Summary Highlights

* 10 - 25% of fine needle aspiration biopsies yield an indeterminate result often labeled as atypia (or follicular lesion) of undetermined significance (AUS/FLUS), follicular lesion of undetermined significance (FLUS) or follicular neoplasm/suspicious for follicular neoplasm (FN/SFN). The risk of malignancy typically varies between 15% and 30% for these categories.

* Though many markers are in development and have been studied in a research setting, two principal tests are currently marketed for use to improve the malignancy risk assessment of "indeterminate" thyroid nodules -- "Rule In" and "Rule Out" tests that attempt to confirm or exclude, respectively, the presence of cancer within a thyroid nodule.

* The Rule In tests assess for the presence of single gene point mutations (BRAF or RAS) or gene rearrangements (RET/PTC, PAX8/PPARγ) which have been shown to increase the ability to predict cancer, while the Rule Out test (Afirma Gene Expression Classifier) utilizes a proprietary gene expression classifier (RNA expression) specifically designed to maximize the ability to define a process as benign.

* None of the presently available tests is associated with a 100% negative or positive predictive value (NPV or PPV). Thus, no currently available molecular test identifies the absence or presence of malignancy in all indeterminate nodules.

* The category of cytologically "indeterminate" nodule (AUS, FLUS, FN/SFN), cytopathology practice patterns, and prevalence of malignancy within the population being tested all impact the NPV and PPV for the tests in question.

* At present, molecular testing is meant to complement and not replace clinical judgment, sonographic assessment, and visual cytopathology interpretation.

* As molecular testing is new and advances in the field are regularly occurring, clinicians need to stay informed as recommendations for use within practice are expected to evolve.
Introduction

Thyroid nodules occur in approximately 50% of the general population and ~5-15% prove to be malignant. Fine needle aspiration biopsy has emerged as the most reliable clinical tool for determining whether a nodule is benign or malignant. However, approximately 10 - 25% of fine needle aspiration biopsies yield an indeterminate result with the risk of malignancy varying from 5% to 75% (typically 15-30%) depending on the indeterminate category according to the Bethesda System for Reporting Thyroid Cytopathology (1). This review focuses on nodules which display atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) and follicular neoplasm/suspicious for follicular neoplasm (FN/SFN), Bethesda categories III and IV, respectively, and not category V, often labeled “suspicious for malignancy”.

Recent advances in thyroid and cancer biology have led to the development and marketing of several tests to assist in determining whether a nodule with indeterminate cytology, AUS/FLUS or FN/SFN, is benign or malignant.

Our goal is to provide a concise review of the commercially available “molecular tests” for thyroid nodule assessment and to provide the clinician with a practical summary of the power and limitations of these tests in the clinical setting.

Available Tests

At present, two main types of molecular tests are marketed for the assessment of thyroid nodules for the presence of cancer. These tests, Veracyte’s Afirma gene-expression classifier, and mutation analysis panels such as Asuragen’s miRInform, Quest Diagnostic’s panel and other similar panels, utilize different methods to assess the risk of malignancy. A third type test from the Cleveland Clinic utilizing serum TSH receptor mRNA levels is also available commercially, but has received less scrutiny in the medical literature.

Technical Aspects

The Afirma gene-expression classifier is based on the gene expression profiles of surgically proven benign and malignant thyroid nodules and evaluates for the presence of a benign gene expression profile. Based on validation studies demonstrating high negative predictive value among nodules with cytology of AUS/FLUS or FN/SFN, it has been employed as a “Rule Out” test to identify nodules that are benign. However, since a substantial number of benign nodules do not have a gene expression profile classified as benign, this test cannot reliably “Rule In” the presence of malignancy.

The Asuragen miRInform mutation analysis panel assesses for BRAF and RAS point mutations as well as common rearrangements of RET/PTC and PAX8/PPARY. It is estimated
that one of these mutations is present in approximately 70% of well-differentiated thyroid cancers. The miRInform assay is considered a “Rule In” assay as nodules harboring these mutations or rearrangements have a high likelihood of cancer giving this test’s high positive predictive value. However, as a significant number of malignant nodules do not harbor one of these genetic markers, miRInform cannot reliably “Rule Out” cancer. Additionally, RAS mutations can be commonly found in benign adenomas. Of note, the components of the miRInform test are available as stand-alone assays or in various combinations by other laboratories, e.g. Quest Diagnostics.

The mutation analysis panels and Afirma test are performed using samples from needle washings collected during fine needle aspiration biopsy. The TSH receptor mRNA test is a serum assay that utilizes quantitative RT-PCR to detect circulating thyroid cancer cells by measuring thyrotropin receptor mRNA in peripheral blood.

A large, prospective, multicenter study was undertaken to validate the clinical utility of the Afirma gene expression classifier (2). It found that a negative result, i.e. a benign gene expression profile, has a NPV ranging from 94- 95% for cytologically AUS/FLUS and FN/SFN lesions (1). Accordingly, because of its high NPV for AUS/FLUS and FN/SFN nodules, the Afirma gene expression classifier has been marketed as a “Rule Out” test with sufficient reliability to defer surgery. It is not recommended for use in nodules with FNA cytology labeled “suspicious for malignancy” given a lower 85% NPV reported in this subgroup. The analytical performance of this test has been validated with respect to sensitivity, specificity, and reproducibility (3). However, as the false negative rate for indeterminate nodules is 5-6%, nodules classified as benign by the gene expression classifier should be followed closely with repeat ultrasound because malignancy is not absolutely excluded. Selective FNA re-assessment or consideration for resection should occur for any significant nodule growth or suspicious change in appearance.

In the case of the Asuragen miRInform and Quest mutation panels, results of validation studies have not been reported or published by the laboratories. However, a large, prospective, non-blinded, single center study has independently validated the utility of testing for markers similar to those used in the miRInform panel in FNA samples with indeterminate cytology (4). This study found that a positive result carries a PPV ranging from 87-95% for predicting cancer. Accordingly, this test is considered to be a “Rule In” test.

Clinical Use

Molecular testing is just one of many factors that must be considered in the evaluation of a thyroid nodule. Patient characteristics, physical examination, thyroid ultrasound and fine-needle aspiration biopsy as well as sound clinical judgment should continue to be applied in a
balanced manner. Molecular testing should only be used to complement and not to replace
cytopathologic evaluation or clinical and imaging assessment.

Molecular testing may be considered in adult patients with thyroid nodules >1 cm in the
presence of a cytological diagnosis of AUS/FLUS or FN/SFN. Clinical validation has not been
performed for thyroid nodules < 1 cm. Of note, neither Afirma nor miRInform testing has been
validated in the pediatric population although preliminary reports suggest promise for use of
mutational markers such as BRAF. Molecular testing is not recommended in instances where
the results are not expected to alter the decision to proceed with surgery or the extent of surgery.
Examples include nodules with a cytological reading of benign, malignant or suspicious for
malignancy, larger nodules (especially > 4-5 cm), presence of compressive manifestations, a
high suspicion of malignancy based on clinical or ultrasonographic features, or patient
preference for or against surgery.

As a “Rule Out” test, the Afirma gene expression classifier has been reported to obviate
the need for surgical excision in ~50% of cases (5). Preliminary cost analysis has suggested
that such an approach may be cost effective based on the stated study parameters (6). However,
the validity of these findings is tempered by the fact that the statistical estimates used for this
cost-effectiveness analysis were based upon preliminary data and not data specifically derived
from the validation study. Moreover, categorization of a nodule as “suspicious” by the Afirma
gene-expression classifier carries a cancer risk of only 38%, so clinicians and patients need to
understand that this result does not establish a diagnosis of cancer (2).

A cost analysis has suggested that use of the mutation panel alone might also be cost
effective by reducing the need for completion thyroidectomy in comparison to lobectomy ±
isthmusectomy (7). At present, there are no data assessing the cost-effectiveness of the
combined use of the Afirma and Asuragen tests.

The importance of obtaining an adequate sample for both FNA cytology and molecular
testing cannot be over emphasized. Collecting samples for cytopathology and molecular testing
during the same procedure may help ensure that these tests are performed on cells which
derive from the same nodule. Alternatively, the second FNA for molecular analysis can be
performed at a later date after an "indeterminate" cytology diagnosis is received. One should
follow the procedures recommended by the testing laboratory regarding the number of needle
passes required and the appropriate processing and packaging of the sample.

The Veracyte Corporation markets the Afirma test in a way that dictates the diagnostic
paradigm that many clinicians must follow. Outside of select academic institutions that have
been allowed to utilize their own internal cytology assessment given their high volume and
internal expertise, four FNA passes are required for each nodule: two for cytological
examination to be performed by Thyroid Cytopathology Partners (TCP) and the other two for
The miRInform test is designed to be applied to patients harboring thyroid nodules with "indeterminate" classification of malignant risk by cytologic assessment. Therefore, the specimen for molecular analysis can be collected at the time of initial FNA for cytology and sent for analysis only if the cytology reading is AUS/FLUS or FN/SFN. The miRInform protocol specifies one dedicated FNA pass for molecular analysis. Alternatively, the second FNA for molecular analysis can be performed later if the first FNA yields an "indeterminate" cytologic diagnosis. Presently, no consensus exists for use of mutation testing results as a guide for surgical decision making. However, when the cytologic interpretation falls into the AUS/FLUS or FN/SFN category, the presence of a mutation such as BRAF might well justify consideration of total thyroidectomy over hemi-thyroidectomy. Based on available data, it is still to be determined whether mutational results should play any role in determining the extent of lymph node dissection.

The need for molecular diagnosis only occurs, in the minority of cases in which the cytological diagnosis is "indeterminate." Therefore, one can consider obtaining the additional sample for molecular diagnosis only in the ~10-25% of cases where this is pertinent. However, a repeat FNA procedure is associated with extra time and effort by both the physician and patient plus additional cost. Use of material from separate FNA passes, especially from procedures done at different times, raises the possibility of specimens not originating from the same lesion examined by the cytopathologist, although the risk of such occurring may be low in association with ultrasound guided aspiration.

Automatically sending all FNA samples for this type of molecular testing is not recommended because in the majority of cases the cytological result will provide actionable diagnostic guidance (benign or malignant) and molecular testing is not required. Afirma testing is not validated for assessment of benign nodules and should be reserved for cases of AUS/FLUS, FN/SFN. With a NPV of 85%, Afirma GEC testing is also not recommended for evaluation of lesions suspicious for malignancy nor for nodules with cytology consistent with cancer. In the case of mutational testing, the false positive rate ranges between 5-13% primarily due to the presence of mutated RAS genes in some follicular adenomas (4).
“thyroid FNA cytomorphology with molecular reflex” as an available test. In their diagnostic paradigm, nodules with indeterminate cytology are subjected to analysis for a selection of common mutations and rearrangements from the list above and are applied selectively based on the diagnostic cytology category.

Limitations

The ultimate goal of molecular testing is the accurate assessment of a thyroid nodule as being benign or malignant prior to surgery. At present, molecular testing should not supplant clinical and ultrasound assessment because none of the available molecular tests can achieve this goal. Available markers that utilize a “Rule-In” cancer method report that between 6-28% of cases of malignancy lack a mutation of one of the included markers thereby leading to a false negative result. This testing paradigm has not yet been validated in a prospective, multi-center clinical trial. Because testing for the BRAF mutation appears to be associated with near 100% specificity, the occurrence of a false-positive result does not appear to be a concern. However, RAS mutations have been reported to be present in both follicular cancer and benign follicular adenomas although there is ongoing debate as to whether or not the latter represents a pre-malignant lesion that should be removed anyway.

The strength of cytology support directly impacts the reliability of molecular testing because it determines the pretest probability of malignancy. The pretest probability of malignancy in turn will influence in an inverse manner the NPV of both types of molecular testing. Accordingly, the potential variance of NPV might well impact physician reliance on test results for determining the need for surgery.

Finally, both physician and patient need to be aware that as molecular testing is in its infancy, continued research and advances may change our understanding of how best to utilize molecular markers in the clinical setting. In addition, this statement has not covered other markers and techniques (galectin-3 immunohistochemical staining, HBME-1, microRNAs, etc) that hold a potential role in this field (8). Prospective multicenter studies are required for validation of all these tests used either singly or in tandem with each other.

References


