Healthy Blue

Clinical UM Guideline

Subject:	Molecular Marker Evaluation of Thyroid Nodules		
Guideline #:	CG-GENE-04	**Publish Date:**	04/15/2020
Status:	Revised	**Last Review Date:**	02/20/2020

Description

This document addresses the use of molecular markers for the evaluation of thyroid nodules to identify genetic mutations (mutation analysis) and to identify benign thyroid nodules preoperatively. Examples of these tests include, but are not limited to:

- Afirma® Thyroid FNA Analysis (Veracyte, South San Francisco, CA)
- ThyGenX® and ThyraMIR™ (Interpace Diagnostics, Parsippany, NJ)
- ThyroSeq® (CBLPath, Rye Brook, NY)
- RosettaGX Reveal™ (Rosetta Genomics, Philadelphia, PA)

Note: Please see the following related document for additional information:

- CG-GENE-03 BRAF Mutation Analysis

Clinical Indications

Medically Necessary:

The use of the Afirma Genomic Sequencing Classifier or the ThyroSeq genomic classifier for molecular marker evaluation of a thyroid nodule is considered **medically necessary** for use with fine needle aspirates, after initial cytopathology is indeterminate (that is, atypia of undetermined significance [AUS], follicular lesion of undetermined significance [FLUS], suspicious for follicular neoplasm [SFN], follicular neoplasm [FN], and suspicious for malignancy [SUS]).

Not Medically Necessary:

The use of the Afirma Genomic Sequencing Classifier or the ThyroSeq genomic classifier for molecular marker evaluation of thyroid nodules is considered **not medically necessary** for repeat testing of the same nodule and all other indications not listed above as medically necessary.

The use of other molecular marker evaluations of thyroid nodules (for example, ThyGenX, ThyraMIR, Rosetta) is considered **not medically necessary.**

Coding

The following codes for treatments and procedures applicable to this guideline are included below for informational purposes. Inclusion or exclusion of a procedure, diagnosis or device code(s) does not constitute or imply member coverage or provider reimbursement policy. 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.

CPT	
81545	Oncology (thyroid), gene expression analysis of 142 genes, utilizing fine needle aspirate, algorithm reported as a categorical result (eg, benign or suspicious) Afirma® Gene Expression Classifier, Veracyte, Inc
81599	Unlisted multianalyte assay with algorithmic analysis [when specified as testing for thyroid molecular markers by other than a gene expression classifier, for example, mutation analysis] **Note:** assays other than Afirma and Thyroseq are considered Not Medically Necessary
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 ThyraMIR™, Interpace Diagnostics, Interpace Diagnostics **Note:** assays other than Afirma and Thyroseq are considered Not Medically Necessary
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 probabilityof malignancy" or "Negative, low probability of malignancy") Thyroseq Genomic Classifier, CBLPath, Inc, University of Pittsburgh Medical Center

ICD-10 Diagnosis	
	All diagnoses

Discussion/General Information

According to the National Cancer Institute (NCI), there were an estimated 52,070 new cases of thyroid cancer and 2170 deaths in 2019. Thyroid cancer is more common in women and typically affects people between 25 and 65 years old. Thyroid cancer usually starts as a nodule; however, thyroid nodules are common and most never become cancerous. While some nodules may be visible or palpable, most are found during imaging of the head and neck for unrelated reasons. Suspicious nodules, such as those that are large or have calcifications, are usually evaluated for cancer by ultrasound and fine needle aspiration (FNA) biopsy with cytological evaluation. For up to 30% of nodules, the results of the FNA are classified as indeterminate by the Bethesda system: atypia of undetermined significance (AUS), follicular lesion of undetermined significance (FLUS), follicular neoplasm (FN), suspicious for follicular neoplasm (SFN), or suspicious for malignancy (SUS). When nodules are considered indeterminate, a cancer diagnosis cannot be established. To make a diagnosis, indeterminate nodules are excised surgically (lobectomy or partial thyroidectomy) for histopathological evaluation, and if the result is positive, a second surgery may be needed to completely remove the thyroid (total thyroidectomy). The majority of surgically resected indeterminate nodules are found to be benign, resulting in many unnecessary surgeries. Molecular marker tests are proposed as an option to help individuals avoid unnecessary surgery by ruling out cancer preoperatively. In

addition, molecular marker tests are used for mutation analysis to help determine the appropriate intervention, such as a total thyroidectomy instead of partial thyroidectomy for individuals with an aggressive cancer type.

Various genetic mutations have been discovered in thyroid cancer. The four gene mutations that are most common and carry the highest impact on tumor diagnosis and prognosis are BRAF, RAS, RET/PTC, and PAX8/PPARγ rearrangements. These non-overlapping genetic alterations are found in more than 70% of papillary and follicular thyroid carcinomas (Nikiforov, 2011).

Studies on the use of molecular markers for thyroid nodules have analyzed single mutations or panels of mutations and compared the preoperative cytologic findings with postoperative histologic diagnosis to determine the diagnostic accuracy of mutation presence and to predict the presence and potential aggressiveness of a malignancy. Additionally, gene expression classifiers have been developed to predict the likelihood that a thyroid lesion with indeterminate cytology is benign, allowing an individual to potentially avoid surgical excision.

Nikiforov and colleagues (2009) prospectively tested a panel of mutations (BRAF, RAS, RET/PTC, and PAX8/PPARγ) in 470 FNA samples of thyroid nodules from 328 individuals. Mutational status was correlated with cytology and either surgical pathology diagnosis or medical follow-up (mean 34 months). A total of 40 individuals (12%) were excluded due to loss at follow-up or for a poor quality specimen. A total of 69 individuals (with 86 thyroid FNA samples) underwent surgery soon after completion of the cytologic evaluation; preoperative cytologic diagnosis was: positive for malignancy in 22 samples, indeterminate (including atypical and suspicious for malignancy) in 52 samples, and negative for malignancy in 12 samples. By FNA, 32 mutations were found (18 BRAF, 8 RAS, 5 RET/PTC, and 1 PAX8/PPARγ); after surgery, 31 mutation-positive nodules (97%) were diagnosed as malignant on pathological examination, and 1 was a benign tumor (3%). A total of 13 of the 32 mutation-positive FNA samples had a definitive cytologic diagnosis of malignancy, whereas the rest were either indeterminate or negative for malignancy. Of the remaining 219 individuals, 147 (229 FNAs) who did not undergo surgery were followed by serial ultrasound with no change in the nodule status (124 cases) or by repeated FNA with cytology negative for malignancy (23 cases) and no mutation found in the FNA material. These nodules were considered negative for malignancy. The remaining 72 individuals who were initially in the follow-up group underwent subsequent surgery. Combining all three groups, the specificity for malignancy was 99.7%; the sensitivity of the molecular test alone was 62%.

Xing and colleagues (2009) investigated the utility of BRAF mutation testing of thyroid FNA biopsy specimens for papillary thyroid cancer (PTC) preoperative risk stratification in 190 individuals. BRAF mutation in preoperative FNA specimens was reported to be associated with poorer clinicopathologic outcomes of PTC. In comparison with the wild-type allele, a BRAF mutation strongly predicted 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% confidence interval [CI], 1.70 to 10.17; p=0.002). The positive predictive values (PPV) and negative predictive values (NPV) for preoperative FNA-detected BRAF mutation to predict PTC persistence/recurrence were 36% and 88%, respectively, for all histologic subtypes of PTC. The authors reported that efforts are needed to define specific indications and practice guidelines for such clinical use of BRAF mutation in the management of PTC.

Moses and colleagues (2010) prospectively analyzed FNA samples from 417 individuals with 455 thyroid nodules for common somatic mutations (BRAF, NRAS, KRAS) and gene rearrangements (RET/PTC1 and 3, RAS and TRK1 mutations). 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%, PPV of 38%, and NPV of 65%.

Ohori and colleagues (2010) performed mutation screening in 117 FNA samples classified as FLUS/AUS. RAF, RAS, RET/PTC, or PAX8/PPARγ mutations were detected in 10% of this category. They demonstrated that the probability of having a malignancy in this cytology category together with a detection of one of the somatic mutations investigated was 100%, whereas the probability of having a thyroid malignancy without molecular alteration detected was 7.6%.

Cantara and colleagues (2010) analyzed panels of mutations from 174 individuals (235 thyroid nodules) undergoing surgery for indeterminate/inadequate/benign thyroid FNA results. The most prevalent mutation was BRAF (49.3% of the positive samples), followed by RAS (34.3%), and RET/PTC (16.4%). The combination of cytology and mutation analysis improved the accuracy for diagnosing cancer from 83% to 93.2% when compared to cytologic analysis alone. Molecular analysis detected 8 thyroid cancers that were missed on cytology from a total of 32 cancers that were diagnosed as indeterminate/inadequate/benign. When the FNA mutation analysis was compared with the mutation analysis of the corresponding histologic material from the surgical sample, in 88.2% of cases, the mutation found in the FNA material was also detected in the histologic samples. The 11.8% discrepant results were reported as being due to the presence of a mutation in the tissue sample that was not found in the cytology sample.

Mathur and colleagues (2010) collected thyroid FNA samples, thyroid tissue, clinical and histopathology data, and tumor genotyping for mutations BRAF V600E, NRAS, KRAS, RET/PTC1, RET/PTC3, and NTRK1 for 341 individuals with 423 dominant thyroid nodules. A cytologic examination of the samples showed that 51% were benign (one-quarter of these were surgically resected), 21% were malignant, 11% were atypical lesions, 12% were follicular or Hürthle cell neoplasms, and 4% were suspicious for malignancy. On final analysis, 165 nodules were benign and 123 were malignant. Of the 423 FNA samples, 24 BRAF V600E mutations, 7 KRAS, 21 NRAS, 4 PAX8-PPARγ rearrangements, 3 RET/PTC1, and 2 RET/PTC3 rearrangements were detected. In all, 17 of 165 (10.3%) benign thyroid nodules had a mutation compared with 26% (32 of 123) for malignant tumors ($p<0.05$).

Adeniran and colleagues (2011) evaluated 157 cases of equivocal thyroid FNA readings (indeterminate and suspicious for PTC) or a positive diagnosis of PTC and concomitant BRAF mutation analysis. Histopathologic follow-up results were 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 both cytology and BRAF testing. 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. 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.

Ferraz and colleagues (2011) evaluated 20 publications that reported on the type and number of mutations in cases of FNA of the thyroid diagnosed as indeterminate and compared the results to final histology after surgical resection. A total of 16 studies analyzed 1 mutation (such as BRAF or RET/PTC) and 4 studies analyzed a panel of several mutations (BRAF, RAS, RET/PTC, and PAX8/PPARγ). The detection of a mutation in a surgically resected benign thyroid lesion was categorized as a false-positive case. Detecting no mutation in an FNA sample from a histologically benign surgical sample was considered a true-negative, and finding no mutation in a histologically malignant lesion was categorized as a false-negative. Based on four studies that examined a panel of mutations, there was a broad sensitivity range of 38-85.7% (mean 63.7%), a mean specificity of 98% (range 95-100%), mean false-positive rate of 1.25% (0-4%), and mean false-negative rate of 9% (1-21%). Based on two studies that examined RET/PTC rearrangements, mean sensitivity was 55% (50-60%), specificity 100%, false-positive rate of 0%, and mean false-negative rate 3.5% (1-6%). Based on three studies that examined BRAF mutations, mean sensitivity was 13% (0-37.5%), mean specificity 92.3% (75-100%), mean false-positive rate 0.5% (0-1%), and mean false-negative rate 6% (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.

Due to the potential for cancer overdiagnosis, an international and multidisciplinary study (Nikiforov, 2016) was done to evaluate and refine the diagnostic criteria for encapsulated follicular variant of papillary thyroid carcinoma (EFVPTC). Previously, all individuals with EFVPTC were considered to have thyroid cancer. Researchers retrospectively analyzed 109 subjects with noninvasive EFVPTC for 10-26 years and 101 subjects with invasive EFVPTC for 1-18 years. At the end of the study, all of the non-invasive EFVPTC subjects were alive with no evidence of thyroid cancer. For the invasive group, 12 subjects had adverse events, including distant metastasis (5 subjects) and death from thyroid cancer (2 subjects). Based on the outcomes, the researchers proposed to reclassify non-invasive EFVPTC to noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP).

ThyroSeq

The ThyroSeq test, developed at the University of Pittsburgh, uses next-generation sequencing (NGS) of DNA and RNA. The test started out as a 7-gene panel in 2007 and expanded to 15 genes in 2013 (version 1), 56 genes in 2014 (version 2), and 112 genes in 2017 (version 3). The test is primarily marketed to rule out cancer for nodules with indeterminate FNA cytology. In addition, ThyroSeq is marketed for cancer prognostication to determine the type of surgery needed for malignant nodules (for example, lobectomy for indolent cancer versus total thyroidectomy for aggressive cancer). The latest version of the test provides information on > 12,000 mutation hotspots and > 120 gene fusion types, while detecting mutations, gene fusions, gene expression alterations, and copy number variations. A proprietary genomic classifier is used to report positive or negative results.

Nikiforova and colleagues (2013) examined the use of targeted NGS for simultaneous testing of multiple mutations in thyroid cancer. The aim of the study was to create an NGS approach to allow for the detection of most point mutations and small insertions or deletions known to occur in thyroid cancer. A custom panel, ThyroSeq, was designed to target 12 thyroid cancer genes with 284 mutational hotspots. DNA from 228 thyroid neoplastic and

non-neoplastic samples, including 105 frozen, 72 formalin-fixed, and 51 FNA samples representing all major types of thyroid cancer, was analyzed. The analytical accuracy for mutation detection was reported to be 100% with the sensitivity of 3–5% of mutant allele. ThyroSeq DNA assay identified mutations in 19/27 (70%) of classic PTC, 25/30 (83%) of follicular variant PTC, 14/18 (78%) of conventional and 7/18 (39%) of oncocytic follicular carcinomas, 3/10 (30%) of poorly differentiated carcinomas, 20/27 (74%) of anaplastic (ATC), and 11/15 (73%) medullary carcinomas. In contrast, 5/83 (6%) of benign nodules were positive for mutations. Most tumors had a single mutation; however, several ATC and PTC demonstrated two or three mutations. The most common mutations detected were BRAF and RAS. Additional studies are needed to support the conclusions.

Additional research by Nikiforov and others (2014, 2015) further studied the use of NGS testing using ThryoSeq. In the 2014 study, the authors evaluated 143 consecutive FNA samples with a cytologic diagnosis of follicular neoplasm or suspicious for follicular neoplasm (FN/SFN) from individuals with known surgical outcomes. Included were 91 retrospective samples and 52 prospective samples. Analyses were performed on a proprietary sequencer using the targeted ThyroSeq v2 NGS panel, simultaneously testing for point mutations in 13 genes and for 42 types of gene fusions that occur in thyroid cancer. The expression of 8 genes was used to assess the cellular composition of FNA samples. Histologic analysis revealed 104 benign nodules and 39 malignant nodules. The most common point mutations involved the neuroblastoma RAS viral oncogene homolog (NRAS), followed by the Kirsten rat sarcoma viral oncogene homolog (KRAS), the telomerase reverse transcriptase (TERT) gene, and the thyroid-stimulating hormone receptor (TSHR) gene. The identified fusions involved the thyroid adenoma associated (THADA) gene; the peroxisome proliferator-activated receptor gamma (PPARG) gene; and the neurotrophic tyrosine kinase, receptor, type 3 (NTRK3) gene. Performance characteristics were similar in the retrospective and prospective groups. Among all FN/SFN nodules, preoperative ThyroSeq v2.0 performed with 90% sensitivity (95% CI, 80% to 99%), 93% specificity (95% CI, 88% to 98%), a PPV of 83% (95% CI, 72% to 95%), a NPV of 96% (95% CI, 92% to 100%), and 92% accuracy (95% CI, 88% to 97%).

In 2015, the authors similarly tested the ThryoSeq v2.1 (v2.1 has an additional gene for analysis when compared to v2.0) for AUS/FLUS. A total of 465 consecutive FNA samples with the cytologic diagnosis of AUS/FLUS underwent prospective molecular testing. A total of 98 (21%) of these nodules had definitive surgical (n = 96) or nonsurgical (n = 2) follow-up and were used to determine the assay performance. Among 465 AUS/FLUS nodules, 3 were found to be composed of parathyroid cells and 462 of thyroid follicular cells. Of the latter, 31 (6.7%) were positive for mutations. The most frequently mutated genes were NRAS and HRAS, and overall point mutations in seven different genes and five types of gene fusions were identified in these nodules. Among 98 nodules with known outcome, histologic analysis revealed 22 (22.5%) cancers. ThyroSeq v2.1 was able to classify 20/22 cancers correctly, showing a sensitivity of 90.9% (95% CI, 78.8 to 100), specificity of 92.1% (95% CI, 86.0 to 98.2), PPV of 76.9% (95% CI, 60.7 to 93.1), and NPV of 97.2% (95% CI, 78.8 to 100), with an overall accuracy of 91.8% (95% CI, 86.4 to 97.3).

Shrestha and colleagues (2016) conducted a retrospective review on the performance of ThyroSeq at a single hospital. The researchers included thyroid nodule surgeries (n=261) from January 2013 to December 2014, and a total of 73 indeterminate nodules had molecular testing, including 5 that were nondiagnostic due to insufficient samples. The remaining 68 nodules included 44 AUS/FLUS, 12 FN, and 12 SUS. Mutational analysis was done at the University of Pittsburgh Medical Center using the 7-gene panel (n=23) or ThyroSeq v1/ThyroSeq v2 (n=45

combined). The AUS/FLUS group had a 85% sensitivity, a 65% specificity, a 50% PPV, and a 91% NPV. The FN group had a 100% sensitivity, 57% specificity, 63% PPV, and 100% NPV. For the SUS group, there were 8 true positives, no false positives, 1 false negative, and 3 true negatives. Limitations of the study included a small sample size and lack of pathological re-review prior to the analysis. The researchers concluded that the test "had a high sensitivity and NPV consistent with published data, but a lower than anticipated PPV and specificity." They recommended additional independent studies.

Valderrabano and colleagues (2017) evaluated the performance of ThyroSeq v2 for indeterminate nodules between September 2014 and April 2016. As part of an institutional pathway, all individuals with indeterminate nodules were offered ThyroSeq v2 testing. In 2015, the researchers began using the NIFTP reclassification for encapsulated FVPTC, and two pathologists blindly re-reviewed previous FVPTC and NIFTP nodules for classification (1 slide was not available for re-review). The researchers evaluated 192 indeterminate nodules in 184 individuals using ThyroSeq v2. For 190 nodules with valid ThyroSeq v2 results, 102 underwent surgical resection. The performance of ThyroSeq v2 was assessed in two scenarios: (1) considering NIFTP malignant and (2) considering NIFTP benign. When considering NIFTP malignant, the overall performance was 70% sensitivity, 77% specificity, 42% PPV, and 91% NPV. When considering NIFTP benign, the overall performance was 73% sensitivity, 75% specificity, 33% PPV, and 95% NPV. ThyroSeq v2 performed significantly better for FN than AUS/FLUS (area under the receiver operating characteristic [AUROC] curve 0.84 vs 0.57 and 0.85 vs 0.55; p=0.03), and the researchers considered the test "essentially uninformative" for AUS/FLUS. For FN, the test had a PPV of 65% and an NPV of 94%. Limitations of the study included the retrospective design, lack of blinding, and small sample size. The researchers concluded:

> Our study suggests that ThyroSeq v2 is informative for nodules in the B-IV category [FN], achieving a NPV robust enough to consider observation in lieu of surgery. However, the PPV of the test was lower than expected; and its performance particularly in B-III [AUS/FLUS] specimens may be variable among centers… Further studies evaluating the performance of ThyroSeq v2 in indeterminate thyroid nodules, particularly in B-III [AUS/FLUS] specimens, are needed.

Taye and colleagues (2018) performed a retrospective analysis on the performance of ThyroSeq v2. The researchers included all indeterminate nodules at three hospitals (156 nodules from 151 individuals) that were profiled between November 2014 and April 2016. A pathologist re-reviewed ultrasound records, FNA cytopathology, and surgical pathology to make sure the FNA biopsied nodule was matched with the correct surgically resected nodule. Incidental carcinomas were excluded, and the pathologist was blinded to the ThyroSeq test results. For a ThyroSeq result to be considered positive, the malignancy probability had to be greater than 30%. After analysis, the ThryoSeq negative call rate was 65% (102/156), and the PPV of a positive result was 22% (27% if counting NIFTP as malignant). The NPV was 96% (one false-negative result). A Bayes theorem analysis revealed an NPV of 96-98% and a PPV of 56-83%, leading the researchers to consider that the specificity of Thyroseq is lower than previously reported. Limitations of the study included a small sample size, retrospective design, re-review by a single pathologist, small number of FN nodules (n=15), and lack of all positive ThyroSeq nodules having surgical resection results. The researchers concluded, "Our experience indicates that Thyroseq is likely to offer high NPV but may have a lower PPV than reported."

Nikiforova and colleagues (2018) reported on the analytical performance of ThyroSeq v3. Study samples included 238 surgically removed tissue samples, 175 FNA indeterminate cytology samples, 16 cell lines, and 4 reference controls. For the surgically removed tissue samples, the test detected over 100 alterations and accurately classified the majority of papillary carcinomas (91.2%), follicular carcinomas (90.9%), Hürthle cell lesions (93.1%), medullary thyroid carcinomas (100%), and parathyroid lesions (100%). The genomic classifier cutoffs for differentiating cancerous and benign nodules had a 93.9% sensitivity, 89.4% specificity, and 92.1% accuracy. For the FNA samples, the validation set had a 98.0% sensitivity, 81.8% specificity, and 90.9% accuracy. The authors concluded that the test is valid for clinical use, and further studies will be done to determine clinical utility.

Steward and colleagues (2018) published a prospective, double-blinded, multicenter study that evaluated the diagnostic accuracy of ThyroSeq v3. Inclusion criteria included 18 years of age and older, 1 or more thyroid nodule, and an FNA procedure to collect samples for cytological examination and molecular analysis. Those who had a cytologic diagnosis of Bethesda III, IV, or V underwent thyroid surgery. The study was conducted across 10 sites (9 in the United States and 1 in Singapore) between January 2015 and December 2016. The primary outcome was the sensitivity, specificity, NPV, and PPV of the test to predict the histopathologic diagnosis of benign nodule versus cancer/NIFTP in indeterminate nodules with a Bethesda III and IV cytology. The secondary outcome was the prediction of cancer/NIFTP by specific genetic alterations in Bethesda III, IV, and V cytology nodules. A total of 286 FNA samples met inclusion criteria. After some samples were found inadequate for molecular analysis, a total of 257 samples (90%) from 232 subjects were in the final study set (Bethesda III [n=154], Bethesda IV [n=93], and Bethesda V [n=10]). For Bethesda III and IV combined, the sensitivity was 94% (95% CI, 86 to 98%), and the specificity was 82% (95% CI, 75 to 87%). Considering a cancer/NIFTP prevalence of 28%, the NPV was 97% (95% CI, 93 to 99%) and the PPV was 66% (95% CI, 56-75%). For Bethesda III and IV nodules, the negative benign call rate was 61%. Of 152 test-negative samples, 5 (3%) were found to be false-negative. For the nodules that tested positive, specific groups of genetic alterations had cancer probabilities from 59 to 100%. The authors concluded:

> The study documents a high sensitivity and correspondingly high NPV of the ThyroSeq GC test for Bethesda III and IV indeterminate cytology nodules, which together with high specificity may prevent diagnostic surgeries in the majority of such patients. The availability of detailed genetic information in test-positive cases may help to further inform individualized treatment for these patients after integration with imaging and other clinical information.

Several additional recently published cohort studies and retrospective reviews also indicate that ThyroSeq testing may contribute to avoidance of surgery for initially indeterminate thyroid nodules (Marcadis, 2019; Ohori, 2019).

Afirma Thyroid FNA Analysis

The Afirma Thyroid FNA Analysis combines specialized cytopathology with the Afirma Gene Expression Classifier (GEC). The GEC analyzes the mRNA expression of 167 genes in aspiration material and reclassifies FNAs with ambiguous cytopathology diagnoses as either benign or suspicious for cancer. Approximately 10% of

FNA samples have inadequate RNA yield or quality and are reported by the Afirma GEC as "no result" (Duick 2012; Ali 2013). The Afirma is intended to rule out thyroid cancer (Ward, 2013).

Chudova and colleagues (2010) set out to develop a molecular test to distinguish between benign and malignant thyroid nodules using FNA. More than 247,000 transcripts in 315 thyroid nodules were measured by mRNA expression analysis. The data set consisted of 178 retrospective surgical specimens, representing the most common benign and malignant histologic subtypes, and 137 prospectively collected aspirate specimens. Two classifiers were trained separately on surgical samples and aspirates. The performance was evaluated using an independent test set of 48 prospective FNA samples which had known surgical pathology diagnoses, including 50% with indeterminate cytopathology. The performance of the classifier was markedly lower in the FNAs than in tissue, likely due to differences in cellular heterogeneity between the two types of specimens. On the test set, NPV and specificity were estimated to be 96% and 84%, respectively.

Walsh and colleagues (2012) sought to verify the analytical performance of the Afirma GEC in the classification of cytologically indeterminate thyroid nodule FNAs. The analytical performance studies were designed to characterize the stability of the RNA in the aspirates during collection and shipment, analytical sensitivity and specificity, and assay performance studies including intra-nodule, intra-assay, inter-assay, and inter-laboratory reproducibility. Analytical sensitivity, specificity, robustness and quality control of the GEC were all verified in the study. Based on these results, the authors concluded that routine testing of FNA specimens is feasible.

Alexander and colleagues (2012) performed a 19-month, prospective, multicenter validation study of the Afirma GEC, involving 49 clinical sites (both academic and community centers), 3789 individuals and 4812 FNAs of thyroid nodules that were at least 1 cm in size. Histopathologic reports of the cytologic diagnosis were collected for all cases, and reports without a definitive benign or malignant diagnosis at the local site were reviewed by three expert cytopathologists, who reclassified them as atypical, follicular neoplasm or suspicious for a follicular neoplasm, or suspicious for malignancy. Corresponding histopathologic diagnoses from excised specimens were available (excisions were performed without knowledge of the results of the GEC). After inclusion criteria were met, 265 FNA samples considered to be cytologically indeterminate were tested with the GEC assay at Veracyte Laboratory. Of the 265 indeterminate samples, 85 were reported as malignant. The GEC correctly identified 78 of the 85 as suspicious (92% sensitivity; 95% CI, 84 to 97%), with a specificity of 52% (95% CI, 44 to 59%). NPVs ranged from 85% for "suspicious cytologic findings" to 95% for "atypia of undetermined clinical significance" with a PPV of 47%. Limitations of this study, as reported by the authors, included imperfect interobserver agreement possibly affecting the sensitivity or specificity of the classifier, since pathological assessment of benign versus malignant disease is not always absolute.

Duick and colleagues (2012) reported the impact of Afirma GEC test results on the decision to operate on thyroid nodules with indeterminate cytology. This retrospective, multicenter study included individuals who were 21 years or older, had one or more thyroid nodules 1 cm in diameter or greater by ultrasound, and had an indeterminate diagnosis by cytology and a GEC from the same nodule that was reported as benign. A total of 51 endocrinologists at 21 practice sites in 11 states contributed to the study. From September 2011 through March 2012, data were collected on 368 persons with 395 thyroid nodules. Surgery was primarily performed on those with indeterminate cytology and a benign GEC due to large or symptomatic nodules, rapidly growing nodules, or a second suspicious

or malignant nodule in the same individual, the same reasons typically given for an operation on cytologically benign nodules. The authors compared this surgical excision rate of the study population (7.6%) to a historical rate of surgical excision of 74% previously reported for cases with an indeterminate cytologic diagnosis without GEC testing.

Alexander and colleagues (2014) reported on a study of Afirma GEC testing performed at five academic medical centers between 2010 and 2013. Nodule and individual characteristics, Afirma GEC results, fine needle aspiration cytology, and clinical or surgical follow-up were obtained and results evaluated. A total of 339 persons underwent Afirma GEC testing of cytologically indeterminate nodules (165 AUS/FLUS; 161 FN; 13 SUS). A total of 174 of 339 (51%) indeterminate nodules were GEC benign and 148 (44%) GEC were suspicious. A total of 4 of 175 GEC benign were recommended for surgery as compared to 141 of 149 GEC suspicious (p=0.01). Of the 121 cytologically indeterminate/GEC suspicious nodules surgically removed, 53 (44%) were malignant. Site-to-site GEC variation was reported. The proportion of GEC benign varied up to 29%, while the malignancy rate of cytologically indeterminate/GEC suspicious nodules varied up to 47%. A total of 71 of 174 GEC benign nodules had documented clinical follow-up for an average of 8.5 months, in which 1 of 71 nodules was confirmed as being malignant. Study limitations included the retrospective nature of the data, incomplete outcome measures, and that only 41% of subjects with Afirma GEC benign results had documented assessment of their subsequent clinical status.

A single hospital-based thyroid surgical practice (Harrell, 2014) evaluated the performance of the Afirma GEC over 27 months from January 2011 through April 2013. A total of 645 FNAs were performed on 519 subjects during the study period. There were 58 FNAs (9%) reported as indeterminate, with 36 of these classified as suspicious by GEC (62%), 20 classified as GEC benign (34%), and 2 considered inadequate due to low mRNA content. Of the 36 subjects with suspicious GEC, 30 received a thyroidectomy, and 21 of the 30 had malignant final pathology. Of the 20 subjects with a benign GEC, 5 underwent thyroid surgery, and 2 had malignancies. The NPV for the Afirma GEC in this practice environment was reported to be 89.6%. The authors concluded that the NPV for the GEC diminished in populations where the indeterminate cytology group contained more than 1 in 4 subjects with thyroid cancer. The study was limited by a relatively small sample size.

In an independent study of the Afirma GEC performed at a large academic medical center, McIver and colleagues (2014) demonstrated lower positive predictive values of the GEC than had been previously reported. All subjects undergoing thyroid fine-needle aspiration during the study period, whose cytology was reported as FN or AUS/FLUS, were offered access to the test and recruited to this study (n=90). From those 90 subjects, 18 (20%) chose treatment with a diagnostic or therapeutic lobectomy, rather than use the GEC. A total of 72 samples were sent for GEC analysis; however, only 60 Afirma results were analyzed as 12 (17%) had insufficient mRNA. Of those, 16 (27%) were benign and 44 (73%) were suspicious. The rate of confirmed malignancy in GEC-suspicious nodules was only 17%. An NPV of benign GEC was 94% (assuming no additional false-negatives) and a PPV of suspicious GEC was 15.6%. The authors concluded that "additional confirmatory studies are necessary to assess the performance characteristics of the Afirma GEC before widespread adoption of this technology can be recommended."

A 3-year single surgical practice, single institution retrospective review of indeterminate fine needle aspiration cytology (FNAC) (Witt, 2015) was performed to evaluate the outcome of GEC in a clinical practice setting. All FNAs from February 2012 to February 2015 (n=520) were reviewed. Of these, 47 (9%) had indeterminate FNAC, and 32/47 (68%) agreed to Afirma GEC testing. Upon GEC testing, 14 cases were reported as benign, 15 suspicious, and 3 with no result due to inadequate RNA content. Two of the three subjects with no results went on to have diagnostic thyroid surgery, and had benign final pathology results. The 14 cases of indeterminate FNAC that had benign GEC results did not require diagnostic surgery. These subjects were observed for a mean and median duration of 14 and 7 months, respectively. Benign GEC testing had an estimated NPV of 100% during the study period of subjects completing the study. This study was limited by a relatively small sample size.

Angell and colleagues (2015) evaluated clinical outcomes and sonographic changes of consecutive cytologically indeterminate-GEC benign (Cyto-I/GEC-B) nodules and compared them to cytologically benign (Cyto-B) nodules. Data from adults (21 years of age or older) with a thyroid nodule at least 1 cm in dimension and a benign Afirma GEC result between November 2010 and April 2014 were retrospectively analyzed. Primary outcomes included nodule growth of 20% or greater in two dimensions or of 50% or greater in volume, change in sonographic features, and rates of repeat fine-needle aspiration, thyroidectomy, and malignancy. A total of 95 Cyto-I/GEC-B nodules (90 subjects) were studied. Of these, 5 individuals received primary surgical resection. From the 90 remaining nodules, 58 (64.4%) had ultrasound follow-up results available at a median of 13 months (range 4-40 months). The comparison group had 1678 Cyto-B nodules evaluated from 2001-2010. A total of 1379 had repeat US assessment from 4-40 months post initial cytologic evaluation. After exclusion of 155 Cyto-B nodules with a greater than 50% cystic component, the final comparison group cohort consisted of 1224 nodules in 873 subjects. The Cyto-I/GEC-B nodules showed similar growth compared to 1224 Cyto-B nodules using either of the following criteria: 20% or greater in two dimensions (8.6% vs. 8.3%, p=0.80) or 50% or greater in volume (17.2% vs. 13.8%, p=0.44). Thyroidectomies were performed more frequently in the Cyto-I/GEC-B group (13.8% vs. 0.9%, p<0.0001), however, cancer was only found in 1 person. The authors concluded that their "data provided a strong indication that treating Cyto-I/GEC-B nodules similarly to those with benign cytology is clinically appropriate." This study was limited by retrospective analysis, and a relatively short and highly variable follow-up period.

In 2016, Santhanam and colleagues performed a meta-analysis that combined individual data from seven studies that examined the GEC test for indeterminate thyroid nodules. The pooled sensitivity of the GEC was reported to be 95% and the specificity was 30.5%. Individuals with benign GEC results were not followed long enough to determine reliable false negative rates. The authors indicated that GEC is useful to rule out malignancy in thyroid nodules that have indeterminate cytology although the long-term benefits are unclear. Specifically, while the GEC "might prevent some unnecessary thyroid surgeries…for many persons it might represent an additional layer of testing prior to diagnostic thyroidectomy." A significant limitation of the meta-analysis was the heterogeneity of the studies.

A 2016 population-based, retrospective cohort study by Singer and colleagues assessed long-term management patterns and rates of thyroid surgery for individuals with benign GEC results as compared to a control group with cytopathology benign results. Individuals who underwent FNA biopsy between January 1, 2011 and July 31, 2013 were included in the study. Study outcomes included rates of thyroid-related follow-up clinic visits, ultrasound examinations, and thyroid-related surgeries. A total of 159 of the 2059 participants who met study inclusion criteria

had their thyroid nodule FNA biopsy diagnosed as indeterminate, and only a molecular sample was sent for GEC testing. Remaining were 1900 participants whose paired cytopathology and GEC samples were sent for evaluation. Nodule cytology pathology results for this group were: non-diagnostic for 157 subjects (8.3%), benign for 1308 subjects (68.8%), indeterminate for 357 subjects (18.8%), suspicious of malignancy in 19 subjects (1.0%), and malignant for 59 subjects (3.1%). Altogether, 532 subjects had samples that underwent GEC testing and 35 of those molecular samples did not produce a result. Of the remaining 497, results were GEC benign for 218 (43.9%) subjects, and GEC suspicious for 279 (56.1%) subjects. Of the 218 GEC-benign subjects, 201 were successfully matched 1:3 to 603 subjects with a cytopathology benign diagnosis. The number of GEC-benign and cytopathology-benign subjects that underwent thyroid surgery (11.4% versus 10.1%, p=0.594), and received a follow-up ultrasound exam (60.2% versus 61.7%, p=0.706), respectively, were not significantly different. The majority of subjects in both groups did not require surgery and were managed with routine care consisting of ultrasounds and clinical follow-up. The authors concluded their data suggested a GEC diagnosis of "benign" could be clinically managed in the manner as an initially cytopathologic "benign" thyroid nodule.

A retrospective analysis by Yang and colleagues (2016) was completed at a single institution that performed the Afirma GEC test between August 24, 2012 (when GEC testing began in the institution) and April 1, 2014. Cases of indeterminate cytology that also had GEC testing were selected and GEC results were compared to the histopathology findings. A total of 1693 thyroid FNAs were performed and of these, 789 (46.7%) had GEC samples collected for potential testing and 217 (13% of the total number of FNAs, rate of indeterminate cytology results in the study institution) had GEC completed. Among the 217 cases, 189 were of indeterminate cytology. Of the indeterminate cytology cases, 42% were benign and 50% were suspicious by GEC. The rate of excision of atypia of undetermined significance-follicular lesion of undetermined significance in the pre-GEC category was 63% with the rate decreasing to 47% in the post-GEC category. The malignancy rate of excised thyroids increased from 35% in the pre-GEC group to 47% in the post-GEC group. Findings were similar for lesions suspicious for a follicular neoplasm-follicular neoplasm lesion. The authors concluded that GEC testing contributed to an avoidance of surgery for initially indeterminate thyroid nodules.

Sipos and colleagues (2016) conducted a retrospective study of non-academic medical practices utilizing the GEC between September 2010 and June 2014.The primary study objective was to evaluate the rate of surgical intervention in subjects with a benign Afirma GEC result during long-term follow-up. The secondary study objective was to determine the opinion of treating physicians regarding the safety of GEC use compared to the hypothetical situation of providing care without the GEC. During the 36-month follow-up period, 17 of 98 subjects (17.3%) with a benign GEC result had surgery. After a benign GEC, 88% of surgeries occurred during the first 2 years. Additionally, a survey was administered to treating physicians to assess their perception of safety in using the GEC. Reports from treating physicians indicated that patient safety was improved by using the GEC compared to not using the GEC in 78 of 91 (86%) cases.

Abeykoon and colleagues (2016) conducted a retrospective cohort study at a single institution comparing the rate ofsurgical recommendations for all cytologically indeterminate thyroid nodules pre-and post-Afirma introduction and found a statistically significant reduction from 81.5% to 50% (p=0.01) after Afirma GEC implementation. Of the individuals who underwent surgery, 85.7% in the post-Afirma cohort showed evidence of malignancy, as compared with 20% of those in the pre-Afirma cohort (p<0.01).

Several additional recently published case series and retrospective reviews also indicate that Afirma GEC testing may contribute to avoidance of surgery for initially indeterminate thyroid nodules (Chaudhary, 2016; Dhingra, 2016; Jug, 2019; Parajuli, 2019).

A more recent systematic review and meta-analysis sought to compare post-marketing findings to the initial clinical validation findings of GEC testing (Valderrabano, 2019). A total of 19 studies were included, comprising a total of 2568 thyroid nodules. Based on a simulation using the sensitivity and specificity reported in the initial validation study, the observed benign call rate and PPV values in post-marketing studies would have to be explained by enormously different underlying prevalence rates of cancer (15% vs 30%); an impossibility. The findings suggest that the initial validation study cohort was not representative of the populations in whom the GEC has been used, raising doubt regarding the accuracy of its reported diagnostic performance, including the negative predictive value.

In 2017, a second-generation assay (Afirma Genetic Sequence Classifier [GSC]) was made available for clinical use. Clinical outcomes after 11 months of use with the GSC (n=146) were compared with 6.5 years of experience with the GEC (n=509). When compared to GEC, GSC identified less indeterminate cytology nodules as suspicious (38.8%; 54/139 vs. 58.4%; 281/481, respectively). A total of 82.7% of oncocytic FNAB subjects were classified as suspicious in the GEC group (86 of 104 oncocytic indeterminates) whereas only 35.3% were classified as suspicious by GSC (12 of 34). There was a 45% reduction in the rate of subsequent surgery in individuals with oncocytic aspirates (56% in the GEC group vs. 31% in the GSC group). Pathology analysis demonstrated a false-negative percentage for an incomplete surgical group of 9.5% for GEC and 1.2% for GSC. Authors conclude that the GSC further reduces surgery in indeterminate thyroid nodules by improving the specificity without compromising sensitivity, especially considering the significant improvement in the specificity of the GSC test in oncocytic FNAB aspirate (Harrell, 2019).

ThyGenX and ThyraMIR

ThyGenX (based on the predicate 17-gene alteration panel miRInform test) is a mutational panel for the detection of 8 genes associated with thyroid papillary carcinoma and follicular carcinoma. ThyraMIR is a micro RNA (miRNA) gene expression classifier that is based on the evaluation and expression of 10 miRNAs. The tests are marketed to use in combination, in that ThyraMIR can identify malignancy when ThyGenX has a negative result. ThyGenX and ThyraMIR are designed to minimize the need for surgery when nodules are indeterminate or to assist in making surgical decisions when nodules are malignant (lobectomy versus total thyroidectomy).

Beaudenon-Huibregtse and colleagues (2014) performed a prospective, double-blind study on the performance of FNA cytology combined with the molecular analysis of a 17-gene alteration panel. The researchers collected 806 FNA specimens from 618 subjects at 5 U.S. clinical sites. A total of 737 nodules from 581 subjects met inclusion criteria, and at the end of the study there were 109 specimens that had post-surgical histopathology. For those 109 specimens, oncogenic mutations were present in 50% of malignant nodules missed by FNA cytology. A total of 14 nodules that were indeterminate after FNA were negative by molecular testing but positive after surgical resection and histopathology (false-negative detection rate of 25%). There were 6 false-positive molecular test results. The researchers concluded that molecular testing compliments cytopathological testing, but “not all malignant tumors

carry one of the 17 genetic alterations evaluated, and it is important to emphasize that, unless the local pretest probability of cancer is known and sufficiently low, a negative molecular result alone should not be used to rule out surgical therapy for nodules with an indeterminate or nondiagnostic cytologic diagnosis."

Labourier and colleagues (2015) evaluated a diagnostic algorithm that combined mutation detection and miRNA expression. A total of 638 surgical specimens and preoperative FNAs were tested for 17 validated gene alterations using the miRInform Thyroid test and a 10-miRNA gene expression classifier that generates positive (malignant) or negative (benign) results. A cross-sectional sampling of 109 thyroid nodules with AUS/FLUS or FN/SFN cytology was performed at 12 endocrinology centers. Mutations were detected in 24 of 35 (69%) nodules with malignant outcome. Of the mutation-negative specimens, miRNA testing identified 64% of malignant cases and 98% of benign cases. The diagnostic sensitivity and specificity of the combined algorithm was 89% and 85%, respectively. At 32% cancer prevalence, 61% of the molecular results were benign with an NPV of 94%. The authors concluded that the diagnostic algorithm "combining miRNA expression and gene mutation detection yields clinically actionable molecular information in thyroid nodules with AUS/FLUS or FN/SFN cytology."

Wylie and colleagues (2016) performed an analysis on the diagnostic potential of miRNA for the development of ThyraMIR. The researchers analyzed 534 archived nodule remnants from FNAs, including 257 nodules that underwent surgical resection. The samples, which were from 14 centers across the United States, were histologically reviewed by a single pathologist and classified according to current World Health Organization schemes. Expression profiling was done by reverse transcription-quantitative polymerase chain reaction (PCR). The researchers extracted total nucleic acids for 17 gene alterations (BRAF, RAS, RET, or PAX8), selected 31 miRNA candidates based on literature review and differential expression analyses, and genotyped over 1500 unique gene alterations using a custom sequencing assay. Eight supervised machine learning algorithms were used to distinguish benign and malignant samples. The AUROC was invariant in cross-validation (0.89) and was optimal for 235 preoperative aspirates (0.94). The models also classified 92% of benign lesions as low risk/negative and 92% of malignant lesions as high risk/positive. miRNA significantly increased the diagnostic performance of the 17-mutation panel ($p<0.001$). For a subset of resected tissue samples (n=54) and an independent indeterminate set of nodules (n=42), miRNA increased sensitivity by 30-39% and was able to classify all the benign nodules as negative. When the researchers compared miRNA and NGS testing, they found that both methods increased sensitivity of the 17-mutation panel; however, NGS decreased the specificity. For 30 nodules reported benign by the 17-mutation panel, miRNA was able to correctly classify 67% of malignant cases and 100% of benign cases. The combination of miRNA expression combined with the 17-gene panel resulted in an 85% sensitivity and 95% specificity. The authors concluded that "a molecular test combining an optimized miRNA classification algorithm with a validated panel of somatic driver mutations displays high diagnostic sensitivity and specificity."

Rosetta GX

The Rosetta GX is a diagnostic assay designed to classify indeterminate thyroid smears as benign or suspicious for malignancy. The assay uses routinely prepared FNA cytology smears and measures a set of miRNAs by quantitative RT-PCR to classify a nodule. Also measured is a miRNA specific to medullary carcinoma.

The analytical validity (Benjamin, 2016) and clinical validity (Lithwick, 2017) of Rosetta GX was evaluated. The RosettaGx reportedly distinguishes benign from malignant thyroid nodules using a single FNA stained smear using quantitative RT-PCR. Initial study results appear promising with a 99% NPV in indeterminate nodules where all three reviewing pathologists were in agreement regarding final diagnosis and 91% for the entire validation set (Lithwick, 2016). However, Lithwick concluded, "Additional cohorts, both academic and nonacademic, could help to further validate the performance of the assay."

Other Considerations

The American Thyroid Association (ATA) *Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer* (Haugen, 2015) indicate that there is currently no single optimal molecular test that can definitively rule in or rule out malignancy in all cases of indeterminate cytology. However, the ATA did offer the following specific recommendations for molecular testing in adults:

- 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 recommendation, Moderate-quality evidence)
- 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 recommendation, Low-quality evidence).

The corresponding ATA pediatric guidelines (2015) indicate that although studies have shown molecular testing aids in the management of thyroid nodules with indeterminate cytopathology in adults, there are no studies determining its usefulness in the evaluation of indeterminate pediatric thyroid nodules. As a result, the ATA was unable to offer a recommendation for the use of molecular diagnostics in children, and reported that additional studies are required prior to a recommendation being made. The ATA also indicated that their pediatric guidelines applied to ages18 years and below.

The ATA (Ferris, 2015) also published a *Statement on Surgical Application of Molecular Profiling for Thyroid Nodules: Current Impact on Perioperative Decision Making* and reported:

> Techniques for molecular profiling of thyroid cytology specimens have evolved as adjuncts to guide the appropriate management of cytologically indeterminate nodules. However, it must be stressed that the utility of any molecular test is only applicable clinically when combined with clinical and sonographic risk factors for malignancy and with understanding of the prevalence of malignancy for the Bethesda cytologic categories at the reporting institution. For example, a ''rule out'' test such as the GEC will perform better in a setting of lower cancer frequency, as well as in a cytologic category of low cancer frequency such as AUS/FLUS or FN, than it will in a setting of higher cancer frequency such as SMC or a site with a high prevalence of malignancy in a given cytologic category.

In 2017, the ATA released a recommendation (Haugen, 2017) on the proposed renaming of non-invasive EFVPTC to NIFTP:

> The histopathologic nomenclature for Encapsulated Follicular Variant Papillary Thyroid Carcinoma (eFVPTC) without invasion may be re-classified as a NIFTP given the excellent prognosis of this neoplastic variant. Prospective studies are needed to validate the observed patient outcomes (and test performance in predicting thyroid cancer outcomes), as well as implications on patients' psychosocial health and economics (Weak Recommendation, Moderate-quality evidence).

The ATA further states:

> The proposed name change will also affect the performance of molecular tests when applied to patients with indeterminate cytology. For example, neoplasms harboring RAS mutations, will likely have a lower positive predictive value for malignancy, while nodules with no genetic mutation or a negative gene expression classifier will likely have a slightly higher negative predictive value. These effects will be dependent on the prevalence of NIFTP in a given population. Since NIFTP, like follicular adenoma, requires surgery for a definitive diagnosis, the changes in PPV and NPV of the molecular tests will not alter the requirement of surgical intervention for these patients. Potential frameworks for addressing cytology and molecular reporting are still evolving.

The *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* (Gharib, 2016) states:

- Because of the insufficient evidence and the limited follow-up, we do not recommend either in favor of or against the use of GECs for cytologically indeterminate nodules [BEL 2, GRADE B].
- Currently, with the exception of mutations such as BRAFV600E that has a PPV approaching 100% for PTC, evidence is insufficient to recommend in favor of or against the use of mutation testing as a guide for the extent of surgery [BEL 2, GRADE A].
- 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 [BEL 3, GRADE B].

National Comprehensive Cancer Network (NCCN) Guidelines on the treatment of thyroid cancer (V2.2019) report that molecular diagnostics to detect individual mutations in BRAF, V600E, RET/PTC, RAS, PAX8/PPAR or pattern recognition approaches using molecular classifiers may be useful in the evaluation of FNA samples that are indeterminate (that is, follicular lesion of undetermined significance). NCCN issued 2A recommendations that molecular diagnostics may be used for reclassification of follicular lesions (FN, AUS, and FLUS) as either more or

less likely to be benign or malignant based on the genetic profile. NCCN stated that historically studies have not shown that molecular diagnostics perform well for Hürthle cell neoplasms. NCCN discusses the encouraging evidence surrounding Afirma and ThyroSeq but does not make mention of ThyGenX, ThyraMIR or Rosetta for molecular testing of cytologically indeterminate thyroid nodules.

The strength of the cytology report impacts the reliability of molecular testing. Cibas and colleagues (2013) assessed inter- and intraobserver variability of preoperative cytopathologic and postoperative histopathologic thyroid diagnoses. A total of 653 subjects with 776 surgically resected thyroid nodules of 1 cm or greater were evaluated at 14 academic centers and 35 clinical sites. Study samples were collected in a prospective, multicenter trial validating a GEC between June 2009 and December 2010. Intraobserver concordance among two or more central histopathologists who independently read histopathology slides was calculated. Interobserver concordance between the diagnoses made by the central histopathologists and those made by local pathologists was also calculated. Intra- and interobserver concordance for cytopathology was similarly calculated by comparing diagnoses made by local pathologists with those made by a central panel of three cytopathologists. Concordance on the histopathologic distinction between benign and malignant diagnoses was 91% comparing local with central histopathologists and 90% comparing two central histopathologists. Sixty four percent of diagnoses made by central and local cytopathologists and 74.7% of intraobserver diagnoses were concordant using the six category Bethesda System for reporting thyroid cytopathology. Central cytopathologists made fewer indeterminate diagnoses than local pathologists (41.2% vs. 55.0%). A significant limitation of this study was that many local pathologists did not use the Bethesda System, and their reports were translated to allow for comparison.

Conclusions

The incremental added value of mutation analysis to an equivocal FNA result is not known, and although mutation 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, at this time, it is not clear how it will impact clinical management of an individual or surgical decision making.

Published data provides sufficient evidence demonstrating that the Afirma GEC and the ThyroSeq genomic classifier contribute to the avoidance of surgery for those with indeterminate initial cytopathology of a thyroid nodule. As such, an improvement in net health outcome may occur as a result of the Afirma GEC and the ThyroSeq genomic classifier.

The commercially available, laboratory-developed molecular marker tests are regulated under the Clinical Laboratory Improvement Amendments (CLIA). At this time, premarket approval from the U.S. Food and Drug Administration (FDA) is not enforced when the assay is performed in a laboratory that is licensed by CLIA.

Definitions

Bethesda System for Reporting Thyroid Cytopathology (TBSRTC): A standardized reporting system used by pathologists to classify FNA specimens.

BRAF: A protein which influences the regulation of the MAP kinase/ERKs signaling pathway, which affects cell division, differentiation, and secretion. BRAF is also known as serine/threonine-protein kinase B-Raf, v-raf murine sarcoma viral oncogene homolog B1.

mRNA: Messenger RNA conveys genetic information from DNA to the ribosome, where they specify the amino acid sequence of the protein products of gene expression.

miRNA: Micro RNA is a small non-coding RNA molecule containing about 22 nucleotides found in plants, animals and some viruses, that functions in RNA silencing and post-transcriptional regulation of gene expression.

Mutation: A permanent, transmissible change in genetic material.

Next-generation sequencing (NGS): Any of the technologies that allow rapid sequencing of large numbers of segments of DNA, up to and including entire genomes.

References

Peer Reviewed Publications:

1. 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. 2016; 7(4):272-278.
2. 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.
3. 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.
4. Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma gene expression classifier. J Clin Endocrinol Metab. 2014; 99(1):119-125.
5. Ali SZ, Fish SA, Lanman R, et al. Use of the Afirma® gene expression classifier for preoperative identification of benign thyroid nodules with indeterminate fine needle aspiration cytopathology. PLoS Curr. 2013; 5.
6. Angell TE, Frates MC, Medici M, et al. Afirma benign thyroid nodules show similar growth to cytologically benign nodules during follow-up. J Clin Endocrinol Metab. 2015; 100(11):E1477-E1483.
7. Beaudenon-Huibregtse S, Alexander EK, Guttler RB, et al. Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. Thyroid. 2014; (10):1479-87.
8. Benjamin H, Schnitzer-Perlman T, Shtabsky A, et al. Analytical validity of a microRNA-based assay for diagnosing indeterminate thyroid FNA smears from routinely prepared cytology slides. Cancer Cytopathol. 2016; 124(10):711-721.
9. Borowczyk M, Szczepanek-Parulska E, Olejarz M, et al. Evaluation of 167 Gene Expression Classifier (GEC) and ThyroSeq v2 diagnostic accuracy in the preoperative assessment of indeterminate thyroid nodules: bivariate/HROC meta-analysis. Endocr Pathol. 2019; 30(1):8-15.
10. Cantara S, Capezzone M, Marchisotta S, et al. Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. J Clin Endocrinol Metab. 2010; 95(3):1365-1369.

11. 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.
12. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. J Clin Endocrin Metab. 2010; 95(12):5296-5304.
13. Cibas ES, Baloch ZW, Fellegara G, et al. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. Ann Intern Med. 2013; 159(5):325-332.
14. Dhingra JK. Office-based ultrasound-guided FNA with molecular testing for thyroid nodules. Otolaryngol Head Neck Surg. 2016; 155(4):564-567.
15. Duick DS, Klopper JP, Diggans JC et al. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. Thyroid. 2012; 22(10):996-1001.
16. 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.
17. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. Endocr Pract. 2014; 20(4):364-369.
18. Harrell RM, Eyerly-Webb SA, Golding AC, et al. Statistical comparison of Afirma GSC and AFIRMA GSC outocomes in a community endocrine surgical practice: early findings. Endocr Pract. 2019; 25(2):161-164.
19. Jug R, Parajuli S, Ahmadi S, Jiang XS. Negative results on thyroid molecular testing decrease rates of surgery for indeterminate thyroid nodules. Endocr Pathol. 2019; 30(2):134-137.
20. Labourier E, Shifrin A, Busseniers AE, et al. Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. J Clin Endocrinol Metab. 2015; 100(7):2743-2750.
21. 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. J Clin Pathol. 2017; 70(6):500-507.
22. Marcadis AR, Valderrabano P, Ho AS, et al. Interinstitutional variation in predictive value of the ThyroSeq v2 genomic classifier for cytologically indeterminate thyroid nodules. Surgery. 2019;165(1):17-24.
23. Mathur A, Weng J, Moses W, et al. A prospective study evaluating the accuracy of using combined clinical factors and candidate diagnostic markers to refine the accuracy of thyroid fine needle aspiration biopsy. Surgery. 2010; 148(6):1170-1177.
24. McIver B, Castro MR, Morris JC, et al. An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. J Clin Endocrinol Metab. 2014; 99(11):4069-4077.
25. Moses W, Weng J, Sansano I, et al. Molecular testing for somatic mutations improves accuracy of thyroid fine-needle aspiration biopsy. World J Surg. 2010; 34(11):2589-2594.
26. Nikiforov YE. Molecular diagnostics of thyroid tumors. Arch Pathol Lab Med. 2011; 135(5):569-577.
27. 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.
28. 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.

29. Nikiforov YE, Seethala RR, Tallini G, et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. JAMA Oncol. 2016; 2(8):1023-1029.
30. Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. J Clin Endocrinol Metab. 2009; 94(6):2092-2098.
31. Nikiforova MN, Mercurio S, Walk AI, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. Cancer. 2018; 124(8):1682-1690.
32. 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. 2013; 98(11):E1852-1860.
33. Ohori NP, Landau MS, Carty SE, et al. Benign call rate and molecular test result distribution of ThyroSeq v3. Cancer Cytopathol. 2019; 127(3):161-168.
34. 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. 2010; 118(1):17-23.
35. Parajuli S, Jug R, Ahmadi S, Jiang XS. Hurthle cell predominance impacts results of Afirma gene expression classifier and ThyroSeq molecular panel performance in indeterminate thyroid nodules. Diagn Cytopathol. 2019; 47(11):1177-1183.
36. Santhanam P, Khthir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Med Oncol. 2016; 33(2):14.
37. 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.
38. Singer J, Hanna JW, Visaria J, Gu T, et al. Impact of a gene expression classifier on the long-term management of patients with cytologically indeterminate thyroid nodules. Curr Med Res Opin. 2016; 32(7):1225-1232.
39. Sipos JA, Blevins TC, Shea HC, et al. Long-term nonoperative rate of thyroid nodules with benign results on the Afirma gene expression classifier. Endocr Pract. 2016; 22(6):666-672.
40. 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. 2019; 5(2):204-212.
41. 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. 2018; 163(1):97-103.
42. Valderrabano P, Khazai L, Leon ME, et al. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. Endocr Relat Cancer. 2017; 24(3):127-136.
43. Valderrabano P, Hallanger-Johnson JE, 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. 2019 Jul 18. [Epub ahead of print].
44. Vargas-Salas, Martínez JR, Urra S, et al. Genetic testing for indeterminate thyroid cytology: review and meta-analysis. Endocr Relat Cancer. 2018; 25(3):R163-R177.
45. 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.
46. Ward LS, Kloos RT. Molecular markers in the diagnosis of thyroid nodules. Arq Bras Endocrinol Metabol. 2013; 57(2):89-97.
47. Witt RL. Outcome of thyroid gene expression classifier testing in clinical practice. Laryngoscope. 2016; 126(2):524-527.

48. Wylie D, Beaudenon-Huibregtse S, Haynes BC, et al. Molecular classification of thyroid lesions by combined testing for miRNA gene expression and somatic gene alterations. J Pathol Clin Res. 2016; 2(2):93-103.
49. 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.
50. Yang SE, Sullivan PS, Zhang J, et al. Has Afirma gene expression classifier testing refined the indeterminate thyroid category in cytology? Cancer Cytopathol. 2016; 124(2):100-109.

Government Agency, Medical Society, and Other Authoritative Publications:

1. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. Thyroid. 2017; 27(11):1341-1346.
2. Ferris RL, Baloch Z, Bernet V, et al. American Thyroid Association Surgical Affairs Committee. American Thyroid Association statement on surgical application of molecular profiling for thyroid nodules: current impact on perioperative decision making. Thyroid. 2015; 25(7):760-768.
3. Francis GL, Waguespack SG, Bauer AJ, et al; American Thyroid Association Guidelines Task Force. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. Thyroid. 2015; 25(7):716-759.
4. 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.
5. 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.
6. Haugen BR, Sawka AM, Alexander EK, et al. American Thyroid Association guidelines on the management of thyroid nodules and differentiated thyroid cancer task force review and recommendation on the proposed renaming of encapsulated follicular variant papillary thyroid carcinoma without invasion to noninvasive follicular thyroid neoplasm with papillary-like nuclear features. Thyroid. 2017; 27(4):481-483.
7. NCCN Clinical Practice Guidelines in Oncology®. © 2019 National Comprehensive Cancer Network, Inc. Thyroid cancer (V.2.2019). For additional information visit the NCCN website: http://www.nccn.org/index.asp. Accessed on January 12, 2020.

Websites for Additional Information

1. National Cancer Institute. Thyroid Cancer. Available at: https://www.cancer.gov/types/thyroid/hp. Accessed on January 12, 2020.

Index

ThyroSeq

The use of specific product names is illustrative only. It is not intended to be a recommendation of one product over another, and is not intended to represent a complete listing of all products available.

History

Status	Date	Action
Revised	02/20/2020	Medical Policy & Technology Assessment Committee (MPTAC) review. Revised MN statement to reflect updated second generation Affirma test, Genomic Sequencing Classifier (GSC). Updated Description/Scope, Background/Overview, References and Websites sections.
Revised	03/21/2019	MPTAC review.
Revised	03/20/2019	Hematology/Oncology Subcommittee review. Added ThyroSeq to the MN statement. Coding, References and Websites sections updated.
Revised	07/26/2018	MPTAC review.
Revised	07/18/2018	Hematology/Oncology Subcommittee review. Updated acronym in Clinical Indications section. Description, Discussion/General Information, and References sections updated. Added Websites for Additional Information section.
New	11/02/2017	MPTAC review.
New	11/01/2017	Hematology/Oncology Subcommittee review. Initial document development. Moved content of GENE.00032 Molecular Marker Evaluation of Thyroid Nodules to new clinical utilization management guideline document with the same title. Addition of Afirma, ThyGenX, ThyraMIR, Thryoseq, and Rosetta to Clinical Indications section. Updated Coding section with 01/01/2018 CPT changes; added code 0026U.

Appendix

Bethesda System for Reporting Thyroid Cytopathology: Recommended Diagnostic Categories

I. NONDIAGNOSTIC OR UNSATISFACTORY
- Cyst fluid only
- Virtually acellular specimen
- Other (obscuring blood, clotting artifact, etc.)

II. BENIGN
- Consistent with a benign follicular nodule (includes adenomatoid nodule, colloid nodule, etc.)
- Consistent with lymphocytic (Hashimoto) thyroiditis in the proper clinical context
- Consistent with granulomatous (subacute) thyroiditis
- Other

III. ATYPIA OF UNDETERMINED SIGNIFICANCE or FOLLICULAR LESION OF UNDETERMINED SIGNIFICANCE
IV. FOLLICULAR NEOPLASM or SUSPICIOUS FOR A FOLLICULAR NEOPLASM
- Specify if Hürthle cell (oncocytic) type

V. SUSPICIOUS FOR MALIGNANCY
- Suspicious for papillary carcinoma
- Suspicious for medullary carcinoma
- Suspicious for metastatic carcinoma
- Suspicious for lymphoma
- Other

VI. MALIGNANT
- Papillary thyroid carcinoma
- Poorly differentiated carcinoma
- Medullary thyroid carcinoma
- Undifferentiated (anaplastic) carcinoma
- Squamous-cell carcinoma
- Carcinoma with mixed features (specify)
- Metastatic carcinoma
- Non-Hodgkin lymphoma
- Other

(Cibas and Ali, 2017)