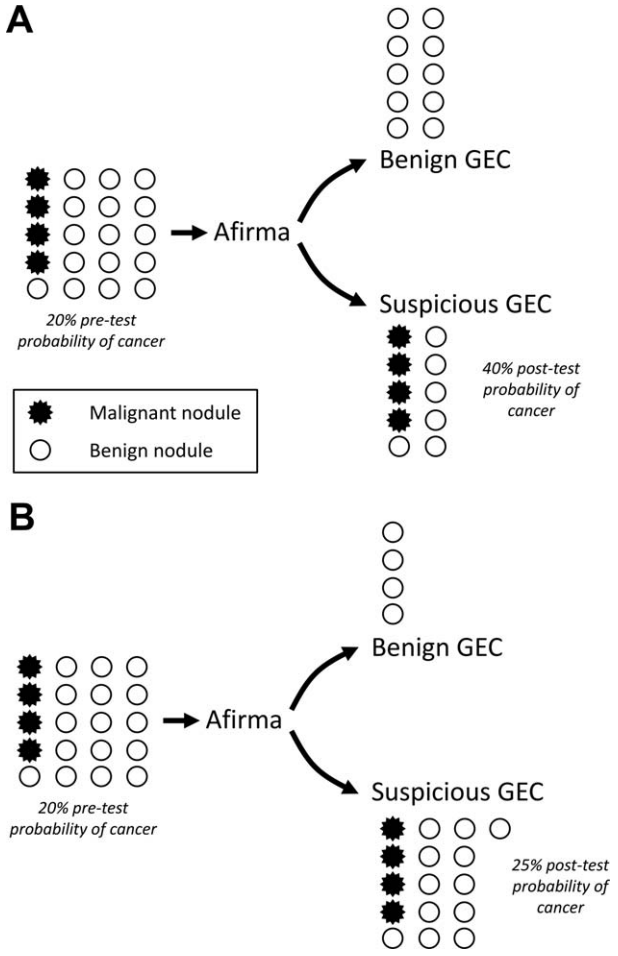

Figure 2. Schematic illustrating how the positive predictive value (posttest probability of malignancy) reflects test performance. (A) In a population with a 20% pretest probability of cancer, the Afirma test is expected to identify nodules with a "benign" gene expression classifier (GEC) profile. By eliminating benign nodules from the pool, the remaining nodules with "suspicious" GEC results are expected to have a higher rate of malignancy (40%). (B) In some studies, the posttest probability of malignancy (25%) has been similar to the pretest probability (20%); one interpretation is that in select settings, the Afirma test may be overcalling benign nodules as having "suspicious" GEC profiles.

In particular, recent reports have emphasized the limitations of the Afirma test for classifying oncocytic follicular (Hurthle cell) nodules. In both the Veracyte clinical validation study as well as in subsequent postvalidation studies, authors have noted a tendency for Afirma to classify a disproportionately high percentage of benign Hurthle cell nodules as having a suspicious GEC.[7,9,11,13,15] Together, these studies indicate that the risk of malignancy for a suspicious GEC result is lower for aspirates with Hurthle cell cytology (19%-23%) compared with those without a prominent population of Hurthle cells. One reason for Afirma's tendency to overcall benign Hurthle cell nodules as having a suspicious GEC stems from the poor concordance among surgical pathologists in classifying Hurthle cell nodules as benign or malignant.[4] Given the lack of a reliably concordant histologic reference, the sensitivity of Afirma's Hurthle cell cassette was augmented to minimize false-negative results.[20] In this way, the Hurthle cell cassette reflects the overarching design of the Afirma GEC: it is optimized to identify patients with benign nodules who can be spared a diagnostic surgical procedure. Whether the test is cost-effective for nodules with Hurthle cell cytology must be examined further, given the smaller percentage of patients spared a diagnostic surgical procedure based on the test results.

Afirma MTC and Afirma *BRAF* tests

In 2014, Veracyte added 2 "malignancy classifiers" to their test menu: Afirma MTC and Afirma BRAF. Both of these tests are mRNA-based classifiers, similar to the Afirma GEC. Afirma MTC has hitherto been a part of the Afirma testing algorithm as 1 of the 6 screening cassettes for uncommon thyroid lesions, before analysis by the main Afirma GEC.[7] Afirma MTC analyzes the expression of 5 genes that are differentially expressed in MTC: calcitonin-related polypeptide-alpha (*CALCA*), carcinoembryonic antigen-related cell adhesion molecule 5 (*CEACAM5*), secretogranin III (*SCG3*), sodium channel voltage-gated type IX alpha subunit (*SCN9A*), and synaptotagmin IV (*SYT4*). Its ability to identify the gene expression signature of MTC with high specificity was recently verified; its recent release as a new test provides a formal reporting mechanism for nodules that trigger the MTC expression cassette.[21] Notably, Afirma MTC does not include testing for mutations in the *RET* oncogene; patients with cytologically or histologically confirmed MTC should undergo a separate genetic test on a blood sample for germline *RET* mutations.[22]

The Afirma BRAF test searches for the gene expression profile of nodules with the *BRAF* V600E mutation. Activating mutations in *BRAF* have been found in approximately 45% of papillary thyroid carcinomas and to a lesser degree in follicular variant papillary carcinomas. Although the absence of a *BRAF* mutation does not exclude malignancy, the detection of an oncogenic *BRAF* mutation rules in malignancy with nearly 100% specificity.[23] A validation study by Veracyte demonstrated that

the Afirma BRAF test has clinical sensitivity and specificity comparable to a polymerase chain reaction-based test for *BRAF* V600E mutations.[24]

Although a suspicious GEC result alone is not particularly helpful due to its low PPV for malignancy, a positive Afirma MTC or Afirma BRAF test in this setting can rule in MTC or papillary thyroid carcinoma, respectively, directing these patients toward appropriate surgical and oncologic management. Veracyte suggests that these malignancy classifiers also may be useful for some nodules that are cytologically "suspicious for malignancy" (SFM) or "positive for malignancy." For nodules that are suspicious but not definitive for MTC by FNA, confirmation of the diagnosis by Afirma MTC can prompt testing for multiple endocrine neoplasia type 2, including germline *RET* mutations and evaluation for pheochromocytoma and hyperparathyroidism.[22] The usefulness of knowing *BRAF* mutational status for a cytologically SFM or "positive" nodule is controversial; although some studies have correlated *BRAF* V600E mutations with aggressive behavior such as increased lymph node metastasis and extrathyroidal extension, data supporting the role of prophylactic central lymph node dissection based on the preoperative detection of *BRAF* mutations are currently lacking.[25–27] The cost-effectiveness of the Afirma malignancy classifiers compared with alternative methods for detecting MTC (eg, cell block with immunohistochemistry or calcitonin measurement in aspirates) and *BRAF* V600E mutations (polymerase chain reaction-based DNA tests) remains to be established.

The "Rule-In" Approach: Mutational Panels

A variety of oncogenic mutations and gene fusions involving the mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K) signaling cascades have been identified in thyroid cancer.[28] For example, approximately 45% of papillary thyroid carcinomas harbor oncogenic mutations in *BRAF*, and an additional 20% are driven by chromosomal translocations involving the *RET* protooncogene; similarly, 40% of follicular thyroid carcinomas possess activating *RAS* mutations and 35% have chromosomal translocations that result in paired box gene 8/peroxisome proliferator-activated receptor gamma (*PAX8/PPARγ*) gene fusion. These genetic changes demonstrate a high specificity for malignancy, ranging from 70% to 80% (*RAS* mutations) to nearly 100% (*BRAF* mutations and *RET/PTC* fusions). Based on these findings, several studies have explored the value of using one or a panel of these markers for "ruling in" malignancy among cytologically indeterminate thyroid nodules.[29–35]

Test design

Clinical versions of the mutation/translocation panel require at least 50 ng of cellular material from 1 dedicated FNA pass to be collected into a tube of nucleic acid preservative solution, in addition to passes required for cytologic examination. For those nodules diagnosed as cytologically indeterminate, the nucleic acid is assayed for the most common oncogenic mutations in *BRAF, KRAS, HRAS, NRAS*, and chromosomal translocations resulting in *RET/PTC1, RET/PTC3*, and *PAX8/PPARγ* fusions. A commercially available version of this 7-gene panel was initially marketed as the miR*Inform* test (Asuragen, Austin, Tex). The test is currently offered by Interpace Diagnostics (Parsippany, NJ) as the ThyGenX test (Table 1). Cytopathologic interpretation is performed locally, with only the material for molecular testing submitted to Interpace Diagnostics.

Test validation

The initial 7-gene panel has been evaluated for thyroid FNAs in several studies (Table 3).[29,30,33,34,36–38] The largest of these was a prospective single-institution study of 513 thyroid nodules with indeterminate cytology and histologic follow-up.[33] The validation set included 247 AUS/FLUS, 214 FN/SFN, and 52 SFM aspirates. The prevalence of malignancy for each of these cytologic interpretive categories was 14%, 27%, and 54%, respectively. The detection of any of the 7 mutations or fusions increased the cancer risk (PPV) to 88%, 87%, and 95%, respectively, for AUS/FLUS, FN/SFN, and SFM aspirates. The absence of a mutation or fusion was associated with a cancer risk of 6% (NPV, 94%), 14% (NPV, 86%), and 28% (NPV, 72%), respectively, for AUS/FLUS, FN/SFN, and SFM aspirates. The pathologists establishing the reference histology diagnoses were aware of the preoperative molecular testing results in most cases, possibly introducing workup bias into this study. Additional validation studies using the same 7-gene panel, including a blinded, prospective, multiinstitutional study sponsored by Asuragen, have confirmed the high specificity and PPV of these oncogenic markers among indeterminate thyroid nodules (Table 3).[29,30,33,34,36–38]

TABLE 3. Comparison of Published Experiences With the 7-Gene Mutation/Fusion Panels on Thyroid Nodules With Indeterminate Cytology[a]

	Nikiforov 2009[34]		Cantara 2010[30b]	Nikiforov 2011[33]		Beaudenon-Huibregtse 2014[29]		Eszlinger 2014[36c]	Eszlinger 2015[37]	Labourier 2015[38d]
Sample collection	Prospective		Prospective	Prospective		Prospective		Retrospective	Retrospective	Prospective
No. of institutions	Single		Single	Single		Multiple		Single	Single	Multiple
Cytologic category	AUS/FLUS	FN/SFN	Indeterminate	AUS/FLUS	FN/SFN	AUS/FLUS	FN/SFN	Indeterminate	Thy3 (FN/SFN)	AUS/FLUS and FN/SFN
No. of cases	21	23	41	247	214	22	19	141	163	109
Prevalence of malignancy based on cytology[e]	14%	52%	17%	14%	27%	50%	32%	16%	28%	32%
Results										
Negative	18	14	34	222	176	16	14	120	127	75
Positive	3	9	7	25	38	6	5	21	31	34
Test results[e]										
True-negative	18	11	33	209	151	9	12	102	104	64
False-negative	0	3	1	13	25	7	2	18	23	11
False-positive	0	0	1	3	5	2	1	17	9	10
True-positive	3	9	6	22	33	4	4	4	22	24
Test performance[e]										
Sensitivity	100%	75%	86%	63%	57%	36%	67%	18%	49%	69%
Specificity	100%	100%	97%	99%	97%	82%	92%	86%	92%	86%
NPV	100%	79%	97%	94%	86%	56%	86%	85%	82%	85%
PPV	100%	100%	86%	88%	87%	67%	80%	19%	71%	71%

Abbreviations: AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; FN/SFN, follicular neoplasm/suspicious for a follicular neoplasm cytology; NPV, negative predictive value; PPV, positive predictive value.

[a] Only data for nodules with indeterminate cytology (AUS/FLUS, FN/SFN, or comparable categories), satisfactory molecular testing results using a 7-gene panel (BRAF, HRAS, KRAS, NRAS, RET/PTC1, RET/PTC3, and PAX8/PPARG), and follow-up reference histology were extracted from the above studies for this table. Cases with nondiagnostic molecular testing results or those lacking follow-up surgical resection were excluded.

[b] The study by Cantara et al also tested for gene fusions involving TRK. The authors used a 4-category cytology reporting system ("inadequate," "benign," "indeterminate," and "suspicious for thyroid cancer").

[c] The malignant cases in the data set in the 2014 study by Eszlinger et al were enriched for follicular thyroid carcinomas relative to papillary carcinomas. The authors indicated that cytologically diagnostic specimens were reported in 3 categories ("benign," "indeterminate," and "malignant").

[d] The study by Labourier et al compared performance of the 7-gene panel with and without a microRNA expression classifier. Only the data for the 7-gene panel were extracted here. Some of the data may overlap with those of Beaudenon-Huibregtse et al.

[e] Prevalence of malignancy, test results, and test performance calculations based on resected nodules only.

Strengths and limitations of the test

By ruling in malignancy preoperatively with its high PPV, the 7-gene panel identifies a subset of patients with cytologically indeterminate nodules who would benefit from a total thyroidectomy upfront, thereby reducing the need for a 2-stage surgical procedure (ie, diagnostic lobectomy followed by completion thyroidectomy) (Fig. 1B). For practices that regularly recommend total thyroidectomy for nodules with "suspicious for malignancy" cytology, the preoperative detection of mutations/rearrangements may not significantly alter management decisions.

In contrast to the proprietary nature of the Afirma GEC, mutational panels offer the advantage of revealing which oncogenic mutations and/or gene fusions are present or absent. Because different mutations are associated with different malignancy risks, such panels may provide more granular information compared with the binary "benign" or "suspicious" Afirma GEC results. In particular, the preoperative detection of *BRAF* V600E mutations and *RET/PTC* rearrangements has been associated with a 100% risk of papillary thyroid carcinoma (PTC), not only in cases with indeterminate cytology but also in cytologically nondiagnostic and benign cases.[29,30,33,34,37,39] The impact of preoperatively detecting *RAS* mutations or *PAX8-PPARγ* fusions is evolving. Among cytologically indeterminate aspirates, the PPVs for *RAS* mutations and *PAX8-PPARγ* rearrangements vary from 57% to 100%.[29,30,33,34] Although a "false-positive" rate has been ascribed to both of these molecular changes, some of these benign resection cases may represent a preinvasive form of carcinoma or minimally invasive carcinomas with subtle evidence of capsular or vascular invasion that was missed on initial histologic examination.[37,40–42] Furthermore, mutational panels have the added advantage of being customizable to different preparation types; several studies have demonstrated the feasibility of DNA-based molecular

analyses on residual ThinPrep material as well as air-dried smears.[36,37,43]

For nodules with indeterminate cytology, a mutation-negative result using this 7-gene panel reduces the probability of malignancy by varying degrees. In cohorts in which the pretest probability of malignancy for indeterminate thyroid FNA was low (14%-17%), a negative test result was associated with a <6% risk of malignancy, a finding that is comparable to that of a cytologically benign FNA or a cytologically indeterminate FNA with a benign Afirma GEC result (Table 3).[30,33,34] However, the results of these validation studies may not be applicable to every practice setting due to differences in the pretest probability of malignancy of tested populations. In cases in which the pretest probability of malignancy is not known or is much higher than 15%, a negative mutation panel may not be sufficient to forgo diagnostic lobectomy. As a case in point, the 22 AUS/FLUS cases with histologic follow-up in the study by Beaudenon-Huibregtse et al demonstrated a pretest probability of malignancy of 50%; the risk of malignancy for a mutation-negative test result was 44% in this cohort.[29]

New Horizons: The Quest for a Comprehensive Rule-In and Rule-Out Test

The initial molecular tests for indeterminate thyroid FNAs were limited by either low PPV (Afirma GEC) or low NPV (7-gene panel). Ideally, an ancillary test for indeterminate thyroid FNAs would have sufficient predictive power to either rule in or rule out cancer with its test result, thereby obviating the need for a diagnostic lobectomy (Fig. 1C). Early studies have indicated 2 strategies toward this all-inclusive test: 1) ThyroSeq, a next-generation sequencing (NGS)-based platform that expands the list of oncogenic mutations and gene fusions; and 2) a hybrid approach combining ThyraMIR, a microRNA gene expression classifier, with the existing 7-gene mutation/fusion panel.

ThyroSeq: A NGS-Based Gene Mutation and Fusion Panel

Test design

NGS technology provides a high-throughput, cost-effective, and analytically sensitive mechanism for sequencing multiple targeted portions of the genome in parallel. This development, in conjunction with the discovery of novel driver mutations in thyroid cancer,[28,44–47] has propelled the addition of numerous oncogenes to the original 7-gene panel, including numerous hotspot mutations in *PIK3CA*, *PTEN, TP53, TSHR, CTNNB1, RET, AKT1*, and *TERT*, as well as a gene fusions involving *RET, BRAF, NTRK1, NTRK3, AKT, PPARγ*, and *THADA* to various partner genes.[48,49] The test panel also included a *GNAS* mutation associated with benignity. The latest version of this expanded panel, ThyroSeq v2, is commercially offered by CBLPath (Rye Brook, NY), with the test performed and interpreted in the Division of Molecular and Genomic Pathology at the University of Pittsburgh Medical Center. CBLPath offers thyroid FNA interpretation, but for practices that choose to use their local cytopathology interpretative services, specimens can be submitted for molecular testing alone. ThyroSeq requires only 10 ng of input DNA, which may be extracted from a sufficiently cellular first FNA pass (Table 1).[50] Although the discussion below focuses on an NGS panel customized for thyroid cancer, the feasibility of using a commercially available primer pool for sequencing generic cancer genes has also been explored in a limited cohort of indeterminate thyroid FNAs.[51]

Test validation

Performance of the ThyroSeq v2 panel in thyroid FNAs was evaluated in a single-institution study involving 143 aspirates with FN/SFN cytology.[48] The cohort tested included a mixture of 91 retrospectively and 52 prospectively collected samples with an overall prevalence of malignancy of 27%. Combining the retrospective and prospective groups, ThyroSeq v2 demonstrated a sensitivity of 90%, a specificity of 93%, an NPV of 96%, and a PPV of 83% (Table 4).[48] Point mutations in *HRAS* (2 cases), *BRAF* V600E (1 case), *TERT* (4 cases), *TP53* (1 case), *PIK3CA* (1 case), and any gene fusion (11 cases) were associated with cancer in 100% of cases. The surgical pathologists making the benchmark histologic diagnoses on the resections were aware of the results of either the 7-gene panel or an earlier version of ThyroSeq, possibly introducing workup bias into the study.

Strengths and limitations of the test

The results of this initial study indicate that ThyroSeq v2 may perform well as both a "rule-out" and "rule-in" test for a subset of nodules with indeterminate cytology. Using Bayes' theorem to extrapolate the NPV and PPV based on cancer prevalence, the authors demonstrated that at pretest probabilities of malignancy between 14% and 34%, ThyroSeq v2 would maintain a high NPV (95%-98%)

TABLE 4. Comparison of NGS-Based Mutational Panels Versus Combined MicroRNA/Mutational Panel[a]

	Nikiforov 2014[48]	Labourier 2015[38]			Le Mercier 2015[51b]
Test	NGS-based, thyroid-specific mutation/gene fusion panel (ThyroSeq)	MicroRNA-based expression classifier (ThyraMIR) and 7-gene mutation panel (ThyGenX)			NGS-based mutation panel of generic cancer genes (AmpliSeq Cancer Hotspot Panel)
Sample collection	Retrospective and prospective	Prospective			Retrospective
No. of institutions	Single	Multiple			Single
Cytologic category	FN/SFN	AUS/FLUS and FN/SFN			Indeterminate ("follicular proliferation")[b]
Indeterminate FNA with molecular test	143	109			34
Prevalence of malignancy based on cytology[c]	27%	32%			21%
Results		miRNA Classifier	7-Gene Panel	Both Tests Combined	
Negative[d]	101 (71%)	83 (76%)	75 (69%)	67 (61%)	26 (76%)
Positive	42 (29%)	26 (24%)	34 (31%)	42 (39%)	8 (24%)
Test results[c]					
True-negative	97	68	64	63	24
False-negative	4	15	11	4	2
False-positive	7	6	10	11	3
True-positive	35	20	24	31	5
Test performance[c]					
Sensitivity	90%	57%	69%	89%	71%
Specificity	93%	92%	86%	85%	89%
NPV	96%	82%	85%	94%	92%
PPV	83%	77%	71%	74%	63%

Abbreviations: AUS/FLUS, atypia of undetermined significance/follicular lesion of undetermined significance; FN/SFN, follicular neoplasm/suspicious for a follicular neoplasm cytology; FNA, fine-needle aspiration; NGS, next-generation sequencing; NPV, negative predictive value; PPV, positive predictive value.
[a] Only data for nodules with indeterminate cytology (AUS/FLUS, FN/SFN, or comparable categories), satisfactory molecular testing results, and follow-up reference histology were extracted from the listed studies.
[b] The study by Le Mercier et al used 4 cytologic interpretive categories: "unsatisfactory," "benign," "follicular proliferation," and "malignant."
[c] Prevalence of malignancy, test results, and test performance calculations based only on resected nodules with reference histologic diagnoses.
[d] Percentages represent the benign call rate, or the estimate of patients with benign nodules who could be spared an unnecessary surgical procedure based on the results of the molecular test. Note that the corresponding NPV should be >94% to be comparable to the NPV of a cytologically benign nodule.

and high PPV (68%-87%). At these predictive values, a negative test could safely triage a patient toward active surveillance of a nodule. Conversely, a detection of a mutation, particularly those that are highly predictive of malignancy (eg, *BRAF* V600E, *TERT, TP53, PIK3CA*, and any gene rearrangement) could direct patients toward total thyroidectomy.

Would ThyroSeq v2 perform as well for entities in the AUS/FLUS category? The answer depends on the pretest probability of malignancy for AUS/FLUS cases within a particular practice. With a pretest probability of malignancy in the range of 5% to 15% for AUS/FLUS, ThyroSeq v2 would be expected to have an NPV of 98% to 99% and a PPV of 40% to 69% based on Bayesian modeling. At the lower end of this range of pretest probabilities, ThyroSeq v2 would remain an effective rule-out test but may be less helpful as a rule-in test due to the sharp decline in PPV. In their earlier validation study of the 7-marker panel, Nikiforov et al achieved an NPV of 94% and a PPV of 88% among AUS/FLUS cases with a 14% pretest probability of malignancy.[33] The superior PPV of the smaller molecular panel can be explained in part by its higher specificity (97%-99%) compared with the ThyroSeq v2 panel (93%). Kennedy et al similarly cautioned that the enhanced sensitivity gained from broadening NGS-based mutational platforms should be balanced with the risk of detecting "false-positive" molecular abnormalities in histologically benign nodules.[52]

Regardless of these limitations, as experience with this expanded NGS panel increases, the data it provides are expected to refine preoperative risk stratification based both on the types of gene mutation or fusions that are detected as well as their allelic frequency. Furthermore, although most of the attention for molecular testing in thyroid FNA has been directed toward improving diagnostic certainty, genomic data can also provide prognostic and predictive information regarding tumors.

ThyraMIR: A MicroRNA-Based GEC

MicroRNAs are small, noncoding ribonucleic acids that regulate gene expression at a posttranscriptional level. Certain microRNAs demonstrate differential expression patterns in benign versus malignant thyroid tumors.[44,53–61] Labourier et al have explored the usefulness of combining the 7-gene mutational panel with a gene expression classifier involving 10 microRNAs: miR-29-b-1-5p, miR-31-5p, miR-138-1-3p, miR-139-5p, miR-146b-5p, miR-155, miR-204-5p, miR-222-3p, miR-375, and miR-551b-3p.[38] Similar in concept to the Afirma GEC, the microRNA classifier renders a "positive" or "negative" result based on an algorithm trained on the microRNA expression profiles of histologically benign and malignant reference thyroid nodules. Interpace Diagnostics offers the ThyraMIR microRNA expression classifier as a reflex test on aspirates that are negative for the 7-gene ThyGenX mutation panel (Table 1).

Test validation

A cross-sectional cohort study of 109 cytologically indeterminate (AUS/FLUS and FN/SFN) thyroid nodules demonstrated that the combination of the 7-gene mutational panel with the microRNA expression classifier achieved a sensitivity of 89%, a specificity of 85%, an NPV of 94%, and a PPV of 74% (Table 4).[38]

Strengths and limitations of the test

Similar to other tests discussed in this review, the test performance of ThyraMIR and ThyGenX will vary with the prevalence of malignancy in the tested population. Using Bayes' theorem, we can predict that the test would be optimal for patients whose nodules fall within a narrow range of pretest probabilities of malignancy. At a pretest probability of malignancy <30%, the tests would achieve an effective "rule-out" NPV of >94%; conversely, a pretest probability of malignancy >20% would be necessary to maintain a PPV of ≥60%.[38]

Considerations for the Current State of Molecular Testing in Thyroid FNAs

As the above studies indicate, achieving a high NPV remains a key goal of molecular testing. In this regard, time remains the biggest limitation in our ability to assess the performance of these tests. Because thyroid nodules with mutation-negative or benign GEC test results are rarely resected, to our knowledge the false-negative rates of these tests in clinical practice are currently unknown. Although the current data indicate that these tests reduce surgical rates for indeterminate thyroid nodules, the vital question of whether these tests are reducing unnecessary surgeries remains unanswered at this early stage. Until longitudinal follow-up studies of these unresected nodules are available, close clinical monitoring of nodules with negative molecular testing results is strongly recommended.

An additional challenge to assessing test performance is the recognition that histology is an imperfect gold standard for validating these molecular studies. The interobserver variation is high in the histologic classification of some thyroid tumors.[4] Reasons for discrepant diagnoses among pathologists include differing thresholds for interpreting capsular and vascular invasion, as well as variability in judging whether a follicular-patterned thyroid tumor has sufficient nuclear atypia for the diagnosis of the follicular variant of PTC (FVPTC). With regard to the latter issue, ongoing efforts to reclassify noninvasive FVPTCs will impact the landscape of molecular testing for indeterminate thyroid FNAs. Reclassification of noninvasive FVPTCs as neoplasms with low malignant potential rather than carcinomas would reduce the prevalence of malignancy among indeterminate FNAs, which would in turn affect the predictive values of these tests.[62] Furthermore, the genotype of FVPTCs appears to correlate with biologic behavior; noninvasive FVPTCs tend to harbor *RAS* mutations, *BRAF* K601E mutations, and *PAX8/PPARγ* fusions, whereas invasive FVPTCs demonstrate an increased rate of *BRAF* V600E mutations.[44,46,63–66] Therefore, test panels that preoperatively distinguish between "*RAS*-like" and "*BRAF* V600E-like" molecular profiles will become increasingly relevant in guiding appropriate treatment options; lobectomy may be appropriate initial management for the former, whereas total thyroidectomy would be indicated for the latter (Fig. 1C).

Molecular testing has shown great promise in reducing the diagnostic uncertainty of cytologically indeterminate thyroid nodules, but it is one of many factors that contribute to the overall probability of malignancy for a patient. Accordingly, the decision to use ancillary molecular testing, the selection of the appropriate molecular test, and the interpretation of its results should always be performed within the context of cytologic, clinical, and ultrasonographic findings.

FUNDING SUPPORT

No specific funding was disclosed.

CONFLICT OF INTEREST DISCLOSURES

The author made no disclosures.

REFERENCES

1. Silverman ME, Murray TJ, Bryan CS. The Quotable Osler. Philadelphia: American College of Physicians; 2008.
2. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol*. 2012;56:333-339.
3. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2009;132:658-665.
4. Cibas ES, Baloch ZW, Fellegara G, et al. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann Intern Med*. 2013;159:325-332.
5. Gerhard R, da Cunha Santos G. Inter- and intraobserver reproducibility of thyroid fine needle aspiration cytology: an analysis of discrepant cases. *Cytopathology*. 2007;18:105-111.
6. Renshaw AA, Wang E, Wilbur D, et al. Interobserver agreement on microfollicles in thyroid fine-needle aspirates. *Arch Pathol Lab Med*. 2006;130:148-152.
7. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;367:705-715.
8. Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma gene expression classifier. *J Clin Endocrinol Metab*. 2014;99:119-125.
9. 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;20:364-369.
10. McIver B, Castro MR, Morris JC, et al. An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2014;99:4069-4077.
11. 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 (Cancer Cytopathol)*. 2014;122:737-744.
12. Marti JL, Avadhani V, Donatelli LA, et al. Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules [published online ahead of print April 11, 2015]. *Ann Surg Oncol*.
13. Brauner E, Holmes BJ, Krane JF, et al. Performance of the Afirma gene expression classifier in Hurthle cell thyroid nodules differs from other indeterminate thyroid nodules. *Thyroid*. 2015;25:789-796.
14. Faquin WC. Can a gene-expression classifier with high negative predictive value solve the indeterminate thyroid fine-needle aspiration dilemma? *Cancer (Cancer Cytopathol)*. 2013;121:116-119.
15. Krane JF. Lessons from early clinical experience with the Afirma gene expression classifier. *Cancer (Cancer Cytopathol)*. 2014;122:715-719.
16. Nishino M, Wang HH. Should the thyroid AUS/FLUS category be further stratified by malignancy risk? *Cancer (Cancer Cytopathol)*. 2014;122:481-483.
17. VanderLaan PA, Marqusee E, Krane JF. Clinical outcome for atypia of undetermined significance in thyroid fine-needle aspirations: should repeated FNA be the preferred initial approach? *Am J Clin Pathol*. 2011;135:770-775.
18. Parangi S, Suh H. The role of genetic markers in the evaluation and management of thyroid nodules. *Surg Clin North Am*. 2014;94:515-528.
19. 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;22:996-1001.
20. Steward DL, Kloos RT. Clinical diagnostic gene expression thyroid testing. *Otolaryngol Clin North Am*. 2014;47:573-593.
21. Kloos RT, Kennedy GC, Monroe R, Traweek T, Lanman R. Validation of a molecular classifier for preoperative identification of medullary thyroid cancer in thyroid nodule fine-needle aspiration biopsies. Presented at: American Association of Clinical Endocrinologists 23rd Annual Scientific and Clinical Congress; May 14-18 2014; Las Vegas, NV.
22. Wells SA Jr, Asa SL, Dralle H, et al; American Thyroid Association Guidelines Task Force on Medullary Thyroid Carcinoma. Revised American Thyroid Association guidelines for the management of medullary thyroid carcinoma. *Thyroid*. 2015;25:567-610.
23. Xing M. BRAFV600E mutation and papillary thyroid cancer: chicken or egg? *J Clin Endocrinol Metab*. 2012;97:2295-2298.
24. Diggans J, Kim SY, Hu Z, et al. Machine learning from concept to clinic: reliable detection of BRAF V600e DNA mutations in thyroid nodules using high-dimensional RNA expression data. *Pac Symp Biocomput*. 2015:371-382.
25. Gouveia C, Can NT, Bostrom A, Grenert JP, van Zante A, Orloff LA. Lack of association of BRAF mutation with negative prognostic indicators in papillary thyroid carcinoma: the University of California, San Francisco, experience. *JAMA Otolaryngol Head Neck Surg*. 2013;139:1164-1170.
26. Papp S, Asa SL. When thyroid carcinoma goes bad: a morphological and molecular analysis. *Head Neck Pathol*. 2015;9:16-23.
27. Xing M, Alzahrani AS, Carson KA, et al. Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *JAMA*. 2013;309:1493-1501.
28. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol*. 2011;7:569-580.
29. 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;24:1479-1487.
30. 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;95:1365-1369.
31. Ferraz C, Rehfeld C, Krogdahl A, et al. Detection of PAX8/PPARG and RET/PTC rearrangements is feasible in routine air-dried fine needle aspiration smears. *Thyroid*. 2012;22:1025-1030.
32. 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;34:2589-2594.
33. 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;96:3390-3397.
34. Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab*. 2009;94:2092-2098.
35. Ohori NP, Nikiforova MN, Schoedel KE, et al. Contribution of molecular testing to thyroid fine-needle aspiration cytology of "follicular lesion of undetermined significance/atypia of undetermined significance." *Cancer (Cancer Cytopathol)*. 2010;118:17-23.
36. Eszlinger M, Krogdahl A, Munz S, et al. Impact of molecular screening for point mutations and rearrangements in routine air-dried fine-needle aspiration samples of thyroid nodules. *Thyroid*. 2014;24:305-313.
37. 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;25:401-409.
38. 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;100:2743-2750.
39. 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;97:2299-2306.
40. Eberhardt NL, Grebe SK, McIver B, Reddi HV. The role of the PAX8/PPARgamma fusion oncogene in the pathogenesis of follicular thyroid cancer. *Mol Cell Endocrinol*. 2010;321:50-56.
41. Giordano TJ, Beaudenon-Huibregtse S, Shinde R, et al. Molecular testing for oncogenic gene mutations in thyroid lesions: a case-control validation study in 413 postsurgical specimens. *Hum Pathol*. 2014;45:1339-1347.
42. Nikiforova MN, Lynch RA, Biddinger PW, et al. RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. *J Clin Endocrinol Metab*. 2003;88:2318-2326.
43. Krane JF, Cibas ES, Alexander EK, Paschke R, Eszlinger M. Molecular analysis of residual ThinPrep material from thyroid FNAs increases diagnostic sensitivity. *Cancer (Cancer Cytopathol)*. 2015;123:356-361.
44. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;159:676-690.
45. Kelly LM, Barila G, Liu P, et al. Identification of the transforming STRN-ALK fusion as a potential therapeutic target in the aggressive forms of thyroid cancer. *Proc Natl Acad Sci U S A*. 2014;111:4233-4238.
46. McFadden DG, Dias-Santagata D, Sadow PM, et al. Identification of oncogenic mutations and gene fusions in the follicular variant of papillary thyroid carcinoma. *J Clin Endocrinol Metab*. 2014;99:E2457-E2462.
47. Ricarte-Filho JC, Li S, Garcia-Rendueles ME, et al. Identification of kinase fusion oncogenes in post-Chernobyl radiation-induced thyroid cancers. *J Clin Invest*. 2013;123:4935-4944.
48. 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;120:3627-3634.
49. 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:E1852-E1860.
50. Division of Molecular and Genomic Pathology, University of Pittsburgh Medical Center. ThyroSeq®-thyroid cancer next-generation sequencing panel. Available at: http://path.upmc.edu/divisions/map/PDF/Thyroseq_details.pdf. Accessed July 22, 2015.
51. Le Mercier M, D'Haene N, De Neve N, et al. Next-generation sequencing improves the diagnosis of thyroid FNA specimens with indeterminate cytology. *Histopathology*. 2015;66:215-224.
52. Kennedy GC, Pagan M, Walsh PS, et al. Detection of increasing numbers of point mutations and fusions in thyroid nodules using next-generation sequencing dramatically reduces marker panel specificity. Presented at: Endocrine Society 97th Annual Meeting and Expo; March 5-8, 2015; San Diego, CA.
53. Eszlinger M, Paschke R. Molecular fine-needle aspiration biopsy diagnosis of thyroid nodules by tumor specific mutations and gene expression patterns. *Mol Cell Endocrinol*. 2010;322:29-37.
54. He H, Jazdzewski K, Li W, et al. The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci U S A*. 2005;102:19075-19080.
55. Lassalle S, Hofman V, Ilie M, et al. Can the microRNA signature distinguish between thyroid tumors of uncertain malignant potential and other well-differentiated tumors of the thyroid gland? *Endocr Relat Cancer*. 2011;18:579-594.
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;93:1600-1608.
57. Pallante P, Visone R, Ferracin M, et al. MicroRNA deregulation in human thyroid papillary carcinomas. *Endocr Relat Cancer*. 2006;13:497-508.
58. Reddi HV, Madde P, Milosevic D, et al. The putative PAX8/PPARγ fusion oncoprotein exhibits partial tumor suppressor activity through up-regulation of micro-RNA-122 and dominant-negative PPARγ activity. *Genes Cancer*. 2011;2:46-55.
59. Rossing M. Classification of follicular cell-derived thyroid cancer by global RNA profiling. *J Mol Endocrinol*. 2013;50:R39-R51.
60. Rossing M, Borup R, Henao R, et al. Down-regulation of microRNAs controlling tumourigenic factors in follicular thyroid carcinoma. *J Mol Endocrinol*. 2012;48:11-23.
61. Weber F, Teresi RE, Broelsch CE, Frilling A, Eng C. A limited set of human microRNA is deregulated in follicular thyroid carcinoma. *J Clin Endocrinol Metab*. 2006;91:3584-3591.
62. 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 [published online ahead of print July 29, 2015]. *Thyroid*.
63. Cho U, Oh WJ, Bae JS, et al. Clinicopathological features of rare BRAF mutations in Korean thyroid cancer patients. *J Korean Med Sci*. 2014;29:1054-1060.
64. Howitt BE, Jia Y, Sholl LM, Barletta JA. Molecular alterations in partially-encapsulated or well-circumscribed follicular variant of papillary thyroid carcinoma. *Thyroid*. 2013;23:1256-1262.
65. Ohori NP, Singhal R, Nikiforova MN, et al. BRAF mutation detection in indeterminate thyroid cytology specimens: underlying cytologic, molecular, and pathologic characteristics of papillary thyroid carcinoma. *Cancer (Cancer Cytopathol)*. 2013;121:197-205.
66. 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;23:1191-1200.