MEDICAL POLICY	Genetic Testing: Thyroid Nodules (All Lines of Business Except Medicare)
Effective Date: 6/1/2021  6/1/2021 Medical Officer Date	Medical Policy Number: 39 Technology Assessment Committee approved date: 4/13; 3/14; 2/15; 2/16 Medical Policy Committee approved date: 9/15; 2/15; 4/17; 2/18; 8/19; 12/19; 09/2020; 11/2020; 12/2020; 3/2021; 5/2021

See Policy CPT CODE section below for any prior authorization requirements

SCOPE:

Providence Health Plan, Providence Health Assurance, Providence Plan Partners, and Ayin Health Solutions as applicable (referred to individually as "Company" and collectively as "Companies").

APPLIES TO:

All lines of business except Medicare

BENEFIT APPLICATION

Medicaid Members

Oregon: Services requested for Oregon Health Plan (OHP) members follow the OHP Prioritized List and Oregon Administrative Rules (OARs) as the primary resource for coverage determinations. Medical policy criteria below may be applied when there are no criteria available in the OARs and the OHP Prioritized List.

POLICY CRITERIA

Gene Expression Testing

I. The use of the Afirma gene expression classifier (GEC) or Afirma genomic sequence classifier (GSC) may be considered **medically necessary and covered** to assess fine needle aspirates (FNAs) of thyroid nodules when **all** of the following criteria are met (A. – D.):

   A. Patient is 18 years of age or older; **and**
   B. Ultrasonographically confirmed thyroid nodule(s) is 1 cm or larger in diameter; **and**
   C. Presence of indeterminate thyroid cytopathology obtained from fine needle aspirates, with **either** of the following classifications (1. or 2.):

1. Bethesda class III: atypia of undetermined significance (AUS) or follicular lesion of undetermined significance (FLUS) (see Policy Guidelines section below for Bethesda classification); **or**
2. Bethesda class IV: Follicular/Hürthle cell neoplasm or suspicious for a follicular/Hürthle cell neoplasm (see Policy Guidelines section below for Bethesda classification); **and**

D. The patient is not undergoing thyroid surgery for diagnostic confirmation.

II. The use of the Afirma GEC or Afirma genomic sequence classifier (GSC) to assess FNAs of thyroid nodules is considered **investigational and not covered** when criterion I. above is not met, including, but not limited to:

A. Patient under the age of 18 years.
B. Evaluation of FNA when cytology has classified the nodule with Bethesda diagnostic categories I, II, V, or VI.
C. Evaluation of specimen other than fine needle aspirate (FNA) of thyroid nodules.

III. The use of all GECs or GSCs (other than Afirma tests listed in criterion I), including miRNA classifiers, is considered **investigational and not covered** for all indications. This includes, but is not limited to, the following classifiers:

A. Afirma BRAF
B. Afirma MTC
C. Afirma Xpression Atlas
D. ThyraMIR
E. Rosetta GX Reveal

Mutational Analysis

IV. Mutation analysis of fine needle aspirates (FNA) of the thyroid, including, but not limited to the following is considered **investigational and not covered**:

A. ThyroSeq (any version)
B. ThyGenX
C. ThyGeNext Thyroid Oncogene Panel
D. Thyroid Cancer Mutation Panel
E. Individual gene mutations (e.g., BRAF V600E, KRAS, NRAS, HRAS)
F. Gene rearrangements (e.g., RET/PTC, PAX8/PPARγ)

Link to Policy Summary

POLICY GUIDELINES

Table 1: The Bethesda System for Reporting Thyroid Cytopathology (Adapted from Cibas et al., 2009.)[1]

Bethesda Classification	Cytological Classification	Risk of Malignancy (%)
I	Nondiagnostic or unsatisfactory	1%-4%
II	Benign	0%-3%
III	Atypia of undetermined significance, or follicular lesion of undetermined significance	5%-15%
IV	Follicular/Hürthle cell neoplasm or suspicious for a follicular/Hürthle cell neoplasm	15%-30%
V	Suspicious for malignancy	60%-75%
VI	Malignant	97%-99%

BILLING GUIDELINES

Afirma Billing Guidelines

- *For claims received prior to 01/01/2016*: CPT code 81479, unlisted molecular pathology procedure.
- *For claims billed on or after 01/01/2016*: 2016 CPT code 81545, Oncology (thyroid), gene expression analysis of 142 genes.

CPT CODES

All Lines of Business Except Medicare	
Prior Authorization Required	
81210	BRAF (B-Raf proto-oncogene, serine/threonine kinase) (eg, colon cancer, melanoma), gene analysis, V600 variant(s)
81311	NRAS (neuroblastoma RAS viral [v-ras] oncogene homolog) (eg, colorectal carcinoma), gene analysis, variants in exon 2 (eg, codons 12 and 13) and exon 3 (eg, codon 61)

81401	Molecular Pathology Procedure, Level 2
81403	Molecular Pathology procedure, Level 4
81404	Molecular Pathology Procedure, Level 5
81405	Molecular Pathology Procedure, Level 6
81406	Molecular Pathology Procedure, Level 7
81445	Targeted genomic sequence analysis panel, solid organ neoplasm, DNA analysis, and RNA analysis when performed, 5-50 genes (eg, ALK, BRAF, CDKN2A, EGFR, ERBB2, KIT, KRAS, NRAS, MET, PDGFRA, PDGFRB, PGR, PIK3CA, PTEN, RET), interrogation for sequence variants and copy number variants or rearrangements, if performed
Afirma (Veracyte)	
~~81545~~	**TERMED 12/31/2020** ~~Oncology (thyroid), gene expression analysis of 142 genes, utilizing fine needle aspirate, algorithm reported as a categorical result (eg, benign or suspicious)~~
81546	Oncology (thyroid), mRNA, gene expression analysis of 10,196 genes utilizing fine needle aspirate, algorithm, reported as a categorical result (eg benign or suspicious)
Not Covered	
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")
0204U	Oncology (thyroid), mRNA, gene expression analysis of 593 genes (including BRAF, RAS, RET, PAX8, and NTRK) for sequence variants and rearrangements, utilizing fine needle aspirate, reported as detected or not detected
0208U	Oncology (medullary thyroid carcinoma), mRNA, gene expression analysis of 108 genes, utilizing fine needle aspirate, algorithm reported as positive or negative for medullary thyroid carcinoma
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
Unlisted Codes All unlisted codes will be reviewed for medical necessity, correct coding, and pricing at the claim level. If an unlisted code is billed related to services addressed in this policy then **prior-authorization is required.**	
81479	Unlisted molecular pathology procedure
81599	Unlisted multianalyte assay with algorithmic analysis
84999	Unlisted chemistry procedure

DESCRIPTION

Thyroid Nodules and Thyroid Cancer

Thyroid nodules (also called tumors or growths) are caused by overproliferation of thyroid gland cells and are three times as prevalent in women as in men. The lifetime risk of developing a thyroid nodule is 5% to 10% in the United States, with most diagnoses being made in individuals over 45 years of age. Thyroid nodules may be benign or malignant. Benign nodules typically pose no immediate danger to the patient and do not usually require treatment. Malignant nodules can invade nearby organs and tissues and metastasize to other parts of the body and may become life-threatening if not surgically removed.

Thyroid nodules have a 5% to 15% risk of malignancy. The initial diagnostic step in determining a nodule's status (benign or malignant) involves fine needle aspiration (FNA) biopsy followed by cytopathologic analysis and classification via the Bethesda system. The Bethesda System contains six diagnostic categories, each with specified cytopathology characteristics, associated risk of malignancy, and typical treatment options.[1] (See Table 1 above regarding the Bethesda system for reporting thyroid cytopathology). Between 15% to 30% of FNA samples result in indeterminate cytologic classification (Bethesda class III or IV). Indeterminate nodules have a roughly 24% risk of malignancy and most patients undergo diagnostic thyroid surgery (lobectomy) followed by histopathologic analysis, which is the current standard for assessing indeterminate nodules.

Since approximately 75% of indeterminate nodules are ultimately classified as benign, a number of diagnostic technologies have emerged that may help reclassify indeterminate samples, thereby sparing a large number of patients with potentially benign nodules from diagnostic thyroid surgery and allowing patients the option to undergo observation. Currently, observation for benign nodules includes repeat ultrasound after 6-12 months, and if stable for one to two years, then subsequent ultrasound at 3- to 5-year intervals.[2]

Genetics of Thyroid Cancer

In recent years, significant advances in understanding the genetic mechanisms of thyroid cancer have changed the way thyroid nodules are treated clinically. Several molecular markers have been associated with malignancy, including point mutations in the BRAF and RAS genes and gene fusions of RET/PTC and PAX8/PPAR. In addition, microRNAs (miRNAs), which are noncoding RNAs that can regulate gene expression, have been found to be dysregulated in malignant thyroid nodules and are currently being evaluated as a potential tool to differentiate between benign and malignant thyroid nodules. There are a number of commercially available tests that are marketed as "gene expression classifier (GEC)" tests intended to aid in the reclassification of indeterminate thyroid nodules as either more or less likely to be benign or malignant based on the genetic profile. These tests are not intended to diagnose different sub-types of thyroid neoplasms. These tests are described below.

Afirma® (Veracyte)

The Afirma® suite of tests include the following:[3]

The Afirma® Gene Expression Classifier (GEC) is a "rule out" test intended to help identify thyroid nodules that are at low risk for malignancy after cytologic analysis from a FNA biopsy returns an indeterminate result. The GEC assay uses RNA extracted from patient FNA samples to quantify the expression levels of 142 cancer-related genes. A proprietary algorithm integrates these measures into a malignancy risk score used to reclassify cytologically indeterminate nodules as "benign" or "suspicious for malignancy." A reclassification of the nodule to "benign" enables low-risk patients to opt for clinical surveillance and avoid diagnostic thyroid surgery.

The Afirma Malignancy Classifier (AMC) test consists of the Afirma GEC plus Afirma BRAF and Afirma Medullary Thyroid Cancer (MTC). The Afirma BRAF is a BRAF V600E RNA classifier and Afirma MTC is an RNA classifier that identifies the presence of medullary thyroid carcinoma (MTC). The AMC test results are designed to inform surgical strategy for those patients headed to surgery based on their cytopathology or Afirma GEC results.

The Afirma Genomic Sequencing Classifier (GSC) test includes an additional malignancy classifier component, the RET/PTC fusion gene.

The Afirma Xpression Atlas is an RNA sequencing-based test added on to the Afirma GSC to provide physicians with genomic alteration information from the same FNA samples as used in GST testing. Afirma Xpression Atlas was designed to enable physicians to tailor surgery strategy or treatment options for patients with cancerous thyroid nodules or nodules suspicious for cancer. The test measures 761 DNA variants and 130 RNA fusions and is intended use is among cytologically indeterminate nodules that are Afirma GSC suspicious, Bethesda V/VI nodules, or known thyroid metastases.

RosettaGX Reveal (Rosetta Genomics Ltd.)

The RosettaGX Reveal test is a quantitative reverse transcription PCR (qRT-PCR)-based profiling test that measures expression levels of 24 microRNAs (miRNAs) in RNA extracted from FNA biopsy smears or cell blocks. The assay is intended to differentiate indeterminate thyroid nodules as benign, suspicious for malignancy, or as having high risk for medullary carcinoma. The assay is intended for nodules that are greater than 0.5 centimeters in patients aged 18 years or older. A proprietary algorithm classifier is used to differentiate thyroid nodules as benign or suspicious for malignancy. In addition, the test measures the hsa-miR-375 marker for medullary carcinoma.

ThyGeNEXT® + ThyraMIR™ / (Interspace Diagnostics)

ThyGeNEXT® (ThyGenX®) Thyroid Oncogene Panel and ThyraMIR™ Theyroid miRNA Classifier are complimentary "rule in" tests intended to identify malignancy in thyroid nodules for which cytopathology analysis after FNA has yielded an indeterminate result.[4]

ThyGeNEXT (and its predecessor, ThyGenX) uses a PCR-based enrichment from FNA biopsy and next-generation sequencing (NGS) analysis of a panel of thyroid cancer relevant markers, including but not limited to mutations in the BRAF, HRAS, KRAS, NRAS, and PIK3CA gene; and testing for the presence of fusion genes. ThyGenX was developed based on older PCR-based miRInform test but the current version

uses NGS and has increased panel content. Although the company has indicated that there is 98-100% concordance between the older miRInform assay and the new ThyGeNEXT assay, no peer-reviewed publications were identified that could confirm this claim.

ThyraMIR is a microRNA (miRNA) gene expression classifier, which evaluates the expression of 10 miRNAs, and is used reflexively when results from the ThyGeNEXT test are negative. The intent of the ThyraMIR GEC is to assess the risk of thyroid nodules being either benign or malignant and help reduce unnecessary surgeries in conjunction with all other available clinical data.

ThyroSeq®, ThyroSeq v2 and ThyroSeq v3 (University of Pittsburgh Medical Center/CBL Path)

Since 2007, the ThyroSeq® Genomic Classifier (GC) has been offered.[5] From the first version (v0) to the 112-gene panel, ThyroSeq is intended for the pre-operative assessment of thyroid nodules with indeterminate cytology. ThyroSeq v3 uses NGS technology to sequence DNA and RNA, evaluating 112 genes and detecting four classes of genetic alterations: mutations, gene fusions, gene expression alterations and copy number variants (CNVs). The test utilizes a proprietary Genomic Classifier (GC) based on the algorithmic analysis of all detected genetic alterations and reports the test result as Positive or Negative.

REVIEW OF EVIDENCE

A review of the ECRI, Hayes, Cochrane, and PubMed databases was conducted regarding the use of molecular marker evaluation of fine needle aspirates of thyroid nodules. Below is a summary of the available evidence identified through October 2020.

Afirma (Veracyte)

Analytical Validity

- In 2018 (archived 2019), Hayes published a review of the Afirma tests, including four studies reporting on the analytical validity.[6] Only two of the four studies included that reported on test performance addressed the main Afirma GEC test, one of which was the initial development and performance study of the GEC test. One study reported on the analytical performance of the Afirma BRAF V600E assay, and one study evaluated the test performance of the Afirma reflexive Medullary Thyroid Carcinoma (MTC) Classifier assay. One study by Chudova et al. reported the sensitivity of the Afirma GEC classifier as 100% (95% CI, 64% to 100%) and the specificity as 73.3% (95% CI, 49% to 89%).[7] The study by Walsh et al. reported on intra- and interassay variability with concordance 93.9 - 97.3% and interlaboratory variability between two laboratories as having 100% concordance on a 22-sample test.[8] Despite the fact that the performance measures were based on small sample sizes (22-37 samples), the review concluded that “there is an established assay process for the Afirma assay that is accurate and reproducible.”

Also included in the Hayes review were two additional studies reporting on the test performance of the Afirma BRAF assay and the Afirma MTC assays, which were separate expression classifiers from the Afirma GEC assay. In 2013, Kloos et al. examined the analytical performance of the Afirma BRAF assay, one part of the Afirma malignancy classifier assay (MCA) in a prospectively designed cohort study. Using 11 FNA specimens, the assay was determined to be reproducible across a range of 3.3% to 60%, with an average coefficient of variant of 20.8%.[9] In 2016, Pankratz et al. examined aspects of analytical validity of the Afirma Medullary Thyroid Carcinoma (MTC) malignancy classifier, reporting a sensitivity estimate on tissue of 96.3% (95% CI, 81% to 99.9%) and 100% concordance between two laboratories assaying 20 specimens.[10]

- A 2019 study by Angell and colleagues investigated the analytical and clinical validity of Afirma Xpression Atlas (XA) in thyroid nodule FNA samples.[11] DNA and RNA were purified from 943 blinded FNAs and compared using whole-transcriptome RNA-seq, targeted RNA-seq, and Targeted DNA-seq. Another 695 blinded FNAs were used to define performance for fusions between whole-transcriptome RNA-seq and targeted RNA-seq. Of the variants detected in the DNA at 20% variant allele frequency, XA detected 88%. XA variant detection was 89% when compared to an alternative RNA-based method. Nearly all variants or fusions detected by XA were confirmed by an alternative detection method. The study found high intra-plate reproducibility (89%-94%), inter-plate reproducibility (86%-91%), and inter-lab accuracy (90%). Among Bethesda III/IV nodules, XA had a sensitivity of 49%. The authors conclude that their findings demonstrate the analytical and clinical validation of XA in thyroid nodule evaluation and that the test is intended to supplement clinical decision-making among patients with Bethesda II-VI nodules.

Clinical Validity

- In 2018, ECRI published an evidence report on the use of Afirma as a "rule out" and "rule in" test for determining thyroid nodule malignancy, including one prospective cohort study that reported on clinical validity.[12] The single clinical validity study included in the review study was a manufacturer-sponsored, prospective, blinded, multicenter clinical validation study that evaluated 265 FNAs from indeterminate nodules 1 cm or larger in diameter with the Afirma GEC, using surgical pathology as the reference standard.[13] This study was determined to be a high quality study with low risk of bias. However, variation was observed in subgroup analyses of diagnostic categories within the study. For example, while the negative predictive value (NPV) for the entire cohort (n=265 nodules) was 93% (95% CI, 86% to 97%), Bethesda class III and IV NPVs were 95% (95% CI, 85% to 99%) and 94% (95% CI, 79% to 99%), respectively, but the NPV for Bethesda class V nodules was 85% (95% CI, 0% to 85%). ECRI concluded that the Afirma GEC test is effective for ruling out malignancy. However, with reported specificities of around 50%, Afirma was deemed to be ineffective for identifying malignant nodules (as a "rule in" test). The evidence for this conclusion was rated moderate strength. Of note, ECRI acknowledged 14 additional studies that reported on the validity of the Afirma test, stating that these studies were not included in the review as they were either not sufficiently powered to contribute additional meaningful data, selected inappropriate patient populations, or had a retrospective study design.

- In the 2018 Hayes review described above, 22 studies were included that evaluated the clinical validity of the Afirma gene expression classifier.[6] These studies included four prospectively designed studies, 17 retrospectively designed studies, and two pooled analysis studies. Of note, the majority of the included studies focused on the clinical validity of the Afirma GEC assay and did not report test performance measures for the reflexive Medullary Thyroid Carcinoma (MTC) Classifier. All studies ranged from fair- to very-poor-quality. Reported sensitivity ranged from 77.6% to 100%, and specificity ranged from 14.7% to 17%. PPV ranged from 14.3% to 57.1%. Reported NPV varied from 75% to 100%. Substantial variation was observed in subgroup analyses diagnostic categories of indeterminate nodules within individual studies as well. For example, in one retrospective case review, while the NPV for the entire cohort (n=178) was 91% (95% CI, 83% to 97%), Bethesda class III and IV NPVs were 88% (95% CI, 63% to 100%) and 92% (95% CI, 85% to 99%), respectively, but the NPV for Bethesda class V nodules was 26% (95% CI, 0% to 85%). However, the Hayes review stated that there was moderate-quality evidence for the clinical validity of the Afirma GEC assay, but noted several limitations including: contradicting data, predominately retrospectively designed studies, and small patient populations.

  Also included in the Hayes review were two additional studies reporting on the clinical validity of the Afirma BRAF assay and the Afirma MTC assays. In 2013 Kloos et al. published the results of a retrospective case review to determine the frequency of BRAF mutations in cytologically classified indeterminate thyroid nodules and assess the diagnostic accuracy of the Afirma GEC with the incorporation of BRAF mutation testing.[14] In indeterminate nodules (AUS/FLUS, FN/SFN, and suspicious for malignancy [SUSP] classifications), 21 (10.1%) were BRAF-positive (BRAF+), with the majority identified in nodules cytologically classified as SUSP (n=18; 41.9%; 95% CI, 27% to 58%). The study also concluded that the addition of the BRAF assay did not impact the results of the Afirma GEC assay.

- In the 2016 Kloos study described above, the performance of the Afirma MTC Classifier was evaluated in 7815 indeterminate Bethesda class III and IV cases from a consecutive series of 50,430 cytologically assessed thyroid nodules.[9] A pooled analysis of three independent validation sets demonstrated a sensitivity of 97.9% (95% CI, 87.3% to 99.9%), a specificity of 99.8% (95% CI, 98.7% to 100%), PPV of 97.9% (95% CI, 87.3% to 99.9%), and a NPV of 99.8% (95% CI, 98.7% to 100%). Additionally, the study reported a MTC Classifier false-positive rate of 2.3% (1/43n=1 positive cases).

No additional studies reporting on the clinical validity of the Afirma assays were identified after the publication of the Hayes review.

Clinical Utility

- In 2020, ECRI published a report addressing the Afirma Genomic Sequencing Classifier (GSC) for evaluating thyroid nodules of indeterminate cytopathologic diagnosis.[15] In total, results from 3 cohort studies indicated that Afirma GSC's sensitivity (≥91%) and negative predictive value (NPV; ≥96%) are capable of ruling out malignancy in thyroid nodules with indeterminate fine needle

aspiration (FNA). Studies also showed high concordance between Afirma GSC and Afirma GEC (the earlier test version). An earlier evidence assessment by ECRI concluded that low-quality evidence showed fewer patients with indeterminate nodules undergo diagnostic surgery when managed with Afirma than when managed without it. ECRI deemed evidence inconclusive regarding whether Afirma reduces surgery rates in nodules classified as benign by cytology. Investigators concluded that evidence supporting the clinical utility of Afirma GSC was "somewhat favorable."[15]

- In the 2018 ECRI report described above, eight retrospective cohort studies reporting measures of Afirma GEC's were identified which assessed clinical utility.[12] Six studies were included that assessed the effect of Afirma test results on treatment decisions, and two studies which compared four-year outcomes of patients with indeterminate nodules that tested "benign" with Afirma by cytological evaluation. Results from one study indicated that differences in nodule growth, sonographic characteristics, and rate of malignancy were similar for patients receiving a "benign" Afirma result and those with a benign cytology result.[16] ECRI concluded that Afirma decreased the proportion of patients with indeterminate thyroid nodules who underwent diagnostic thyroid surgery, and that the incidence of malignancy and growth of benign nodules was similar whether a patient had received a classification of benign by either Afirma or cytology. The evidence for this conclusion was deemed to be of low strength since all of the clinical utility studies were of retrospective design and had large differences in comparison group sizes, indicating a high risk of bias.

- In the 2018 Hayes review described above, nine studies were included that evaluated the clinical utility of the Afirma gene expression classifier.[6] All studies were retrospective in design and focused primarily on surgical intervention rates with the use of the Afirma GEC assay. Three studies reported on changes in malignancy rates with the use of the Afirma assay, one study focused on physician-reported patient safety with the use of the Afirma GEC test, and one study reported on the extent of surgery performed as a result of the Afirma assay. Overall, the majority of studies reported significant decreases in surgical rates with the use of the Afirma GEC assay. However, there is a lack of long-term follow-up data on this and other patient outcomes. In addition, of the four studies that compared pre-GEC and post-GEC malignancy rates, three reported an increase in malignancy of excised nodules post-GEC. Limitations of the clinical studies include the fact that all studies were retrospective and consisted of small patient populations. However, the Hayes review concluded that the evidence of clinical utility for the Afirma GEC test is consistent, giving the test a "B" rating.

No additional studies reporting on the clinical utility of the Afirma assays were identified after the publication of the Hayes review. No studies on the clinical utility of Afirma Xpression Atlas were identified.

<u>RosettaGX Reveal (Rosetta Genomics Ltd.)</u>

In 2016, Hayes published a review on the use of RosettaGX Reveal for all proposed purposes. Only two peer-reviewed studies were identified, one reporting analytical validity and one reporting on both

analytical and clinical validity. These two studies are described below. No studies were identified which reported on clinical utility. As a result, Hayes rated the RosettaGX Reveal assay a "D2" rating for all proposed purposes.[17]

Analytical Validity

In 2016, Benjamin et al. published a study that examined the analytical performance of the RosettaGX Reveal test (24 miRNAs) on routinely prepared FNA slides.[17] The investigators reported an intranodule concordance of 92.81% for a set of 139 slides and 91.25% concordance for 80 slides using two different types of stains, suggesting that stain type does not impact classification accuracy.[18] Analytical sensitivity and specificity were presented in bar graphs and exact values were not reported but indicated that sensitivity was <80% and specificity was approximately 94% to 100%. Interlaboratory reproducibility, assaying between 25 and 82 samples, ranged from 90.9% to 96%.

Clinical Validity

In 2016, Lithwick-Yanai et al. published the results of a multicenter retrospective study that evaluated the clinical validity of the RosettaGX Reveal test, using a training set and a validation set of FNA smears from indeterminate nodules.[19] In the training set, the test was trained on 262 indeterminate samples (Bethesda categories III and IV) with reported estimates of 78% (95% CI, 65%-88%) sensitivity and 76% (95% CI, 68%-83%) specificity. The 201 indeterminate samples used in the validation set had corresponding histological information of excised thyroid nodules and were evaluated in a blinded fashion. In the validation set, 12 samples failed the quality control, resulting in 189 samples for performance characteristics assessment. The validation set (Bethesda III and IV samples only) reported a sensitivity of 74% (95% CI, 55%–88%), specificity of 74% (95% CI, 65%–82%), NPV of 92% (95% CI, 84%–96%), and a PPV of 43% (95% CI, 29%–657%). Nine malignant samples were misclassified as benign in the validation set, indicating a false negative rate of 4.8% (9 of 189). The study has several limitations, including a highly selected patient population, retrospective study design and small sample size.

Clinical Utility

No studies were identified that reported on the clinical utility of the RosettaGX Reveal miRNA GEC.

ThyGenX® / ThyraMIR™ (Interspace Diagnostics)

In 2019, Hayes published a review on the use of ThyGenX/ThyraMIR for all proposed purposes. Only six peer-reviewed studies were identified, all of which reported test performance measures (clinical validity). However, only one study was included (described below). No studies were identified which reported on analytical validity or clinical utility of these tests. As a result, Hayes rated ThyGenX and/or ThyraMIR a "D2" rating for all proposed purposes.[20]

Clinical Validity

In 2020, ECRI published an updated a review on the use of ThyGeNEXT/ThyGenX/ThyraMIR as a "rule in" and a "rule out" test for determining thyroid nodule malignancy, including 17 studies that investigated the diagnostic performance (clinical validity) of either or both the miRInform/ThyGenX and ThyraMIR tests.[21] Studies on miRInform were included "because the manufacturer provided data from an unpublished study showing high concordance between miRInform and ThyGenX (positive and negative agreements of 95.7% and 100%, respectively)". Of the included studies, only five were articles published in peer-reviewed journals, and 12 were conference abstracts. Four studies reported mean values for miRInform/ThyGenX test performance (89.8% specificity and 59.9% sensitivity. Mean sensitivity and specificity of ThyraMIR, reported by two individual studies, were 58.5% and 89.8%, respectively. The review reported that with combined use of both tests, sensitivity and specificity were "both moderately high (mean of 87% and 90%, respectively), suggesting that ThyraMIR, despite its own limited sensitivity, can complement the rule-out performance of ThyGenX", and that use of both tests may accurately identify nodules at low risk or high risk for malignancy ("rule in" and "rule out" capability). However, the review concluded that the data was limited regarding the clinical diagnostic performance of the combined or individual use of these tests. Limitations of the evidence base include the small sample sizes used in the included studies (N=42-138 FNA samples from indeterminate nodules) and only one study published on the newer-generation version of the mutation analysis test, ThyGenX. ECRI did not identify any studies that reported on analytical validity or clinical utility of these tests.

Included in the 2017, Hayes review described above, in 2017, Young et al. reported the clinical validity of the ThyGenX and ThyraMIR assays using indeterminate thyroid nodules in an urban academic community hospital setting.[22] Twenty-four indeterminate nodules (Bethesda III, n=23; Bethesda IV, n=1) were tested, 12.5% of which (three of 24) were found to have single point mutations. None of the samples had miRNA expression abnormalities as determined by ThyraMIR. The Hayes review stated that "the quality of this study was very poor, as no patient or nodule characteristics were reported and the sample size was very small. Additionally, only one patient underwent surgery during the study time frame, limiting further assessment of the tests' performance in this clinical setting."[20]

Clinical Utility

No studies were identified that reported on the clinical utility of either the ThyGenX mutation analysis test or the ThyraMIR miRNA GEC.

<u>ThyroSeq®, ThyroSeq v2 and ThyroSeq v3 (University of Pittsburgh Medical Center, CBLPath)</u>

In 2016, Hayes published a review on the use of ThyroSeq assay for all proposed purposes. Only five peer-reviewed cross-sectional studies were identified, one reported on analytical validity and four studies reported on clinical validity. These studies are described below. The review stated that the evidence for ThyroSeq suggested that when applied to indeterminate thyroid nodules, the test appeared to be a better "rule-out" test than it is a "rule-in" test, with sensitivities ranging from 85% to 100% and specificities ranging from 57% to 94%. In addition, while the PPVs ranged from 50% to 85% and the NPVs ranged from 90% to 100%, the accuracy of the pretest probabilities used to calculate these measures

was questionable and likely varied between studies. No studies were identified which reported on clinical utility. As a result, Hayes rated the ThyroSeq assay a "D2" rating for all proposed purposes based on evidence determined to be of very-low quality.

In 2020, Hayes published an updated review of the use of ThyroSeq v3 assay for the preoperative assessment of thyroid nodules with indeterminate cytology to assess the cancer probability in a given nodule.[23] The results of 1 analytical validity study did not report analytical sensitivity and specificity for all 112 genes, nor genetic alteration. Results from one clinical validity study reported a 97% negative predictive value; no clinical utility studies have been published to date. Hayes ultimately assigned the ThyroSeq v3 assay a "D2" rating (insufficient evidence) for assessing cancer probability in a given nodule and improving resultant patient management.

Analytical Validity

In 2013, Nikiforova et al. published a study to evaluate the diagnostic utility of the original version of the ThyroSeq next-generation sequencing (NGS) panel test, including 12 genes (284 mutations).[24] The panel test was used to analyze 228 thyroid neoplastic and nonneoplastic samples including 51 FNA samples to aid in cancer detection. The investigators reported 100% analytical accuracy of mutation detection. The analytical sensitivity was determined using serial dilutions of thyroid tumor DNA carrying a BRAF, TP53, or NRAS mutation in normal DNA, and mutations were detected down to a 3% allele frequency. Among the analyzed 228 samples, 115 mutations were detected. The mutation detection rate varied by cancer type, ranging from 30% in poorly differentiated and anaplastic tumors to 83% in the follicular variant papillary thyroid carcinomas (PTCs). The most common mutation found in PTCs was the BRAF V600E mutant, with an allele frequency of up to 48% in PTCs. Six percent of benign nodules were also found to have a mutation. Repeatability and precision were not reported.

Clinical Validity

- In 2014, Nikiforov et al. reported on the diagnostic utility of ThyroSeq V2 (13 genes) in 143 FNA samples of indeterminate nodules of the FN/FSN type (Bethesda IV), who subsequently had diagnostic surgery and a determination of whether the patient had thyroid cancer or not.[25] Pathogenic mutations or gene rearrangements were confirmed in 42 samples, 35 of which were malignant. The sensitivity and specificity of the ThyroSeq v2 were reported as 90% and 93%, respectively. The PPV and NPV were 83% and 96%, respectively. The overall accuracy of the test was reported as 92%.

- In 2015, Nikiforov et al. performed the ThyroSeq v.2.1 test on 465 thyroid nodule FNA samples with an indeterminate FNA cytopathologic result of the AUS/FLUS type (Bethesda III). [26] The ThyroSeq v2.1 panel included 14 genes analyzed for point mutations and 42 types of gene fusions occurring in thyroid cancer. Among 98 nodules with known outcome, ThyroSeq v2.1 was able to classify 20/22 cancers correctly, showing a sensitivity of 90.9% (CI 78.8%-100%), specificity of 92.1% (CI 86.0%-98.2%), PPV of 76.9% (CI 60.7%-93.1%), and negative predictive value of 97.2% (CI 78.8%-100%), with an overall accuracy of 91.8% (CI 86.4%-97.3%).

- In 2016, Shrestha et al. reported on a retrospective review of the diagnostic utility of the ThyroSeq panel (12 genes) in 261 FNA samples (68 of which had an indeterminate cytopathologic result) of patients who had also had thyroid surgery and histological diagnosis.[27,28] The sensitivity and specificity were 85% and 65%, respectively, for the analysis of AUS (Bethesda III) samples and 100% and 57%, respectively, for FN (Bethesda IV) samples. The specificity reported in this analysis was lower than previous studies, which could indicate a high false positive rate. However, it is unclear if this result was due to the samples included in this version of the panel test.

- In 2016, Valderrabano et al. published the results of a retrospective study that estimated predictive values for ThyroSeq v.2 and two other assays (Afirma and miRInform/ThyGenX) based on calculated prevalence of malignancy (PoM) for indeterminate FNA samples (Bethesda categories III and IV).[29] The authors reported that ThyroSeq v2 achieved the highest theoretical NPV and PPV in Bethesda III (98% and 75%, respectively) and Bethesda IV categories (96% and 83%, respectively). However, the authors noted that the false-negative rate of ThyroSeq v2 for the Bethesda III specimens may be underestimated, and therefore, the NPV might be actually lower.

- In 2016, Picarsic et al. published the results of a study that evaluated a 60-gene version of the ThyroSeq panel in 18 sporadic pediatric carcinomas, in an effort to identify additional genetic markers, including gene fusions, in pediatric differentiated thyroid carcinomas (DTC).[30] Six cases were mutation-negative using a smaller (7-gene panel) and three cases had not been tested previously. The use of the ThyroSeq assay revealed new gene fusions in four of six previously negative cases (67%) and one point mutation in a previously untested sample. The authors concluded that while a 7-gene panel identified only molecular alterations in 60% of cases, with the addition of the ThyroSeq v2 NGS, this increased to 87% (n=13/15).

Clinical Utility

No studies were identified that reported on the clinical utility of any version of the ThyroSeq mutation analysis panel test.

<u>BRAF Mutation Testing</u>

Clinical Validity

- In 2011, Adeniran et al. reported on a case series that evaluated whether BRAF V600E mutation testing improved diagnostic accuracy and provided prognostic information for 157 FNA samples that were either indeterminate/suspicious for PTC, or had a positive diagnosis of PTC.[31] Sixty-four (40.8%) FNAs were found to have BRAF V600E mutation, including 12 of 89 (13.5%) FNAs with an indeterminate diagnosis, four of 20 (20.0%) with a suspicious diagnosis, and 48 of 57 (84.2%) with a positive diagnosis. The percentage of FNAs with BRAF mutation was significantly higher in cases with a "positive" diagnosis than those with equivocal cases ($p < 0.0001$). The sensitivity and specificity of BRAF as a marker for PTC in thyroid nodules with an equivocal

cytologic diagnosis were 45.5% and 100%, respectively. All surgically resected nodules with a positive cytologic diagnosis were found to be PTC regardless of their BRAF status. Overall, the sensitivity of diagnosing PTC was 63.3% and 80.0% with cytology alone and combined cytology and BRAF testing, respectively. The specificity was 100% in both instances. No false positives were noted with either cytology or BRAF mutation analysis. All PTCs with extrathyroidal extension or aggressive histologic features were positive for BRAF mutation, however, the authors did not find any significant differences between PTCs with and without BRAF mutation in term of the tumor size, nodal status, and pathological stages. The authors concluded that individuals with an equivocal cytologic diagnosis and BRAF V600E mutation could be candidates for total thyroidectomy and central lymph node dissection.

- In 2016, Jinih et al. published the results of a systemic review that evaluated the diagnostic utility of the BRAFV600E mutation in indeterminate nodules (Bethesda class III and IV), including 32 studies involving 3150 indeterminate nodules.[32] The overall sensitivity and specificity for BRAFV600E for the diagnosis of thyroid malignancy was 40% (95% CI: 32%-48%) and 100% (95% CI: 98%-100%) respectively. The diagnostic odds ratio (DOR) was 205.4 (95% CI: 40.1-1052). The post-test probability of thyroid cancer, given a negative mutation was 6% and 92% with a positive result. The reviews concluded that "despite a high specificity for thyroid cancer, BRAFV600E mutation has a low overall sensitivity and therefore has a limited diagnostic value as a single screening test."

Clinical Utility

- In 2009, Xing et al. published a case series that reported on the clinical validity and utility of BRAF mutation testing of 190 FNA samples for papillary thyroid cancer (PTC) and if it could be used as a novel tool for preoperative risk stratification.[33] Seventy three samples were BRAF positive and 117 were BRAF negative (wild type). The authors reported that a BRAF mutation in preoperative FNA specimens was associated with poorer clinicopathologic outcomes when compared to wild type alleles, including extrathyroidal extension (23% vs. 11%; p=0.039), thyroid capsular invasion (29% vs. 16%; p=0.045), and lymph node metastasis (38% vs. 18%; p=0.002). During a median follow-up of 3 years (range, 0.6 to 10 years), PTC persistence/recurrence was observed in 36% of BRAF mutation-positive cases versus 12% of BRAF mutation-negative cases, with an odds ratio of 4.16 (95% CI: 1.70 to 10.17; p=0.002). The PPVs and NPVs for preoperative FNA-detected BRAF mutation to predict PTC persistence/recurrence were 36% and 88%, respectively, for all histologic subtypes of PTC. The authors concluded that BRAF mutation analysis, may be useful in appropriately tailoring the initial surgical extent for patients with PTC, but did not report on if the presence of a mutation changed actual management decisions. However, they also noted that BRAF testing has limited diagnostic value because of the low sensitivity of BRAF mutation when used in cytologically indeterminate specimens.

- In 2011, Kim et al. published a case series that reported on the clinical validity and utility of prospective BRAF mutation testing of 865 FNA samples for papillary thyroid cancer (PTC) and evaluated the surgical results of thyroid nodules based on BRAF V600E mutation status.[34] The

authors reported that of the 141 patients who had a diagnosis of atypical cells of undetermined significance (ACUS), 45 patients were BRAF-positive and were recommended to undergo surgery. Of the thirty BRAF-positive ACUS patients, 29 had PTC. However, 54 patients diagnosed as suspicious for malignancy were advised to undergo an operation, and they turned out to have PTCs regardless of their BRAFV600E mutation status. The authors reported that BRAF mutation testing was more specific (95.5% versus 36.4%) and had a higher PPV (99.4% versus 92.9%) than cytology testing, but was less sensitive (89.6% versus 100%) and had a lower NPV (52.5% versus 100%), but statistical analyses were not reported. There was no significant association between extrathyroidal extension, and lymph node metastasis and the BRAFV600E mutation

- status, which conflicts with the previous study and indicates a lack of prognostic value for BRAF testing.

Other Mutation Analysis Tests (Single- and/or Multi-gene)

Clinical Validity

- In 2010, Moses et al. reported the results of a large prospective series FNA samples from 417 individuals with 455 thyroid nodules that tested for common somatic mutations (BRAF, NRAS, KRAS) and gene rearrangements (RET/PTC1, RET/PTC3, RAS, and TRK1).[35] A total of 125 of 455 biopsies were found to be suspicious or indeterminate on cytologic exam. Overall, 50 mutations (23 BRAF V600E, 21 NRAS and 4 RET/PTC1 and 2 RET/PTC3 rearrangements) were detected. There were significantly more mutations identified in malignant nodules than in benign (p=0.0001). For thyroid FNA biopsies that were indeterminate or suspicious, genetic testing had a sensitivity of 12%, specificity of 98%, positive predictive value (PPV) of 38% and negative predictive value (NPV) of 65%. The authors noted that their results only moderately improved diagnostic accuracy in cases with indeterminate and/or suspicious thyroid FNA cytologic results.

- In 2011, Ferraz et al. published a review of 20 studies that reported on the type and number of mutations in cases of FNA of the thyroid diagnosed as indeterminate and compared the results to histological analysis y after surgical resection.[36] Sixteen studies analyzed single mutations (e.g., the BRAF V600E mutation or the RET/PTC gene rearrangement) and four studies analyzed a multi-gene or muti-mutation panels (e.g., BRAF, RAS, RET/PTC, and PAX8/PPARγ). Based on four studies that examined a panel of mutations, sensitivity ranged from 38-85.7% (mean of 63.7%) and specificity ranged from 95-100% (mean of 98%). The false-positive rates ranged from 0-4% (mean of 1.25%) and false negative rates ranged from 1-21% (mean of 9%). Based on two studies that examined RET/PTC rearrangements, mean sensitivity was 55% (range: 50-60%), specificity 100%, with a false-positive rate of 0% and mean false-negative rate of 3.5% (range: 1-6%). Based on three studies that examined BRAF mutations, mean sensitivity was 13% (range: 0-37.5%), mean specificity was 92.3% (range: 75-100%), mean false-positive rate was 0.5% (range: 0-1%) and mean false-negative rate was 6% (range 3-12%). The authors concluded that testing for a panel of mutations leads to an improvement in the sensitivity and specificity for indeterminate FNA of the thyroid; however, further standardizations and further molecular markers are needed before broad application of molecular FNA cytology for the diagnosis of thyroid nodules.

Evidence Summary

Afirma

There is sufficient evidence that the Afirma gene expression classifier (GEC) assay is an effective test to rule-out malignancy in FNA samples of thyroid nodules cytologically classified as indeterminate (Bethesda class III and IV), including its use on FNAs from Hurthle cell neoplasms. The evidence base consistently reports high sensitivity and high negative predictive values, indicating its use as a rule-out test. In addition, the majority of clinical utility studies reported significant decreases in surgical rates with the use of the Afirma GEC assay.

Other GEC Assays

There is a paucity of evidence on the clinical utility of GEC assays, other than Afirma, as either "rule-in" or "rule-out" tests for determining surgical treatment vs. watchful waiting for patients with thyroid nodules. Although other GECs like the Rosetta and the ThyraMIR miRNA GECs have moderate evidence of clinical validity, it is unclear if these tests perform as well as the Afirma GEC in terms of test sensitivity and NPV, due to the evidence base consisting entirely of small validation cohorts. In addition, there have been no studies published comparing other GECs to Afirma, to assess whether testing would be useful in making treatment decisions.

Mutational Analysis

There is a paucity of evidence on the clinical utility of any single-mutation, single-gene, or multi-gene tests including the ThyroSeq NGS panels. Although mutational analysis tests have been reported to have high specificities, NPV and/or PPV in certain subtypes of thyroid cancers and/or nodules with certain cytological classifications, these tests lack diagnostic value due to low sensitivities in indeterminate nodules. Furthermore, test performance values can vary widely depending on the test methodology used and the mutations and/or genes being tested.

Regarding BRAF V600E mutation testing for papillary thyroid cancer, there is conflicting evidence on test performance measures including sensitivity, specificity, PPV and NPV. In addition, there is conflicting evidence on whether the presence of the mutation is associated with poorer clinicopathological outcomes such as extrathyroidal extension, and lymph node metastasis, and a paucity of evidence on whether testing for the BRAF V600E mutation altered management decisions or improved health outcomes.

CLINICAL PRACTICE GUIDELINES

National Comprehensive Cancer Network (NCCN)

The NCCN thyroid carcinoma guidelines (v.2.2020) state that molecular diagnostic testing, by, "individual mutation testing or pattern recognition approaches using molecular classifiers" of fine needle aspirates from thyroid nodules *may be considered* in the following situations:[2]

- When results from FNA are indicative of any of the following:
    - atypia of undetermined significance (AUS)
    - follicular lesion of undetermined significance (FLUS)
    - follicular cell neoplasm (FCN)

NCCN states that in the above situations, molecular diagnostics *may be useful* to allow for reclassification of follicular lesions as either more or less likely to be benign or malignant, but that "molecular markers should be interpreted with caution and in the context of other clinical, radiographic and cytologic features of each individual patient." However, the panel indicated that the diagnostic utility of molecular tests has only been determined in adults.

In addition, NCCN stated that patients with indeterminate FNA results can be followed by observation, as opposed to immediate surgical resection, if the molecular test results in a predicted risk of malignancy comparable to the rate seen in cytologically benign FNAs (approximately 5%).

NCCN recommends *against* the use of molecular diagnostics for Hurthle cell neoplasms. However, this was stated in the footnotes of the recommendation section and was not supported by any evidence.

Lastly, the NCCN panel was not in consensus regarding the testing of BRAF V600E for papillary thyroid cancer.

<u>American Association of Clinical Endocrinologists (AACE), American College of Endocrinology (ACE), and the Associazione Medici Endocrinologi (AME)</u>

AACE/ACE/AME published a 2016 update to their joint evidence-based Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules.[37] The AACE/ACE/AME guidelines recommend the following:

When molecular testing should be considered

- To complement not replace cytologic evaluation.
- The results are expected to influence clinical management.
- As a general rule, not recommended in nodules with established benign or malignant cytologic characteristics.

These were strong recommendations, based on RCTs, and well-conducted prospective cohort studies and meta-analyses of those studies.

Molecular testing for cytologically indeterminate nodules

- 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.

- 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.

These were recommendations of moderate strength, based on primary on RCTs, and well-conducted prospective cohort studies and meta-analyses of those studies.

Role of molecular testing for deciding the extent of surgery

- 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.

This was a strong recommendation, based on RCTs, and well-conducted prospective cohort studies and meta-analyses of those studies.

How should patients with nodules that are negative at mutation testing be monitored?

- Since the false-negative rate for indeterminate nodules is 5 to 6% and the experience and follow-up for mutation-negative nodules or nodules classified as *benign* by a GEC are still insufficient, close follow-up is recommended.

This was a recommendation of moderate strength, based on low-grade evidence including poor quality RCTs, observational studies and case series.

The panel stated that molecular testing *may be considered*, if available, to reinforce the choice of a conservative strategy in select patients, but molecular analyses does not provide a conclusive diagnosis. In addition, the panel stated that a "benign" result from a GEC expression assay would be useful clinically when the NPV is approximately 95% or greater for any given test. Furthermore, sequencing-based genetic "testing for a panel of mutations may provide a diagnosis of malignancy in 20 to 40% of FN cases, but nodules that are negative for mutation still carry a substantial risk of malignancy."

American Thyroid Association (ATA)

The 2015 ATA Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer recommended the following regarding molecular testing for indeterminate nodules:[38]

- Regarding indeterminate nodules (AUS/FLUS and FN/SFN) on cytopathology "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."
- "Molecular testing may be used to supplement malignancy risk assessment data in lieu of proceeding directly with surgery."

The above were weak recommendations based on moderate-quality evidence. The panel concluded by stating that "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."

POLICY SUMMARY

There is sufficient evidence that the Afirma gene expression classifier (GEC) assay is an effective test to rule-out malignancy from FNA samples of 1 cm or greater thyroid nodules cytologically classified as indeterminate (Bethesda class III and IV). In addition, significant decreases in surgical rates with the use of the Afirma GEC assay indicate its clinical utility.

There is a paucity of evidence regarding the clinical utility of GEC assays other than Afirma, including the Rosetta and the ThyraMIR miRNA GECs. There is also a paucity of evidence on the clinical utility of any single-mutation, single-gene, or multi-gene tests including the ThyroSeq and ThyGenX NGS panels. Of note, when the ThyraMIR GEC and the ThyGenX mutation panel are used in combination, the test performance appears to be high, but there is no evidence of clinical utility, or how results from these tests alter treatment decisions.

Current clinical practice guidelines do not strongly recommend molecular diagnostic testing of any type as part of the standard evaluation for any sub-type of thyroid cancer.

INSTRUCTIONS FOR USE

Company Medical Policies serve as guidance for the administration of plan benefits. Medical policies do not constitute medical advice nor a guarantee of coverage. Company Medical Policies are reviewed annually and are based upon published, peer-reviewed scientific evidence and evidence-based clinical practice guidelines that are available as of the last policy update. The Companies reserve the right to determine the application of Medical Policies and make revisions to Medical Policies at any time. Providers will be given at least 60-days notice of policy changes that are restrictive in nature.

The scope and availability of all plan benefits are determined in accordance with the applicable coverage agreement. Any conflict or variance between the terms of the coverage agreement and Company Medical Policy will be resolved in favor of the coverage agreement.

REGULATORY STATUS

Mental Health Parity Statement

Coverage decisions are made on the basis of individualized determinations of medical necessity and the experimental or investigational character of the treatment in the individual case. In cases where medical necessity is not established by policy for specific treatment modalities, evidence not previously considered regarding the efficacy of the modality that is presented shall be given consideration to determine if the policy represents current standards of care.

REFERENCES

1. Cibas ES, Ali SZ. The Bethesda System For Reporting Thyroid Cytopathology. *American journal of clinical pathology.* 2009;132(5):658-665.https://academic.oup.com/ajcp/article/132/5/658/1765741.
2. National Comprehensive Cancer Network (NCCN). Thyroid Carcinoma. v.2.2020. July 15, 2020. https://www.nccn.org/professionals/physician_gls/pdf/thyroid.pdf. Accessed 10/5/2020.
3. Veracyte. Afirma® product page. https://www.veracyte.com/our-products/afirma. Accessed 10/5/2020.
4. Interpace Biosciences, Inc. Interpace Diagnostics. ThyGeNEXT | ThyraMIR. https://thygenext-thyramir.com/. Accessed 10/5/2020.
5. University of Pittsburgh Medical Center. TheyroSeq | Test Description. https://thyroseq.com/physicians/test-details/test-description. Accessed 10/5/2020.
6. Hayes. Genetic Test Evaluation (GTE) Report: Afirma Thyroid FNA Analysis (Veracyte). 2018 (archived 2019). https://www.hayesinc.com/subscribers/displaySubscriberArticle.do?articleId=14250. Accessed 10/5/2020.
7. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab.* 2010;95(12):5296-5304.https://academic.oup.com/jcem/article/95/12/5296/2835246.
8. Walsh PS, Wilde JI, Tom EY, et al. Analytical performance verification of a molecular diagnostic for cytology-indeterminate thyroid nodules. *J Clin Endocrinol Metab.* 2012;97(12):E2297-2306.https://academic.oup.com/jcem/article/97/12/E2297/2536299.
9. Kloos RT, Monroe RJ, Traweek ST, Lanman RB, Kennedy GC. A Genomic Alternative to Identify Medullary Thyroid Cancer Preoperatively in Thyroid Nodules with Indeterminate Cytology. *Thyroid.* 2016;26(6):785-793.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4913490/.
10. Pankratz DG, Hu Z, Kim SY, et al. Analytical Performance of a Gene Expression Classifier for Medullary Thyroid Carcinoma. *Thyroid.* 2016;26(11):1573-1580.https://www.ncbi.nlm.nih.gov/pubmed/?term=27605259.
11. Angell TE, Wirth LJ, Cabanillas ME, et al. Analytical and Clinical Validation of Expressed Variants and Fusions From the Whole Transcriptome of Thyroid FNA Samples. *Frontiers in Endocrinology.* 2019;10(612).https://www.frontiersin.org/article/10.3389/fendo.2019.00612.
12. ECRI Institute. ECRIgene Evidence Report. Thyroid Gene Expression Classifier (Afirma) for Assessing Indeterminate Thyroid Nodules. May 2018. https://www.ecri.org/components/ECRIgene/Documents/EG0128.pdf. Accessed 10/5/2020.
13. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med.* 2012;367(8):705-715.https://www.ncbi.nlm.nih.gov/pubmed/22731672.
14. Kloos RT, Reynolds JD, Walsh PS, et al. Does addition of BRAF V600E mutation testing modify sensitivity or specificity of the Afirma Gene Expression Classifier in cytologically indeterminate thyroid nodules? *J Clin Endocrinol Metab.* 2013;98(4):E761-768.https://academic.oup.com/jcem/article/98/4/E761/2537248.
15. ECRI Institute. Afirma Genomic Sequencing Classifier (Veracyte, Inc.) for Evaluating Thyroid Nodules of Indeterminate Cytopathologic Diagnosis.

https://www.ecri.org/components/ECRIgene/Documents/EG0213.pdf. Published 2020. Accessed 12/11/2020.

16. Angell TE, Frates MC, Medici M, et al. Afirma Benign Thyroid Nodules Show Similar Growth to Cytologically Benign Nodules During Follow-Up. *J Clin Endocrinol Metab.* 2015;100(11):E1477-1483.https://www.ncbi.nlm.nih.gov/pubmed/26353010.
17. Hayes. Genetic Test Evaluation (GTE) Report: RosettaGX Reveal (Rosetta Genomics Ltd.). Published: 12/29/2016. https://www.hayesinc.com/subscribers/displaySubscriberArticle.do?articleUid=gte.rosettagx3907. Accessed last: 12/5/2019.
18. Benjamin H, Schnitzer-Perlman T, Shtabsky A, et al. Analytical validity of a microRNA-based assay for diagnosing indeterminate thyroid FNA smears from routinely prepared cytology slides. *Cancer cytopathology.* 2016;124(10):711-721.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5096036/.
19. Lithwick-Yanai G, Dromi N, Shtabsky A, et al. Multicentre validation of a microRNA-based assay for diagnosing indeterminate thyroid nodules utilising fine needle aspirate smears. *Journal of Clinical Pathology.* 2017;70(6):500-507.http://jcp.bmj.com/content/jclinpath/70/6/500.full.pdf.
20. Hayes. Genetic Test Evaluation (GTE) Report: ThyGenX and ThyraMIR (Interspace Diagnostics). Published: 11/2/2017. https://www.hayesinc.com/subscribers/displaySubscriberArticle.do?articleUid=gte.thygenx4186. Accessed last: 12/5/2019.
21. ECRIgene Genetic Test Product Brief. ThyGeNEXT Oncogene Panel Plus ThyraMIR microRNA Classifier (Interpace Diagnostics, LLC) for Evaluating Cytologically Indeterminate Thyroid Nodules. Published: January 2020. https://www.ecri.org/components/ECRIgene/Documents/EG0089.pdf. Accessed 10/05/2020.
22. Young NA, Win KK, Pomo L, Anastasopoulou C, Minimo C, Mayrin J. An academic community hospital experience using commercially available molecular testing in the management of indeterminate thyroid nodules. *Journal of the American Society of Cytopathology.* 2017.https://www.sciencedirect.com/science/article/pii/S2213294517301060.
23. Hayes, Inc. ThyroSeq v3 (University of Pittsburgh Medical Center, CBLPath Inc.) | Annual Review: Jun 16, 2020. https://evidence.hayesinc.com/report/gte.thryoseq4640. Accessed 10/5/2020.
24. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab.* 2013;98(11):E1852-1860.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3816258/.
25. Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer.* 2014;120(23):3627-3634.https://www.ncbi.nlm.nih.gov/pubmed/25209362.
26. Nikiforov YE, Carty SE, Chiosea SI, et al. Impact of the Multi-Gene ThyroSeq Next-Generation Sequencing Assay on Cancer Diagnosis in Thyroid Nodules with Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance Cytology. *Thyroid.* 2015;25(11):1217-1223.https://www.ncbi.nlm.nih.gov/pubmed/26356635.
27. Shrestha RT, Evasovich MR, Amin K, et al. Correlation Between Histological Diagnosis and Mutational Panel Testing of Thyroid Nodules: A Two-Year Institutional Experience. *Thyroid.* 2016;26(8):1068-1076.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4976225/.

28. Shrestha RT, Karunamurthy A, Amin K, Nikiforov YE, Caramori ML. Multiple Mutations Detected Preoperatively May Predict Aggressive Behavior of Papillary Thyroid Cancer and Guide Management--A Case Report. *Thyroid.* 2015;25(12):1375-1378
29. Valderrabano P, Leon ME, Centeno BA, et al. Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. *European journal of endocrinology.* 2016;174(5):621-629.http://www.eje-online.org/content/174/5/621.long.
30. Picarsic JL, Buryk MA, Ozolek J, et al. Molecular Characterization of Sporadic Pediatric Thyroid Carcinoma with the DNA/RNA ThyroSeq v2 Next-Generation Sequencing Assay. *Pediatric and developmental pathology : the official journal of the Society for Pediatric Pathology and the Paediatric Pathology Society.* 2016;19(2):115-122.https://www.ncbi.nlm.nih.gov/pubmed/26367451.
31. 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.* 2011;21(7):717-723.https://www.ncbi.nlm.nih.gov/pubmed/?term=Reflex+BRAF+testing+in+thyroid+fine-needle+aspiration+biopsy+with+equivocal+and+positive+interpretation%3A+a+prospective+study.
32. Jinih M, Foley N, Osho O, et al. BRAF(V600E) mutation as a predictor of thyroid malignancy in indeterminate nodules: A systematic review and meta-analysis. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology.* 2017;43(7):1219-1227.https://www.ncbi.nlm.nih.gov/pubmed/27923591.
33. 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.* 2009;27(18):2977-2982.https://www.ncbi.nlm.nih.gov/pubmed/19414674.
34. Kim SK, Hwang TS, Yoo YB, et al. Surgical results of thyroid nodules according to a management guideline based on the BRAF(V600E) mutation status. *J Clin Endocrinol Metab.* 2011;96(3):658-664.https://academic.oup.com/jcem/article/96/3/658/2596610.
35. Moses W, Weng J, Sansano I, et al. Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. *World J Surg.* 2010;34(11):2589-2594.https://www.ncbi.nlm.nih.gov/pubmed/20703476.
36. Ferraz C, Eszlinger M, Paschke R. Current state and future perspective of molecular diagnosis of fine-needle aspiration biopsy of thyroid nodules. *J Clin Endocrinol Metab.* 2011;96(7):2016-2026.https://academic.oup.com/jcem/article/96/7/2016/2834027.
37. 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.* 2016;22(5):622-639.https://www.ncbi.nlm.nih.gov/pubmed/27167915.
38. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid.* 2016;26(1):1-133.https://www.ncbi.nlm.nih.gov/pubmed/26462967.