Clinical Validation of the ThyroSeq v3 Genomic Classifier in Thyroid Nodules With Indeterminate FNA Cytology

Yuri E. Nikiforov, MD, PhD [1]; and Zubair W. Baloch, MD, PhD[2]

INTRODUCTION

Thyroid nodules are common and can be encountered in up to 60% of individuals aged >60 years by ultrasound, and in approximately 5% to 7% by palpation.[1-3] The incidence of cancer in these thyroid nodules remains low and the majority are benign.[4-6] An accurate distinction between benign and cancerous thyroid nodules is important to offer the most appropriate treatment to patients. Since its introduction in the United States in the 1980s, fine-needle aspiration (FNA) biopsy, which currently is performed in the majority of cases with ultrasound guidance, represents the most effective triage procedure to reliably distinguish benign from malignant thyroid nodules in approximately 60% to 80% of cases,[7,8] whereas the remaining nodules are diagnosed as "indeterminate for malignancy" on FNA, mainly due to the lack of specific cytomorphologic features needed for a definitive diagnosis.[9]

In the updated version of The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC), the indeterminate cytologic diagnoses based on cytologic features are categorized as either atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) (TBSRTC category III), follicular or oncocytic (Hurthle cell) neoplasm/suspicious for a follicular or oncocytic (Hurthle cell) neoplasm (FN/SFN) (TBSRTC category IV), or suspicious for malignancy (TBSRTC category V).[9] TBSRTC categories III and IV are most common because each of these diagnoses is rendered in approximately 10% of all FNA biopsies.[7]

Molecular testing of thyroid FNA specimens has been developed with the primary goals of resolving the uncertainty of indeterminate cytology and avoiding the diagnostic surgeries generally recommended for these patients.[8] With the advent of next-generation sequencing technology and the accelerated discovery of thyroid cancer markers, molecular testing for indeterminate thyroid nodules has become increasingly common and is widely used as an adjunct diagnostic tool for thyroid nodules diagnosed as indeterminate on FNA.

Evolution of the ThyroSeq Test

ThyroSeq (which stands for "*Thyro*id *Seq*uencing") represents one of the molecular approaches to thyroid nodules that is based on the detection of thyroid cancer–associated molecular alterations in cell DNA and RNA. It has a substantial history of evolution, with its first and most limited version introduced into routine clinical practice in 2007 as a 7-gene panel (ThyroSeq v0).[10,11] The next versions of the test migrated to the next-generation sequencing platforms and included a 13-gene panel that was launched in 2013 (ThyroSeq v1)[12] and

Corresponding author: Yuri E. Nikiforov, MD, PhD, Department of Pathology, University of Pittsburgh, 3477 Euler Way, Rm 8031, Pittsburgh, PA 15213; nikiforovye@upmc.edu

[1]Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania; [2]Department of Pathology and Laboratory Medicine, Hospital of the University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania.

Received: February 1, 2019; **Accepted:** February 2, 2019

Published online February 27, 2019 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.22112, wileyonlinelibrary.com

a 56-gene panel launched in 2014 (ThyroSeq v2).[13,14] The latter version of this test not only used DNA and RNA for detecting point mutations/indels and gene fusions found in approximately 90% of papillary carcinomas,[15] but also used a limited gene expression panel to monitor the cellularity and cell lineage of the tested samples. This allowed an overall high positive predictive value (PPV) and negative predictive value (NPV) for cancer detection in TBSRTC categories III and IV nodules and facilitated the detection of parathyroid nodules and medullary carcinomas.[14,16] More recent advances in the genetic science related to the molecular mechanisms of less common types of thyroid cancer, and particularly Hurthle cell carcinomas,[17,18] offered an opportunity for the even more comprehensive and accurate detection of all main types of thyroid cancer, which led to the creation of the latest version of the ThyroSeq test.

ThyroSeq v3 Genomic Classifier

The ThyroSeq v3 genomic classifier (GC) is the most recent and advanced version of the test and was launched for clinical use toward the end of 2017/beginning of 2018. The ThyroSeq v3 GC uses next-generation sequencing technology to analyze 112 genes, providing information regarding >12,000 mutation hotspots and >120 fusion types. It detects 5 different classes of genetic alterations: 1) mutations (single nucleotide variants); 2) insertions and deletions; 3) gene fusions; 4) gene expression alterations; and 5) copy number alterations. Compared with the previous version of the test, the main advances of the ThyroSeq v3 GC include the analysis of a larger number of genes, mutation hotspots, and gene fusions; the detection of DNA copy number alterations, which is particularly important for Hurthle cell and follicular carcinomas; and the use of a GC.[19] This classifier is based on the algorithmic analysis of all identified genetic alterations and their level (mutant allele frequency) to report the test result as negative (including currently negative) or positive, thereby providing a more specific cancer probability assessment for nodules with a positive test result. The test has been validated with respect to its analytical performance in a study of 238 resected tissue samples and 175 FNA samples.[19] It has been optimized for fresh thyroid FNA samples collected into nucleic acid preservative solution and for formalin-fixed and paraffin-embedded sections of thyroid tissue or cytology cell block.[19]

Clinical Validation of the ThyroSeq v3 GC

A prospective, double-blinded, multicenter study recently published by Steward et al reported the results of clinical validation of the ThyroSeq v3 GC in thyroid nodules with indeterminate FNA cytology.[20] This study initially recruited 782 patients with 1013 thyroid nodules at 9 institutions in the United States and 1 institution in Singapore. Samples were collected by either rinsing the residual material in the aspiration needle from all passes or by placing a dedicated pass into the collection tube containing nucleic acid preservative solution. Of all the samples, 257 met all study requirements, and were diagnosed as one of the indeterminate TBSRTC categories (ie, category III, IV, or V), had informative ThyroSeq results, and had a final surgical pathology diagnosis. Surgical pathology results were reviewed centrally by a panel of expert thyroid pathologists. However, by study design, cytology smears were not reviewed centrally so that the study represented the "real-world" experience of thyroid nodules diagnosed as indeterminate cytology being evaluated by this test. The study was double-blinded: molecular testing personnel were blinded to all FNA and histopathology diagnoses, and both cytologists and pathologists were blinded to molecular testing results. In addition, pathologists were blinded to the diagnoses made by local and other panel pathologists. The discrepant cases were reviewed by the panel to arrive at a consensus diagnosis. There was no post-unblind sample exclusion in the current study.

The primary outcome of the study was assessing the sensitivity, specificity, NPV, and PPV of the ThyroSeq v3 GC in predicting the histopathologic diagnosis of a benign nodule versus cancer/noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) in thyroid nodules with TBSRTC category III (AUS/FLUS) and category IV (SFN/FN) cytology results. Although not classified as carcinoma, NIFTP is considered to be a precancer, borderline tumor that requires surgical excision,[21,22] and therefore it was grouped with cancer for data analysis.

The final study cohort included 154 thyroid nodules and 93 thyroid nodules diagnosed as AUS/FLUS and SFN/FN, respectively. As shown in Table 1,[20] the ThyroSeq V3 GC demonstrated a combined sensitivity of 94% and a specificity of 82% for cases diagnosed as AUS/FLUS and SFN/FN. With a 28% prevalence of

TABLE 1. Performance of the ThyroSeq v3 GC in Cytologically Indeterminate Thyroid Nodules Based on the Results of a Prospective Multicenter Study[20]

ThyroSeq Result	Cancer plus NIFTP n = 35	Benign n = 119	Test Performance (95% CI)
Performance in Bethesda Category III Nodules (154 Nodules; Disease Prevalence, 23%)			
Positive	32	18	Sensitivity: 91% (77%-97%) Specificity: 85% (77%-90%)
Negative	3	101	NPV: 97% (92%-99%) PPV: 64% (50%-77%)
Performance in Bethesda Category IV Nodules (93 Nodules; Disease Prevalence, 35%)			
Positive	32	15	Sensitivity: 97% (85%-100%) Specificity: 75% (63%-84%)
Negative	1	45	NPV: 98% (89%-100%) PPV: 68% (54%-80%)
Performance in Bethesda Category III and Category IV Nodules (247 Nodules; Disease Prevalence, 28%)			
Positive	64	33	Sensitivity: 94% (86%-98%) Specificity: 82% (75%-87%)
Negative	4	146	NPV: 97% (93%-99%) PPV: 66% (56%-75%)

Abbreviations: 95% CI, 95% confidence interval; GC, genomic classifier; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; NPV, negative predictive value; PPV, positive predictive value.

TABLE 2. ThyroSeq GC Performance in Specific Histopathologic Types of Thyroid Lesions Based on the Results of a Prospective, Multicenter Study[20]

Histopathology Diagnosis	No. of Nodules (%)	Test Positive	Test Negative	Correctly Classified (95% CI)
Benign				
Hyperplastic follicular cell nodule	95 (37%)	11	84	88% (80%-93%)
Hyperplastic Hurthle cell nodule	5 (2%)	0	5	100% (57%-100%)
Follicular adenoma	47 (18%)	10	37	79% (65%-88%)
Hurthle cell adenoma	34 (13%)	13	21	62% (45%-76%)
NIFTP	11 (4%)	11	0	100% (74%-100%)[a]
Malignant				
Papillary thyroid carcinoma	49 (19%)	45	4	92% (81%-97%)
Follicular thyroid carcinoma	4 (2%)	3	1	75% (30%-99%)
Hurthle cell carcinoma	10 (4%)	10	0	100% (72%-100%)
Medullary thyroid carcinoma	1 (0.5%)	1	0	100% (5%-100%)
Metastatic carcinoma[b]	1 (0.5%)	1	0	100% (5%-100%)
Total	257 (100%)	105	152	85% (80%-89%)

Abbreviations: 95% CI, 95% confidence interval; GC, genomic classifier; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features.
[a]Considering positive test result for NIFTP as the correct classification.
[b]Metastatic renal cell carcinoma.

cancer/NIFTP, the test demonstrated a NPV of 97% among these samples. There were 5 false-negative test results in the entire study cohort, representing 4 papillary carcinomas and 1 minimally invasive follicular carcinoma (Table 2).[20] All of the missed cancers were intrathyroidal, low-risk, and low-stage tumors with no vascular invasion, extrathyroidal extension, or other aggressive features.

The frequency of a negative test result or "benign call rate" in the current study was 61% for thyroid FNA specimens diagnosed as AUS/FLUS and SFN/FN. A residual cancer risk in these test-negative nodules was 3%. According to the National Comprehensive Cancer Network guidelines, if molecular testing in combination with imaging and clinical information predicts a malignancy risk of ≤5%, which is comparable to the malignancy risk of benign FNA cytology, clinical management may include observation.[23] It is important to note that although the sensitivity and specificity are independent

TABLE 3. Probability of Cancer/NIFTP in Specific Molecular Alteration Groups Based on the Results of a Prospective, Multicenter Study[20]

Alteration Group	Most Prevalent Specific Alterations (No.)	Prevalence in Test-Positive Samples, No.	Probability of Cancer/NIFTP, %	Most Prevalent Cancer Type/NIFTP, %
High-risk	*TERT* and *HRAS* (1) *TP53* and *MEN1* (1)	2 (2%)	100	Papillary carcinoma, 50% Follicular carcinoma, 50%
BRAF-like	*BRAF* V600E (9)	13 (12%)	100	Classic papillary carcinoma, 92%
RAS-like	*NRAS* (21)	60 (57%)	62	Follicular variant papillary carcinoma, 22%
	HRAS (18)			NIFTP, 15%
Copy number alterations	Copy number alterations	22 (21%)	59	Hurthle cell carcinoma, 32%
Gene expression alterations	Gene expression alterations	8 (8%)	75	Classic papillary carcinoma, 37%

Abbreviation: *MEN1*, multiple endocrine neoplasia type 1; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; *TERT*, telomerase reverse transcriptase.

characteristics of every test, the NPV and PPV can vary according to the prevalence of disease in the studied population. By using the Bayes theorem and the sensitivity and specificity demonstrated in the multicenter study, the NPV of the ThyroSeq v3 GC was modeled for varying disease prevalences. It demonstrated that the NPV of the ThyroSeq GC is expected to remain at ≥95% in populations with a disease prevalence of up to 40% for thyroid nodules classified as AUS/FLUS and up to 60% for nodules classified as SFN/FN.[20] This is well within the range of cancer/NIFTP prevalence expected in these TBSRTC diagnostic categories.[7,9,24] Therefore, a ThyroSeq V3 GC–negative result may prevent diagnostic surgery for patients with thyroid nodules diagnosed as either AUS/FLUS or SFN/FN in most clinical settings.

The ThyroSeq V3 GC demonstrated a fairly high PPV of 66% in thyroid FNA specimens diagnosed as AUS/FLUS and SFN/FN. As shown in Table 2,[20] the test correctly identified all Hurthle cell (oncocytic) carcinomas and NIFTP nodules as positive. There were 34 test-positive samples diagnosed by pathology as benign after surgery. It is interesting to note that of these samples, 32 (94%) had ≥1 clonal molecular alterations identified in a large percentage of cells, which indicate that these nodules represent neoplasia, not hyperplasia.

The secondary outcome of the current clinical validation study was to evaluate the prediction of cancer/NIFTP by specific genetic alterations. The results of the study demonstrated that among 105 test-positive samples, the probability of a histological diagnosis of cancer/NIFTP varied by 57% to 100% based on a specific group of genetic alterations (Table 3).[20] The high-risk (*TERT* and *TP53*) and *BRAF* ^{V600E}-like mutations conferred a cancer probability of 100%. Furthermore, the *BRAF* ^{V600E}-like mutations were found to be strongly predictive of classic papillary carcinoma in these nodules. The *RAS*-like mutations were associated with a 57% probability of cancer/NIFTP and were predictive of the follicular variant of papillary carcinoma. The finding of copy number alterations conferred a 59% probability of cancer/NIFTP, and this group was found to be enriched in Hurthle cell carcinomas. Finally, gene expression alterations were associated with a 75% probability of cancer; this finding was made among a heterogeneous pathology group that included, in addition to papillary carcinomas, nonfollicular thyroid cell tumors such as medullary carcinomas and metastatic/secondary tumors. The latter were reliably predicted by the classifier based on the expression pattern of cell lineage markers. A more granular prediction of cancer probability and tumor phenotype by specific genetic alterations may help to inform the more individualized management of patients with indeterminate cytology nodules.

Benign Call Rate and the Avoidance of Diagnostic Surgeries

The benign call rate is a valuable characteristic for "rule-out tests" that allows the estimation of the percentage of patients with indeterminate cytology nodules who can avoid undergoing diagnostic surgeries and be managed similarly to patients with benign cytology results. The study by Steward et al found a ThyroSeq V3 GC benign call rate of 61% of cases diagnosed as AUS/FLUS and SFN/FN, indicating that up to 61% of surgeries may be avoided for patients with these nodules.[20] This includes 82% of all histologically benign thyroid nodules with diagnoses of AUS/FLUS and SFN/FN. This study also compared the performance of ThyroSeq V3

TABLE 4. Study Characteristics and Performance of the ThyroSeq GC and Afirma GSC in Bethesda Categories III and IV Indeterminate Cytology Thyroid Nodules[a]

	ThyroSeq GC[20]	Afirma GSC[25]
Study type	Multicenter, prospective, double-blind	Multicenter, retrospective, double-blind
Total no. of samples	247	191
Median nodule size on ultrasound (range), cm	2.1 (0.5-7)	2.6 (1.0-9.1)
Disease prevalence, %	27.5	23.7
Sensitivity (95% CI)	94.1% (86%-98%)	91.1% (79%-98%)
Specificity (95% CI)	81.6% (75%-87%)	68.3% (60%-76%)
NPV	97.3% (93%-99%)	96.1% (90%-99%)
PPV	65.9% (56%-75%)	47.1% (36%-58%)
Benign call rate	61%	54%
Avoidable surgeries for patients with histologically benign nodules with indeterminate cytology	82%	68%

Abbreviations: 95% CI, 95% confidence interval; GC, genomic classifier; GSC, genomic sequencing classifier; NPV, negative predictive value; PPV, positive predictive value.

[a]Adapted from 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 [published online ahead of print November 8, 2018]. *JAMA Oncol*. 10.1001/jamaoncol.2018.4616.[20]

GC with that of another commonly used test for thyroid nodules, the Afirma GSC (Veracyte Inc, South San Francisco, California), as reported elsewhere.[25] Based on these 2 multicenter studies of approximately similar size, the benign call rate for the ThyroSeq V3 GC was 61% compared with a benign call rate of 54% for the Afirma GSC, with 82% of histologically benign thyroid nodules classified as negative by the ThyroSeq V3 GC compared with 68% by Afirma GSC (Table 4).[20,25]

It is interesting to note that a recent study by Ohori et al evaluated the benign call rate of the ThyroSeq V3 GC during the routine clinical use of molecular testing at the University of Pittsburgh Medical Center.[26] This study included 224 consecutive cases with indeterminate cytology for which the benign call rate was 74%, which is even higher than that found in the multicenter study. Although the study was retrospective and represented a single institution, it provided evidence that the benign call rate of the ThyroSeq v3 GC found in the multi-institutional study[20] should be expected to be higher when the test is used routinely. Further studies are required to evaluate the "real-world" performance of this test in different clinical populations to fine-tune its application for the management of patients with thyroid nodules.

FUNDING SUPPORT

Supported in part by the P50CA097190 Head and Neck Cancer SPORE grant to the University of Pittsburgh from the National Institutes of Health.

CONFLICT OF INTEREST DISCLOSURES

Yuri E. Nikiforov holds intellectual property rights related to ThyroSeq (licensee is CBLPath/Sonic Healthcare) and will receive royalties associated with the commercial use of ThyroSeq.

REFERENCES

1. Gharib H. Changing trends in thyroid practice: understanding nodular thyroid disease. *Endocr Pract*. 2004;10:31-39.
2. Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med*. 1993;328:553-559.
3. Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest*. 2009;39:699-706.
4. Frates MC, Benson CB, Doubilet PM, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin Endocrinol Metab*. 2006;91:3411-3417.
5. Kim DL, Song KH, Kim SK. High prevalence of carcinoma in ultrasonography-guided fine needle aspiration cytology of thyroid nodules. *Endocr J*. 2008;55:135-142.
6. Brito JP, Yarur AJ, Prokop LJ, McIver B, Murad MH, Montori VM. Prevalence of thyroid cancer in multinodular goiter versus single nodule: a systematic review and meta-analysis. *Thyroid*. 2013;23:449-455.
7. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol*. 2012;56:333-339.
8. Durante C, Grani G, Lamartina L, Filetti S, Mandel SJ, Cooper DS. The diagnosis and management of thyroid nodules: a review. *JAMA*. 2018;319:914-924.
9. Ali SZ, Cibas ES. The Bethesda System for Reporting Thyroid Cytopathology. 2nd ed. New York: Springer; 2018.
10. 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;94:2092-2098.
11. Nikiforov YE, Ohori NP, Hodak SP, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metab*. 2011;96:3390-3397.
12. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. *J Clin Endocrinol Metab*. 2013;98:E1852-E1860.
13. 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;120:3627-3634.
14. Nikiforov YE, Carty SE, Chiosea SI, et al. Impact of the multigene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid*. 2015;25:1217-1223.
15. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;159:676-690.
16. Cho M, Oweity T, Brandler TC, Fried K, Levine P. Distinguishing parathyroid and thyroid lesions on ultrasound-guided fine-needle aspiration: a correlation of clinical data, ancillary studies, and molecular analysis. *Cancer Cytopathol*. 2017;125:674-682.

17. Ganly I, Makarov V, Deraje S, et al. Integrated genomic analysis of Hurthle cell cancer reveals oncogenic drivers, recurrent mitochondrial mutations, and unique chromosomal landscapes. *Cancer Cell*. 2018;34:256-270.e5.
18. Gopal RK, Kubler K, Calvo SE, et al. Widespread chromosomal losses and mitochondrial DNA alterations as genetic drivers in Hurthle cell carcinoma. *Cancer Cell*. 2018;34:242-255.e5.
19. Nikiforova MN, Mercurio S, Wald AI, et al. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer*. 2018;124:1682-1690.
20. 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 [published online ahead of print November 8, 2018]. *JAMA Oncol*. doi: 10.1001/jamaoncol.2018.4616.
21. Baloch ZW, Seethala RR, Faquin WC, et al. Noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP): a changing paradigm in thyroid surgical pathology and implications for thyroid cytopathology. *Cancer Cytopathol*. 2016;124:616-620.
22. 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;27:481-483.
23. NCCN Clinical Practice Guidlines in Oncology: *Thyroid Carcinoma*. National Comprehensive Cancer Network; 2017.
24. Ohori NP, Schoedel KE. Variability in the atypia of undetermined significance/follicular lesion of undetermined significance diagnosis in The Bethesda System For Reporting Thyroid Cytopathology: sources and recommendations. *Acta Cytol*. 2011;55:492-498.
25. Patel KN, Angell TE, Babiarz J, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg*. 2018;153:817-824.
26. Ohori NP, Landau MS, Carty SE, et al. Benign call rate and molecular test result distribution of ThyroSeq v3 [published online ahead of print December 18, 2018]. *Cancer Cytopathol*. doi: 10.1002/cncy.22088.