Diagnostic Potential of Ancillary Molecular Testing in Differentiation of Benign and Malignant Thyroid Nodules

PARISHA BHATIA1, AHMED DENIWAR1, PAUL FRIEDLANDER2, RIZWAN ASLAM2 and EMAD KANDIL1,2

Departments of 1Surgery and 2Otolaryngology-Head and Neck Surgery, Tulane University School of Medicine, New Orleans, LA, U.S.A.

Abstract. Fine needle aspiration (FNA) cytology, being the mainstay to diagnose thyroid nodules, does not provide definitive results in a subset of patients. The use of molecular markers testing has been described as a useful aid in differentiation of thyroid nodules that present with an indeterminate cytodiagnosis. Molecular tests, such as the Afirma gene classifier, mutational assay and immunohistochemical markers have been increasingly used to further increase the accuracy and defer unnecessary surgeries for benign thyroid nodules. However, in light of the current literature, their emerging roles in clinical practice are limited due to financial and technical limitations. Nevertheless, their synergistic implementation can predict the risk of malignancy and yield an accurate diagnosis. This review discusses the clinical utility of various molecular tests done on FNA indeterminate nodules to avoid diagnostic thyroidectomies and warrant the need of future multi-Institutional studies.

Thyroid nodules are being increasingly identified during screening studies. According to the American Thyroid Association, 5%-10% of thyroid nodules are detected by palpation. However, in women, approximately 50% are detected by ultrasound. Although a majority of these nodules are benign, 10%-15% can harbor malignancy (1, 2). To detect the type of cancer in these nodules, current guidelines focus on performing a fine needle aspiration (FNA) biopsy on nodules meeting specific criteria, including solid nodules of >1cm and complex nodules of >1.5cm (3). The diagnostic results of FNA can be improved when performed on nodules with suspicious features of malignancy on ultrasound. A substantial cellular smear prepared can yield an accurate diagnosis in 62%-85% of benign nodules; however, 6% of these nodules still prove to be malignant on surgical pathology. Nevertheless, a major concern to clinicians occurs when cytology results are documented as indeterminate (15%-30%) (4). According to the Bethesda System for Reporting Thyroid Cytology (TBSRTC), these are classified as Class III: atypia of undetermined significance/ follicular lesion of undetermined significance (AUS/FLUS), Class IV: follicular neoplasm (FN) and Class V: suspicious for malignancy. Subsequently, these nodules undergo diagnostic surgical resection where final pathological diagnosis is deemed malignant in 5%-15% cases of Class III nodules and 15%-30% of Class IV nodules (4, 5). Therefore, a vast majority of patients suffer a risk of unnecessary, potentially dangerous and costly surgical intervention. Clearly, a test is needed that is sensitive enough to prevent unnecessary surgeries in patients with benign nodules. The inherent limitations of FNA have led to the proposal of newly-evolved diagnostic strategies to evaluate cytologically-indeterminate thyroid nodules. Because these nodules are associated with an increased risk of carcinoma, efforts have been made to increase the pre-surgical diagnostic power of tests in detection of malignancy.

Molecular testing is being increasingly recommended to detect cancer in indeterminate thyroid nodules (5-8). Over the past few years, an increased emphasis has been placed on a large number of biomarkers known to frequently occur in thyroid cancer. Research has focused mainly on detection of papillary thyroid carcinoma (PTC), especially the follicular variant, and follicular neoplasms since they make up a large part of histological subtypes of thyroid cancer. When a final diagnosis of benign adenoma or FLUS is reached on FNA, patients are directed to thyroid surgery. Attempts to reveal a synergistic role of molecular analysis and cytological tests have been studied and reported by several groups. However, the accumulated evidence on their clinical validity varies significantly. Current literature supports the use of distinct
types of biomarkers for thyroid cancer molecular analysis. These can be divided into: i) mutational assay, such as the BRAF\textsuperscript{V600E} mutation that is believed to be associated with aggressive forms of papillary thyroid cancer (PTC), ii) RNA microarray assay such as Afirma, where cellular aspirate from thyroid nodules is evaluated for malignancy by measuring the expression of gene transcripts and iii) immunohistochemical stains, such as Galectin-3, asparagine synthase (ASNS) and tyrosine related protein 1 (TYRP2) (6-8). Their utility has been demonstrated for diagnostic, as well as prognostic purposes in thyroid cancer. A number of studies have been performed to resolve the inconsistency in pre-surgical diagnosis, which carries a considerable risk of malignancy.

Our review was conducted to resolve this dilemma and analyze the need for synergistic diagnostic tests, in addition to the gold-standard FNA, to provide a firm diagnosis for nodules with THY3 cytology. This review also discusses various tests that have recently been proven efficient and feasible options to distinguish benign from malignant thyroid nodules.

**Types of Molecular Testing on Biopsy Samples**

1. **Afirma\textregistered Gene Expression Classifier (AGEC).** This is a recently commercialized tool (owned by the Veracyte Corporation; South San Francisco, CA, USA) that assists in pre-surgical diagnosis of thyroid nodules with indeterminate cytology. It has been proposed and advocated as an ancillary test to FNA, thereby increasing the accuracy in determination of malignancy in cytological-indeterminate nodules. It is crucial for physicians to establish the analytical validity, clinical validity and clinical utility of all commercially available tests and only Afirma has successfully demonstrated its robustness. One of the first studies to successfully verify analytical validity was performed by Walsh et al. who demonstrated that gene expression classifier (GEC) can tolerate variations in RNA input (5-25 ng) and dilutions of malignant FNA material down to 20% confirming its analytical sensitivity (9). Negligible assay interference was observed for FNA material admixed blood (up to 83%) and genomic DNA (up to 30%). However, it was observed that a higher proportion of blood in benign FNA samples led to false-negative results.

A large group prospective clinical validation study was conducted by Alexander and et al. who assessed the performance of gene expression classifiers on 4812 FNA-aspirates (5). Of the 577 indeterminate samples, GEC was applied to 265 indeterminate nodules. The specimen histopathology discovered 85 malignant nodules. After application of GEC, 78 of the 85 malignant nodules were correctly diagnosed and labeled as “suspicious” producing a sensitivity of 92% and specificity of 52%. With Afirma results classified as “benign” in 52% of the nodules, unnecessary surgery could be avoided for patients with cytologically indeterminate/GEC-benign status nodules. Additionally, the highest negative predictive value (NPV) of 95% was observed for Bethesda Class 3 nodules, while for Class IV and Class V NPV was 94% and 85%, respectively. An important finding of this study was that in 6 out of the 7-false negative nodules, the aspirated material was declared sub-optimal for performing GEC (5). Therefore, ensuring the adequacy of the sample to perform this test could lead to fewer false negatives and better selection of candidates for surgery.

**Method of Implementation and Validation**

The Afirma test works on the principle of identification of mRNA gene expression extracted from FNA aspirate of thyroid nodules and comparing it against a panel of 167 molecular genes (5, 9). The molecular classifier is applied to six cassettes to filter the sample from rare neoplasms that are otherwise difficult to detect by FNA biopsy, namely, malignant melanoma, renal cell carcinoma, breast carcinoma, parathyroid tissue, medullary thyroid carcinoma and Hurthle cell neoplasm. Subsequently, using the proprietary algorithm for gene expression, FNA samples of these nodules are labeled “benign” or “suspicious” while only 10% of these samples with low RNA yield are labeled as “No result” (10, 11). Figure 1 displays the stepwise manner of GEC implementation. According to National Cancer Comprehensive Network Thyroid Carcinoma guidelines, indeterminate FNA samples that undergo GEC and are designated as “benign” have a malignancy risk of <5% (10). Therefore, these nodules can be followed up carefully with ultrasound avoiding the risk of diagnostic thyroid surgery. As a rebuttal, another validation study by McIver et al. found the performance of GEC on indeterminate nodules to be lower than previously reported (12). The reported sensitivity and specificity were 83% and 10%, respectively. However, of the 44 GEC-suspicious samples, only 32 patients underwent surgery. There was 1 false negative result, which revealed follicular carcinoma on histopathology, but the overall malignancy rate was reported to be approximately 12.5%, much lower than the previously published data. Their lower positive predictive value (16%) may be attributed to lower pre-test probability of malignancy in the chosen population of indeterminate nodules. However, it should be noted that GEC may have a lower clinical utility to detect other forms of thyroid carcinoma, such as follicular carcinoma. The major concern regarding its implementation is false negative results. A malignant nodule labeled GEC-benign would alter a surgeon’s management plan, unless there is a high suspicion of malignancy from ultrasound. The results of McIver and colleagues raise questions about the clinical validity of GEC and demonstrate the need to amend the application of this test with additional diagnostic measures. A retrospective study to determine the clinical utility of GEC was performed on 368 patient contributed by 51 endocrino-
logists from 21 practice sites (11). They evaluated the impact of GEC results on patients’ management plan where they observed a significant decline of 74% to 7.6% operative rate on indeterminate nodules after GEC implementation. At the same time, they recommended close observation for GEC benign nodules, thus avoiding the need of diagnostic surgeries and the associated financial burden that prevail.

2. DNA-based Mutational Assay. The high volume of unnecessary diagnostic surgeries performed for benign nodules has led researchers to establish additional methods to improve the accuracy of FNA biopsies. Several mutations such as BRAF, RET/PTC, PAX8/PPARγ and RAS are frequently associated with different types of thyroid cancer. These genetic mutations, seen in up to 70%-80% of carcinomas, may display aggressive behavior and can be used as diagnostic markers in assessment of thyroid nodules (6, 13-16).

**BRAF**, a proto-oncogene that regulates cell growth and functioning through the RAS/MAPK pathway, has been increasingly detected in more than half of PTC cases. It is found in classic PTC, which is associated with lymph node metastases, extrathyroidal extension and distant metastases. However, its prognostic role is still controversial (17). The most common type of BRAF mutation results in replacement of glutamic acid by valine at the 600 position (BRAFV600E), which activates the MAPK pathway leading to uncontrolled proliferation of cells. A study by Cohen and colleagues studied mutations in **BRAF** for pre-operative assessment in thyroid cancer. It was observed that a **BRAF** mutation was present in 72% of cytological-malignant nodules and that it was significantly more prevalent in conventional PTC (6). Subsequent studies showed that the **BRAF** mutation is also associated with poorer prognosis of PTC (7). One of the largest prospective studies was conducted by Nikiforova et al., who studied the role of tumor-specific mutations in 470 FNAs of thyroid nodules (16). A comprehensive panel of mutations (**BRAF, RET/PTC, PAX8/PPARγ and RAS**) were tested for their feasibility to improve diagnostic accuracy of FNA cytology. Their promising results indicated 100% validity for malignancy in TBSRTC Class III, with higher performance of molecular testing for Class IV and Class V. These results strongly suggest the necessity of surgical intervention for FNA-indeterminate/mutation-positive nodules. Furthermore, **BRAF** (PPV 100%) and **RAS** (87.5%) proved to be significantly important in the diagnosis of an underlying malignancy in the nodules. This study also depicted a 30% decrease in the need for a second surgery in patients who tested positive for mutations.

**RAS** mutations are found in 10-20% of PTCs, mostly the follicular variant, and 40-50% of FTCs. In addition to malignant nodules, these are also increasingly found in benign adenomas (20-40%) (18). The increased prevalence of these mutations in thyroid carcinoma, follicular adenomas and hyperplastic nodules may also be of prognostic value to predict metastatic disease and, thus, direct the extent of surgical intervention. However, it is not recommended for physicians to integrate mutational analysis for risk stratification into routine practice as not all **BRAF**-positive tumors behave aggressively. The biomarker **PAX8/PPARγ** is a growth inhibitor mutation, which is responsible for cancerous growth, particularly follicular cancers (16). Furthermore, its comparatively higher prevalence in benign follicular
adenomas precludes its use in strictly diagnosing cancerous nodules. Such mutational assays can increase the probability of pre-surgical diagnosis of malignancy, especially in nodules with cytological diagnosis of Bethesda reporting Class V.

3. Immunohistochemical Stains. The immunophenotypical assay is based on the detection of markers expressed by thyrocytes during normal circumstances and altered cell growth. Several molecules have been identified that can improve the accuracy of pre-surgical diagnosis of malignant thyroid nodules. Due to the poor ability of aspiration cytology to distinguish a follicular neoplasm as benign or malignant, attempts were made to find an optimal way to accurately characterize these nodules. To date, various immunohistochemical molecules expressed by thyrocytes have been identified; however, affirmation of their clinical utility is still warranted. A large number of markers, such as CD44, galectin-3, cytokeratin-19 (CK-19) and HBME-1 are identified by measures that can be reliably interpreted by cytopathologists, such as immunochemistry, western blot and polymerase chain reaction (PCR) (19-22).

Evolving Roles of Various Immunohistochemical Biomarkers

CK-19, expressed by normal thyroid epithelium, has shown an increased sensitivity for PTC. It is also expressed in a number of non-neoplastic cases, such as papillary thyroid hyperplasia, Hashimoto’s thyroiditis and benign adenomas. Raphael et al. discovered a diffuse positivity of CK-19 and high-molecular weight cytokeratin in PTC (20). Out of the 116 resected tumors, there were 48 PTCs where these markers were successful in differentiation of PTC from follicular neoplasms and nodular hyperplasia. These neoplasms are difficult to interpret on conventional aspiration cytology. Further successful results were obtained by Casey and colleagues where CK-19 exhibited 100% sensitivity for PTC. HBME-1 also showed 100% sensitivity and 96.7% specificity for the same (19). Galectin-3, believed-to-be expressed only by malignant cells, showed moderate to severe immunohistochemical staining in 24/30 cases of PTC and had the lowest expression in a majority of cases of papillary hyperplasia (sensitivity 100%, specificity 40%). Conversely, Sahoo and colleagues acknowledged that CK-19 was not specific to PTC as 100% of the follicular adenomas showed positive staining with the same biomarker. However, only 20% of the nodules showed a 3+ staining (25%-75% positively stained cells) (21). Therefore, the intensity of these staining methods is also of considerable value in differentiation of tumors. HBME-1 has shown promising results for a wide range of tumors including thyroid tumors. Another study evaluating the diagnostic potential of these biomarkers was performed, in which all cases of PTC and follicular carcinomas showed positivity to HBME-1, while benign tumors were only focally positive with the same antigen. CD15, an oncofetal antigen, showed comparatively lower positivity for PTC and follicular carcinomas with none for high grade malignancy, such as anaplastic carcinomas (22). Other potential markers, such as type 1 receptor for transferrin (TfR1/CD71), p53, CD44, carbonic anhydrase 4 and crystalline α-B have been observed with over-expression in malignant and metastatic tumors of the thyroid (23, 24) that help distinguish these tumors from benign adenomas.

A recently reported newer cytochemical marker, CD117, a type III tyrosine kinase receptor that is variably expressed in several tumors of gastrointestinal tract, has demonstrated its significance in thyroid tumors as well. It showed 100% staining for benign thyroid nodules but lacked immunoreactivity in 89% of PTCs (25). The activating mutation c-KIT, which regulates cell proliferation, apoptosis and chemotaxis, is down-regulated in cancer of the thyroid gland, thus CD117 is not expressed in PTC and metastases from thyroid carcinoma.

Although the diagnostic accuracy of individual markers is not yet definitive, an integrated use of these markers on indeterminate cytological samples can help limit diagnostic surgeries. However, these tests may require laboratory techniques that are time-consuming and prohibitively expensive, which may preclude their use in clinical practice.

Conclusion

Molecular testing, when used in addition to the routine gold-standard FNA cytology, will further improve diagnostic accuracy for indeterminate thyroid nodules. As evidenced by current studies, these ancillary tests are highly beneficial when a panel of biomarkers is used instead of a single marker for pre-surgical diagnosis. The difficulty in differentiation of follicular neoplasia, Hurthle-cell neoplasia and the follicular variant of PTC makes it important to integrate these tests into clinical practice. With the aim to reduce the number of false-negative FNA cytology results and subsequent diagnostic surgeries, molecular testing holds great promise in the management of patients with thyroid nodules suspicious of malignancy. However, further prospective studies are warranted to perfect clinical protocol for this challenging scenario, the impact of which can alter the scope of clinical decisions.

References


Received November 30, 2014
Revised December 11, 2014
Accepted December 15, 2014

1241