2.04.78	Molecular Markers in Fine Needle Aspiration of the Thyroid		
Original Policy Date:	November 26, 2014	**Effective Date:**	October 1, 2025
Section:	2.0 Medicine	**Page:**	Page 1 of 53

Policy Statement

I. For individuals who have thyroid nodules without strong clinical or radiologic findings suggestive of malignancy in whom surgical decision making would be affected by test results, the use of **either** of the following types of molecular marker testing or gene variant analysis in fine needle aspirates of thyroid nodules with indeterminate cytologic findings (i.e., Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/ suspicion for a follicular neoplasm]) may be considered **medically necessary**:
   A. Afirma® Genomic Sequencing Classifier
   B. ThyroSeq®

II. The use of **any** of the following types of molecular marker testing or gene variant analysis in fine needle aspirates of thyroid nodules with indeterminate findings (Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/suspicion for a follicular neoplasm]) or suspicious findings (Bethesda diagnostic category V [suspicious for malignancy]) to rule in malignancy to guide surgical planning for initial resection rather than a 2-stage surgical biopsy followed by definitive surgery may be considered **medically necessary**:
   A. ThyroSeq
   B. ThyraMIR® microRNA/ThyGenX®
   C. Afirma BRAF after Afirma Genomic Sequencing Classifier
   D. Afirma MTC after Afirma Genomic Sequencing Classifier

III. Gene expression classifiers, genetic variant analysis, and molecular marker testing in fine needle aspirates of the thyroid not meeting criteria outlined above, including but not limited to use of single-gene telomerase reverse transcriptase (*TERT*) testing, are considered **investigational**.

NOTE: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

Policy Guidelines

In individuals who do not undergo surgical biopsy or thyroidectomy on the basis of gene expression classifier or molecular marker results, regular active surveillance is indicated.

Use of molecular marker testing based on fine needle aspirate of a thyroid nodule to rule in malignancy prior to surgical biopsy may guide surgical planning, particularly factors such as choice of surgical facility provider to ensure that the capability is available to conduct a frozen section pathologic reading during surgical biopsy so that surgical approach may be adjusted accordingly in a single surgery.

Genetic Counseling

Experts recommend formal genetic counseling for individuals who are at risk for inherited disorders and who wish to undergo genetic testing. Interpreting the results of genetic tests and understanding risk factors can be difficult for some individuals ; genetic counseling helps individuals understand the impact of genetic testing, including the possible effects the test results could have on the individual or their family members. It should be noted that genetic counseling may alter the utilization of genetic testing substantially and may reduce inappropriate testing; further, genetic counseling should be performed by an individual with experience and expertise in genetic medicine and genetic testing methods.

Coding

See the Codes table for details.

Description

To determine which patients need thyroid resection, many physicians will perform a cytologic examination of fine needle aspirate (FNA) samples from a thyroid lesion; however, this method has diagnostic limitations. As a result, assays using molecular markers have been developed to improve the accuracy of thyroid FNA biopsies.

Summary of Evidence

For individuals with thyroid nodule(s) and indeterminate findings on fine needle aspiration (FNA) who receive FNA sample testing with molecular tests to rule out malignancy and to avoid surgical biopsy or resection, the evidence includes prospective clinical validity studies with the Afirma GSC (Gene Sequencing Classifier), a systematic review of prospective and retrospective clinical validity studies, a meta-analysis of real-world postvalidation data for the Afirma GSC platform with comparison to the validation study, and a chain of evidence to support clinical utility. Relevant outcomes are disease-specific survival, test accuracy and validity, morbid events, and resource utilization. A systematic review of 1 prospective and 6 retrospective trials demonstrated a high negative predictive value (NPV; 96%; 95% confidence interval [CI], 94% to 98%). In a multicenter validation study, the Afirma GSC was also reported to have a high NPV (96%; 95% CI, 90% to 99%). The meta-analysis of real-world Afirma GSC data indicated significantly higher NPV (as well as specificity and positive predictive value [PPV]) than in the validation study. These results are consistent with an earlier study on the Afirma GEC (Gene Expression Classifier) in the same study population and a randomized controlled trial of Afirma GSC in a similar population. In other multicenter and single-center studies, there is suggestive evidence that rates of malignancy are low in Afirma GSC or ThyroSeq v3 patients who are classified as benign or negative, with high NPVs (>90%) in a prospective trial with 31.8 months of post-testing imaging surveillance. The available evidence suggests that the decisions a physician makes regarding surgery are altered by Afirma GSC or ThyroSeq v3 results. A chain of evidence can be constructed to establish the potential for clinical utility with Afirma GSC and ThyroSeq v3 testing in cytologically indeterminate lesions, but evidence of improved outcomes must be demonstrated through at least 5 years of surveillance as recommended by the American College of Radiology. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with thyroid nodule(s) and indeterminate findings on FNA who receive FNA sample testing with molecular tests to rule in malignancy and to guide surgical planning, the evidence includes prospective and retrospective studies of clinical validity. Relevant outcomes are disease-specific survival, test accuracy and validity, morbid events, and resource utilization. Variant analysis has the potential to improve the accuracy of an equivocal FNA of the thyroid and may play a role in preoperative risk stratification and surgical planning. Single-center studies have suggested that testing for a panel of genetic variants associated with thyroid cancer may allow for the appropriate selection of patients for surgical management for the initial resection. Prospective studies in additional populations are needed to validate these results. Although the presence of certain

variants may predict more aggressive malignancies, the management changes that would occur as a result of identifying higher risk tumors, are not well-established. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with thyroid nodule(s) and indeterminate findings on FNA who receive FNA sample testing with molecular tests to rule out malignancy and avoid surgical biopsy or to rule in malignancy for surgical planning, the evidence includes multiple retrospective and prospective clinical validation studies for the ThyroSeq test, a systematic review of retrospective and prospective studies, and 3 retrospective studies of the combined ThyGen X + ThyraMIR assay. Relevant outcomes are disease-specific survival, test accuracy and validity, morbid events, and resource utilization. In a retrospective validation study on FNA samples, the 17-variant panel (miR*Inform*) test and ThyraMIR had a sensitivity of 89%, and an NPV of 94%. A prospective clinical validation study of ThyroSeq v3 reported an NPV of 97% and PPV of 68%. Similarly, a systematic review including 3 prospective and 3 retrospective clinical validity studies reported an NPV of 92% and PPV of 70%. One retrospective study found that combined ThyGenX + ThyraMIR had an NPV of 99%, while another study reported an NPV of 76.9%. No prospective studies were identified demonstrating evidence of direct outcome improvements. A chain of evidence for the ThyroSeq v3 test and combined ThyGenX and ThyraMIR testing would rely on establishing clinical validity. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Additional Information

2017 Input

Clinical input was sought to help determine whether testing for molecular markers in fine needle aspirates of the thyroid for management of individuals with thyroid nodule(s) with an indeterminate finding on the fine needle aspirates would provide a clinically meaningful improvement in net health outcome and whether the use is consistent with generally accepted medical practice. In response to requests, clinical input on 7 tests for molecular markers was received from 9 respondents, including 1 specialty society-level response, 1 physician from an academic center, and 7 physicians from 2 health systems

Clinical input supports that the following uses provide a clinically meaningful improvement in net health outcome and indicates the uses are consistent with generally accepted medical practice:
For individuals who have fine needle aspirate (FNA) of thyroid nodules with indeterminate cytologic findings (i.e., Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/suspicion for a follicular neoplasm]) who receive the following types of molecular marker testing to rule out malignancy and to avoid surgical biopsy:

- Afirma Gene Expression Classifier; or
- ThyroSeq v2

For individuals who have FNA of thyroid nodules with indeterminate cytologic findings or Bethesda diagnostic category V (suspicious for malignancy) who receive the following types of molecular marker testing to rule in the presence of malignancy to guide surgical planning for the initial resection rather than a 2 stage surgical biopsy followed by definitive surgery:

- ThyroSeq v2;
- ThyraMIR microRNA/ThyGenX;
- Afirma BRAF after Afirma Gene Expression Classifier; or
- Afirma MTC after Afirma Gene Expression Classifier.

Further details from clinical input are included in the Supplemental Information section and in the Appendix.

Related Policies

- N/A

Benefit Application

Benefit determinations should be based in all cases on the applicable member health services contract language. To the extent there are conflicts between this Medical Policy and the member health services contract language, the contract language will control. Please refer to the member's contract benefits in effect at the time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal law may prohibit health plans from denying FDA-approved Healthcare Services as investigational or experimental. In these instances, Blue Shield of California may be obligated to determine if these FDA-approved Healthcare Services are Medically Necessary.

Regulatory Status

SB 496
SB 496 requires health plans licensed under the Knox-Keene Act ("Plans"), Medi-Cal managed care plans ("MCPS"), and health insurers ("Insurers") to cover biomarker testing for the diagnosis, treatment, appropriate management, or ongoing monitoring of an enrollee's disease or condition to guide treatment decisions, as prescribed. The bill does not require coverage of biomarker testing for screening purposes. Restricted or denied use of biomarker testing for these purposes is subject to state and federal grievance and appeal processes. Where biomarker testing is deemed medically necessary, Plans and Insurers must ensure that the testing is provided in a way that limits disruptions in care.

Clinical Laboratory Improvement Amendments (CLIA) and FDA Regulatory Overview
Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments. Thyroid variant testing and gene expression classifiers are available under the auspices of the Clinical Laboratory Improvement Amendments. Laboratories that offer laboratory-developed tests must be licensed by the Clinical Laboratory Improvement Amendments for high-complexity testing. To date, the U.S. Food and Drug Administration has chosen not to require any regulatory review of this test.

In 2013, the THxID™-BRAF kit (bioMérieux), an in vitro diagnostic device, was approved by the U.S. Food and Drug Administration through the premarket approval process to assess specific *BRAF* variants in melanoma tissue via real-time PCR. However, there are currently no diagnostic tests for thyroid cancer mutation analysis with approval from the U.S. Food and Drug Administration. Table 1 provides a summary of commercially available molecular diagnostic tests for indeterminate thyroid pathology.

Table 1. Summary of Molecular Tests for Indeterminate Thyroid Cytopathology FNA Specimens

Test	Predicate	Methodology	Analyte(s)	Report
Afirma® GSC	Afirma®GEC	mRNA gene expression	1115 genes	Benign/suspicious
Afirma® BRAF		mRNA gene expression	1 gene	Negative/positive
Afirma® MTC		mRNA gene expression		Negative/positive
ThyroSeq v3	ThyroSeq v2	Next-generation sequencing	112 genes	Specific gene variant/translocation

ThyGeNEXT®	ThyGenX®[a], miR*Inform*®[a]	Next-generation sequencing	10 genes and 32 gene fusions	Specific gene variant/translocation
ThyraMIR™ v2	ThyraMIR™	microRNA expression	11 microRNAs	Low, moderate, or high-risk
NeoTYPE® Thyroid Profile		Next-generation sequencing	26 genes and 2 biomarkers	Specific gene variant/translocation

FNA: fine needle aspirate; GEC: Gene Expression Classifer; GSC: Gene Sequencing Classifier; mRNA: messenger RNA; MTC: medullary thyroid carcinoma; PCR: polymerase chain reaction.
[a] The miR*Inform*® test is the predicate test to ThyGenX™ and is not commercially available.

Rationale

Background

Thyroid Nodules

Thyroid nodules are common, present in 5% to 7% of the U.S. adult population; however, most are benign, and most cases of thyroid cancer are curable surgically when detected early.

Diagnosis

Sampling thyroid cells by fine needle aspirate (FNA) is currently the most accurate procedure to distinguish benign thyroid lesions from malignant ones, reducing the rate of unnecessary thyroid surgery for patients with benign nodules and triaging patients with thyroid cancer to appropriate surgery.

About 60% to 70% of thyroid nodules are classified cytologically as benign, and 4% to 10% of nodules are cytologically deemed malignant.[1,] However, the remaining 20% to 30% have equivocal findings, usually due to overlapping cytologic features between benign and malignant nodules; these nodules usually require surgery for a final diagnosis. Thyroid FNA cytology is classified by Bethesda System criteria into the following groups: nondiagnostic; benign; follicular lesion of undetermined significance or atypia of undetermined significance; follicular neoplasm (or suspicious for follicular neoplasm); suspicious for malignancy; and malignant. Lesions with FNA cytology in the atypia of undetermined significance or follicular neoplasm of undetermined significance or follicular neoplasm categories are often considered indeterminate.

Management

There is some individualization of management for patients with FNA-indeterminate nodules, but many patients will require a surgical biopsy, typically thyroid lobectomy, with intraoperative pathology. Consultation would typically be the next step in the diagnosis. Approximately 80% of patients with indeterminate cytology undergo surgical resection; postoperative evaluation has revealed a malignancy rate ranging from 6% to 30%, making this a clinical process with very low specificity.[2,] Thus, if an analysis of FNA samples could reliably identify the risk of malignancy as low, there is potential for patients to avoid surgical biopsy.

Preoperative planning of optimal surgical management in patients with equivocal cytologic results is challenging, because different thyroid malignancies require different surgical procedures (e.g., unilateral lobectomy vs. total or subtotal thyroidectomy with or without lymph node dissection) depending on several factors, including histologic subtype and risk-stratification strategies (tumor size, patient age). If a diagnosis cannot be made intraoperatively, a lobectomy is typically performed, and, if on postoperative histology the lesion is malignant, a second surgical intervention may be necessary for completion of thyroidectomy.

Thyroid Cancer

Most thyroid cancers originate from thyroid follicular cells and include well-differentiated papillary thyroid carcinoma (PTC; 80% of all thyroid cancers) and follicular carcinoma (15%). Poorly

differentiated and anaplastic thyroid carcinomas are uncommon and can arise de novo or from preexisting well-differentiated papillary or follicular carcinomas. Medullary thyroid carcinoma originates from parafollicular or C cells and accounts for about 3% of all thyroid cancers.

The diagnosis of malignancy in the case of PTC is primarily based on cytologic features. If FNA in a case of PTC is indeterminate, surgical biopsy with intraoperative pathology consultation is most often diagnostic, although its efficacy and therefore its use will vary across institutions, surgeons, and pathologists. In 2016, reclassification of encapsulated follicular-variant PTC as a noninvasive follicular tumor with papillary-like nuclei was proposed and largely adopted; this classification removes the word *carcinoma* from the diagnosis to acknowledge the indolent behavior of these tumors.[3,]

For follicular carcinoma, the presence of invasion of the tumor capsule or blood vessels is diagnostic, and cannot be determined by cytology, because tissue sampling is necessary to observe these histologic characteristics. Intraoperative diagnosis of follicular carcinoma is challenging and often not feasible because extensive sampling of the tumor and capsule is usually necessary and performed on postoperative, permanent sections.

New approaches for improving the diagnostic accuracy of thyroid FNA include variant analysis for somatic genetic alterations, to more accurately classify which patients need to proceed to surgery (and may include the extent of surgery necessary), and a gene expression classifier to identify patients who do not need surgery and can be safely followed.

Genetic Variants Associated With Thyroid Cancer

A number of genetic variants have been discovered in thyroid cancer. The most common 4 gene variants are *BRAF* and *RAS* single nucleotide variants (SNVs) and *RET/PTC* and *PAX8/PPAR*$_{\gamma}$ rearrangements.

Papillary carcinomas carry SNVs of the *BRAF* and *RAS* genes, as well as *RET/PTC* and *TRK* rearrangements, all of which can activate the mitogen-activated protein kinase pathway.[4,] These mutually exclusive variants are found in more than 70% of papillary carcinomas. *BRAF* SNVs are highly specific for PTC. Follicular carcinomas harbor either *RAS* SNVs or *PAX8/PPAR*$_{\gamma}$ rearrangements. These variants have been identified in 70% to 75% of follicular carcinomas. Genetic alterations involving the PI3K/AKT signaling pathway also occur in thyroid tumors, although they are rare in well-differentiated thyroid cancers and have a higher prevalence in less differentiated thyroid carcinomas. Additional variants known to occur in poorly differentiated and anaplastic carcinomas involve the *TP53* and *CTNNB1* genes. Medullary carcinomas, which can be familial or sporadic, frequently possess SNVs located in the *RET* gene.

Studies have evaluated the association between various genes and cancer phenotype in individuals with diagnosed thyroid cancer.[5,6,7,]

Telomerase reverse transcriptase (*TERT*) promoter variants occur with varying frequency in different thyroid cancer subtypes. Overall, *TERT* C228T or C250T variants have been reported in approximately 15% of thyroid cancers, with higher rates in the undifferentiated and anaplastic subtypes compared with the well-differentiated subtypes.[8,] *TERT* variants are associated with several demographic and histopathologic features such as older age and advanced TNM stage. *TERT* promoter variants have been reported to be independent predictors of disease recurrence and cancer-related mortality in well-differentiated thyroid cancer.[9,10,11,] Also, the co-occurrence of *BRAF* or *RAS* variants with *TERT* or *TP53* variants may identify a subset of thyroid cancers with unfavorable outcomes.[12,13,14,]

Molecular Diagnostic Testing

Variant Detection and Rearrangement Testing

SNVs in specific genes, including *BRAF*, *RAS*, and *RET*, and evaluation for rearrangements associated with thyroid cancers can be accomplished with Sanger sequencing or pyrosequencing or with real-

time polymerase chain reaction (PCR) of single or multiple genes or by next-generation sequencing (NGS) panels. Panel tests for genes associated with thyroid cancer, with varying compositions, are also available. For example, Quest Diagnostics offers a Thyroid Cancer Mutation Panel, which includes *BRAF* and *RAS* variant analysis and testing for *RET/PTC* and *PAX8/PPAR*$_\gamma$ rearrangements.

The ThyroSeq v3 Next-Generation Sequencing panel (Sonic Healthcare) is an NGS panel of 112 genes.

The test is indicated when FNA cytology suggests atypia of uncertain significance or follicular lesion of undetermined significance, follicular neoplasm or suspicious for follicular neoplasm, or suspicious for malignancy.[15,] In particular, it has been evaluated in patients with follicular neoplasm and/or suspicious for follicular neoplasm on FNA as a test to increase both sensitivity and specificity for cancer diagnosis. ThyGenX is an NGS panel that sequences 8 genes and identifies specific gene variants and translocations associated with thyroid cancer. ThyGenX is intended to be used in conjunction with the ThyraMIR microRNA expression test when the initial ThyGenX test is negative.

Gene Expression Profiling

Genetic alterations associated with thyroid cancer can be assessed using gene expression profiling, which refers to the analysis of messenger RNA (mRNA) expression levels of many genes simultaneously. Several gene expression profiling tests are available and stratify tissue from thyroid nodules biologically.

The Afirma Gene Expression Classifier (Afirma GEC; Veracyte) analyzed the expression of 142 different genes to determine patterns associated with benign findings on surgical biopsy. It was designed to evaluate thyroid nodules that have an "indeterminate" classification on FNA as a method to select patients ("rule out") who are at low-risk for cancer. In 2017, Veracyte migrated the Afirma GEC microarray analysis to a next-generation RNA sequencing platform and now markets the Afirma Gene Sequencing Classifier (Afirma GSC) which evaluates 10,196 genes with 1115 core genes.

Other gene expression profiles have been reported in investigational settings, but have not been widely validated or used commercially (e.g., Barros-Filho et al [2015],[16,] Zheng et al [2015][17,]); they are not addressed in this review.

ThyraMIR is a microRNA expression-based classifier intended for use in thyroid nodules with indeterminate cytology on FNA following a negative result from the ThyGenX Thyroid Oncogene Panel.

Algorithmic Testing

Algorithmic testing involves the use of 2 or more tests in a prespecified sequence, with a subsequent test automatically obtained depending on results of an earlier test.

Algorithmic Testing Using Afirma GEC With Afirma MTC and Afirma BRAF

In addition to Afirma GSC, Veracyte also markets 2 "malignancy classifiers" that use mRNA expression-based classification to evaluate for *BRAF* variants (Afirma BRAF) or variants associated with medullary thyroid carcinoma (Afirma MTC). Table 1 outlines the testing algorithms for Afirma MTC and Afirma BRAF.

Table 2. Afirma MTC and Afirma BRAF Testing Algorithms

Test 1	Test 1 Result	Reflex to Test 2
Thyroid nodule on fine needle aspirate	"Indeterminate"	Afirma MTC
Afirma GSC	"Malignant" or "suspicious"	Afirma MTC
Afirma GSC	"Suspicious"	Afirma BRAF

Afirma GSC: Afirma Gene Sequencing Classifier; Afirma MTC: Afirma medullary thyroid carcinoma

In a description of the Afirma BRAF test, the following have been proposed as benefits of the mRNA-based expression test for *BRAF* variants: (1) PCR-based methods may have low sensitivity, requiring that a large proportion of the nodule have a relevant variant; (2) testing for only 1 variant may not detect patients with low-frequency variants that result in the same pattern of pathway activation; and (3) PCR-based approaches with high analytic sensitivity may require a large amount of DNA that is difficult to isolate from small FNA samples.[18,]

The testing strategy for both Afirma MTC and Afirma BRAF is to predict malignancy from an FNA sample with increased pretest probability for malignancy. A positive result with Afirma MTC or Afirma BRAF would inform preoperative planning such as planning for a hemi- versus a total thyroidectomy or performance of central neck dissection.

Algorithmic Testing Using ThyGenX and ThyraMIR

The ThyGenX Thyroid Oncogene Panel (Interpace Diagnostics; testing is done at Asuragen Clinical Laboratory) is an NGS panel designed to assess patients with indeterminate thyroid FNA results. It includes sequencing of 8 genes associated with PTC and follicular carcinomas. ThyGenX has replaced the predicate miR*Inform* Thyroid test that assesses for 17 validated gene alterations.

ThyraMIR (Interpace Diagnostics) is a microRNA expression-based classifier intended for use in thyroid nodules with indeterminate cytology on FNA following a negative result from the ThyGenX Thyroid Oncogene Panel.

The testing strategy for combined ThyGenX and ThyraMIR testing is first to predict malignancy. A positive result on ThyGenX would "rule in" patients for surgical resection. The specific testing results from a ThyGenX positive test would be used to inform preoperative planning when positive. For a ThyGenX negative result, the reflex testing involves the ThyraMIR microRNA expression test to "rule out" for a surgical biopsy procedure given the high negative predictive value of the second test. Patients with a negative result from the ThyraMIR test would be followed with active surveillance and avoid a surgical biopsy.

Literature Review

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

Molecular Tests to Rule Out Malignancy

Clinical Context and Test Purpose

One purpose of molecular testing in individuals with indeterminate findings on fine needle aspirate(s) (FNA) of thyroid nodules is to rule out malignancy and eliminate the need for surgical biopsy or resection.

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with indeterminate findings on FNAs of thyroid nodules who would be willing to undergo watchful waiting, depending on the results of their

molecular testing. Individuals with indeterminate findings after FNA of thyroid nodule presently proceed to surgical biopsy or resection.

Interventions
The test being considered is molecular testing, which includes Afirma GSC (Gene Sequencing Classifier) (predicate Afirma GEC [Gene Expression Classifier]).

Comparators
The following practice is currently being used: standard surgical management through surgical biopsy or resection for biopsy.

Outcomes
The potential beneficial outcome of primary interest would be avoiding an unneeded surgical biopsy or resection (e.g., lobectomy or hemithyroidectomy) in a true-negative thyroid nodule that is benign. Potential harmful outcomes are those resulting from false-negative test results, which may delay diagnosis and surgical resection of thyroid cancer. For small, slow-growing tumors, it is uncertain that a delay in diagnosis would necessarily worsen health outcomes.

The time frame for evaluating the performance of the test is the time from the initial FNA to surgical biopsy or resection measured in weeks to months following an indeterminate result. Papillary thyroid cancer (PTC) is indolent, and a nodule could be observed for many years to ensure no clinical change. Specifically, the American College of Radiology Thyroid Imaging, Reporting and Data System (TI-RADS) recommends surveillance of suspicious nodules through 5 years.[19,]

Study Selection Criteria
For the evaluation of clinical validity of the molecular testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores)
- Included a suitable reference standard
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

Afirma GSC
Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Systematic Reviews
Lee et al (2022) performed a systematic review and meta-analysis on the diagnostic performance of molecular tests in the assessment of indeterminate thyroid nodules.[20,] Inclusion criteria for trials included indeterminate thyroid results via FNA that included Bethesda categories III and IV, conclusive histopathological results in a group of benign and suspicious changes, and the use of Afirma GSC, ThyroSeq v3, and ThyGeNext as index tests. Investigators identified 7 studies on Afirma GSC: 1 prospective study by Livhits et al (2021), described below, and 6 retrospective studies. Pooled data for GSC studies on 472 thyroid nodules demonstrated a sensitivity of 96.6% (95% confidence interval [CI], 89.7% to 98.9%), specificity of 52.9% (95% CI, 23.4% to 80.5%), positive predictive value (PPV) of 63% (95% CI, 51% to 74%), and negative predictive value (NPV) of 96% (95% CI, 94% to 98%).

Limitations of this meta-analysis include the scarcity of available cohort analyses of the molecular tests and the lack of long-term findings.

Nasr et al (2023) performed a meta-analysis of 13 real-world postvalidation studies (N=1976 patients with indeterminate thyroid nodules) of the Afirma GSC platform and compared results to the validation study by Patel et al (2018, described below).[21,] Studies performed prior to publication of the validation study and commercial availability of Afirma GSC were excluded. Among 11 studies reporting histopathological results for patients who underwent surgery, sensitivity was 97.2% (95% CI, 1.7% to 99.1%; I^2=0%), specificity was 87.7% (95% CI, 83.2% to 91.0%; I^2=63%), PPV ranged from 49.3% (including patients with suspicious molecular testing results who did not undergo surgery; 95% CI, 41.3% to 57.4%; I^2 not reported) to 64.9% (excluding patients with suspicious molecular testing results who did not undergo surgery; 95% CI, 54.4% to 74.1%; I^2=79%), and NPV was 99.5% (95% CI, 98.0% to 99.9%; I^2=0%). Specificity, PPV (excluding patients with suspicious results who did not undergo surgery), and NPV were significantly improved compared to the values reported in the validation study (p<.05 for each comparison).

Vardarli et al (2024) conducted a systematic review and meta-analysis to assess the diagnostic accuracy of Afirma GEC/GSC and ThyroSeq v2/v3 for Bethesda III and IV indeterminate thyroid nodules.[22,] A total of 53 studies (71 diagnostic-accuracy samples and 6,490 FNAs) met prespecified criteria requiring histology or structured follow-up and excluding Bethesda V nodules. Pooled across all testing platforms, the sensitivity reached 95% (95% CI, 94% to 97%) and specificity 35% (28% to 43%). The authors found that ThyroSeq v3 demonstrated the strongest overall test performance, with an area under the ROC curve (AUC) of 0.95 (95% CI.93 to.96), followed by ThyroSeq v2 at 0.90 (95% CI, 0.87 to 0.92), Afirma GSC at 0.86 (95% CI,.82 to.88), and Afirma GEC at 0.82 (95% CI,.78 to.85).

Prospective Clinical Validation

Patel et al (2018) reported a validation study for the Afirma GSC test. The study included 210 thyroid nodules from 183 patients that had indeterminate results (Bethesda III or IV) on FNA, see Table 3.[23,] All FNA samples had been previously used in the validation of the Afirma GEC test as reported by Alexander et al (2012) in a 19-month, prospective, multicenter (49 academic and community sites) study.[24,] Patel et al (2018) used the banked samples which were reassayed with next-generation sequencing (NGS) for the Afirma GSC validation study.[23,] The previous central, blinded postoperative consensus histopathological diagnosis was used as the reference standard (210 samples) and all personnel were blinded to the other outcomes. The sensitivity of the Afirma GSC study was 91.1% with a specificity of 68.3% and NPV of 96.1% (see Table 4). There were 4 false negatives in patients with malignant nodules who would have been assigned for active observation. In comparison, Afirma GEC correctly identified 78 of 85 malignant nodules as suspicious (92% sensitivity; 95% CI, 84% to 97%) with specificity of 52% (95% CI, 44% to 59%). The NPV ranged from 85% for "suspicious cytologic findings" to 95% for "atypia of undetermined clinical significance." With sensitivity that was similar to the Afirma GEC test, the Afirma GSC improved specificity. There were no notable study limitations.

Livhits et al (2021) published a randomized, controlled study that compared the Afirma GSC test to the ThyroSeq v3 test in patients with thyroid nodules with indeterminate FNA results (Bethesda III or IV).[25,] The study reported clinical validity for both tests; the results of the Afirma GSC test are summarized in Table 3 and Table 4. The study used histopathologic review by expert thyroid pathologists as the reference standard. The study included 201 nodules in the Afirma GSC group. The sensitivity of Afirma GSC was 100%, specificity was 79.6%, and the NPV was 100%. A limitation of the study is that the pathologists who interpreted the histopathologic diagnosis were not blinded to the results of the molecular test. Patients in this trial who were managed nonoperatively were prospectively surveilled via ultrasound for 12 to 60 months, with results of surveillance reported with median follow-up of 31.8 months.[26,] Among the nodules initially managed nonoperatively, 44 patients were lost to follow-up without surveillance imaging and were excluded from the analysis, with surveillance data available for 195 nodules. Over the course of surveillance, 84% of nodules with benign or negative molecular testing remained stable. Among the 26 nodules with benign or negative molecular testing that exhibited growth on ultrasound, 12 underwent surgery, with 11 histopathologically diagnosed as benign; the 1 malignant nodule was diagnosed as a minimally invasive Hürthle cell carcinoma. Among 33 nodules with suspicious or positive molecular testing that

were initially managed nonoperatively (due to patient preference or other reasons), 15 were ultimately resected, 6 of which were benign. In surgically-confirmed cases, the sensitivity of the Afirma GSC and ThyroSeq v3 tests was 100% and 97%, respectively; specificity was 40% and 38%, PPV was 57% and 64%, and NPV was 100% and 92%, respectively (p>.05 for all comparisons between test platforms).

Table 3. Study Characteristics for Afirma GSC

Study	Study Population	Design	Reference Standard	Threshold for Positive Index Test	Timing of Reference and Index Tests	Blinding of Assessors	Comment
Patel et al (2018)[23,]	183 patients with 210 indeterminate thyroid nodules by FNA	Multicenter, non-concurrent prospective validation trial	Consensus histopathology diagnosis		Central, blinded histopathological review from Alexander et al (2012)	Assessors were blinded to the pathology	Samples were previously used to validate Afirma GEC
Livhits et al (2021)[25,]	201 indeterminate thyroid nodules by FNA (Afirma GSC)*	Multicenter, randomized controlled trial	Histopathologic diagnosis	Classified as malignant or benign	Samples were tested after surgery	Assessors were unblinded to results of molecular testing	

FNA: Fine needle aspirate; Afirma GEC: gene expression classifier; Afirma GSC: gene sequencing classifier.
*Study included a comparator group assigned to ThyroSeq (reported below)

Table 4. Clinical Validity for Afirma GSC

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Clinical Validity (95% Confidence Interval)			
					Sensitivity	Specificity	PPV	NPV
Patel et al (2018)[23,]	210 nodules	191 nodules	19 with insufficient residual RNA		91.1 (79 to 98)	68.3 (60 to 76)	47.1 (36 to 58)	96.1 (90 to 99)
Livhits et al (2021)[25,]	201 assigned to Afirma GSC	180 nodules	21 were excluded		100 (88.8 to 100)	79.6 (71.7 to 86.1)	53.5 (39.9 to 66.7)	100 (96.6 to 100)

Afirma GSC: gene sequencing classifier; NPV: negative predictive value; PPV: positive predictive value; RNA: ribonucleic acid.

Retrospective Clinical Validation

Meta-analyses have been performed with studies reporting on the performance of the predicate Afirma GEC in cytologically indeterminate nodules.[27,28,] Retrospective studies are subject to ascertainment bias because a large proportion of individuals with Afirma benign reports did not undergo surgery, which makes determining the sensitivity and specificity of the GEC assay impossible.

Supportive information on the accuracy of benign results can be obtained from studies that report long-term follow-up of individuals with indeterminate FNA cytology and Afirma benign results. There are several studies that reported long-term follow-up of Afirma GEC.[29,30,31,] Valderrabano et al (2019) used the benign call rate and PPV of post-marketing studies for a simulation study, concluding that the initial validation study cohort of Afirma GEC was not representative of the populations in whom the test has been used, raising questions regarding its diagnostic performance.[32,] Because the Afirma GSC used the same validation study, these findings would also apply to Afirma GSC.

Harrell et al (2019) reported a retrospective comparison of Afirma GEC (2011 to July 2017) and Afirma GSC (August 2017 through June 2018) for indeterminate FNA.[33,] Afirma GSC identified fewer indeterminate nodules as suspicious (54/139, 38.8%) compared to GEC (281/481, 58.4%) and led to a lower surgery rate, decreasing from 56% in the GEC group to 31% in the GSC group. A similar retrospective comparison was conducted by Polavarapu et al (2021), comparing Afirma GEC and Afirma GSC for indeterminate FNA between January 2013 through December 2019.[34,] Of the 468 indeterminate thyroid nodules included, no molecular testing was performed in 273, 71 had GEC, and 124 had GSC. Use of Afirma GSC led to a lower surgery rate (39.5%; p=.0001) compared to GEC (59.2%) and no molecular testing (67.8%). Additionally, malignancy rate was 20% with no molecular testing, 22% in GEC, and 39% in GSC (p=.022). Afirma GEC benign cell rate was 46%; sensitivity was 100%, specificity was 61%, NPV was 100%, and PPV was 28%. With Afirma GSC, benign cell rate was 60%, sensitivity was 94%, specificity was 76%, NPV was 97%, and PPV was 41%. In conclusion, Afirma GSC testing had a significant reduction in surgical rates and increase in malignancy rates. Sensitivity and NPV were high for both GEC and GSC. A 2023 retrospective analysis of 408 indeterminate thyroid nodules compared the Afirma GSC + XA (n=40), Afirma GEC + GSC (n=255), and Interpace Diagnostics ThyGeNEXT + ThyraMIR platforms (n=113).[35,] Patients either underwent surgery (56.4%) or were monitored for at least 6 months with ultrasound imaging. Sensitivity of the GSC + XA platform was greater than the GEC + GSC platform (80.0% vs 75.81%; p<.001) but not the ThyGeNEXT + ThyraMIR platform (47.4%; p=.08); this may be attributable to the relatively small size of the GSC + XA group. Specificity of the Afirma GSC + Xa (91.4%) and ThyGeNEXT + ThyraMIR platforms (88.3%) was greater than the GEC + GSC platform (45.1%; p<.001 for both comparisons). NPV was >85% for all cohorts and was highest with the GSC + XA platform (97.0%).

Azaryan et al (2024) conducted a single-center retrospective analysis of 237 Bethesda III/IV indeterminate thyroid nodules that underwent Afirma GSC testing between July 2017 and December 2019.[36,] Among the nodules, 195 (82.3%) were classified as benign and 42 (17.6%) as suspicious by GSC testing. Histopathological confirmation of cancer or noninvasive follicular thyroid neoplasm with papillary-like nuclear features was identified in 20 of the suspicious cases, and benign pathology in the remaining 22. The sensitivity, specificity, NPV, and PPV of the test in this center were 87%, 89.7%, 98.5%, and 47.6%, respectively.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

No evidence directly demonstrating improved outcomes in patients managed with the Afirma GEC was identified.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.
Because no direct evidence of utility was identified, a chain of evidence was developed, which addresses 2 key questions:

1. Does use of the Afirma GEC in individuals with cytologically indeterminate thyroid nodules change clinical management (in this case, reduced thyroid resections)?
2. Do those management changes improve outcomes?

Changes in Management

The clinical setting in which the Afirma GEC is meant to be used is well-defined: individuals with atypia of undetermined significance (AUS) or follicular lesion of undetermined significance (FLUS) or follicular neoplasm or who are suspicious for follicular neoplasm (SFN) on FNA, who do not have other indications for thyroid resection (i.e., in whom the GEC results would play a role in surgical decision making). Decision impact studies, most often reporting on clinical management changes but not on outcomes after surgical decisions were made, have suggested that, in at least some cases, surgical decision making changed.[37,38,39,] [40,41,] It cannot be determined from these studies whether the changes in management improved health outcomes.

Improved Outcomes

A simplified decision model was developed for use with Afirma GEC (which can also be applied to use of the Afirma GSC) in individuals with cytologically indeterminate FNA samples. It is shown in Appendix Figure 1. It is assumed that when Afirma GEC/GSC is not used, patients with cytologically indeterminate FNA results undergo thyroid resection. When Afirma GEC/GSC is used, those with Afirma suspicious lesions undergo resection, while those who have Afirma benign lesions do not. In this case, compared with the standard care plan, some patients without cancer will have avoided a biopsy, which is weighed against the small increase in missed cancers, in patients who had cancer but tested as Afirma benign.

Assuming that the rate of cancer in cytologically indeterminate thyroid nodules is approximately 20%,[42,] in the standard care plan, 80% of patients with cytologically indeterminate FNA samples will undergo an unnecessary biopsy. Applying the test characteristic values from Alexander et al (2012),[24,] it is estimated that approximately 1.6% of individuals with true cancer would be missed, but approximately 38%, instead of 80%, would undergo unneeded surgery. The study by Kim et al (2023), described previously above, reported only 1 false-negative case among 15 patients with nodules demonstrating growth on surveillance imaging over 3 years who underwent delayed surgery, suggesting that the rate of false-negative results and avoided unnecessary surgeries may be further improved with the Afirma GSC and ThyroSeq v3 platforms.[26,]

Whether the tradeoff between avoiding unneeded surgeries and the potential for missed cancer is worthwhile depends, in part, on patient and physician preferences. However, some general statements may be made by considering the consequences of a missed malignancy and the consequences of unnecessary surgery. Most missed malignancies will be PTCs, which have an indolent course. Thyroid nodules are amenable to ongoing surveillance (clinical, ultrasound, and with repeat FNAs), with minimal morbidity.

Thyroid resection is a relatively low-risk surgery. However, the consequences of surgery can be profound. Patients who undergo a hemi- or subtotal thyroidectomy have a risk of recurrent laryngeal nerve damage and parathyroid gland loss. The standard of care for thyroid nodules is based on an intervention that is stratified by FNA cytology results, which are grouped into categories with differing prognosis. Avoiding invasive surgery in situations where patients are at very low likelihood of having an invasive tumor is likely beneficial. Among the low-risk population, the alternative to surgical biopsy is ongoing active surveillance.

While the Kim et al (2023) study is encouraging, evidence of improved outcomes through 5 years of surveillance is needed as recommended by the American College of Radiology.[19,]

Section Summary: Molecular Tests to Rule Out Malignancy

A systematic review of 1 prospective and 6 retrospective trials demonstrated a high NPV (96%; 95% CI, 94% to 98%), with a recent meta-analysis of real-world postvalidation data indicating significantly better diagnostic performance of the Afirma GSC platform than in its validation study. In a multicenter validation study, Afirma GSC was also reported to have a high NPV (96%; 95% CI, 90% to 99%). These results are consistent with an earlier study on the Afirma GEC in the same study

population and with a randomized controlled trial of Afirma GSC in a similar study population. In other multicenter and single-center studies, there is suggestive evidence that rates of malignancy are low in Afirma patients who are classified as benign. One prospective study with long-term imaging surveillance of 195 nodules initially managed nonoperatively based on negative/benign Afirma GSC or ThyroSeq v3 testing only indicated 1 false-negative case over 31.8 months of follow-up. Another single-center retrospective analysis found a sensitivity of 87% and an NPV of 98.5%. The available evidence suggests that physician decision making about surgery is altered by Afirma GSC or ThyroSeq v3 results. A chain of evidence can be constructed to establish the potential for clinical utility with Afirma GSC and ThyroSeq v3 testing in cytologically indeterminate lesions, but evidence of improved outcomes must be demonstrated through at least 5 years of surveillance as recommended by the American College of Radiology.

Molecular Tests to Rule in Malignancy

Clinical Context and Test Purpose

The purpose of testing for molecular markers (e.g., single nucleotide variants and gene rearrangements) in individuals with indeterminate findings on FNA of thyroid nodules is to rule in malignancy and to guide surgical approach or management.

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with indeterminate findings on FNA(s) of thyroid nodules. Individuals with indeterminate findings would presently proceed to surgical biopsy perhaps with intraoperative pathology consultation (i.e., intraoperative frozen section) if available.

Interventions

The test being considered is testing for molecular markers (e.g., single nucleotide variants and gene rearrangements) with Afirma BRAF and Afirma MTC (medullary thyroid carcinoma) to guide surgical planning to ensure the capability for intraoperative pathologic confirmation of malignancy to adjust to definitive surgery for initial resection if appropriate.

Comparators

The following practices are currently being used: standard surgical management through surgical resection, including a 2-stage surgical biopsy (i.e., lobectomy) followed by definitive surgery (i.e., hemithyroidectomy or thyroidectomy).

Outcomes

The potential beneficial outcome of primary interest is appropriate surgical planning in the preoperative period (e.g., hemithyroidectomy or thyroidectomy when malignancy is predicted). This has the potential benefit of reducing the likelihood of having the individual repeating surgery if a diagnosis is not made on frozen pathology section during the initial surgery if lobectomy is done as a first procedure.

Potential harmful outcomes are those resulting from false-positive results. However, the use of intraoperative confirmation of malignancy through frozen pathology section in individuals with positive molecular marker testing would mitigate any risk of inappropriately performing more extensive thyroidectomy in the absence of malignancy.

The time frame for evaluating the performance of the test varies from the initial FNA to surgical resection to weeks to months following an indeterminate result.

Study Selection Criteria

For the evaluation of clinical validity of the molecular testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores)
- Included a suitable reference standard
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

Gene Expression Classifiers to Predict Malignancy

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Less evidence exists on the validity of gene expression profiling to rule in malignancy (specifically, the Afirma BRAF and Afirma MTC tests). Genetic variants can be used to improve the sensitivity and specificity for diagnosing indeterminate FNA of the thyroid, with the goal of identifying variants that predict malignancy in FNA samples.

Fnais et al (2015) conducted a systematic review and meta-analysis of studies reporting on the test accuracy of *BRAF* variant testing in the diagnosis of PTC.[43,] Reviewers included 47 studies with 9924 FNA samples. For all cytologically indeterminate nodules, the pooled sensitivity estimate for *BRAF* variant testing was 31% (95% CI, 6% to 56%). Among nodules suspicious for malignancy on FNA, the pooled sensitivity estimate for *BRAF* variant testing was 52% (95% CI, 39% to 64%; I^2=77%).

Afirma BRAF and Afirma MTC

Diggans et al (2015), described the development and validation of the Afirma BRAF test, for a subset of 213 thyroid nodule FNA samples for which histopathology was available, Afirma BRAF test results were compared with pathologic findings.[18,] Afirma BRAF classified all histopathologically benign samples as *BRAF* V600E-negative (specificity, 100%; 95% CI, 97.4% to 100%). Of the 73 histopathologically malignant samples, the Afirma BRAF test identified 32 as *BRAF*-positive (sensitivity, 43.8%; 95% CI, 32.2% to 55.9%).

In a study describing the development and validation of the Afirma MTC classifier, Kloos et al (2016) evaluated the MTC classifier in a sample of 10,488 thyroid nodule FNA samples referred for GEC testing.[44,] In this sample, 43 cases were Afirma MTC-positive, of which 42 were considered to be clinically consistent with MTC on pathology or biochemical testing, for a PPV of 97.7% (95% CI, 86.2% to 99.9%).

Genetic Variants Association With Tumor Behavior

The presence of *BRAF* or telomerase reverse transcriptase (*TERT*) variants is strongly associated with malignancy in thyroid nodule FNA samples. *BRAF* or *TERT* variants have also been associated with more aggressive clinicopathologic features in individuals diagnosed with PTC.

Adeniran et al (2011) assessed 157 cases with equivocal thyroid FNA readings (indeterminate and suspicious for PTC) or with a positive diagnosis for PTC and concomitant *BRAF* variant analysis.[1,] The results of histopathologic follow-up correlated with the cytologic interpretations and *BRAF* status. Based on the follow-up diagnosis after surgical resection, the sensitivity for diagnosing PTC was 63.3% with cytology alone and 80.0% with the combination of cytology and *BRAF* testing. No false-positives were noted with either cytology or *BRAF* variant analysis. All PTCs with an extrathyroidal extension or aggressive histologic features were positive for a *BRAF* variant. The authors concluded that patients with an equivocal cytologic diagnosis and a *BRAF* V600E variant could be candidates for total thyroidectomy and central lymph node dissection.

Xing et al (2009) investigated the utility of *BRAF* variant testing of thyroid FNA specimens for preoperative risk stratification of PTC in 190 patients.[45,] A *BRAF* variant in preoperative FNA

specimens was associated with poorer clinicopathologic outcomes for PTC. Compared with the wild-type allele, a *BRAF* variant strongly predicted extrathyroidal extension (23% vs 11%; p=.039), thyroid capsular invasion (29% vs 16%; p=.045), and lymph node metastasis (38% vs 18%; p=.002). During a median follow-up of 3 years (range, 0.6 to 10 years), PTC persistence or recurrence was seen in 36% of *BRAF* variant-positive patients and 12% of *BRAF* variant-negative patients, with an odds ratio (OR) of 4.16 (95% CI, 1.70 to 10.17; p=.002). The PPV and NPV for preoperative FNA-detected *BRAF* variant to predict PTC persistence or recurrence were 36% and 88%, respectively, for all histologic subtypes of PTC. The authors concluded that preoperative *BRAF* variant testing of FNA specimens might provide a novel tool to preoperatively identify PTC patients at higher risk for extensive disease (extrathyroidal extension and lymph node metastases) and those more likely to manifest disease persistence or recurrence.

Yin et al (2016) reported on a systematic review and meta-analysis evaluating *TERT* promoter variants and aggressive clinical behaviors in PTC.[46,] Eight eligible studies (N =2035 patients; range, 30 to 507) were included. Compared with wild-type, *TERT* promoter variant status was associated with lymph node metastasis (OR, 1.8; 95% CI, 1.3 to 2.5; p=.001), extrathyroidal extension (OR, 2.6; 95% CI, 1.1 to 5.9; p=.03), distant metastasis (OR, 6.1; 95% CI, 3.6 to 10.3; p<.001), advanced TNM stages III or IV (OR, 3.2; 95% CI, 2.3 to 4.5; p<.001), poor clinical outcome (persistence or recurrence; OR, 5.7; 95% CI, 3.6 to 9.3; p<.001), and mortality (OR, 8.3; 95% CI, 3.8 to 18.2; p<.001).

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

Testing for specific variants associated with thyroid cancer (e.g., *BRAF* V600E, *TERT,* and *RET* variants, *RET/PTC* and *PAX8/PPAR*$_{\gamma}$ rearrangements) is generally designed to "rule in" cancer in nodules with indeterminate cytology on FNA.[47,] (Of note, some gene panels, such as the ThyroSeq panel, may have a high enough NPV that their clinical use could also be considered as a molecular marker to predict benignancy; see next section.) A potential area for clinical utility for this type of variant testing would be in informing preoperative planning for thyroid surgery following initial thyroid FNA, such as planning for a hemi- versus a total thyroidectomy or performance of central neck dissection.

In a retrospective analysis, Yip et al (2014) reported on outcomes after implementation of an algorithm incorporating molecular testing of thyroid FNA samples to guide the extent of initial thyroid resection.[48,] The study included a cohort of patients treated at a single academic center at which molecular testing (*BRAF* V600E, *BRAF* K601E, *NRAS* codon61, *HRAS* codon 61, and *KRAS* codon 12 and 13 single nucleotide variants; *RET/PTC1*, *RET/PTC3*, and *PAX8/PPAR*$_{\gamma}$ rearrangements) was prospectively obtained for all FNAs with indeterminate cytology (FLUS, follicular neoplasm, suspicious for malignancy), and for selective FNAs at the request of the managing physician for selected nodules with benign or nondiagnostic cytology. The study also included a second cohort of patients who did not have molecular testing results available. For patients treated with a molecular diagnosis, a positive molecular diagnostic test was considered an indication for an initial total thyroidectomy. Patients with FLUS and negative molecular diagnostic results were followed with repeat FNA, followed by lobectomy or total thyroidectomy if indeterminate pathology persisted. Patients with a follicular neoplasm or suspicious for malignancy results on cytology and a negative molecular diagnostic result were managed with lobectomy or total thyroidectomy.

The sample included 671 patients, 322 managed with and 349 without molecular diagnostics. Positive molecular testing results were obtained in 56 (17% of those managed with molecular diagnostics) patients, most commonly *RAS* variants (42/56 [75%]), followed by *BRAF* V600E (10/56 [18%]) and *BRAF* K601E (2/56 [4%]) variants, and *PAX8/PPAR*$_{\gamma}$ rearrangements (2/56 [4%]). Compared with those managed without molecular diagnostics (63%), patients managed with molecular diagnostics (69%) were nonsignificantly less likely to undergo total thyroidectomy as an initial procedure (p=.08). However, they had nonsignificantly higher rates of central compartment lymph node dissection (21% vs. 15%, p=.06). Across both cohorts, 25% (170/671) of patients had clinically significant thyroid cancer, with no difference in thyroid cancer rates based on the type of initial surgery (26% for total thyroidectomy vs. 22% for lobectomy, p=.3). The incidence of clinically significant thyroid cancer after initial lobectomy (i.e., requiring a 2-stage surgery) was significantly lower for patients managed with molecular diagnostics (17% vs. 43%, p<.001). An indeterminate FNA result had a sensitivity and specificity for the diagnostic of thyroid cancer of 89% and 27%, respectively, with a PPV of 29% and an NPV of 88%. The addition of molecular diagnostics to FNA results increased the specificity for a cancer diagnosis to 95% and the PPV to 82%.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.
A task force from the American Thyroid Association (2015) published a review with recommendations for the surgical management of FNA-indeterminate nodules using various molecular genetic tests.[49,] This review reported on the estimated likelihood of malignancy in an FNA-indeterminate nodule depending on results of the Afirma GEC test (described above) and other panels designed to rule in malignancy. Depending on the estimated prebiopsy likelihood of malignancy, recommendations for surgery included observation, active surveillance, repeat FNA, diagnostic lobectomy, or oncologic thyroidectomy.

Section Summary: Molecular Tests to Predict Malignancy

The available evidence has suggested that the use of variant testing in thyroid FNA samples is generally associated with high specificity and PPV for clinically significant thyroid cancer. The most direct evidence related to the clinical utility of variant testing for genes associated with malignancy in thyroid cancer comes from a single-center retrospective study that reported surgical decisions and pathology findings in patients managed with and without molecular diagnostics. There is a potential clinical utility for identifying malignancy with higher certainty on FNA if such testing permits better preoperative planning at the time of thyroid biopsy, potentially avoiding the need for a separate surgery. A statement from the American Thyroid Association provides some guidelines for surgeons managing patients with indeterminate nodules. However, adoption of these guidelines in practice and outcomes associated with them is uncertain.

Molecular Tests to Rule Out and Rule in Malignancy

Clinical Context and Test Purpose

The purpose of the ThyroSeq v3 test and the combined ThyGeNEXT Thyroid Oncogene Panel plus ThyraMIR microRNA classifier in individuals with indeterminate findings on FNA(s) of thyroid nodules is to predict malignancy and inform surgical planning decisions with positive results using ThyroSeq v3 or the ThyGeNEXT, and if negative, to predict benignancy using ThyraMIR microRNA classifier to eliminate or necessitate the need for surgical biopsy and guide surgical planning.

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with indeterminate findings on FNA(s) of thyroid nodules. Individuals with indeterminate findings presently proceed to surgical resection.

Interventions
The tests being considered are either: (a) the ThyroSeq v3 test or (b) the combined ThyGeNEXT Thyroid Oncogene Panel and ThyraMIR microRNA classifier testing.

Comparators
The following practices are currently being used: surgical biopsy and/or standard surgical management through surgical resection.

Outcomes
The potential beneficial outcomes of primary interest are using a true-negative result to avoid an unneeded surgical biopsy or using a true-positive result to guide surgical resection (e.g., hemithyroidectomy or thyroidectomy).

Potential harmful outcomes are those resulting from false-positive or false-negative test results. False-positive test results can lead to unnecessary surgical biopsy or resection and procedure-related complications. False-negative test results can lead to lack of surgical biopsy or resection for thyroid cancer and delay in diagnosis.

The time frame for evaluating the performance of the test varies from the initial FNA to surgical resection to weeks to months following an indeterminate result.

Study Selection Criteria
For the evaluation of clinical validity of the molecular testing, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores)
- Included a suitable reference standard
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

ThyroSeq v3 Test
Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence
Systematic Review
Lee et al (2022) performed a systematic review and meta-analysis on the diagnostic performance of molecular tests in the assessment of indeterminate thyroid nodules (described above).[20,] Inclusion criteria for trials included indeterminate thyroid results via FNA that included Bethesda categories III and IV, conclusive histopathological results in a group of benign and suspicious changes, and the use of Afirma GSC, ThyroSeq v3, and ThyGeNext as index tests. Investigators identified 6 studies on Thyroseq v3: 3 prospective, including Livhits et al (2021) and Steward et al (2019), described below, and 3 retrospective. Only 2 studies on ThyGeNext were identified and were excluded from meta-analysis due to the small sample size. Pooled data for ThyroSeq studies on 560 thyroid nodules demonstrated a sensitivity of 95.1% (95% CI, 91.1% to 97.4%), specificity of 49.6% (95% CI, 29.3% to 70.1%), PPV of 70% (95% CI, 55% to 83%), and NPV of 92% (95% CI, 86% to 97%). Limitations of this meta-analysis include the scarcity of available cohort analyses of the molecular tests and the lack of long-term findings.

Prospective Clinical Validation
Nikiforova et al (2018) reported on the performance of ThyroSeq v3 with 112 genes.[50,] The training sample included 238 surgically removed tissue samples consisting of 205 thyroid tissue samples representing all main types of benign and malignant tumors and nontumoral conditions. The validation sample included an independent set of 175 FNA samples of indeterminate cytology (see

Table 6). Using the cutoff identified in the training set, the ThyroSeq v3 sensitivity was 98% (95% CI, 93% to 99%), specificity was 82% (95% CI, 72% to 89%), with accuracy of 91% (95% CI, 86% to 94%) (see Table 7).

Steward et al (2019) conducted a multicenter validation study of ThyroSeq v3 in 256 patients with an indeterminate FNA who had surgery with histopathology (see Table 6).[51] Histopathology was reviewed by a central pathology panel and both cytologists and pathologists were blinded to the molecular results. For a benign result, ThyroSeq v3 had a sensitivity of 93%, a specificity of 81%, PPV of 68%, and NPV of 97% (see Table 7). Out of 152 test-negative samples, 5 (3%) were false-negatives. There were 105 cases with positive results, defined as cancer or noninvasive follicular thyroid neoplasm with papillary-like features. Two nodules had high-risk *TERT* or *TP53* variants (both positive for cancer), 13 had variants in *BRAF* V600E or *NTRK3*, or *BRAF*, or RET fusions (all positive for cancer), and 60 nodules were positive for variants in *RAS, BRAF K601E, PTEN, IDH2,* or *DICER1* or *PPARF-THADA* fusion (37 [62%] positive for cancer). No major limitations in study design and conduct of this validation study were identified. Because the nodules with low cancer probability genetic alterations were removed stological analysis, the long-term clinical impact of the genetic alterations could not be determined.

Livhits et al (2021) published a randomized controlled study that compared the ThyroSeq v3 test to the Afirma GSC test in patients with thyroid nodules with indeterminate results (Bethesda III or IV) (as described above).[25] The study reported clinical validity for both tests; the results of the ThyroSeq v3 test are summarized in Tables 6 and 7. The study included 171 nodules in the ThyroSeq v3 group. The sensitivity of ThyroSeq v3 was 96.9%, specificity was 84.8%, and the NPV was 99%. Long-term surveillance follow-up of nonoperatively-managed nodules in this trial, described in the section above, continued to support high NPV.[26] A limitation of the study is that pathologists that interpreted the histopathologic diagnosis were unblinded to the molecular test results. Additionally, the median length of surveillance did not reach 5 years as recommended by the American College of Radiology.

Table 6. Study Characteristics of Clinical Validity ThyroSeq v3

Study	Study Population	Design	Reference Standard	Threshold for Positive Index Test	Timing of Reference and Index Tests	Blinding of Assessors
Nikiforov et al (2018)[50]	175 samples with indeterminate cytology and known surgical follow-up	Retrospective	Histopathologic diagnosis	Cutoffs determined in the training sample	Samples were tested after surgical outcome was known	Unclear
Steward et al (2019) [51]	256 patients (286 nodules) with an indeterminate FNA (Bethesda III, IV, or V) and underwent thyroid surgery	Multicenter (10 sites) prospective validation study	Central pathology review	Classified as malignant or NIFPT or benign	Cross-sectional	Yes
Livhits et al (2021)[25]	171 nodules with indeterminate FNA (Bethesda III, IV) assigned to ThyroSeq v3*	Multicenter, randomized controlled trial	Histopathologic diagnosis	Classified as malignant or benign	Samples were tested after surgery	Assessors were unblinded to results of molecular testing

FNA: fine needle aspirate; Afirma GSC: Gene Sequencing Classifier; NIFPT: noninvasive follicular thyroid neoplasm with papillary-like features.
*Study included a comparator group assigned to Afirma GSC (reported previously)

Table 7. Clinical Validity of ThyroSeq v3

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Clinical Validity (95% Confidence Interval)			
					Sensitivity	Specificity	PPV	NPV
Nikiforov et al (2018)[50,]		175			98 (93 to 100)	81 (72 to 89)		
Steward et al (2019) [51,]	286	57	29 (10%)	30%	93 (86 to 97)	81 (75 to 86)	68 (58 to 76)	97 (93 to 99)
Livhits et al (2021)[25,]	171	163	8		96.9 (83.8 to 100)	84.8 (77 to 90.7)	63.3 (48.3 to 76.6)	99 (94.6 to 100)

NPV: negative predictive value; PPV: positive predictive value.

Additional studies describing the clinical validity of the ThyroSeq v3 panel in external settings (outside of the institution where it was developed) have reported on the diagnostic performance to predict malignancy in thyroid nodules that are indeterminate on FNA have been reported (see Table 8).

These studies differed from the previous studies in that noninvasive follicular thyroid neoplasm with papillary-like nuclear features was classified as not malignant for calculation of performance characteristics.

Table 8. Additional Clinical Validity Studies of ThyroSeq to Predict Malignancy in Indeterminate Thyroid FNA Samples

Study	Population	Genes and Rearrangements Tested	Insufficient or Inadequate for Analysis	Measures of Agreement (95% CI), %			
				Sen	Spec	PPV	NPV
Valderrabano et al (2017)[52,]	190 indeterminate thyroid nodules	ThyroSeq v2 (60+ genes)	2	70 (46 to 88)	77 (66 to 85)	42 (25 to 61)	91 (82 to 97)
Taye et al (2018)[53,]	156 indeterminate thyroid nodules	ThyroSeq v2 (60+ genes)	3	89 (52 to 100)	43 (29 to 58)	22 (10 to 38)	96 (78 to 99)

CI: confidence interval; FNA: fine needle aspiration; NPV: negative predictive value; PPV: positive predictive value; Sen: sensitivity; Spec: specificity.

Additional studies describing the clinical validity of the genes that comprise the ThyroSeq panel or other individual variants and combinations of variants to predict malignancy in thyroid nodules that are indeterminate on FNA have been reported. The results that pertain to the use of gene testing in indeterminate thyroid nodules are summarized in Table 9.

Table 9. Clinical Validity of Molecular Markers to Predict Malignancy in Indeterminate Thyroid FNA Samples

Study	Population	Genes and Rearrangements Tested	Insufficient or Inadequate for Analysis	Measures of Agreement, %				
				Sen	Spec	PPV	NPV	Acc
Moses et al (2010)[54,]	110 indeterminate thyroid nodules	*BRAF, KRAS, NRAS, RET/PTC1, RET/PTC3, NTRK1*	2	38	95	67	79	77
Ohori et al (2010)[55,]	100 patients with 117 atypia or follicular	*BRAF, NRAS, HRAS, KRAS, RET/PTC1,RET/PTC 3, PAX8/PPARγ*	NR	60	100	100	92	93

Study	Population	Genes and Rearrangements Tested	Insufficient or Inadequate for Analysis	Measures of Agreement, %			
	lesions of uncertain significance						
Beaudenon-Huibregtse et al (2014)[56,]	53 nodules with indeterminate or nondiagnostic FNA	*BRAF, HRAS, KRAS, NRAS, PAX8-PPARγ, RET-PTC1, RET-PTC3*		48	89	81	64

Acc: accuracy; FNA: fine needle aspiration; NPV: negative predictive value; NR: not reported; PPV: positive predictive value; PTC: papillary thyroid carcinoma; Sen: sensitivity; Spec: specificity.
[a] FNA-indeterminate nodules.
[b] FNA suspicious nodules.
[c] Atypia of indeterminate significance.
[d] Follicular neoplasm or suspicious for follicular neoplasm.
[e] Suspicious for malignancy.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials. Randomized controlled studies were not identified; however, a retrospective, single-center study found that use of ThyroSeq v3 in a cohort of patients with indeterminate thyroid nodules reduced the surgical resection rate compared to a cohort of patients without molecular testing.[57,] In addition, the risk of malignancy in thyroid nodules with a positive molecular test was higher than those without molecular testing.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

ThyGenX Thyroid Oncogene Panel and ThyraMIR microRNA Classifier

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Review of Evidence

Labourier et al (2015) evaluated the diagnostic algorithm combining a 17-variant panel with ThyraMIR on a cross-sectional cohort of thyroid nodules comprised of 109 FNA samples with AUS/FLUS or follicular neoplasm or SFN across 12 endocrinology centers.[58,] A summary of the sensitivity and specificity of the combined test is listed in Table 10.

Table 10. Summary of Clinical Validity for 17-Variant Panel and ThyraMIR on FNA Samples

Groups	No. of Cases	Sensitivity	Specificity	PPV	NPV	Odds Ratio
Cohort (95% CI), %	109	89 (73 to 97)	85 (75 to 92)	74 (58 to 86)	94 (85 to 98)	44 (13 to 151)
AUS/FLUS (95% CI), %	58	94 (73 to 100)	80 (64 to 91)	68 (46 to 85)	97 (84 to 100)	68 (8 to 590)

Groups	No. of Cases	Sensitivity	Specificity	PPV	NPV	Odds Ratio
FN/SFN (95% CI), %	51	82 (57 to 96)	91 (76 to 98)	82 (57 to 96)	91 (76 to 98)	48 (9 to 269)

Adapted from Labourier et al (2015).[58,]
AUS: atypia of undetermined significance; CI: confidence interval; FLUS: follicular lesion of undetermined significance; FN: follicular neoplasm; FNA: fine needle aspiration; NPV: negative predictive value; PPV: positive predictive value; SFN: suspicious for a follicular neoplasm.

Tumati et al (2024) retrospectively examined 387 Bethesda III/IV thyroid nodules that underwent ThyGenX+ ThyraMIR testing across 3 U.S. tertiary centers from 2017 to 2021.[59,] Utilizing a ≥10 % risk threshold, 262 cases (67.7%) were deemed benign, whereas the other 125 (32.3%) tested positive.

Surgery was performed more frequently for positive versus negative nodules (74.4% vs. 14.9%; p <.0001), and the corresponding pathology was more often malignant or non-invasive follicular thyroid neoplasm with papillary-like nuclear features in the positive cohort (46.4% vs. 3.4%; p <.0001). Comparing test-based risk stratification with surgical histopathology yielded a sensitivity of 86.6%, a specificity of 46.2%, an NPV of 76.9%, and a PPV of 62.4%.

Verma et al (2024) performed a comparison of the original ThyGenX + ThyraMIR test and the updated version (ThyGenX + ThyraMIR v2) in 338 Bethesda III/IV indeterminate thyroid nodules sampled from 2016 to 2020.[60,] The updated test classified fewer nodules as moderate-risk than the prior version of the test (33 [9.8%] vs. 55 [16.3%]; p <.001), and raised the benign call rate from 75.1% to 79.3%. At the study's 14% malignancy prevalence, the sensitivity, specificity, NPV, and PPV were 96%, 99%, 99%, and 89%, respectively, for the updated v2 test. The authors also noted the ThyraMIR v2 test showed significant improvements over the predicate test based on AUC analysis (p =.028).

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials.

Direct evidence for the clinical utility for the ThyroSeq v2 test and the combined ThyGenX and ThyraMIR diagnostic testing algorithm is lacking.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

A chain of evidence may be constructed to infer the potential clinical utility of the combined diagnostic testing algorithm. No studies using ThyGenX NGS panel in FNA samples were identified. However, available evidence has suggested that the use of variant testing using NGS in thyroid FNA samples is generally associated with high specificity and PPV for clinically significant thyroid cancer.

There is the potential clinical utility for identifying malignancy with higher certainty on FNA if such testing permits better preoperative planning at the time of thyroid biopsy, potentially avoiding the need for a separate surgery. However, the variant analysis does not achieve an NPV sufficiently high enough to identify which patients can undergo active surveillance over thyroid surgery. In the diagnostic algorithm that reflexes to the ThyraMIR after a negative ThyGenX result, patients receiving reflex testing could identify who may undergo active surveillance over thyroid surgery. A single study using a 17-variant panel with ThyraMIR showed an NPV of 94%. Therefore, the high NPV

of ThyraMIR has the potential to accurately predict benignancy and triage patients to active surveillance.

Section Summary: Molecular Markers to Rule Out and Rule in Malignancy

Evidence for the clinical validity of the ThyroSeq v3 NGS panel comes from a systematic review of prospective and retrospective studies and a major prospective clinical validity study. In a systematic review including 3 prospective and 3 retrospective clinical validity studies, sensitivity of ThyroSeq v3 was 95.1%, specificity was 49.6%, PPV was 70%, and NPV was 92%. In the prospective clinical validity study, the performance characteristics were sensitivity, 93%; specificity, 81%; PPV, 68%; NPV, 97%. A randomized controlled trial found similar results with ThyroSeq v3. In 2 independent validation studies with a predicate test (ThyroSeq v2) in which noninvasive follicular thyroid neoplasm with papillary-like nuclear features was categorized as not malignant, performance characteristics were lower and variable (sensitivity, 70% to 89%; specificity, 43% to 77%; PPV, 22% to 42%; NPV, 91% to 96%).

Evidence for the clinical validity of combined testing for miRNA gene expression using ThyraMIR and a targeted 17-variant panel comes from 2 retrospective studies using archived surgical specimens and FNA samples. One study combined a 17-variant panel with ThyraMIR testing on archived surgical specimens and resulted in a sensitivity of 85% and specificity of 95%. The second study combined a 17-variant panel (miR*Inform*) with ThyraMIR testing on FNA samples and resulted in a sensitivity of 89%, a specificity of 85%, PPV of 74%, and NPV of 94%. No studies were identified that demonstrated the clinical validity of a combined ThyGenX and ThyraMIR test on FNA samples.

Direct evidence for the clinical utility for the ThyroSeq v2 test and the combined ThyGenX and ThyraMIR reflex testing is lacking. However, available evidence has suggested that testing for gene variants and rearrangements can predict malignancy and inform surgical planning decisions when the test is positive. Pooled retrospective and prospective clinical validation studies of ThyroSeq v2 have reported a combined NPV of 96% (95% CI, 92% to 95%) and PPV of 83% (95% CI, 72% to 95%) and might potentially assist in selecting patient to avoid surgical biopsy if negative and guide surgical planning if positive. One retrospective study found that combined ThyGenX + ThyraMIR had an NPV of 99%, while another study reported an NPV of 76.9%, indicating a high but variable ability to identify nodules suitable for surveillance. In a reflex testing setting, the high NPV for a microRNA gene expression test used on the subset of patients with a negative result from a variant and gene rearrangement testing may provide incremental information in identifying patients appropriately for active surveillance, but improvements in health outcomes are still uncertain.

Supplemental Information

The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

Clinical Input From Physician Specialty Societies and Academic Medical Centers

While the various physician specialty societies and academic medical centers may collaborate with and make recommendations during this process, through the provision of appropriate reviewers, input received does not represent an endorsement or position statement by the physician specialty societies or academic medical centers, unless otherwise noted.

2017 Input

Clinical input was sought to help determine whether testing for molecular markers in fine needle aspirates of the thyroid for management of individuals with thyroid nodule(s) with an indeterminate finding on the fine needle aspirates (FNAs) would provide a clinically meaningful improvement in net health outcome and whether the use is consistent with generally accepted medical practice. In response to requests, clinical input on 7 tests for molecular markers was received from 9 respondents, including 1 specialty society-level response, 1 physician from an academic center, and 7 physicians from 2 health systems.

Clinical input supports that the following uses provide a clinically meaningful improvement in net health outcome and indicates the uses are consistent with generally accepted medical practice:
For individuals who have FNA of thyroid nodules with indeterminate cytologic findings (i.e., Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/suspicion for a follicular neoplasm]) who receive the following types of molecular marker testing to rule out malignancy and to avoid surgical biopsy:

- Afirma Gene Expression Classifier; or
- ThyroSeq v2

For individuals who have FNA of thyroid nodules with indeterminate cytologic findings or Bethesda diagnostic category V (suspicious for malignancy) who receive the following types of molecular marker testing to rule in the presence of malignancy to guide surgical planning for the initial resection rather than a 2 stage surgical biopsy followed by definitive surgery:

- ThyroSeq v2;
- ThyraMIR microRNA/ThyGenX;
- Afirma BRAF after Afirma Gene Expression Classifier; or
- Afirma MTC after Afirma Gene Expression Classifier.

Further details from clinical input are included in the Appendix.

Practice Guidelines and Position Statements

Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

American Association of Clinical Endocrinologists et al

The American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazone Medici Endocrinologi (2016) updated their joint guidelines on molecular testing for cytologically indeterminate thyroid nodules, stating[61]:

- "Cytopathology expertise, patient characteristics, and prevalence of malignancy within the population being tested impact the negative predictive values (NPVs) and positive predictive values (PPVs) for molecular testing."
- "Consider the detection of *BRAF* and *RET/PTC* and, possibly, *PAX8/PPARG* and *RAS* mutations if such detection is available."
- "*TERT* [Telomerase reverse transcriptase] mutational analysis on FNA, when available, may improve the diagnostic sensitivity of molecular testing on cytologic samples."
- "Because of the insufficient evidence and the limited follow-up, we do not recommend either in favor of or against the use of gene expression classifiers (GECs) for cytologically indeterminate nodules."

For the role of molecular testing for deciding the extent of surgery the following recommendations were made:

- "Currently, with the exception of mutations such as BRAFV600E that have a PPV approaching 100% for papillary thyroid carcinoma (PTC), evidence is insufficient to recommend in favor of or against the use of mutation testing as a guide to determine the extent of surgery."

American College of Radiology

The American College of Radiology (ACR; 2017) Thyroid Imaging, Reporting, and Data System (TI-RADS) Committee published a white paper with expert consensus recommendations for FNA biopsy thresholds and imaging surveillance.[19,] Regarding timing of follow-up sonograms, the publication

states: "We advocate timing on the basis of a nodule's ACR TI-RADS level, with additional sonograms for lesions that are more suspicious. For a TR5 lesion, we recommend scans every year for up to 5 years. For a TR4 lesion, scans should be done at 1, 2, 3, and 5 years. For a TR3 lesion, follow-up imaging may be performed at 1, 3, and 5 years. Imaging can stop at 5 years if there is no change in size, as stability over that time span reliably indicates that a nodule has a benign behavior. There is no published evidence to guide management of nodules that enlarge significantly but remain below the FNA size threshold for their ACR TI-RADS level at 5 years, but continued follow-up is probably warranted. If a nodule's ACR TI-RADS level increases on follow-up, the next sonogram should be done in 1 year, regardless of its initial level."

American Thyroid Association

The American Thyroid Association (ATA; 2016) updated its guidelines on the management of thyroid nodules and differentiated thyroid cancer in adults.[62,] These guidelines made the following statements on molecular diagnostics in thyroid nodules that are atypia of undetermined significance or follicular lesion of undetermined significance on cytology and follicular neoplasm or suspicious for follicular neoplasm on cytology (see Table 11).

Table 11. Molecular Diagnostics in Thyroid Nodules on Cytology

Recommendation	SOR	QOE
AUS or FLUS		
"For nodules with AUS/FLUS cytology, after consideration of worrisome clinical and sonographic features, investigations such as repeat FNA or molecular testing may be used to supplement malignancy risk assessment in lieu of proceeding directly with a strategy of either surveillance or diagnostic surgery. Informed patient preference and feasibility should be considered in clinical decision-making."	Weak	Moderate
"If repeat FNA cytology, molecular testing, or both are not performed or inconclusive, either surveillance or diagnostic surgical excision may be performed for an AUS/FLUS thyroid nodule, depending on clinical risk factors, sonographic pattern, and patient preference."	Strong	Low
FN or SFN		
"Diagnostic surgical excision is the long-established standard of care for the management of FN/SFN cytology nodules. However, after consideration of clinical and sonographic features, molecular testing may be used to supplement malignancy risk assessment data in lieu of proceeding directly with surgery. Informed patient preference and feasibility should be considered in clinical decision-making."	Weak	Moderate

AUS: atypia of undetermined significance; FLUS: follicular lesion of undetermined significance; FN: follicular neoplasm; FNA: fine needle aspirate; QOE: quality of evidence; SFN: suspicious for follicular neoplasm; SOR: strength of recommendation.

The guidelines also stated: "there is currently no single optimal molecular test that can definitively rule in or rule out malignancy in all cases of indeterminate cytology, and long-term outcome data proving clinical utility are needed."

National Comprehensive Cancer Network

National Comprehensive Cancer Network (v1.2025) guidelines on the treatment of thyroid cancer comment on the use of molecular diagnostics in thyroid cancer.[63,] For thyroid nodules evaluated with FNA, molecular diagnostics may be employed when lesions are suspicious for:

- Follicular or oncocytic neoplasms.
- Atypia of undetermined significance or follicular lesions of undetermined significance.

The guidelines state that molecular diagnostics have not performed well historically for oncocytic carcinoma. The guideline also endorses the ATA and ACR recommendations for nodule surveillance, described previously above.

U.S. Preventive Services Task Force Recommendations

Not applicable.

Medicare National Coverage

There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Ongoing and Unpublished Clinical Trials

Some currently ongoing and unpublished trials that might influence this review are listed in Table 12.

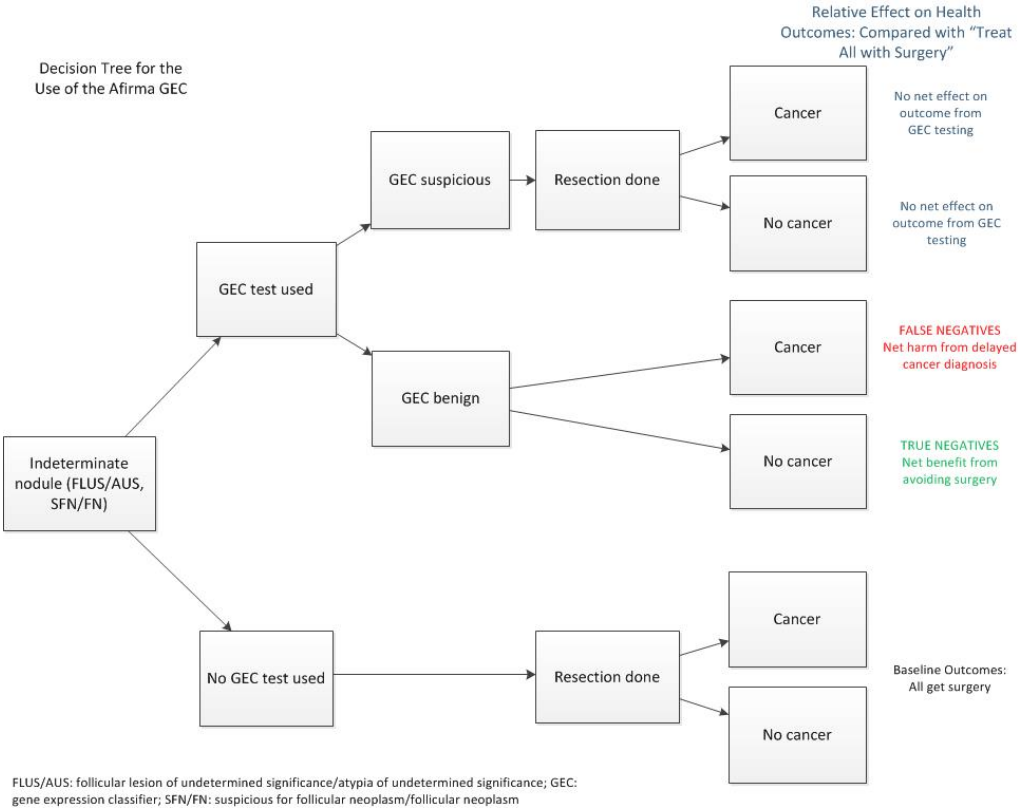
Table 12. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT02681328	Randomized Trial Comparing Performance of Molecular Markers for Indeterminate Thyroid Nodules	328	Dec 2026
Unpublished			
NCT05025046[a]	Prospective, Blinded, Multi-center Clinical Study of NGS-based Thyroscan Genomic Classifier in the Diagnosis of Thyroid Nodules	400	Jun 2022 (unknown)
NCT03170804	Registry for Genomic Profiling of Nodular Thyroid Disease and Thyroid Cancer	200	Jan 2020 (unknown)
NCT02947035	Molecular Testing to Direct Extent of Initial Thyroid Surgery	100	Jun 2023 (completed)

NCT: national clinical trial.
[a] Denotes industry-sponsored or cosponsored trial.

Appendix 1

Appendix Figure 1. Decision Model for the Afirma GEC Use



Appendix 2: Clinical Input

CI - Objective

In 2017, clinical input was sought to help determine whether the evidence and clinical experience support a clinical benefit of testing for molecular markers in fine needle aspirates of the thyroid for

management of individuals with thyroid nodule(s) with an indeterminate finding on the fine needle aspirates.

Respondents

Clinical input was provided by the following specialty societies and physician members identified by a specialty society or clinical health system:

- American Thyroid Association
- Anonymous, MD, Endocrine Surgery, Associate Professor of Surgery; Identified by Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine[a]
- Anonymous, DO, Hematology, Medical Hematology; Identified by Cancer Treatment Centers of America (CTCA)
- Anonymous, MD, Pathology and Laboratory Medicine; Identified by CTCA
- Anonymous, MD, Endocrine; Identified by CTCA
- Bradley R. Mons, DO; Otolaryngology; Identified by CTCA
- Helen Yoo Bowne, MD, Otolaryngology - Head and Surgery; Identified by CTCA
- Asha Karippot, MD, Medical Oncology and Hematology; Identified by CTCA
- Jerome B. Myers, MD, PhD, Pathology; Identified by Catholic Health Initiatives

[a] Indicates that information on conflicts of interest related to the topic where clinical input is being soughtwere identified by this respondent (see Appendix 1).

Clinical input provided by the specialty society at an aggregate level is attributed to the specialty society. Clinical input provided by a physician member designated by the specialty society or health system is attributed to the individual physician and is not a statement from the specialty society or health system. Specialty society and physician respondents participating in the Evidence Street® clinical input process provide a review, input, and feedback on topics being evaluated by Evidence Street. However, participation in the clinical input process by a special society and/or physician member designated by the specialty society or health system does not imply an endorsement or explicit agreement with the Evidence Opinion published by BCBSA or any Blue Plan.

Appendix 2

Clinical Input Responses

Appendix 2 – Figure 1:

Use of molecular testing in FNA of thyroid nodules with indeterminate findings to rule out malignancy.

Molecular Marker	Respondent	Specialty	Identified by	Confidence Level That Clinical Use Expected to Provide Meaningful Clinical Benefit: Yes or No	NO: 5 (High)	4	3	2	1 (Low)	YES: 1 (Low)	2	3	4	5 (High)	Confidence Level that Clinical Use is Consistent with Generally Accepted Medical Practice: Yes or No	NO: 5 (High)	4	3	2	1 (Low)	YES: 1 (Low)	2	3	4	5 (High)
Afirma Gene Expression Classifier	ATA	Thyroid Disease		YES									■		YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES									■		YES									■	
	Anonymous	Hematology/Oncology	CTCA	YES									■		YES									■	
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											YES									■	
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES									■		YES									■	
	Anonymous	Endocrinology	CTCA	YES										■	YES										■
	Dr. Karippot	Hematology/Oncology	CTCA	YES									■		NO		■								
	Dr. Myers	Pathology	CHI	YES									■		YES									■	
ThyroSeq v.2	ATA	Thyroid Disease		YES									■		YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES									■		YES									■	
	Anonymous	Hematology/Oncology	CTCA	YES									■		YES									■	
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											NO										
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES									■		YES									■	
	Anonymous	Endocrinology	CTCA	YES										■	YES										■
	Dr. Karippot	Hematology/Oncology	CTCA	YES							■				NO			■							
	Dr. Myers	Pathology	CHI	YES									■		YES								■		
ThyraMIR microRNA	ATA	Thyroid Disease		YES								■			YES							■			
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	NO	■										NO				■						
	Anonymous	Pathology & Lab Med.	CTCA	YES								■			YES									■	
	Anonymous	Hematology/Oncology	CTCA	NO					■						NO				■						
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											YES									■	
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES						■					NO			■							
	Anonymous	Endocrinology	CTCA	YES									■		YES										■
	Dr. Karippot	Hematology/Oncology	CTCA	YES								■			NO			■							
	Dr. Myers	Pathology	CHI	YES									■		YES								■		
Rosetta GX Reveal	ATA	Thyroid Disease		YES								■			YES							■			
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	NO	■										NO				■						
	Anonymous	Pathology & Lab Med.	CTCA	NR											NR										
	Anonymous	Hematology/Oncology	CTCA	NO					■						NO					■					
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											NO										
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	NO					■						NO										
	Anonymous	Endocrinology	CTCA	YES									■		YES										■
	Dr. Karippot	Hematology/Oncology	CTCA	NR											NR										
	Dr. Myers	Pathology	CHI	YES									■		YES								■		

Gray shading means no rating provided.
ATA: American Thyroid Association; CHI: Catholic Health Initiatives; CTCA: Cancer Treatment Centers of America.
* Indicates that information on conflicts of interest related to the topic where clinical input is being sought was identified by this respondent.

Appendix 2 - Figure 2

Use of molecular testing in FNA of thyroid nodules with indeterminate findings or Bethesda diagnostic category 5 to rule in malignancy.

Molecular Marker	Respondent	Specialty	Identified by	Confidence Level That Clinical Use Expected to Provide Meaningful Clinical Benefit: Yes or No	NO: High 5	NO: 4	NO: 3	NO: 2	NO: Low 1	YES: Low 1	YES: 2	YES: Intermediate 3	YES: 4	YES: High 5	Confidence Level that Clinical Use is Consistent with Generally Accepted Medical Practice: Yes or No	NO: High 5	NO: 4	NO: Intermediate 3	NO: 2	NO: Low 1	YES: Low 1	YES: 2	YES: Intermediate 3	YES: 4	YES: High 5
ThyroSeq v.2	ATA	Thyroid Disease		YES								■			YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES									■		YES									■	
	Anonymous	Hematology/Oncology	CTCA	YES									■		YES									■	
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											NO										
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES									■		YES									■	
	Anonymous	Endocrinology	CTCA	YES										■	YES										■
	Dr. Karippot	Hematology/Oncology	CTCA	YES								■			NO			■							
	Dr. Myers	Pathology	CHI	YES									■		YES								■		
ThyGenX	ATA	Thyroid Disease		YES								■			YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES									■		YES								■		
	Anonymous	Hematology/Oncology	CTCA	NO			■								NO			■							
	Dr. Mons	Otolaryngology (ORL)	CTCA	YES									■		YES								■		
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	NO			■								NO			■							
	Anonymous	Endocrinology	CTCA	YES									■		YES									■	
	Dr. Karippot	Hematology/Oncology	CTCA	YES								■			NO		■								
	Dr. Myers	Pathology	CHI	YES									■		YES								■		
Afirma BRAF	ATA	Thyroid Disease		YES								■			YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES							■				YES							■			
	Anonymous	Hematology/Oncology	CTCA	YES									■		YES									■	
	Dr. Mons	Otolaryngology (ORL)	CTCA	YES										■	YES								■		
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES									■		YES									■	
	Anonymous	Endocrinology	CTCA	YES						■					YES							■			
	Dr. Karippot	Hematology/Oncology	CTCA	YES									■		NO	■									
	Dr. Myers	Pathology	CHI	YES									■		YES									■	
Afirma MTC	ATA	Thyroid Disease		YES								■			YES									■	
	Anonymous*	Endocrine Surgery	Baylor College of Medicine	YES										■	YES										■
	Anonymous	Pathology & Lab Med.	CTCA	YES							■				YES							■			
	Anonymous	Hematology/Oncology	CTCA	YES									■		YES									■	
	Dr. Mons	Otolaryngology (ORL)	CTCA	NO											NO										
	Dr. Yoo Bowne	ORL/Head&Neck Surg.	CTCA	YES									■		YES									■	
	Anonymous	Endocrinology	CTCA	YES						■					YES							■			
	Dr. Karippot	Hematology/Oncology	CTCA	YES									■		NO	■									
	Dr. Myers	Pathology	CHI	YES									■		YES									■	

Grey shading means no rating provided.
ATA: American Thyroid Association; CHI: Catholic Health Initiatives; CTCA: Cancer Treatment Centers of America.
* Indicates that information on conflicts of interest related to the topic where clinical input is being soughtwas identified by this respondent.

Appendix 2 - Figure 3:

For the indication using molecular marker testing in FNA of thyroid nodules with indeterminate findings[1] to rule out malignancy, this testing may provide a meaningful clinical benefit defined by avoiding surgical biopsy with an acceptably low trade off in missed malignancy.

What negative predictive value from molecular marker testing would be needed for a clinically meaningful reduction in the frequency of negative surgical biopsies?

Respondent	Identified by
ATA	
Anonymous*	Baylor College of Medicine
Anonymous	Cancer Treatment Centers of America
Anonymous	Cancer Treatment Centers of America
Dr. Mons	Cancer Treatment Centers of America
Dr. Yoo Bowne	Cancer Treatment Centers of America
Anonymous	Cancer Treatment Centers of America
Dr. Karippot	Cancer Treatment Centers of America
Dr. Myers	Catholic Health Initiatives

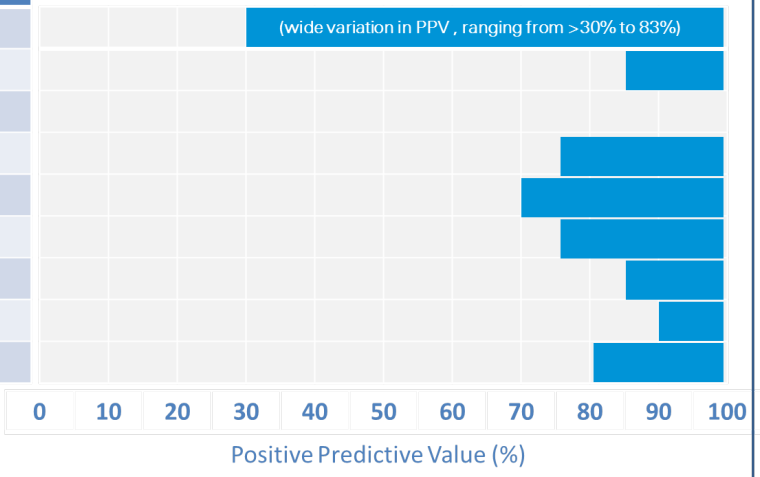
[1]Indeterminate finding on the fine needle aspirate means Bethesda diagnostic category III = atypia/follicular lesion of undetermined significance AFUS/FLUS or Bethesda diagnostic category IV = follicular neoplasm/suspicion for a follicular neoplasm FN/SFN.

Figure 4:

For the indication using molecular marker testing in FNA of thyroid nodules with indeterminate findings[1] or Bethesda diagnostic category V (suspicious for malignancy) to rule in the presence of malignancy, this testing may provide a meaningful clinical benefit defined by guiding definitive surgery planning for the initial resection rather than a two-stage surgical biopsy followed by definitive surgery.

What positive predictive value from molecular marker testing would be needed for a clinically meaningful prediction?

Respondent	Identified by
ATA	
Anonymous*	Baylor College of Medicine
Anonymous	Cancer Treatment Centers of America
Anonymous	Cancer Treatment Centers of America
Dr. Mons	Cancer Treatment Centers of America
Dr. Yoo Bowne	Cancer Treatment Centers of America
Anonymous	Cancer Treatment Centers of America
Dr. Karippot	Cancer Treatment Centers of America
Dr. Myers	Catholic Health Initiatives



[1]Indeterminate finding on the fine needle aspirate means Bethesda diagnostic category III = atypia/follicular lesion of undetermined significance AFUS/FLUS or Bethesda diagnostic category IV = follicular neoplasm/suspicion for a follicular neoplasm FN/SFN.

Appendix Table 1. Respondent Profile

	Specialty Society	
No.	Name of Organization	Clinical Specialty
1	American Thyroid Association	Thyroid disease

	Specialty Society				
	Physician				
No.	Name	Degree	Institutional Affiliation	Clinical Specialty	Board Certification and Fellowship Training
Identified by Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine					
2	Anonymous	MD	Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine	Endocrine surgery	American Board of Surgery, General Surgery. Fellowship training Endocrine Surgery
Identified by Cancer Treatment Centers of America					
3	Anonymous	MD	Cancer Treatment Centers of America	Pathology and laboratory medicine	American Board of Pathology
4	Anonymous	DO	Cancer Treatment Centers of America - ERMC	Hematology, medical oncology	Hematology, Medical Oncology, Internal Medicine
5	Bradley R. Mons	DO	Southwestern Regional Medical Center	Otolaryngology	American Osteopathic Board of Otolaryngology
6	Helen Yoo Bowne	MD	Cancer Treatment Centers of America	Otolaryngology - head and surgery	American Academy of Otolaryngology - Head and Neck Surgery
7	Anonymous	MD	Cancer Treatment Centers of America	Endocrine	ABIM/Endocrine; Rosalind Franklin University of Medicine and Science, Chicago Medical School
8	Asha Karippot	MD	Cancer Treatment Centers of America	Medical oncology and hematology	Medical Oncology, Hematology & Internal Medicine - American Board of Internal Medicine; Fellowship Training: Hematology/Oncology - Seton Hall University, St. Joseph Regional Medical Center, Paterson, NJ
Identified by Catholic Health Initiatives					
9	Jerome B. Myers	MD, PhD	Catholic Health Initiatives	Pathology	Anatomic/Clinical Pathology, Hematopathology

Appendix Table 2. Respondent Conflict of Interest Disclosure

No.	1. Research support related to the topic where clinical input is being sought		2. Positions, paid or unpaid, related to the topic where clinical input is being sought		3. Reportable,more than $1000, healthcare-related assets or sources of income for myself, my spouse, or my dependent children related to the topic where clinical input is being sought		4. Reportable, more than $350, gifts or travel reimbursements for myself, my spouse, or my dependent children related to the topic where clinical input is being sought	
	Yes/No	Explanation	Yes/No	Explanation	Yes/No	Explanation	Yes/No	Explanation
1	No		No		No		No	
2	Yes	I am a co-PI on a Texas cancer research fund (CPRIT) project to use mass spectrometry to develop next-generation diagnostics in thyroid cancer	No		No		No	
3	No		No		No		No	
4	No		No		No		No	
5	No		No		No		No	
6	No		No		No		No	

No.	1. Research support related to the topic where clinical input is being sought	2. Positions, paid or unpaid, related to the topic where clinical input is being sought	3. Reportable,more than $1000, healthcare-related assets or sources of income for myself, my spouse, or my dependent children related to the topic where clinical input is being sought	4. Reportable, more than $350, gifts or travel reimbursements for myself, my spouse, or my dependent children related to the topic where clinical input is being sought
7	No	No	No	No
8	No	No	No	No
9	No	No	No	No

Individual physician respondents answered at individual level. Specialty Society respondents provided aggregate information that may be relevant to the group of clinicians who provided input to the Society-level response.

Appendix 2: Clinical Input Responses

CI - Objective

Clinical input is sought to help determine whether the evidence and clinical experience support a clinical benefit of testing for molecular markers in fine needle aspirates (FNA) of the thyroid for management of individuals with thyroid nodule(s) with an indeterminate finding on the FNA.

Responses

- For the indication using molecular marker testing in FNA of thyroid nodules with indeterminate findings[a] to rule out malignancy, this testing may provide a meaningful clinical benefit defined by avoiding surgical biopsy with an acceptably low trade-off in missed malignancy. What negative predictive value from molecular marker testing wouldbe needed for a clinically meaningful reduction in the frequency of negative surgical biopsies? Please include relevant references to support your clinical input.

No.	Rationale
1	The majority of the panel preferred an NPV of 95%with one response for 90%. The entire panel responded YES to all of the tests for **Clinically Meaningful Benefit** with the confidence in the evidence as follows, listing average and range: Afirma GEC: 3.5 (3-4); ThyroSeq v.2 (note v.3 available as of 1 November 2017): 3.5 (2-4); ThyraMIR microRNA: 1.5 (1-2); Rosetta: 2 (1-3) but with only 2 providing response. All responded YES to all the tests for **Generally Accepted Medical Practice** with confidence in evidence: Afirma GEC: 4 (3-5); ThyroSeq: 4 (4); ThyraMIR microRNA: 1.5 (1-2); Rosetta: all panelists left blank [BCBSA insertion: ATA also noted corrections to the indication that indeterminate findings would be Bethesda III, IV, or V to rule out malignancy.]
2	An NPV of 90% or greater is necessary for a meaningful clinical reduction in frequency of negative surgeries. This is because there is a 5% to even 10% error rate associated with fine needle aspiration biopsy in the diagnosis of malignant lesions. • 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.* Nov 2014;99(11):4069-4077. PMID 24780044.
3	The commercially available molecular tests for indeterminate thyroid FNAs have been found to be good at ruling out malignancy. From Afirma to ThyroSeq to ThyGenX, their negative predictive value has been reported to be very good. I would like to see their NPV to be at least in 90%, preferably in 95%. • Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med.* Aug 23 2012;367(8):705-715. PMID 22731672

No.	Rationale
	• 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.* Nov 2015;22(12):3996-4001. PMID 25862581 • 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.* Jul 2015;100(7):2743-2750. PMID 25965083 • Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer.* Dec 01 2014;120(23):3627-3634. PMID 25209362
4	Negative predictive value should be 95%. • Rago T, Scutari M, Latrofa F, et al. The large majority of 1520 patients with indeterminate thyroid nodule at cytology have a favorable outcome, and a clinical risk score has a high negative predictive value for more cumbersome cancer disease. *J Clin Endocrinol Metab.* Oct 2014;99(10):3700-3707. PMID 24708101
5	No negative predictive value is needed if the patient desires surgery. Some patients undergo undue stress knowing that there is a tumor in their body and having a physician just want to watch it.
6	90 - 95% NPV • Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab.* Dec 2010;95(12):5296-5304. PMID 20826580
7	Preferable NPV of 95%, > 90% acceptable • ATA differentiated thyroid cancer and thyroid nodular guidelines 2016. https://www.thyroid.org/professionals/ata-professional-guidelines/ • 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.* Jan 2016;26(1):1-133. PMID 26462967
8	A high NPV is needed ~ 90 - 95% Afirma gene expression classifier may provide clinical benefit • Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab.* Dec 2010;95(12):5296-5304. PMID 20826580
9	90 - 95%

ATA: American Thyroid Association; NPV: negative predictive value.

- Based on the evidence and your clinical experience for the indications described in Question 1:
  - Respond Yes or No for each clinical indication whether each of the following tests would be expected to provide a clinically meaningful benefit in the net health outcome.
  - Use the 1 to 5 scale outlined below to indicate your level of confidence that there is adequate evidence that supports your conclusions.

No.	Indications	Yes/No	Low Confidence		Intermediate Confidence		High Confidence
			1	2	3	4	5
1	Afirma Gene Expression Classifier	Yes				X	
	ThyroSeq v.2	Yes				X	
	ThyraMIR microRNA	Yes			X		
	RosettaGX Reveal	Yes			X		
2	Afirma Gene Expression Classifier	Yes					X
	ThyroSeq v.2	Yes					X
	ThyraMIR microRNA	No					X
	RosettaGX Reveal	No					X

No.	Indications	Yes/No	Low Confidence	Intermediate Confidence	High Confidence
3	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes		X	
	RosettaGX Reveal	NR	No rating provided		
4	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	No	X		
	RosettaGX Reveal	No	X		
5	Afirma Gene Expression Classifier	No	No rating provided		
	ThyroSeq v.2	No	No rating provided		
	ThyraMIR microRNA	No	No rating provided		
	RosettaGX Reveal	No	No rating provided		
6	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes	X		
	RosettaGX Reveal	No	X		
7	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes			X
	RosettaGX Reveal	Yes			X
8	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes		X	
	ThyraMIR microRNA	Yes		X	
	RosettaGX Reveal	NR	No rating provided		
9	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes			X
	RosettaGX Reveal	Yes			X

NR: not reported.

- Based on the evidence and your clinical judgment for the indications described in Question 1:
  - Respond Yes or No for each indication whether each of the following tests is consistent with generally accepted medical practice.
  - Use the 1 to 5 scale outlined below to indicate your level of confidence in your conclusions.

No.	Indications	Yes/No	Low Confidence		Intermediate Confidence		High Confidence
			1	2	3	4	5
1	Afirma Gene Expression Classifier	Yes				X	
	ThyroSeq v.2	Yes				X	
	ThyraMIR microRNA	Yes		X			
	RosettaGX Reveal	Yes		X			
2	Afirma Gene Expression Classifier	Yes					X
	ThyroSeq v.2	Yes					X
	ThyraMIR microRNA	No		X			
	ThyraMIR microRNA	No		X			
3	Afirma Gene Expression Classifier	Yes				X	

No.	Indications	Yes/No	Low Confidence	Intermediate Confidence	High Confidence
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes			X
	RosettaGX Reveal	NR	No rating provided		
4	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	No		X	
	RosettaGX Reveal	No	X		
5	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	No	No rating provided		
	ThyraMIR microRNA	Yes			X
	RosettaGX Reveal	Yes	No rating provided		
6	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	No		X	
	RosettaGX Reveal	No	No rating provided		
7	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes			X
	ThyraMIR microRNA	Yes			X
	RosettaGX Reveal	Yes			X
8	Afirma Gene Expression Classifier	No			X
	ThyroSeq v.2	No		X	
	ThyraMIR microRNA	No		X	
	RosettaGX Reveal	NR	No rating provided		
9	Afirma Gene Expression Classifier	Yes			X
	ThyroSeq v.2	Yes		X	
	ThyraMIR microRNA	Yes		X	
	RosettaGX Reveal	Yes		X	

NR: not reported.

- For the indication using molecular marker testing in FNA of thyroid nodules with indeterminate findings[b] or Bethesda diagnostic category V (suspicious for malignancy) to rule in the presence of malignancy, this testing may provide a meaningful clinical benefit defined by guiding definitive surgery planning for the initial resection rather than a 2-stage surgical biopsy followed by definitive surgery. What positive predictive value from molecular marker testing would be needed for a clinically meaningful prediction? Please include relevant references to support your clinical input.

No.	Rationale
1	There was a wide variation in PPV needed for clinically meaningful prediction, ranging from > 30% to 83%. Most panelists responded Yes to **Clinically Meaningful Benefit** for the following mutations with the confidence in evidence as: ThyroSeq: 3.3 (3-4); ThyGenX: 2.3 (1-3); Afirma BRAF: 3.7 (2-5); Afirma MTC: 3.7 (2-5) with one NO response and confidence in evidence as ThyroSeq: 4; ThyGenX: 4; Afirma BRAF: 3; Afirma MTC: 3. Most responded Yes to **Generally Accepted Medical Practice** with confidence in evidence as ThyroSeq: 4.3 (3-4); ThyGenX: 3.3 (1-5); Afirma BRAF: 4 (2-5); Afirma MTC: 4 (2-5) with one NO response and confidence in evidence as ThyroSeq: 2; ThyGenX: 3; Afirma BRAF: 3; Afirma MTC: 3.
2	A positive predictive value of 85% or greater would be clinically meaningful. This will help guide extensive surgery regarding total thyroidectomy versus partial thyroidectomy.

No.	Rationale
3	Studies done on ThyroSeq v2 have shown that this test is better at ruling in malignancy when compared to the other known molecular tests. In addition, the use of specific single gene variant of BRAF and the translocation would significantly help in the planning of the surgery on whether to go for a lobectomy versus a total thyroidectomy with central lymph node dissection. • Zhang M, Lin O. Molecular Testing of Thyroid Nodules: A Review of Current Available Tests for Fine-Needle Aspiration Specimens. *Arch Pathol Lab Med.* Dec 2016;140(12):1338-1344. PMID 27557410 • Nikiforova MN, Wald AI, Roy S, et al. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab.* Nov 2013;98(11):E1852-1860. PMID 23979959 • 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.* Feb 2014;24(2):305-313. PMID 23837487
4	Positive predictive value should be 75 to 100% • Nikiforov YE. Molecular diagnostics of thyroid tumors. *Arch Pathol Lab Med.* May 2011;135(5):569-577. PMID 21526955
5	A PPV of over 70% would be needed for me to have a clinically meaningful prediction. I do not have any references besides personal preferences.
6	75-100% PPV • Nikiforov YE. Molecular diagnostics of thyroid tumors. *Arch Pathol Lab Med.* May 2011;135(5):569-577. PMID 21526955 • Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer.* Dec 01 2014;120(23):3627-3634. PMID 25209362 • 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.* Nov 2015;25(11):1217-1223. PMID 26356635 • 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.* Nov 2011;96(11):3390-3397. PMID 21880806 • Ferraz C, Eszlinger M, Paschke R. Current state and future perspective of molecular diagnosis of fine-needle aspiration biopsy of thyroid nodules. *J Clin Endocrinol Metab.* Jul 2011;96(7):2016-2026. PMID 21593119
7	PPV ~ 85% ATA guidelines and clinical experience
8	PPV of 90 - 95% needed for meaningful prediction.
9	80 - 85%

ATA: American Thyroid Association.

- Based on the evidence and your clinical experience for the indications described in Question 4:
  - Respond Yes or No for each clinical indication whether each of the following tests would be expected to provide a clinically meaningful benefit in the net health outcome.
  - Use the 1 to 5 scale outlined below to indicate your level of confidence that there is adequate evidence that supports your conclusions.

No.	Indications	Yes/No	Low Confidence		Intermediate Confidence		High Confidence
			1	2	3	4	5
1	ThyroSeq v.2	Yes			X		
	ThyGenX	Yes			X		
	Afirma BRAF	Yes			X		
	Afirma MTC	Yes			X		
2	ThyroSeq v.2	Yes					X

No.	Indications	Yes/No	Low Confidence		Intermediate Confidence		High Confidence
	ThyGenX	Yes					X
	Afirma BRAF	Yes					X
	Afirma MTC	Yes					X
3	ThyroSeq v.2	Yes				X	
	ThyGenX	Yes				X	
	Afirma BRAF	Yes		X			
	Afirma MTC	Yes		X			
4	ThyroSeq v.2	Yes				X	
	ThyGenX	No			X		
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	
5	ThyroSeq v.2	No	No rating provided				
	ThyGenX	Yes				X	
	Afirma BRAF	Yes					X
	Afirma MTC	No	No rating provided				
6	ThyroSeq v.2	Yes				X	
	ThyGenX	No			X		
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	
7	ThyroSeq v.2	Yes					X
	ThyGenX	Yes				X	
	Afirma BRAF	Yes	X				
	Afirma MTC	Yes	X				
8	ThyroSeq v.2	Yes			X		
	ThyGenX	Yes			X		
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	
9	ThyroSeq v.2	Yes				X	
	ThyGenX	Yes				X	
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	

- Based on the evidence and your clinical judgment for the indications described in Question 4:
  - Respond Yes or No for each indication whether each of the following tests is consistent with generally accepted medical practice.
  - Use the 1 to 5 scale outlined below to indicate your level of confidence in your conclusions.

No.	Indications	Yes/No	Low Confidence		Intermediate Confidence		High Confidence
			1	2	3	4	5
1	ThyroSeq v.2	Yes				X	
	ThyGenX	Yes				X	
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	
2	ThyroSeq v.2	Yes					X
	ThyGenX	Yes					X
	Afirma BRAF	Yes					X
	Afirma MTC	Yes					X
3	ThyroSeq v.2	Yes				X	
	ThyGenX	Yes			X		
	Afirma BRAF	Yes		X			
	Afirma MTC	Yes		X			
4	ThyroSeq v.2	Yes				X	
	ThyGenX	No			X		
	Afirma BRAF	Yes				X	
	Afirma MTC	Yes				X	
5	ThyroSeq v.2	No	No rating provided				

No.	Indications	Yes/No	Low Confidence	Intermediate Confidence	High Confidence
	ThyGenX	Yes		X	
	Afirma BRAF	Yes		X	
	Afirma MTC	No	No rating provided		
6	ThyroSeq v.2	Yes		X	
	ThyGenX	No		X	
	Afirma BRAF	Yes		X	
	Afirma MTC	Yes		X	
7	ThyroSeq v.2	Yes			X
	ThyGenX	Yes		X	
	Afirma BRAF	Yes	X		
	Afirma MTC	Yes	X		
8	ThyroSeq v.2	No		X	
	ThyGenX	No		X	
	Afirma BRAF	No			X
	Afirma MTC	No			X
9	ThyroSeq v.2	Yes		X	
	ThyGenX	Yes		X	
	Afirma BRAF	Yes		X	
	Afirma MTC	Yes		X	

- Please provide in the box below comments/rationale and any citations supporting your clinical input on the use of molecular markers in FNA of the thyroid for management of individuals with thyroid nodule(s) and indeterminate finding on the FNA for each type of molecular test below
  - Afirma Gene Expression Classifier

No.	Response
1	Afirma GEC provides an acceptable NPV allowing it to be useful as a test to "rule out" malignancy. A negative GEC, in a FNAB cytologically indeterminate nodule without suspicious sonographic features, can lead to the decision of clinical observation rather than surgery. The GEC results must be interpreted with the knowledge of your institutional risk of malignancy for indeterminate thyroid nodules to accurately assess the NPV for one's respective patient population being tested. For practices where initial prevalence (risk of malignancy) is higher than 30%, NPV may be lower than the often quoted 95%. It has a low PPV making it unreliable as a test to "rule in" malignancy. It is particularly unreliable and should be avoided in oncocytic (Hürthle cell) neoplasms. • Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. Aug 23 2012;367(8):705-715. PMID 22731672 • Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with Afirma gene expression classifier. J Clin Endocrinol Metab. Jan 2014;99(1):119-25. PMID 24152684 • 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. Nov 2014:99(11):4069-77. PMID 24780044 • Brauner E, Holmes BJ, Krane JF, et al. Performance of the Afirma gene expression classifier in Hürthle Cell thyroid nodules differs from other indeterminate thyroid nodules. Thyroid. Jul 2015;25(7):789-96. PMID 25962906 • Santhanam P, Khthir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Med Oncol. Feb 2016;33(2):14. PMID 26749587
2	This test is optimally used to keep patients out of the operating room in an effort to avoid unnecessary surgery. The test performs well in cases of Bethesda 3 and some Bethesda 4 results. It is poorly performing in cases where Hürthle Cell neoplasm is suspected as it uniformly indicates suspicious characteristics and warrants surgery.
3	Afirma using mRNA gene expression is most helpful in ruling out malignancy. It has a high NPV in indeterminate thyroid FNA cases. Its biggest limitation is its low PPV. Numerous data have been published on this test with most of them supporting its use as a molecular marker in indeterminate cases.

No.	Response
	• Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. Aug 23 2012;367(8):705-715. PMID 22731672 • Zhang M, Lin O. Molecular Testing of Thyroid Nodules: A Review of Current Available Tests for Fine-Needle Aspiration Specimens. Arch Pathol Lab Med. Dec 2016;140(12):1338-1344. PMID 27557410 • Chaudhary S, Hou Y, Shen R, et al. 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. PMID 27344463
4	92% sensitivity and 52% specificity. • Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. Aug 23 2012;367(8):705-715. PMID 22731672
5	I am familiar with this from residency training and NCCN guidelines.
6	Demonstrated analytic validity • Walsh PS, Wilde JI, Tom EY, et al. Analytical performance verification of a molecular diagnostic for cytology-indeterminate thyroid nodules. J Clin Endocrinol Metab. Dec 2012;97(12):E2297-2306. PMID 23087323 Demonstrated clinical validity • Alexander et al (2012) prospective, multicenter study, 85%-95% NPV • Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. Aug 23 2012;367(8):705-715. PMID 22731672 Retrospective clinical validation • Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. J Clin Endocrinol Metab. Dec 2010;95(12):5296-5304. PMID 20826580 • Alexander et al (2014) • Santhanam P, Khthir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Med Oncol. Feb 2016;33(2):14. PMID 26749587 GEC results altered management of 50% of patients in 1 study, less in other studies • Alexander et al (2014) Does change in management improve outcome? Needs long-term analysis, but altered management and decreased number of patients proceeding with surgery
7	Excellent NPV but very low PPV. Does not help with decision making about total thyroidectomy versus hemithyroidectomy.
8	To predict benignancy and to eliminate need for surgical resection. Potential harm - false-positive or false-negative Retrospective single-center studies Harrell and Bernsten 2014 Lester et al (2014)
9	Widely accepted by endocrinologists within our clientele for follow-up on indeterminate FNAs of thyroid nodules. Unfortunately, we have not prospectively tracked follow-up surgeries or any subsequent FNAs to determine correlation of Afirma results with final diagnoses.

- ThyroSeq v.2

No.	Response
1	ThyroSeq v.2 is a useful test to evaluate for the risk of malignancy, it holds potential for serving as both a "rule-in" and "rule-out" test, particular for Bethesda IV lesions. An external validation study suggests a lower PPV and less applicability (SAMPLE SIZE WAS LOW) in Bethesda III and oncocytic (Hürthle cell) lesions. A prospective multi-institutional study of ThyroSeq v3 has promising results but is not yet published. With appropriate patient selection, the results of this test can be very helpful in determining the need for surgical

No.	Response
	intervention. The ability of some of the specific measured gene mutations to predict aggressiveness may aid in determining prognosis as well as the extent of surgery. • Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasms/suspicious for follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. Cancer. Dec 2014;120(23):3627-34. PMID 25209362 • 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. Nov 2015;25(11):1217-1223. PMID 26356635 • Valderrabano P, Khazai L, Leon ME, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. Endocr Relat Cancer. Mar 2017;24(3):127-136. PMID 28104680
2	This test attempts to improve results over those of Afirma. Further data will be necessary to determine if it truly improves results of Afirma. The initial data are promising.
3	ThyroSeq uses NGS [next-generation sequencing] with their final report showing specific gene mutation/translocation. Among the commercially available tests, this is the most helpful in both ruling in and ruling out of indeterminate FNA thyroid cases. It has both high NPV and PPV. The release of version 2.1 is additional support to its usefulness in directing the appropriate care of the patient. • Nikiforov YE. Molecular diagnostics of thyroid tumors. Arch Pathol Lab Med. May 2011;135(5):569-577. PMID 21526955 • Zhang M, Lin O. Molecular Testing of Thyroid Nodules: A Review of Current Available Tests for Fine-Needle Aspiration Specimens. Arch Pathol Lab Med. Dec 2016;140(12):1338-1344. PMID 27557410 • Nikiforova MN, Wald AI, Roy S, et al. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. J Clin Endocrinol Metab. Nov 2013;98(11):E1852-1860. PMID 23979959 • Valderrabano P, Khazai L, Leon ME, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. Endocr Relat Cancer. Mar 2017;24(3):127-136. PMID 28104680
4	90.9% sensitivity and 92.1% specificity • 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. Nov 2015;25(11):1217-1223. PMID 26356635
5	I am not familiar with this test.
6	Multigene NGS panel for thyroid cancer demonstrated analytic validity (100% accuracy in evaluating samples of tumors with no variations). • Nikiforov et al Clinical validity Testing for variants increased sensitivity and specificity. • Ferraz et al NPV 72-94% PPV 80-85% • Nikiforov et al, 2011 Can be used to "rule in" as well as to "rule out" malignancy. Molecular test results from ThyroSeq alter management of these undetermined nodules. • Yip et al study concluded that addition of molecular testing increased the diagnosis of FNA specificity of cancer to 95% and PPV to 82% from specificity of 28% with indeterminate FNA only. It also decreased the incidence of 2 stage surgeries 17% vs. 43%.
7	Excellent NPV and very good PPV. Very helpful with decision making. In general, I would consider this most cost-effective, as it will save time, repeated biopsies and multiple surgeries. Would be nice if this is more affordable and easily available.
8	Uses more than 60 genes. Indicated when FNA cytology indicates atypia of uncertain significance or follicular lesion of undetermined significance, follicular neoplasm or suspicious for malignancy.
9	No experience with this assay.

- ThyGenX/ThyraMIR microRNA

No.	Response
1	The ThyGenX/ThyraMIR test should be able to be used as both a "rule-in" and a "rule-out" malignancy test. This combination is attractive, as is the potential to help indicate tumor aggressiveness, but the paucity of published trial data and validation studies is limiting. The current evidence does suggest more reliability with Bethesda IV lesions and less with Bethesda III lesions. Most of the ATA [American Thyroid Association] panel members indicate that they have no clinical experience with this test. • 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*. Jul 2015;100(7):2743-2750. PMID 25965083
2	This test is best used to guide extent of surgery in patients undergoing biopsy. It is best incorporated into a reflex algorithm where patients will be targeted for diagnostic lobectomy surgery and if this test is positive then those patients would receive a total thyroidectomy at one operation instead of a completion thyroidectomy at a second operation. It is not designed to predict benign behavior of the lesion, rather it is designed to rule in malignancy. It is important that clinicians understand this very significant difference as if it is applied in the same manner that ThyroSeq or Afirma are used it will not function well.
3	ThyGenX uses multiplex PCR by sequence-specific probes with their report showing specific gene mutation and translocation. It requires low DNA requirement. It has a high PPV.ThyraMIR reports as negative and positive. It is reflex testing for a negative ThyGenX. It requires more data to support this format of reporting.This test can be used as an alternative to ThyroSeq. • 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*. Jul 2015;100(7):2743-2750. PMID 25965083 • Zhang M, Lin O. Molecular Testing of Thyroid Nodules: A Review of Current Available Tests for Fine-Needle Aspiration Specimens. *Arch Pathol Lab Med*. Dec 2016;140(12):1338-1344. PMID 27557410
4	Negative predictive value 94% and positive predictive value 74% • Labourier E, Beaudenon A, Wylie D, Giordano TJ. Multi-categorical testing for miRNA, mRNA, and DNA on fine needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. ENDO 2015. Presented at the 97th Meeting and Expo of the Endocrine Society on March 5-8, 2015. SAT-344
5	I am familiar with this from residency training and conference resources.
6	Combined use demonstrated analytic validity: Hadd et al and Wylie et al (2010) PPV: Up to 82% NPV: 94% Await further investigations to determine clinical validity and utility.
7	Excellent NPV but mildly low PPV for AUS/FLUS (<70%). I have less experience with this testing, as it is not utilized in our institution.
8	Combined testing is to predict malignancy ThyGenX Positive Test - Rule in for surgical resection ThyGenX Negative Test - Reflex ThyraMIR micro RNA
9	No experience with this assay.

- RosettaGX Reveal

No.	Response
1	This is a genetic test that can be performed on the prepared cytology slides and obviates the need for dedicated passes or repeat biopsy. The small validation study reported an adequate NPV but limited PPV and more studies are needed. None of the members of the ATA panel had clinical experience with this test.

No.	Response
	• Lithwick-Yanai G, Dromi N, Shtabsky A, et al. Multicenter validation of a micro-RNA based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate smears. *J ClinPathol*, June 2016;70(6);500-507. PMID 27798083
2	Further data are needed to demonstrate the clinical utility of this test and it is unlikely to be useful by itself. It is likely best used in combination with other molecular tests as it has only the ability to predict malignancy. In general cytopathology is fairly good at diagnosing malignancy in Bethesda 5 nodules and this test may be used similar to ThyGenX as a rule in molecular marker to guide extent of excision. It does not help to keep patients out of the operating room.
3	Not familiar with the test
4	99% NPV and 62% PPV • Bar D, Meiri E, et al., International Thyroid Congress, Orlando, October 18-23, 2015
5	I am not familiar with this test.
6	Unknown
7	Excellent NPV, low PPV. Good resource as repeat sampling is not needed. Would definitely use it for "Rule Out" test where patient prefers no repeat biopsy.
8	No response
9	No experience with this assay.

- Afirma BRAF

No.	Response
1	This provides a reflex option to test for BRAF V600E, the most common genomic abnormality in differentiated thyroid cancer. A positive BRAF mutation confers a 99% risk of PTC and has been shown in some studies to be associated with more aggressive disease and higher mortality. This would help "rule in" malignancy in the 45% of PTC lesions that are BRAF mutated and help in the clinical decision process by identifying FNAB cytologically indeterminate nodules that are most concerning for malignancy. However, all of the panel members feel that more clinical data on its performance is needed. • Xing M, Alzahrani AS, Carson KA, et al. Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. *JAMA*. Apr 2013;309(14):1493-501. PMID 23571588
2	This test by itself is not useful however in combination with the gene expression classifier it is. Data are mixed regarding whether BRAF positivity really indicates a more aggressive tumor. More aggressive thyroid cancers likely have multiple mutations triggering aggressive behavior outside of isolated BRAF positivity.
3	Both Afirma BRAF and MTC have limited value as an individual test. Its analytical accuracy is good, but its clinical utility is of concern. They have to be part of a panel. • Jara SM, Bhatnagar R, Guan H, et al. Utility of *BRAF* mutation detection in fine-needle aspiration biopsy samples read as "suspicious for papillary thyroid carcinoma". *Head Neck.* Dec 2015;37(12):1788-93. PMID 24989827 • Lin JD, Fu SS, Chen JY, et al. Clinical Manifestations and Gene Expression in Patients with Conventional Papillary Thyroid Carcinoma Carrying the *BRAF(V600E)* Mutation and *BRAF* Pseudogene. *Thyroid.* May 2016;26(5):691-704. PMID 26914762
4	Positive mutation has high specificity for cancer. • Diggans, J, et al. Pac Symp Biocomput 2015Diggans 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.* 2015:371-82. PMID 25592597
5	I am familiar with this from residency training and NCCN guidelines.
6	Diggans (2015) PPA: 90.4% NPA: 99% with 100% specificity for papillary thyroid cancer (If positive proceed with surgery)
7	Excellent PPV, but there are some cases where Afirma was suspicious BRAF not reported. Unsure why. Need for proper sampling.

No.	Response
8	To predict malignancy from a FNA sample with increased pretest probability for malignancy. Positive would inform preoperative planning such as planning for hemi-vs. a total thyroidectomy or performance of a central neck dissection.
9	No experience with this BRAF assay.

- Afirma MTC

No.	Response
1	This provides a reflex test option to test for Medullary Thyroid Cancer (MTC), MTC can be difficult to diagnose on FNA (cytology) specimens and is often called indeterminate (Bethesda 4) on FNA cytology. The cytopathologist may not think of this rare cancer, thus the MTC classifier is a helpful addition to the Afirma GEC. Identifying MTC preoperatively is important to guide appropriate surgery. While the PPV of this test for MTC is 98%, it is unknown if it performs better than measuring calcitonin level on blood or in cytology specimen, in addition to the regular Afirma GEC without the MTC cassette. • Trimboli P, Treglia G, Guidobaldi L, et al. Detection rate of RNA cytology in medullary thyroid carcinoma: a meta-analysis. *Clinical Endocrinology.* Feb 2015;82(2):280-5. PMID 25047365 • Pankratz D Analytical validation of a gene expression classifier for medullary thyroid carcinoma. American Association of Clinical Endocrinologist Annual Meeting 2014.
2	Again, this test is useful in combination with the gene expression classifier but is not useful as an isolated test by itself.
3	Both Afirma BRAF and MTC have limited value as an individual test. Its analytical accuracy is good, but its clinical utility is of concern. They have to be part of a panel. • Jara SM, Bhatnagar R, Guan H, et al. Utility of *BRAF* mutation detection in fine-needle aspiration biopsy samples read as "suspicious for papillary thyroid carcinoma". *Head Neck.* Dec 2015;37(12):1788-93. PMID 24989827 • Lin JD, Fu SS, Chen JY, et al. Clinical Manifestations and Gene Expression in Patients with Conventional Papillary Thyroid Carcinoma Carrying the *BRAF(V600E)* Mutation and *BRAF* Pseudogene. *Thyroid.* May 2016;26(5):691-704. PMID 26914762
4	Positive predictive value of 97.9%, Negative predictive value of 99.8% If positive, high likelihood for cancer. • Kloos RT, Monroe RJ, Traweek ST, et al. A Genomic Alternative to Identify Medullary Thyroid Cancer Preoperatively in Thyroid Nodules with Indeterminate Cytology. *Thyroid.* Jun 2016;26(6):785-793. PMID 26992356
5	I am not familiar with this test.
6	Kloos, et al (2016) Pan Krutz, et al (2016) PPV of 97.7% (If positive proceed with surgery)
7	Useful. Rarely positive.
8	To predict malignancy from FNA sample with increased pretest probability for malignancy.
9	Widely accepted by endocrinologists within our clientele for follow-up on indeterminate FNAs of thyroid nodules. Unfortunately, we have not prospectively tracked follow-up surgeries or any subsequent FNAs to determine correlation of Afirma results with final diagnoses.

- Additional comments and/or any citations supporting your clinical input on this topic.

No.	Additional Comments
1	NPV and PPV are not inherent to a test and are largely influenced by the prevalence of disease in the patient cohort. Molecular testing should be considered for Bethesda categories III, IV and V.

No.	Additional Comments
	It is frequently said that the absence of evidence of benefit is not the same as evidence of absence of a benefit. With the rapid changes in technology and available testing for molecular markers, there is an increasing recognition that papillary thyroid cancer may be a disease that is indolent for many patients. Prior approaches to patients with PTC have been regarded as too aggressive regarding surgery, surgical extent, and post-surgical therapy and surveillance. These factors make it unlikely that we will see well done prospective trials in which pathology will be available as a gold standard to adjudicate the true presence of disease in patients with a reported low likelihood of malignancy. It is important that this technology is permitted to advance and mature and that clinical experience informs the optimum use of these methodologies in the future. Greatest utility of molecular testing of thyroid nodules is to accurately diagnose benign nodules; most of the above tests can diagnose benign nodules and have helped reduce the number of patients who historically have undergone surgery that was ultimately deemed unnecessary because of a benign final pathology result. All have their limitations; for all of the tests, the validation cohorts are small test populations.
2	Molecular testing of fine needle aspirates for thyroid tissue is an important adjunct in the management of these patients. Because the majority of disease is benign every effort should be undertaken to keep patients out of the operating room. Surgery is both expensive and can result in disastrous lifelong complication and impaired quality of life for a patient. Even though surgical complications after thyroidectomy are very rare because the incidence of malignancy is so low any type of impairment in quality of life as a result of surgical therapy must be weighed carefully against the opportunity for improved diagnostics to eliminate the need for diagnostic surgery in cases of indeterminate thyroid nodules. The biggest opportunity for improvement regarding the use of these tests is to work with institutions so that ThyroSeq or Afirma GEC is performed on Bethesda 3 and 4 nodules. The latest recommendations from the American Thyroid Association indicate that unilateral excision can be performed for malignancy up to 4 cm. As a result, the extent of surgery just based on ultra characteristics has changed with an emphasis now being placed on unilateral resection for many of these lesions. Therefore performing RosettaGX, or ThyGenX, Afirma Braf is not as important because many of these patients only need unilateral resection regardless of the genetics of the Bethesda 5 lesion. Again data are mixed whether or not the molecular markers at this point can truly predict more aggressive carcinoma and guide extent of resection. Clinically if I tell the patient they have greater than a 70% chance of cancer in a lesion (Bethesda 5) the patient will want the lesion removed regardless of additional molecular testing. During surgery for Bethesda 5 lesions, frozen section is able to accurately diagnose papillary thyroid carcinoma at the time of surgery, and as such if the surgeon and patient have agreed to a total thyroidectomy then that can be pursued during the index operation without need for completion directly. At the current time, the data are not sufficient to recommend total thyroidectomy on the basis of BRAF positivity alone, and while initial data seem to be promising, certainly more data needs to emerge before molecular testing alone guides extent of surgery. As a result, these tests are very important and should continue to be offered to patients.
3	All these tests need to be supported because they help triage the care of the patient and will ultimately lower the overall healthcare cost.
4	None.
5	None.
6	For all thyroid nodules with diagnosis of FLUS, AUS and follicular neoplasm, molecular testing with ThyroSeq v.2 is utilized most of the time. Results from this analysis have played a crucial role in determining which patients are recommended to have surgery and which patients are placed under active surveillance. The number of "surgical biopsies" has decreased in number as a result of the availability of molecular testing. Clinical utility remains to be seen after long term follow-up. Although previous studies have demonstrated more aggressive nature of tumors with BRAF mutations, this behavior has not been observed in other studies. Thus BRAF mutation is not clinically utilized in determining or changing the type of planned surgery. Clinical staging still plays a major role in determining the extent of surgical intervention (i.e., presence or absence of nodal disease on exam and radiography, extracapsular extension of primary tumor on ultrasound, vocal fold paralysis, fixation of tumor to anatomical structures, etc).
7	Would be nice to have one generalized test and one generalized recommendation for thyroid nodules of indeterminate category.

No.	Additional Comments
8	Surgical excision is the standard of care for the management of equivocal cytologic results. Optimal molecular test that has high NPV or high PPV - that can definitely rule in or rule out malignancy is needed. Long term outcome data proving clinical utility is mandated.
9	None.

9. Is there any evidence missing from the attached draft review of evidence that demonstrates clinical benefit?

No.	Yes/No	Citations of Missing Evidence
1	Yes	Afirma developed a new GSC (gene sequence classifier) test with improved PPV. The results have not yet been published, but have been presented at medical conferences.ThyroSeq is testing a new version (ThyroSeq v3) that includes 112 gene mutations, fusions, etc.
2	No	
3	No	
4	No	
5	No	
6	Yes	Results from long term follow up of those patients (longer than 13 months) placed under active surveillance from benign category to determine the true negative results from false negatives, and assessment of whether there has been harm from delay in diagnosis would be needed to demonstrate true clinical benefit from molecular testing.
7	No	
8	No	
9	No	

[a] Indeterminate finding on the FNA means Bethesda diagnostic category III (atypia/follicular lesion of undetermined significance) or Bethesda diagnostic category IV (follicular neoplasm/suspicion for a follicular neoplasm).
[b] Indeterminate finding on the FNA means Bethesda diagnostic category III (atypia/follicular lesion of undetermined significance) or Bethesda diagnostic category IV (follicular neoplasm/suspicion for a follicular neoplasm).

References

1. Adeniran AJ, Theoharis C, Hui P, et al. Reflex BRAF testing in thyroid fine-needle aspiration biopsy with equivocal and positive interpretation: a prospective study. Thyroid. Jul 2011; 21(7): 717-23. PMID 21568726
2. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. J Clin Endocrinol Metab. Dec 2010; 95(12): 5296-304. PMID 20826580
3. 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. Aug 01 2016; 2(8): 1023-9. PMID 27078145
4. Nikiforov YE. Molecular diagnostics of thyroid tumors. Arch Pathol Lab Med. May 2011; 135(5): 569-77. PMID 21526955
5. Han PA, Kim HS, Cho S, et al. Association of BRAF V600E Mutation and MicroRNA Expression with Central Lymph Node Metastases in Papillary Thyroid Cancer: A Prospective Study from Four Endocrine Surgery Centers. Thyroid. Apr 2016; 26(4): 532-42. PMID 26950846
6. Yip L, Nikiforova MN, Yoo JY, et al. Tumor genotype determines phenotype and disease-related outcomes in thyroid cancer: a study of 1510 patients. Ann Surg. Sep 2015; 262(3): 519-25; discussion 524-5. PMID 26258321
7. Lin JD, Fu SS, Chen JY, et al. Clinical Manifestations and Gene Expression in Patients with Conventional Papillary Thyroid Carcinoma Carrying the BRAF(V600E) Mutation and BRAF Pseudogene. Thyroid. May 2016; 26(5): 691-704. PMID 26914762
8. Alzahrani AS, Alsaadi R, Murugan AK, et al. TERT Promoter Mutations in Thyroid Cancer. Horm Cancer. Jun 2016; 7(3): 165-77. PMID 26902827

9. 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. Sep 2013; 98(9): E1562-6. PMID 23833040
10. Liu X, Bishop J, Shan Y, et al. Highly prevalent TERT promoter mutations in aggressive thyroid cancers. Endocr Relat Cancer. Aug 2013; 20(4): 603-10. PMID 23766237
11. Liu T, Wang N, Cao J, et al. The age- and shorter telomere-dependent TERT promoter mutation in follicular thyroid cell-derived carcinomas. Oncogene. Oct 16 2014; 33(42): 4978-84. PMID 24141777
12. 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. Sep 01 2014; 32(25): 2718-26. PMID 25024077
13. 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. May 01 2016; 122(9): 1370-9. PMID 26969876
14. Nikiforova MN, Wald AI, Roy S, et al. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. J Clin Endocrinol Metab. Nov 2013; 98(11): E1852-60. PMID 23979959
15. Sonic Healthcare USA. Thyroseq. Thyroid Genomic Classifier. 2024. https://www.thyroseq.com/physicians/. Accessed June 13, 2024.
16. Barros-Filho MC, Marchi FA, Pinto CA, et al. High Diagnostic Accuracy Based on CLDN10, HMGA2, and LAMB3 Transcripts in Papillary Thyroid Carcinoma. J Clin Endocrinol Metab. Jun 2015; 100(6): E890-9. PMID 25867809
17. Zheng B, Liu J, Gu J, et al. A three-gene panel that distinguishes benign from malignant thyroid nodules. Int J Cancer. Apr 01 2015; 136(7): 1646-54. PMID 25175491
18. 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: 371-82. PMID 25592597
19. Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol. May 2017; 14(5): 587-595. PMID 28372962
20. Lee E, Terhaar S, McDaniel L, et al. Diagnostic performance of the second-generation molecular tests in the assessment of indeterminate thyroid nodules: A systematic review and meta-analysis. Am J Otolaryngol. 2022; 43(3): 103394. PMID 35241290
21. Nasr CE, Andrioli M, Endo M, et al. Real-World Performance of the Afirma Genomic Sequencing Classifier (GSC)-A Meta-analysis. J Clin Endocrinol Metab. May 17 2023; 108(6): 1526-1532. PMID 36470585
22. Vardarli I, Tan S, Görges R, et al. Diagnostic accuracy of Afirma gene expression classifier, Afirma gene sequencing classifier, ThyroSeq v2 and ThyroSeq v3 for indeterminate (Bethesda III and IV) thyroid nodules: a meta-analysis. Endocr Connect. Jul 01 2024; 13(7). PMID 38771544
23. Patel KN, Angell TE, Babiarz J, et al. Performance of a Genomic Sequencing Classifier for the Preoperative Diagnosis of Cytologically Indeterminate Thyroid Nodules. JAMA Surg. Sep 01 2018; 153(9): 817-824. PMID 29799911
24. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med. Aug 23 2012; 367(8): 705-15. PMID 22731672
25. Livhits MJ, Zhu CY, Kuo EJ, et al. Effectiveness of Molecular Testing Techniques for Diagnosis of Indeterminate Thyroid Nodules: A Randomized Clinical Trial. JAMA Oncol. Jan 01 2021; 7(1): 70-77. PMID 33300952
26. Kim NE, Raghunathan RS, Hughes EG, et al. Bethesda III and IV Thyroid Nodules Managed Nonoperatively After Molecular Testing With Afirma GSC or Thyroseq v3. J Clin Endocrinol Metab. Aug 18 2023; 108(9): e698-e703. PMID 36995878

27. Santhanam P, Khthir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Med Oncol. Feb 2016; 33(2): 14. PMID 26749587
28. Liu Y, Pan B, Xu L, et al. The Diagnostic Performance of Afirma Gene Expression Classifier for the Indeterminate Thyroid Nodules: A Meta-Analysis. Biomed Res Int. 2019; 2019: 7150527. PMID 31531363
29. 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. Nov 2015; 100(11): E1477-83. PMID 26353010
30. 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. Jun 2016; 22(6): 666-72. PMID 26789352
31. Deaver KE, Haugen BR, Pozdeyev N, et al. Outcomes of Bethesda categories III and IV thyroid nodules over 5 years and performance of the Afirma gene expression classifier: A single-institution study. Clin Endocrinol (Oxf). Aug 2018; 89(2): 226-232. PMID 29791966
32. Valderrabano P, Hallanger-Johnson JE, Thapa R, et al. Comparison of Postmarketing Findings vs the Initial Clinical Validation Findings of a Thyroid Nodule Gene Expression Classifier: A Systematic Review and Meta-analysis. JAMA Otolaryngol Head Neck Surg. Sep 01 2019; 145(9): 783-792. PMID 31318389
33. Harrell RM, Eyerly-Webb SA, Golding AC, et al. STATISTICAL COMPARISON OF AFIRMA GSC AND AFIRMA GEC OUTCOMES IN A COMMUNITY ENDOCRINE SURGICAL PRACTICE: EARLY FINDINGS. Endocr Pract. Feb 2019; 25(2): 161-164. PMID 30383497
34. Polavarapu P, Fingeret A, Yuil-Valdes A, et al. Comparison of Afirma GEC and GSC to Nodules Without Molecular Testing in Cytologically Indeterminate Thyroid Nodules. J Endocr Soc. Nov 01 2021; 5(11): bvab148. PMID 34708178
35. Kandil E, Metz TA, Issa PP, et al. Diagnostic Performance of Afirma and Interpace Diagnostics Genetic Testing in Indeterminate Thyroid Nodules: A Single Center Study. Cancers (Basel). Mar 31 2023; 15(7). PMID 37046759
36. Azaryan I, Endo M, Sipos JA, et al. Ultrasound Features and Performance of Afirma Gene Sequencing Classifier in Cytologically Indeterminate Thyroid Nodules. J Endocr Soc. Jan 16 2024; 8(3): bvae010. PMID 38348302
37. 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. Oct 2012; 22(10): 996-1001. PMID 22873825
38. Aragon Han P, Olson MT, Fazeli R, et al. The impact of molecular testing on the surgical management of patients with thyroid nodules. Ann Surg Oncol. Jun 2014; 21(6): 1862-9. PMID 24522987
39. 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. Dec 2015; 141(12): 1082-8. PMID 26606459
40. Abeykoon JP, Mueller L, Dong F, et al. The Effect of Implementing Gene Expression Classifier on Outcomes of Thyroid Nodules with Indeterminate Cytology. Horm Cancer. Aug 2016; 7(4): 272-8. PMID 27102883
41. Chaudhary S, Hou Y, Shen R, et al. 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): 205-10. PMID 27344463
42. Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. Thyroid. Nov 2009; 19(11): 1159-65. PMID 19888858
43. Fnais N, Soobiah C, Al-Qahtani K, et al. Diagnostic value of fine needle aspiration BRAF(V600E) mutation analysis in papillary thyroid cancer: a systematic review and meta-analysis. Hum Pathol. Oct 2015; 46(10): 1443-54. PMID 26232865

44. Kloos RT, Monroe RJ, Traweek ST, et al. A Genomic Alternative to Identify Medullary Thyroid Cancer Preoperatively in Thyroid Nodules with Indeterminate Cytology. Thyroid. Jun 2016; 26(6): 785-93. PMID 26992356
45. Xing M, Clark D, Guan H, et al. BRAF mutation testing of thyroid fine-needle aspiration biopsy specimens for preoperative risk stratification in papillary thyroid cancer. J Clin Oncol. Jun 20 2009; 27(18): 2977-82. PMID 19414674
46. Yin DT, Yu K, Lu RQ, et al. Clinicopathological significance of TERT promoter mutation in papillary thyroid carcinomas: a systematic review and meta-analysis. Clin Endocrinol (Oxf). Aug 2016; 85(2): 299-305. PMID 26732020
47. Bernet V, Hupart KH, Parangi S, et al. AACE/ACE disease state commentary: molecular diagnostic testing of thyroid nodules with indeterminate cytopathology. Endocr Pract. Apr 2014; 20(4): 360-3. PMID 24727662
48. 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. Jul 2014; 260(1): 163-8. PMID 24901361
49. Ferris RL, Baloch Z, Bernet V, et al. American Thyroid Association Statement on Surgical Application of Molecular Profiling for Thyroid Nodules: Current Impact on Perioperative Decision Making. Thyroid. Jul 2015; 25(7): 760-8. PMID 26058403
50. Nikiforova MN, Mercurio S, Wald AI, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. Cancer. Apr 15 2018; 124(8): 1682-1690. PMID 29345728
51. Steward DL, Carty SE, Sippel RS, et al. Performance of a Multigene Genomic Classifier in Thyroid Nodules With Indeterminate Cytology: A Prospective Blinded Multicenter Study. JAMA Oncol. Feb 01 2019; 5(2): 204-212. PMID 30419129
52. Valderrabano P, Khazai L, Leon ME, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. Endocr Relat Cancer. Mar 2017; 24(3): 127-136. PMID 28104680
53. Taye A, Gurciullo D, Miles BA, et al. Clinical performance of a next-generation sequencing assay (ThyroSeq v2) in the evaluation of indeterminate thyroid nodules. Surgery. Jan 2018; 163(1): 97-103. PMID 29154079
54. 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. Nov 2010; 34(11): 2589-94. PMID 20703476
55. 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 Cytopathol. Feb 25 2010; 118(1): 17-23. PMID 20099311
56. 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. Oct 2014; 24(10): 1479-87. PMID 24811481
57. Li W, Justice-Clark T, Cohen MB. The utility of ThyroSeq ® in the management of indeterminate thyroid nodules by fine-needle aspiration. Cytopathology. Jul 2021; 32(4): 505-512. PMID 33914382
58. 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. Jul 2015; 100(7): 2743-50. PMID 25965083
59. Tumati A, Egan CE, Lee-Saxton YJ, et al. Clinical utility of a microRNA classifier in cytologically indeterminate thyroid nodules with RAS mutations: A multi-institutional study. Surgery. Jan 2024; 175(1): 234-240. PMID 37907382
60. Verma T, Marshall C, Dantey KE, et al. The role of the ThyGeNEXT oncogene panel used in combination with the expanded miRNA panel ThyraMIRv2 in Indeterminate thyroid nodules: A large, blinded, real-world, observational study. Cancer Cytopathol. Sep 2024; 132(9): 556-563. PMID 38814813
61. Gharib H, Papini E, Garber JR, et al. AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS, AMERICAN COLLEGE OF ENDOCRINOLOGY, AND ASSOCIAZIONE

MEDICI ENDOCRINOLOGI MEDICAL GUIDELINES FOR CLINICAL PRACTICE FOR THE DIAGNOSIS AND MANAGEMENT OF THYROID NODULES--2016 UPDATE. Endocr Pract. May 2016; 22(5): 622-39. PMID 27167915
62. 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. Jan 2016; 26(1): 1-133. PMID 26462967
63. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Thyroid Carcinoma. Version 1.2025. Updated March 27, 2025. https://www.nccn.org/professionals/physician_gls/pdf/thyroid.pdf. Accessed June 13, 2025.

Documentation for Clinical Review

Please provide the following documentation:

- History and physical and/or consultation notes including:
  - Preliminary diagnosis and prognosis
  - Operative reports
  - Specific test(s) requested and clinical reason/justification for testing
  - How test result will impact clinical decision making
  - Treatment plan
- Laboratory and/or Pathology report(s) (e.g., Fine Needle Aspiration (FNA) of the thyroid cytology reports)
- Laboratory invoice/order indicating specific test(s)/panel(s) and associated procedure codes

Post Service (in addition to the above, please include the following):

- Results/reports of tests performed
- Procedure report(s)

Coding

The list of codes in this Medical Policy is intended as a general reference and may not cover all codes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy.

Type	Code	Description
CPT®	0018U	Oncology (thyroid), microRNA profiling by RT-PCR of 10 microRNA sequences, utilizing fine needle aspirate, algorithm reported as a positive or negative result for moderate to high risk of malignancy
	0026U	Oncology (thyroid), DNA and mRNA of 112 genes, next-generation sequencing, fine needle aspirate of thyroid nodule, algorithmic analysis reported as a categorical result ("Positive, high probability of malignancy" or "Negative, low probability of malignancy")
	0245U	Oncology (thyroid), mutation analysis of 10 genes and 37 RNA fusions and expression of 4 mRNA markers using next-generation sequencing, fine needle aspirate, report includes associated risk of malignancy expressed as a percentage
	81345	TERT (telomerase reverse transcriptase) (e.g., thyroid carcinoma, glioblastoma multiforme) gene analysis, targeted sequence analysis (e.g., promoter region)
	81445	Solid organ neoplasm, genomic sequence analysis panel, 5-50 genes, interrogation for sequence variants and copy number variants or

Type	Code	Description
		rearrangements, if performed; DNA analysis or combined DNA and RNA analysis
	81546	Oncology (thyroid), mRNA, gene expression analysis of 10,196 genes, utilizing fine needle aspirate, algorithm reported as a categorical result (e.g., benign or suspicious)
HCPCS	None	

Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

Effective Date	Action
11/26/2014	BCBSA Medical Policy adoption
09/30/2015	Policy revision without position change
02/01/2017	Policy revision without position change
03/01/2018	Policy revision without position change
05/01/2018	Coding update
09/01/2018	Policy revision without position change
02/01/2019	Coding update
10/01/2019	Policy revision without position change
10/01/2025	Policy reactivated. Previously archived from 08/01/2020 to 09/30/2025.

Definitions of Decision Determinations

Healthcare Services: For the purpose of this Medical Policy, Healthcare Services means procedures, treatments, supplies, devices, and equipment.

Medically Necessary: Healthcare Services that are Medically Necessary include only those which have been established as safe and effective, are furnished under generally accepted professional standards to treat illness, injury or medical condition, and which, as determined by Blue Shield of California, are: (a) consistent with Blue Shield of California medical policy; (b) consistent with the symptoms or diagnosis; (c) not furnished primarily for the convenience of the patient, the attending Physician or other provider; (d) furnished at the most appropriate level which can be provided safely and effectively to the member; and (e) not more costly than an alternative service or sequence of services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of the member's illness, injury, or disease.

Investigational or Experimental: Healthcare Services which do not meet ALL of the following five (5) elements are considered investigational or experimental:

A. The technology must have final approval from the appropriate government regulatory bodies.
   - This criterion applies to drugs, biological products, devices and any other product or procedure that must have final approval to market from the U.S. Food and Drug Administration ("FDA") or any other federal governmental body with authority to regulate the use of the technology.
   - Any approval that is granted as an interim step in the FDA's or any other federal governmental body's regulatory process is not sufficient.
   - The indications for which the technology is approved need not be the same as those which Blue Shield of California is evaluating.

B. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.
   - The evidence should consist of well-designed and well-conducted investigations published in peer-reviewed journals. The quality of the body of studies and the consistency of the results are considered in evaluating the evidence.
   - The evidence should demonstrate that the technology can measure or alter the physiological changes related to a disease, injury, illness, or condition. In addition, there should be evidence, or a convincing argument based on established medical facts that such measurement or alteration affects health outcomes.

C. The technology must improve the net health outcome.
   - The technology's beneficial effects on health outcomes should outweigh any harmful effects on health outcomes.

D. The technology must be as beneficial as any established alternatives.
   - The technology should improve the net health outcome as much as, or more than, established alternatives.

E. The improvement must be attainable outside the investigational setting.
   - When used under the usual conditions of medical practice, the technology should be reasonably expected to satisfy Criteria C and D.

Feedback

Blue Shield of California is interested in receiving feedback relative to developing, adopting, and reviewing criteria for medical policy. Any licensed practitioner who is contracted with Blue Shield of California or Blue Shield of California Promise Health Plan is welcome to provide comments, suggestions, or concerns. Our internal policy committees will receive and take your comments into consideration. Our medical policies are available to view or download at www.blueshieldca.com/provider.

For medical policy feedback, please send comments to: MedPolicy@blueshieldca.com

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department at (800) 541-6652, or the Transplant Case Management Department at (800) 637-2066 ext. 3507708 or visit the provider portal at www.blueshieldca.com/provider.

Disclaimer: Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as member health services contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member health services contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.

Appendix A

POLICY STATEMENT	
BEFORE	**AFTER** **Blue font: Verbiage Changes/Additions**
Reactivated Policy **Policy Statement:** N/A	**Molecular Markers in Fine Needle Aspiration of the Thyroid 2.04.78** **Policy Statement:** I. For individuals who have thyroid nodules without strong clinical or radiologic findings suggestive of malignancy in whom surgical decision making would be affected by test results, the use of **either** of the following types of molecular marker testing or gene variant analysis in fine needle aspirates of thyroid nodules with indeterminate cytologic findings (i.e., Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/ suspicion for a follicular neoplasm]) may be considered **medically necessary**: A. Afirma® Genomic Sequencing Classifier B. ThyroSeq® II. The use of **any** of the following types of molecular marker testing or gene variant analysis in fine needle aspirates of thyroid nodules with indeterminate findings (Bethesda diagnostic category III [atypia/follicular lesion of undetermined significance] or Bethesda diagnostic category IV [follicular neoplasm/suspicion for a follicular neoplasm]) or suspicious findings (Bethesda diagnostic category V [suspicious for malignancy]) to rule in malignancy to guide surgical planning for initial resection rather than a 2-stage surgical biopsy followed by definitive surgery may be considered **medically necessary**: A. ThyroSeq B. ThyraMIR® microRNA/ThyGenX® C. Afirma BRAF after Afirma Genomic Sequencing Classifier D. Afirma MTC after Afirma Genomic Sequencing Classifier III. Gene expression classifiers, genetic variant analysis, and molecular marker testing in fine needle aspirates of the thyroid not meeting criteria outlined above, including but not limited to use of single-

POLICY STATEMENT	
BEFORE	AFTER Blue font: Verbiage Changes/Additions
	gene telomerase reverse transcriptase (*TERT*) testing, are considered **investigational**.