Molecular Markers in Fine Needle Aspirates of the Thyroid

Description

Cytologic examination of fine needle aspiration (FNA) samples from a thyroid lesion to identify which patients need to undergo surgery has diagnostic limitations. Assays using molecular markers have been developed in an attempt to improve the accuracy of thyroid FNA biopsies.

Background

Fine needle aspiration (FNA) of the thyroid

Thyroid nodules are common, present in 5% to 7% of the U.S. adult population. Most are benign and most cases of thyroid cancer are curable by surgery if detected early. FNA of the thyroid is currently the most accurate procedure to distinguish benign thyroid lesions and malignant ones, reducing the rate of unnecessary thyroid surgery for patients with benign nodules and triaging patients with thyroid cancer to appropriate surgery.

About 60% to 70% of thyroid nodules are classified cytologically as benign, and 4% to 10% of nodules are cytologically deemed malignant. (1) However, the remaining 20% to 30% have equivocal findings (inclusive, indeterminate, atypical or suspicious), usually due to overlapping cytologic features between benign and malignant nodules; these nodules usually require surgery for a final diagnosis. Guidelines from the American Thyroid Association recommend repeat FNA for patients with a diagnosis of “atypia of undetermined significance” and lobectomy with or without intraoperative pathology consultation for those with a suspicious diagnosis. (2)

Approximately 80% of patients with indeterminate cytology undergo surgical resection; postoperative evaluation reveals a malignancy rate ranging from 6% to 30%, making this clinical process one with very low specificity. (3)

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

**Thyroid cancer**

Most thyroid cancers originate from thyroid follicular cells and include well-differentiated papillary carcinoma (PTC) (80% of all thyroid cancers) and follicular carcinoma (15%). Poorly differentiated and anaplastic thyroid carcinomas are uncommon and can arise de novo or from preexisting well-differentiated papillary or follicular carcinomas. Medullary thyroid carcinoma originates from parafollicular or C cells and accounts for ~3% of all thyroid cancers.

The diagnosis of malignancy in the case of PTC is primarily based on cytologic features. If a fine-needle aspiration in a case of PTC is indeterminate, intraoperative consultation is most often diagnostic, although its efficacy and therefore use will vary between institutions, surgeons, and pathologists.

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

New approaches for improving the diagnostic accuracy of thyroid FNA include mutation analysis for somatic genetic alterations, in order to more accurately classify which patients need to proceed to surgery (and may include the extent of surgery necessary) versus those patients who do not need surgery and can be safely followed.

**Mutations associated with thyroid cancer**

Various mutations have been discovered in thyroid cancer. The 4 gene mutations that are the most common and carry the highest impact on tumor diagnosis and prognosis and *BRAF* and *RAS* point mutations and *RET/PTC* and *PAX8/PPARγ* rearrangements.

Papillary carcinomas carry point mutations of the *BRAF* and *RAS* genes as well as *RET/PTC* and *TRK* rearrangements, all of which are able to activate the mitogen-activated protein kinase (MAPK) pathway. (4) These mutually exclusive mutations are found in more than 70% of papillary carcinomas. (4) *BRAF* mutations are highly specific for PTC. Follicular carcinomas harbor either *RAS* mutations or *PAX8/PPARγ* rearrangement. These mutations are also mutually exclusive and identified in 70-75% of follicular carcinomas. (4) Genetic alterations involving the *PI3K/AKT* signaling pathway also occur in thyroid tumors, although they are rare in well-differentiated thyroid cancer and have higher prevalence in less differentiated thyroid carcinomas. (4) Additional mutations known to occur in poorly differentiated and anaplastic carcinomas involve the *TP53* and *CTNNB1* genes. Medullary carcinomas, which can be familial or sporadic, frequently possess point mutations located in the *RET* gene.
Available Molecular Diagnostic Testing

Mutation Testing

Point mutations in specific genes, including BRAF, RAS, and RET, and evaluation for rearrangements associated with thyroid cancers can be accomplished by gene sequencing with Sanger sequencing or pyrosequencing or by real-time polymerase chain reaction (rtPCR). Panels of tests for mutations associated with thyroid cancer are also available. For example, Quest Diagnostics offers a Thyroid Cancer Mutation Panel, which includes BRAF and RAS mutation analysis and testing for RET/PTC and PAX8/PPARγ rearrangements.

In addition to standard Sanger sequencing or rtPCR-based mutation testing for genes associated with thyroid cancer, next-generation sequencing (NGS) panels that simultaneously evaluate for point mutations or gene fusions in multiple genes have been developed. For example, the ThyroSeq® v.2 Next Generation Sequencing panel (CBLPath, Ocala, FL) includes sequencing of more than 60 genes. According to the ThyroSeq's manufacturer's website, the test is indicated when FNA cytology indicates atypia of uncertain significance or follicular lesion of undetermined significance, follicular neoplasm or suspicious for follicular neoplasm, or suspicious for malignancy. (5) In particular, it has been evaluated in patients with follicular neoplasm/suspicious for follicular neoplasm on FNA as a test to increase both sensitivity and specificity for cancer diagnosis.

The ThyGenX™ Thyroid Oncogene Panel (formerly miRInform® Thyroid; Interpace Diagnostics, Parsippany, NJ; testing done at Asuragen Clinical Laboratory) is another NGS sequencing panel designed to be used in patients with indeterminate thyroid FNA results. It includes sequencing of 8 genes associated with papillary thyroid carcinoma and follicular carcinomas.

Gene Expression Profiling

Genetic alterations associated with thyroid cancer can be assessed through the use of gene expression profiling, which refers to analysis of messenger RNA (mRNA) expression levels of many genes simultaneously. Several gene expression profiling tests are now available to biologically stratify tissue from thyroid nodules. The Afirma® Gene Expression Classifier (Afirma GEC; Veracyte, South San Francisco, CA) analyzes the expression of 142 different genes to determine patterns associated with benign findings on surgical biopsy. It is designed to be used for thyroid nodules that have an "indeterminate" classification on FNA as a method to select patients who are at low risk for cancer ("rule out").

Veracyte also markets 2 “malignancy classifiers” that use mRNA expression-based classification to evaluate for BRAF mutations or mutations associated with medullary thyroid carcinoma (Afirma BRAF and Afirma MTC, respectively). In a description of the generation of the Afirma BRAF test, the authors outline the following proposed benefits of the mRNA-based expression test for BRAF mutations: 1) PCR based methods may have low sensitivity, requiring that a large proportion of the nodule have a relevant mutation; 2) testing for only 1 mutation may not detect patients with low-frequency mutations that result in the same pattern of pathway activation; and 3) PCR-based approaches with high analytic sensitivity may require a large of amount of DNA that is difficult to isolate from small FNA samples. (6) The Afirma MTC is an option when the Afirma GEC is ordered for thyroid nodules with an “intermediate” classification on FNA, and can also be used for thyroid nodules with “malignant” or “suspicious” results
on Afirma GEC. The Afirma BRAF is designed to be used for nodules with “suspicious” results on Afirma GEC.

**Regulatory Status**

Testing for mutations associated with thyroid cancer via sequencing or rtPCR are laboratory-developed tests (LDTs). Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; LDTs must meet the general regulatory standards of the Clinical Improvement Act (CLIA). Laboratories that offer LDTs must be licensed by CLIA for high-complexity testing.

In 2013, the U.S. Food and Drug Administration (FDA) approved through the premarket approval process the THxID™-BRAF kit, which is an in vitro diagnostic device to assess specific BRAF mutations in melanoma tissue via rtPCR. However, there are currently no diagnostic tests for thyroid cancer mutation analysis with approval from FDA.

**Related Policies**

None

**Policy**

*This policy statement applies to clinical review performed for pre-service (Prior Approval, Precertification, Advanced Benefit Determination, etc.) and/or post-service claims.*

Mutation analysis in fine-needle aspirates of the thyroid is considered to be **investigational**.

The use of a gene expression classifier in fine-needle aspirates of the thyroid that are cytologically considered to be indeterminate, atypical or suspicious for malignancy, is considered to be **investigational**.

**Rationale**

The literature on the use of molecular markers for thyroid nodules diagnosed by fine needle aspiration (FNA) as indeterminate, atypical, or suspicious consists of approximately 20 publications. These studies have analyzed either panels of mutations or a single mutation in these fine needle aspirates and compared the preoperative cytologic diagnosis and mutation status to postoperative final histologic diagnosis to determine diagnostic accuracy of the presence of a mutation, to predict the presence of malignancy. Some authors have also reported that the presence of certain mutations may predict more aggressive behavior in a malignant thyroid lesion. A gene expression classifier (GEC) has been developed to predict the likelihood that a thyroid lesion with indeterminate cytology is benign, allowing a patient to avoid surgical excision if that action is deemed to be clinically appropriate.
Molecular Markers to Predict Malignancy

Analytic Validity

Point Mutation Testing

Point mutations in specific genes associated with thyroid cancer, such as the BRAF V600E gene, and the detection of genetic rearrangements associated with thyroid cancer, such as the RET/PTC rearrangement, are typically detected with real-time PCR (rtPCR) sequencing methods. In the case of mutation testing for genes associated with thyroid cancer malignancy, analytic validity refers to a test’s technical accuracy in detecting a mutation that is present or in excluding a mutation that is absent. Realtime PCR-based methods are generally considered to have high accuracy. For example, Smith et al reported technical performance characteristics for BRAF mutation detection by qualitative PCR in thyroid FNA samples with high within- and between-run reproducibility. (7)

Next-generation sequencing (NGS) is expected to have high accuracy for detecting a mutation that is present. However, with increasing numbers of tested mutations, there is increased risk of detection of variants of uncertain significance (VUS). The VUS rate for currently available NGS panels for thyroid cancer is not well-characterized. Nikiforova et al described the development and validation of a multigene NGS panel for thyroid cancer, the ThyroSeq panel. (8) The authors developed a custom library of gene sequence variants based on mutations previously reported in the literature. The assay demonstrated 100% accuracy in evaluating samples of 15 thyroid tumors and 3 cell lines with known genetic alterations and 15 DNA samples with no mutations. In analysis of 229 DNA samples from frozen tissues, formalin fixed, paraffin-embedded tissues, and FNAs (n=105, 72, and 52, respectively), the panel identified mutations in 19 of 27 (70%) of classic papillary thyroid carcinomas (PTCs), 25 of 30 (83%) follicular variant PTCs, 14 of 18 (78%) conventional and 7/18 (39%) Hürthle cell carcinomas, 3 of 10 (30%) poorly differentiated carcinomas, 20 of 27 (74%) anaplastic thyroid carcinomas, and 11 of 15 (73%) medullary thyroid carcinomas. Of 83 benign nodules, 5 (6%) were positive for mutations.

Beaudenon-Huibregtset al reported the results of a prospective evaluation of an NGS panel that evaluates for 14 single nucleotide substitutions in the BRAF, HRAS, KRAS, or NRAS genes and 3 fusion transcripts, PAX8-PPARG, RET-PTC1, and RET-PTC3 (ThyGenX panel) in 806 nodule aspirates from 618 subjects. (9) A single genetic alteration was detected in 80% of cytology malignant cases, 21% of indeterminate, 7.8% of nondiagnostic, and 3.5% of benign cases.

Gene Expression Profiling

In 2015, Diggans et al described the development and validation Afirma BRAF malignancy classifier. (6) The study included FNA biopsies from 716 thyroid nodules. Biopsies were evaluated with quantitative PCR (qPCR) for the BRAF V600E gene, with 181 used as a training sample and 535 used as a validation sample. The Afirma BRAF malignancy classifier was generated using robust multichip average normalized gene expression summaries, and the classifiers were evaluated for positive percent agreement (PPA) and negative percent agreement (NPA) with the PCR-derived gene classification. The highest scoring classification method and gene set were then used in a final round of model building. The maximum PPA and NPA for all cytology categories was observed when the threshold for BRAF-positive status was 5% or more BRAF mutations. At 5% analytic sensitivity, Afirma BRAF demonstrates a PPA with PCR results of 90.4% (95% exact binomial confidence interval [CI],
83.5% to 95.1%) and an NPA of 99% (95% CI, 97.6% to 99.7%). There were 2 samples in the training set and 4 samples in the validation set that were Afirma BRAF positive but negative (0% mutation) on PCR, which the authors attribute to either technical variability in either assay or mutations other than the BRAF V600E mutation that cause similar gene expression changes.

Intra- and interrun reproducibility of the classifier was evaluated using 9 FNA biopsies (FNABs) and 3 tissue controls selected from among training samples with high (BRAF-positive) or low (BRAF-negative) classifier scores and scores near the classifier decision boundary. Each FNAB and tissue was processed from total RNA in triplicate in each of three different runs across days, operators and reagent lots. The intraassay standard deviation (SD) of Afirma BRAF scores was 0.171 (95% CI, 0.146 to 0.204). Of the 106 Afirma BRAF calls produced (2 arrays failed quality control requirements), 106 resulted in concordant calls across all 3 runs (100% concordance). The interassay SD of scores was 0.204 (95% CI, 0.178 to 0.237) for scores measured on a 6-point scale. These results suggest low intra- and interrun variability. No studies describing the analytic validity of the Afirma MTC test were identified.

Clinical Validity

Point Mutation/Gene Fusion Testing

A number of studies have evaluated whether testing for point mutations or gene fusions (either single mutation or panels of mutations) can be used to improve the sensitivity and specificity for diagnosing indeterminate FNA of the thyroid, with the goal of identifying mutations that predict malignancy in FNA samples.

Mutation Panel Testing: Standard Sequencing

Ferraz et al 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. (10) Sixteen studies analyzed one mutation (eg, 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 histologically (surgically resected) benign thyroid lesion was categorized as a false positive (FP) 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 4 studies that examined a panel of mutations, there was a broad sensitivity range of 38% to 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 2 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% (91-6%). Based on 3 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 of 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 but that further standardizations and further molecular markers are needed before broad application of molecular FNA cytology for the diagnosis of thyroid nodules.

Nikiforov et al prospectively tested a panel of mutations (BRAF, RAS, RET/PTC and PAX8/PPARγ) in 470 FNA samples of thyroid nodules from 328 consecutive patients. (11) Mutational status was
correlated with cytology, and either surgical pathology diagnosis or follow-up (mean, 34 months). A total of 40 patients were excluded for poor quality of specimen or loss to follow-up. Sixty-nine patients (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 one was a benign tumor (3%). Thirteen 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 patients, 147 (229 FNAs) who did not undergo surgery were followed by serial ultrasound with no change in the nodule status (124 patients) or by repeated FNA with cytology negative for malignancy (23 patients) and no mutation found in the FNA material. These nodules were considered as negative for malignancy. The remaining 72 patients that were initially in the follow-up group underwent subsequent surgery. Combining all 3 groups, the specificity for malignancy was high (99.7%), but the sensitivity of the molecular test alone was 62%.

Moses et al prospectively tested FNA samples from 417 patients with 455 thyroid nodules for BRAF, NRAS, KRAS and RET/PTC 1 and 3 and TRK1 mutations. (12) Overall, 50 mutations (23 BRAF V600E, 21 NRAS and 4 RET/PTC1 and 2 RET/PTC3 rearrangements) were detected. There were significantly more mutations detected in malignant nodules than in benign (p<0.001). For thyroid FNA biopsies that were indeterminate or suspicious (n=137), genetic testing had a sensitivity of 12%, specificity of 98%, positive predictive value of 38% and negative predictive value (NPV) of 65%.

Ohori et al performed mutation screening in 117 FNA samples classified as a follicular lesion of indeterminate significance/atypia of indeterminate significance. (13) BRAF, 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 a mutation detected was 7.6%.

In 2011, Nikiforov et al reported results of a prospective study to assess the clinical utility of a panel of mutations to predict the likelihood of malignancy in thyroid nodules that were indeterminate on FNA.(14) The authors included 1056 consecutive FNA samples with indeterminate cytology on FNA that underwent mutation testing, with 967 of those adequate for molecular analysis (653 follicular lesion of undetermined significance/atypia of undetermined significance; 247 follicular or Hürthle cell neoplasm or suspicious for follicular neoplasm; 67 suspicious for malignant cells). One hundred seventeen of the samples were included in the Ohori et al analysis previously described. Eighty-seven BRAF, RAS, RET/PTC, or PAX8/PPARγ mutations were detected. At the time of analysis, 479 patients had undergone thyroidectomy for further evaluation, providing a histopathologic diagnosis for 513 samples. The presence of a mutation had low sensitivity for predicting malignant histology (63%, 57%, 68% for samples with follicular lesion of undetermined significance/atypia of undetermined significance, follicular or Hürthle cell neoplasm/suspicious for follicular neoplasm, and suspicious for malignant cells on cytology, respectively), but high specificity (99%, 97%, 96%, respectively). The NPV for the mutation analysis results was 94%, 86%, and 72% for samples with follicular lesion of undetermined
significance/atypia of undetermined significance, follicular or Hürthle cell neoplasm/suspicious for follicular neoplasm, and suspicious for malignant cells on cytology, respectively. The authors conclude that mutation analysis may be useful in surgical planning, such as determining whether patients should undergo a thyroid lobectomy or a total thyroidectomy as a first surgery.

In a subsequent study, Nikiforov et al evaluated the accuracy of an NGS panel that included tests for point mutations in 13 genes and for 42 types of gene fusions (ThyroSeq v2 NGS panel) in a series of 143 consecutive thyroid FNA samples with a cytologic diagnosis of follicular or Hürthle cell neoplasm/suspicious for follicular or Hürthle cell neoplasm. Molecular testing was retrospectively performed for 91 samples and prospectively performed for the remaining 52. The prevalence of cancer on histology was 27.5% and 26.9% in the retrospective and prospective cohorts, respectively. In the retrospective cohort, of the 25 malignant nodules, 22 were PTCs, and 3 were follicular thyroid carcinomas (FTCs). In the prospective cohort, of the 14 malignant nodules, 11 were PTCs and 3 were FTCs. The performance of the ThyroSeq in both cohorts is shown in Table 1.

### Table 1: Performance of ThyroSeq in Nikiforov et al (2014)\(^\text{15}\)

<table>
<thead>
<tr>
<th>Mutation testing results</th>
<th>Retrospective (n=91)</th>
<th>Prospective (n=52)</th>
<th>Overall (N=143)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative</strong></td>
<td>64 (n=2 with cancer; n=62 benign)</td>
<td>37 (n=2 with cancer; n=35 benign)</td>
<td></td>
</tr>
<tr>
<td><strong>Positive</strong></td>
<td>27 (n=23 with cancer; n=4 benign)</td>
<td>15 (n=12 with cancer; n=3 benign)</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92%</td>
<td>86%</td>
<td>90% (95% CI, 80% to 99%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>94%</td>
<td>92%</td>
<td>93% (95% CI, 88% to 98%)</td>
</tr>
<tr>
<td>PPV</td>
<td>85%</td>
<td>80%</td>
<td>83% (95% CI, 72% to 95%)</td>
</tr>
<tr>
<td>NPV</td>
<td>97%</td>
<td>95%</td>
<td>96% (95% CI, 92% to 95%)</td>
</tr>
</tbody>
</table>

CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

The authors noted that, compared with the panel of mutations used in their 2011 study, the NGS panel was associated with marked increase in NPV, with a similar PPV. In this case, the authors proposed that the panel could be used to both “rule in” and “rule out” invasive cancers.

Cantara et al analyzed a panel of mutations in samples of 174 patients undergoing thyroid surgery for indeterminate/inadequate/benign FNA results. (16) The most prevalent mutation was \(\text{BRAF}\) (49.3% of the positive samples), followed by \(\text{RAS}\) (34.3%) and \(\text{RET/PTC}\) (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 due to the presence of a mutation in the tissue sample that was not found in the cytology sample.
Mathur et al 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 patients with 423 dominant thyroid nodules. (17) 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 Hurlth 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% (32/123) malignant tumors (p<0.05).

Eszlinger et al retrospectively analyzed a panel of mutations (*BRAF* and *RAS* point mutations and *PAX8-PPARγ* and *RET/PTC* rearrangements) in a sample of 310 thyroid air-dried FNA specimens with available corresponding formalin-fixed, paraffin embedded thyroid biopsy samples (164 indeterminate, 57 malignant, and 89 benign on FNA). (18) A total of 47 mutations were detected on FNA: 22 *BRAF* mutations, 13 *NRAS* mutations, 3 *HRAS* mutations, 8 *PAX8-PPARγ* rearrangements, and 1 *RET/PTC* rearrangement. The addition of mutation analysis to cytology results was associated with a sensitivity and specificity of 75.3% and 90.4% for the detection of malignancy, respectively, with a PPV and NPV of 77.2% and 89.4%, respectively. The presence of a *BRAF* or *RET/PTC* mutation was associated with cancer in 100% of samples.

**Mutation Panel Testing: NGS**

The 2014 Beauden-Huibregtse et al (9) study previously described, which evaluated the technical performance of an NGS panel also reported on the sensitivity and specificity of the panels in detecting cancer in subsets of nodules with available postresection histology. These results are summarized in Table 2.

<table>
<thead>
<tr>
<th>Table 2: Diagnostic Performance of NGS Panels in Thyroid FNA Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Patients With Histology Available</strong></td>
</tr>
<tr>
<td>Sample N</td>
</tr>
<tr>
<td>Sensitivity 66% (95% CI, 52% to 78%)</td>
</tr>
<tr>
<td>Specificity 89% (95% CI, 77% to 96%)</td>
</tr>
<tr>
<td>PPV 86% (95% CI, 72% to 95%)</td>
</tr>
<tr>
<td>NPV 71% (95% CI, 59% to 82%)</td>
</tr>
</tbody>
</table>

CI: confidence interval; FNA: fine needle aspiration; NGS: next-generation sequencing; NPV: negative predictive value; PPV: positive predictive value.

**BRAF**

Adeniran et al conducted a study of 157 cases with equivocal thyroid FNA readings (indeterminate and suspicious for papillary thyroid carcinoma [PTC]) or a positive diagnosis for PTC, and concomitant *BRAF* mutation analysis. (1) The results of histopathologic follow-up 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 the combination of 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 \textit{BRAF} mutation. The authors concluded that patients with an equivocal cytologic diagnosis and \textit{BRAF V600E} mutation could be candidates for total thyroidectomy and central lymph node dissection.

Xing et al investigated the utility of \textit{BRAF} mutation testing of thyroid FNA specimens for preoperative risk stratification in PTC in 190 patients. (19) A \textit{BRAF} mutation in preoperative FNA specimens was associated with poorer clinicopathologic outcomes of PTC. In comparison with the wild-type allele, a \textit{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, 0.6-10 years), PTC persistence/recurrence was seen in 36\% of \textit{BRAF} mutation-positive patients versus 12\% of \textit{BRAF} mutation-negative patients, with an odds ratio of 4.16 (95\% confidence interval [CI], 1.70 to 10.17; \(P = 0.002\)). The positive and NPVs for preoperative FNA-detected \textit{BRAF} mutation to predict PTC persistence/recurrence were 36\% and 88\% for all histologic subtypes of PTC. The authors concluded that preoperative \textit{BRAF} mutation testing of FNA specimens may provide a novel tool to preoperatively identify PTC patients at higher risk for extensive disease (extrathyroidal extension and lymph node metastases) and those who are more likely to manifest disease persistence/recurrence.

Jara et al retrospectively evaluated the utility of \textit{BRAF} mutation testing in 66 thyroid nodules with "suspicious for PTC" on FNA and available histopathologic samples from thyroid biopsy. (20) Forty-two subjects (62.6\%) had PTC diagnosed on final histopathology. A positive \textit{BRAF} mutation test was associated with a sensitivity and specificity for PTC of 45.5\% and 87.5\%, respectively, and a PPV and NPV of 88.2\% and 43.8\%, respectively.

Several other studies have also reported high specificity (100\%) of \textit{BRAF} specificity in the detection of PTC in indeterminate thyroid nodules. (21, 22)

\textbf{Gene Expression Profiling}

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

In the Diggans et al study describing the development and validation of the Afirma BRAF test, previously described, for a subset of 213 thyroid nodule FNA samples for which histopathology was available, the Afirma BRAF test results were compared with pathologic findings. (6) The Afirma BRAF classified all histopathologically benign samples as \textit{BRAF V600E}–negative (specificity, 100\%; 95\% CI, 97.4\% to 100\%). Of the 73 histopathologically malignant samples, the Afirma BRAF test identified 32 as \textit{BRAF} positive (sensitivity, 43.8\%; 95\% CI, 32.2\% to 55.9\%).

\textbf{Clinical Utility}

Testing for specific mutations associated with thyroid cancer (eg., \textit{BRAF V600E} mutations, \textit{RET} mutations, and \textit{RET/PTC} and \textit{PAX8/PPAR\textsubscript{\gamma}} rearrangements) are generally designed to “rule in” cancer in nodules that have indeterminate cytology on FNA. (23) Of note, some mutation panels, such as the ThyroSeq panel, may have a high enough NPV that their clinical use could also be considered as a

<table>
<thead>
<tr>
<th>Section</th>
<th>Medicine</th>
<th>Effective Date:</th>
<th>October 15, 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsection</td>
<td>Pathology/Laboratory</td>
<td>Original Policy Date:</td>
<td>September 20, 2013</td>
</tr>
<tr>
<td>Subject</td>
<td>Molecular Markers in Fine Needle Aspirates of the Thyroid</td>
<td>Page:</td>
<td>10 of 20</td>
</tr>
</tbody>
</table>
molecular marker to predict benignancy; see next section). A potential area for clinical utility for this type of mutation testing would be in informing preoperative planning for thyroid surgery following initial thyroid FNA, such as planning for a hemi- versus a total thyroidectomy or performance of a central neck dissection.

In a retrospective analysis, Yip et al reported outcomes after implementation of an algorithm incorporating molecular testing of thyroid FNA samples to guide the extent initial thyroid resection.24 The study included a cohort of patients treated at a single academic center at which molecular testing (BRAF V600E, BRAF K601E, NRAS codon 61, HRAS codon 61, and KRAS codon 12 and 13 point mutations; RET/PTC1, RET/PTC3, and PAX8/PPARγ rearrangements) was prospectively obtained for all FNAs with indeterminate cytology (follicular lesion of undetermined significance, follicular neoplasm, and suspicious for malignancy), and for selective FNAs at the request of the managing physician for selected nodules with either benign or nondiagnostic cytology. The study also included a second cohort of patients who did not have molecular testing results available. For the patients treated with molecular diagnosis, a positive molecular diagnostic test was considered to be an indication for an initial total thyroidectomy. Patients with follicular lesion of undetermined significance and negative molecular diagnostic results were followed with repeat FNA, followed by a lobectomy or total thyroidectomy if indeterminate pathology persisted. Patients with follicular neoplasm or suspicious for malignancy results on cytology and a negative molecular diagnostic result were managed with lobectomy or total thyroidectomy.

The sample included 671 patients, 322 and 349 managed with and without molecular diagnostics, respectively. Positive molecular testing results were obtained in 56 patients (17% of those managed with molecular diagnostics), most commonly RAS mutations (42/56 [75%]), followed by BRAF V600E (10/56 [18%]), BRAF K601E (2/56 [4%]), and PAX8/PPARγ rearrangements (2/56 [4%]). Compared with those managed without molecular diagnostics, patients managed with molecular diagnostics were nonsignificantly less likely to undergo total thyroidectomy as an initial procedure (63% vs 69%, p=0.08). However, they had nonsignificantly higher rates of central compartment lymph node dissection (21% vs 15%, p=0.06). Across both cohorts, 25% of patients (170/671) were found to have clinically significant thyroid cancer, with no difference in thyroid cancer rates based on the type of initial operation (26% for total thyroidectomy vs 22% for lobectomy, p=0.3). The incidence of clinically significant thyroid cancer after initial lobectomy (ie, requiring a 2-stage surgery) was significantly lower for patients managed with molecular diagnostics (17% vs 43%, p<0.001). An indeterminate FNA result had sensitivity and specificity for the diagnostic of thyroid cancer of 89% and 27%, respectively, with PPV and NPV of 29% and 88%, respectively. The addition of molecular diagnostics to FNA results increased the specificity for a cancer diagnosis to 95% and the PPV to 82%.

Section Summary
The available evidence suggests that the use of mutation testing in thyroid FNA samples is generally associated with a high specificity and PPV for clinically significant thyroid cancer. The most direct evidence related to the clinical utility of mutation testing for genes associated with malignancy in thyroid cancer comes from a single-center retrospective study which reported surgical decisions and pathology findings in patients managed with and without molecular diagnostics. This study suggests that testing for a panel of mutations associated with thyroid cancer malignancy may allow the appropriate selection of patients for an initial total thyroidectomy. However, prospective validation of these study findings in additional settings is needed.
Molecular Markers to Predict Benignancy (GEC)

Analytic validity

Walsh et al verified the analytical performance of the Afirma gene expression classifier (GEC) in the classification of cytologically indeterminate FNAs from thyroid nodules. (25) 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. The authors concluded that the analytical sensitivity and specificity, robustness and quality control of the GEC were successfully verified.

Chudova et al developed a molecular test to distinguish between benign and malignant thyroid nodules using FNAs. (3) The authors used mRNA analysis to measure more than 247,000 transcripts in 315 thyroid nodules. 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, and included 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, negative predictive value and specificity were estimated to be 96% and 84%, respectively.

Clinical validity

Alexander et al reported on a 19-month, prospective, multicenter validation study of the Afirma GEC, which involved 49 clinical sites (both academic and community centers), 3789 patients and 4812 FNAs from thyroid nodules that were at least 1 cm in size. (26) Local pathology reports of the cytologic diagnosis were collected for all patients, and reports without a definitive benign or malignant diagnosis at the local site were reviewed by 3 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 deemed to be cytologically indeterminate were successfully tested with the GEC assay at Veracyte Laboratory. Of the 265 samples, 85 were malignant; the GEC correctly identified 78 of the 85 as suspicious (92% sensitivity; 95% CI: 84%-97%), with a specificity of 52% (95% CI: 44%-59%). NPV ranged from 85% for “suspicious cytologic findings” to 95% for “atypia of undetermined clinical significance”. There were 7 FNAs with false-negative results, 6 of which were thought to be due to hypocellular aspirate specimens.

Several single-center studies, including Harrell and Bimston (2013), (27) Lastra et al (2014), (28) and McIver et al (2014) (29) have reported the diagnostic accuracy of the Afirma GEC, which are summarized in Table 3. Compared with earlier studies, in a sample of thyroid FNAs form an academic medical center, McIver et al reported a lower than expected rate of Afirma “benign” reports in follicular or Hürthle cell neoplasms and a lower than expected malignancy rate among Afirma “suspicious” FNAs.
Table 3: Single-Center Studies Reporting Afirma GEC Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FNA samples</td>
<td>645 FNA samples</td>
<td>132 FNA samples</td>
<td>1207 FNA samples</td>
</tr>
<tr>
<td>FNA samples classified as indeterminate</td>
<td>58 FLUS/AUS or FN (8.9%)</td>
<td>69 FLUS/AUS (51.5%)</td>
<td>12 FLUS/AUS (1%)</td>
</tr>
<tr>
<td>Afirma GEC results</td>
<td>Suspicious</td>
<td>Benign</td>
<td>Suspicious</td>
</tr>
<tr>
<td>Total N</td>
<td>36(^a)</td>
<td>20</td>
<td>62</td>
</tr>
<tr>
<td>N undergoing thyroidectomy</td>
<td>30</td>
<td>5</td>
<td>48</td>
</tr>
</tbody>
</table>

\(^a\)Two samples inadequate due to low mRNA content.  
\(^b\)GEC results were available for 60 subjects.

Clinical utility

Duick et al reported on the impact of Afirma GEC test results on physician and patient decision-making to operate on thyroid nodules with indeterminate cytology. (30) This retrospective, multicenter study included patients who were 21 years or older, had one or more thyroid nodules 1 cm 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 endocrinology practices in 11 states participated. Data were collected on 368 patients with 395 nodules. The data collection period was September 2011 through March 2012. Surgery was performed in 7.6% of the patients with indeterminate cytology and a benign GEC. (Surgery was primarily performed on those patients with indeterminate cytology and a benign GEC because of large or symptomatic nodules, rapidly growing nodules or a second suspicious or malignant nodule in the same patient, the same reasons typically given for 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 patients with an indeterminate cytologic diagnosis (but no GEC test).

In 2014, Alexander et al reported results from a retrospective analysis of 339 thyroid nodules which underwent Afirma GEC testing for indeterminate cytology on FNA (follicular lesion of undetermined significance/atypia of undetermined significance, follicular neoplasm, or suspicious for malignancy) at 5 academic medical centers. (31) Most of the nodules sent for GEC testing were follicular lesions of undetermined significance/atypia of undetermined significance or follicular neoplasm. The distribution of GEC testing results for each cytologic classification is shown in Table 1.

Table 4: GEC Testing Results for Alexander et al

<table>
<thead>
<tr>
<th>Cytologic Classification</th>
<th>Total</th>
<th>GEC Testing Results, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>Follicular lesion of undetermined significance/atypia of undetermined significance</td>
<td>165</td>
<td>91 (55%)</td>
</tr>
<tr>
<td>Follicular neoplasm</td>
<td>161</td>
<td>79 (49%)</td>
</tr>
<tr>
<td>Suspicious for malignancy</td>
<td>13</td>
<td>4 (31%)</td>
</tr>
<tr>
<td>Total</td>
<td>339</td>
<td>174</td>
</tr>
</tbody>
</table>
A subset of patients whose nodules underwent GEC testing underwent a subsequent thyroid resection. Among 148 cases with suspicious Afirma GEC findings, surgery (thyroid resection) was recommended for 141 (95%). For the 174 cases with benign Afirma GEC findings, surgery was recommended for 4 (2%; p<0.01). Using the assumption that, in the absence of the GEC results, thyroid surgery would be recommended for patients with cytologically indeterminate FNA results, the authors report that the GEC results altered management in 50% of patients. Table 5 shows thyroidectomy biopsy results for the subset of patients shown in Table 4 who underwent surgery.

### Table 5: Thyroidectomy Biopsy Results for Alexander et al

<table>
<thead>
<tr>
<th>GEC Testing Results</th>
<th>Total, n</th>
<th>Surgery Recommended, n</th>
<th>Surgery Completed, n</th>
<th>Pathology Malignant, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspicious</td>
<td>148</td>
<td>141</td>
<td>121</td>
<td>53 (44% of those with completed surgery)</td>
</tr>
<tr>
<td>Benign</td>
<td>174</td>
<td>4</td>
<td>11</td>
<td>1 (9% of those with completed surgery)</td>
</tr>
</tbody>
</table>

Seventeen patients who had indeterminate cytology, benign Afirma GEC results, and did not undergo surgery had follow-up beyond 1 year. Of those, 3 patients underwent surgical removal of the nodule because of compressive symptoms (n=2) or nodule growth (n=1); all nodules were benign on final histology. The remaining 14 patients had ongoing follow up with ultrasound with no ongoing evidence of malignancy. The study demonstrated site-to-site variation in the proportion of samples that were GEC benign.

This study suggests that the Afirma GEC may alter clinical management of patients with indeterminate thyroid nodules. While the treating physicians presumably elected to obtain the GEC testing with the intent of altering management recommendations, the magnitude of the difference in surgical recommendations for patients with GEC suspicious or benign results was large. A limitation of this study is its retrospective, unblinded nature; thus, factors other than GEC testing may have contributed to either the recommendation for surgery or patients’ decisions to undergo surgery. A benign GEC result did not completely rule out malignant pathology. Long-term follow up was available for only a small proportion of patients with benign GEC findings who did not undergo surgery.

In a single-center study, Aragon Han et al reported surgical management decision making outcomes among 114 patients with thyroid nodules who underwent molecular testing. (32) Of 114 patients, 87 underwent thyroid surgery. Testing included a combination of the Afirma gene-expression classifier (n=37), a DNA-based somatic mutation panel (n=21), and testing for BRAF mutations (n=29), BRAF/NRAS (n=1), and BRAF/RET/PTC (n=1). A surgical decision-making algorithm that did not include mutation testing was developed by consensus among 4 thyroid surgeons. If the surgeon performed the same surgery as anticipated by the management algorithm, then the molecular test was considered to have no impact. If the surgeon performed a different surgery than anticipated by the management algorithm, the molecular test was considered to effect a change in management. The authors report that surgical management was not changed by molecular testing in 89.7% of cases. This study is limited by its use of multiple types of molecular testing, along with a nonstandardized incorporation of molecular genetic testing results into the surgical decision making. As such, the study has limited implications for the clinical utility of molecular diagnostics for thyroid cancer.
Section Summary
There is 1 commercially available GEC that is designed to exclude malignancy in individuals with indeterminate thyroid FNA results. The GEC has been reported to have a high NPV in a limited number of studies. The available evidence suggests that physician decision making about surgery is altered by GEC results, although long-term follow-up of patients with thyroid nodules who avoided surgery based on GEC results is limited.

Ongoing and Unpublished Clinical Trials
Some currently unpublished trials that might influence this policy are listed in Table 6.

Table 6. Summary of Key Trials

<table>
<thead>
<tr>
<th>NCT No.</th>
<th>Trial Name</th>
<th>Planned Enrollment</th>
<th>Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unpublished</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT00316823</td>
<td>Gene Expression Analysis in Thyroid Nodule FNA Samples</td>
<td>400</td>
<td>Dec 2009</td>
</tr>
</tbody>
</table>

NCT: national clinical trial.

Practice Guidelines and Position Statements
American Thyroid Association
In 2013, the American Thyroid Association (ATA) issued a report from its clinical affairs committee regarding commercially available molecular diagnostic testing in the evaluation of thyroid nodule FNA specimens, which stated that the currently available mutation analysis panel and gene expression classifier have promising roles but that at this time, experience with them remains limited and that until an expert consensus review of existing data (currently underway) can be completed, no evidence-based recommendation for or against the use of these methods can be made. (33)

The most recent ATA guidelines on the management of thyroid nodules are from 2009, before the widespread clinical availability of mutation analysis or gene expression profiles for thyroid cancer. (34)
Updated guidelines were previewed in 2014, but final guidelines have not been published.

National Comprehensive Cancer Network (NCCN)
NCCN guidelines on the treatment of thyroid cancer make the following comments about the use of molecular diagnostic in thyroid cancer (35):
For thyroid nodules evaluated with FNA, molecular diagnostics may be employed in the following cases (category 2B recommendation):
- Follicular or Hürthle cell neoplasm without high clinical suspicion of malignancy.
- Atypia of undetermined significance/Follicular lesion of undetermined significance without high clinical suspicion of malignancy.

The guidelines state: “Molecular diagnostics to detect individual mutations (eg, BRAF, RET/PTC, RAS, PAX8/PPAR [peroxisome proliferator-activated receptors] gamma) or pattern recognition approaches using molecular classifiers may be useful in the evaluation of FNA samples that are indeterminate to assist in management decisions. The choice of the precise molecular test depends on the cytology and the clinical question being asked.” (35)
Summary

Mutation analysis

Analysis for mutations associated with thyroid cancer in fine needle aspirates (FNA) of the thyroid that are cytologically indeterminate has a high positive predictive value for malignancy. However, patients with an equivocal FNA result would likely proceed to surgery regardless of mutation status, with intraoperative consultation to guide the necessity and extent of surgery. 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. One retrospective single-center study suggested that testing for a panel of mutations associated with thyroid cancer may allow the appropriate selection of patients for surgical management with an initial total thyroidectomy. Prospective studies in additional populations are needed to validate these results. Mutation analysis does not achieve a high enough negative predictive value (NPV) to identify which patients can undergo watchful waiting over thyroid surgery. Although the presence of certain mutations may predict more aggressive malignancies, the clinical utility of identifying these mutations preoperatively has not been established. Given the limitations in the evidence base, the use of molecular diagnostics in thyroid FNA samples for mutations associated with thyroid cancer is considered investigational.

Gene expression classifier (GEC)

There is 1 commercially available gene expression classifier (GEC) that has been developed to predict benignancy in thyroid nodules. The reported NPV of the GEC in predicting which thyroid nodules with indeterminate cytology are benign is high. Two relevant retrospective studies on the clinical utility of the GEC have been published and suggest that treatment recommendations for patients with indeterminate cytology are affected by the results of the GEC test. For patients who avoided surgery based on GEC results, limited longer term follow-up data are available. Although the available evidence suggests that this group of patients does well, longer term follow-up has been reported for only a small number of patients. Therefore, the use of a GEC to predict which thyroid nodules with indeterminate cytology are benign is considered investigational.

Medical National Coverage

There is no national coverage determination (NCD).

References


33. Hodak SP, Rosenthal For The American Thyroid Association Clinical Affairs Committee DS. Information for clinicians: commercially available molecular diagnosis testing in the evaluation of thyroid nodule fine-needle aspiration specimens. Thyroid. Feb 2013; 23(2):131-134. PMID 22984796


**Policy History**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action</th>
<th>Reason</th>
</tr>
</thead>
</table>

**Keywords**

miRInform
Veracyte
Mutation analysis
Fine needle aspirate
Thyroid
This policy was approved by the FEP® Pharmacy and Medical Policy Committee on September 18, 2015 and is effective October 15, 2015.

Signature on file
Deborah M. Smith, MD, MPH