The Role of Molecular Diagnostics in the Management of Indeterminate Thyroid Nodules

Joanna Klubo-Gwiezdzinska[1] and Leonard Wartofsky[2]

[1]National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20814; and [2]Thyroid Cancer Unit, MedStar Health Research Institute, Washington, District of Columbia 20010

In a recent issue of the Journal, Livhits *et al.* (1) addressed the relative performance of two molecular diagnostic techniques for thyroid nodules indeterminate on fine-needle aspiration (FNA) cytology. The study was designed as a classic "real life" effectiveness trial, and as such was tailored toward external validation of the most commonly clinically applied molecular tests. Compared with strictly controlled efficacy trials characterized by higher internal validation achieved by strict design, *e.g.*, blinding of the pathologist to the results of the molecular tests, the availability of the final pathology for all enrolled patients, the advantage of effectiveness—"real life" study design—is that it provides an evaluation of the intervention in a routine practice experience. Consequently, generalizability of its findings is increased, as the study accounts for patient-provider-system-level factors. The study aim was particularly important, since >500,000 FNAs are performed yearly in the United States, with around 100,000 nodules characterized by indeterminate cytology (2). Prior to the availability of molecular diagnostics, patients were referred for lobectomy to achieve a specific pathologic diagnosis. This approach is still followed widely throughout the world today and results in a large number of unnecessary surgeries, perhaps as many as 50% of surgeries for benign nodules. As a result, overtreatment is associated with substantial side effects and incremental cost contributing to the cost of treating and monitoring of patients with thyroid cancer, which reached $1.6 billion in the United States in 2013 (3).

Analysis of molecular signature of thyroid nodules with indeterminate cytology has been evolving over the past two decades. Initially, the only available test was PCR-based screening for *BRAFV600E* mutation, characterized by high specificity and positive predictive value (PPV) for detection of thyroid cancer, but very low sensitivity. The sensitivity was slightly improved by the introduction of a seven-gene mutation panel (ThyGenX) for testing for the most common genetic alterations present in up to 70% of thyroid cancers, namely *BRAFV600E*, *NRAS*, *HRAS*, and *KRAS* mutations and gene fusions *RET/PTC1, RET/PTC3*, and *PAX8-PPARG* (4–7). However, despite slightly increased sensitivity compared with single gene mutation testing, the performance of this assay was still characterized by very low negative predictive value (NPV), precluding its use as a rule-out test. Conversely, the high specificity and high PPV of this assay has been validated in several studies and the seven-gene mutation panel has been used as a rule-in test in clinical practice (4–7). To increase the accuracy and utility of this panel, combination with a panel of 10 miRNA markers was introduced (ThyGenX/ThyraMIR) (8). Based on a small multi-institutional study, the NPV of the assay was 94% and PPV was 74% (8). However, ThyGenX/ThyraMIR has not been appropriately validated in a large prospective study.

An alternative method applied to stratification of thyroid nodules into high vs low cancer risk used a combination of the ultrasonographic features with the molecular signature of the nodules. De Napoli *et al.* (9) documented as high as a 100% PPV for the combination of high risk ultrasonographic features such as hypoechogenicity, microcalcifications, irregular margins, and taller than wide shape of the nodules, with the presence of the most commonly observed thyroid cancer mutations in *BRAF* and *NRAS* genes (9). However, the diagnostic utility of this approach was tested on a population with a

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA
Received 17 May 2018. Accepted 12 July 2018.
First Published Online 18 July 2018

Abbreviations: FNA, fine-needle aspiration; NPV, negative predictive value; PPV, positive predictive value.

relatively high cancer prevalence of 35%, thus increasing the likelihood for a high PPV. The shortcoming of this approach was a low sensitivity with associated low NPV precluding its use as a rule out test.

The Afirma assay by Veracyte was introduced in 2012 as a response to a need for a good rule-out test (10). The assay was designed to analyze the expression profiles of 142 genes, based on a proprietary algorithm that classified the nodules as either benign or suspicious for malignancy. Based on several clinical validation studies performed to date, Afirma serves as a relatively good rule-out test with high sensitivity and NPV ranging from 75% to 100% (9–11). Unfortunately, the specificity is relatively low and as such is associated with a PPV ranging only between 14% and 44%, limiting its use as a rule-in test (10–12). Livhits *et al.* (1) consistently documented a sensitivity and NPV of 100%, but low specificity and a PPV of 15.8% and 38.5%, respectively. Of note, NPV and sensitivity of the assay described in this effectiveness trial might be lower. The estimate of false negative results in this study design was impossible, as most patients with benign or negative molecular tests were treated conservatively. Nevertheless, the results are valid and generalizable as they resemble a "real life experience," where a nonsurgical approach is generally preferred and implemented in patients in whom a molecular signature is suggestive of a benign lesion.

In 2014, the Cancer Genome Atlas, based on DNA and RNA sequencing data, identified genetic alterations in 97% of papillary thyroid cancers (13). The Cancer Genome Atlas formed a basis for the construction of thyroid cancer-tailored next generation sequencing panels designed by Dr. Yuri Nikiforov and colleagues at the University of Pittsburgh. The rationale for utilizing the next generation sequencing panels platforms was to improve diagnostic accuracy. The increased number of mutations examined should lead to a higher assay sensitivity and be associated with a higher NPV, whereas an anticipated detection of cancer-specific mutations would result in maintaining relatively high specificity and PPV. An initial custom panel (ThyroSeq) was designed to target 12 cancer genes with 284 mutational hot spots (14). The second panel, ThyroSeq v2, included 56 thyroid-related genes and analyzed for point mutations and small insertions/deletions in 14 genes and 42 types of gene fusions found in 90% of thyroid cancers. In addition, ThyroSeq v2 evaluated the expression levels of 16 genes to provide an internal control for an adequacy of cellularity in FNA material, as well as to assure appropriate screening for medullary thyroid cancer by measurement of calcitonin expression and parathyroid adenomas by assessing parathyroid hormone expression (15, 16). ThyroSeq v2 has been validated as characterized by high sensitivity and NPV of 96% to 97% with reasonably high specificity and a PPV of 77% to 83% (15–17), enabling its use as both a rule-out and a rule-in test. And Livhits *et al.* (1) reported that ThyroSeq v2 outperformed Afirma in its specificity and PPV, thereby affecting clinical management and leading to conservative rather than a surgical therapeutic approach in a larger number of patients. However, the study is limited by its analysis of two different cohorts of patients, thus rendering impossible the direct comparison of the accuracy of Afirma and ThyroSeq v2 in the same nodules. Moreover, it is important to recognize that although the sensitivity and specificity of any diagnostic test depends only on test performance, the NPV and PPV depend on the prevalence of disease in the tested population. Thus, in a recently published study analyzing a population with cancer prevalence of 16%, the NPV of ThyroSeq v2 remained high at 96%, but the PPV was only 22% (18). One of the reasons for low PPV in mutation positive samples is the well-known phenomenon of the presence of certain mutations in benign lesions. Indeed, there are mutations associated with close to 100% cancer risk such as *BRAFV600E* mutation or *PPARG*, *NTRK1*, *NTRK3*, and *ALK* fusions, whereas the presence of *RAS*, *PTEN*, and *EIF1AX* mutations or *THADA* fusions entails a significantly lower risk of cancer, and *GNAS* mutation is almost exclusively associated with benign lesions (15, 16). This phenomenon has been recognized in the newest version of the ThyroSeq v3 (112-gene test). ThyroSeq v3 was designed to increase sensitivity by adding recently discovered genetic markers related to thyroid nodules and cancer, including the copy number alteration observed in up to 7% of cancers (19). Most importantly, this version of the test introduces a scoring system to account for molecular signatures typical for cancer and for benign adenomas. Another advantage of ThyroSeq v3 is the inclusion of benign and malignant Hurthle cell lesions thereby increasing the accuracy of classification of Hurthle cell tumors (19). Based on the analysis of a training set and validation set, Nikiforov *et al.* (19) found that the test is characterized by a sensitivity of 94% to 98% and specificity of 81.8% to 89.4% (19). A new version of Afirma has also been recently introduced and this genomic sequencing classifier demonstrated a sensitivity of 91% and specificity of 68% (20). Both of the latter tests need to be validated in prospective cohort studies.

Finally, the recently introduced Rosetta GX test is based exclusively on detection of 24 miRNA markers. Based on a retrospective study involving 150 nodules with indeterminate cytology, its NPV of 92% and PPV of only 43% suggests its potential utility as a rule-out test with limited utility as a rule-in test. Moreover, its performance in Hurthle cell lesions remains unknown (21). It is worthwhile

to underscore that the performance of all of the above mentioned molecular tests is also unknown for noninvasive follicular thyroid neoplasm with papillary-like nuclear features—NIFTP, as it has been categorized previously by pathologists as either benign or malignant noninvasive encapsulated follicular variant of papillary thyroid cancer.

Given the above-mentioned limitations and the lack of appropriate independent validation of the most recently developed molecular tests, current American Thyroid Association guidelines suggest patient-tailored individualized decision making for patients with thyroid nodules with indeterminate cytology. Recommendation 13 states: "If molecular testing is being considered, patients should be counseled regarding the potential benefits and limitations of testing, and about the possible uncertainties in the therapeutic and long-term clinical implications of results" (22).

Another potential limitation of the routine utilization of molecular diagnostics is their relatively high cost. However, that cost needs to be balanced against the cost of unnecessary surgery and potential postsurgical complications, which are likely, given that the majority of thyroidectomies in the United States are performed by low volume surgeons, with 51% of surgeons performing just one thyroidectomy per year (23). Lee *et al.* (24), in an analysis published 4 years ago, and as such not including the newest molecular tools, concluded that Afirma gene expression profiling combined with a seven-gene mutation panel was associated with beneficial cost-effectiveness in the US health care system, but not in the Canadian system (24).

It is clear that the molecular signature of thyroid nodules may serve as one of the useful tools in guiding decision making regarding a conservative vs surgical approach to management. Interpretation of the results of molecular screening of cytologically indeterminate nodules needs to account for cancer prevalence at a given institution, as low cancer prevalence increases NPV of the test, whereas high cancer prevalence increases the PPV of the test. Appropriate clinical judgment and individualized risk stratification is best based on patient age, presence of comorbidities, ultrasonographic features of the nodules, and determination of molecular signature to identify candidates for active surveillance or surgery.

Acknowledgments

Correspondence and Reprint Requests: Leonard Wartofsky, MD, MACP, Georgetown University School of Medicine, MedStar Health Research Institute, 110 Irving Street, NW, Washington, District of Columbia 20010-2975. E-mail: Leonard.Wartofsky@medstar.net.

Disclosure Summary: The authors have nothing to disclose.

References

1. Livhits MJ, Kuo EJ, Leung AM, Rao J, Levin M, Douek ML, Beckett KR, Zanocco KA, Cheung DS, Gofnung YA, Smooke-Praw S, Yeh MW. Gene expression classifier vs targeted next-generation sequencing in the management of indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2018;**103**(6):2261–2268.
2. Sosa JA, Hanna JW, Robinson KA, Lanman RB. Increases in thyroid nodule fine-needle aspirations, operations, and diagnoses of thyroid cancer in the United States. *Surgery*. 2013;**154**(6): 1420–1426, discussion 1426–1427.
3. Lubitz CC, Kong CY, McMahon PM, Daniels GH, Chen Y, Economopoulos KP, Gazelle GS, Weinstein MC. Annual financial impact of well-differentiated thyroid cancer care in the United States. *Cancer*. 2014;**120**(9):1345–1352.
4. Nikiforov YE, Steward DL, Robinson-Smith TM, Haugen BR, Klopper JP, Zhu Z, Fagin JA, Falciglia M, Weber K, Nikiforova MN. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. *J Clin Endocrinol Metab*. 2009;**94**(6):2092–2098.
5. Cantara S, Capezzone M, Marchisotta S, Capuano S, Busonero G, Toti P, Di Santo A, Caruso G, Carli AF, Brilli L, Montanaro A, Pacini F. Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. *J Clin Endocrinol Metab*. 2010;**95**(3): 1365–1369.
6. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, Yip L, Seethala RR, Tublin ME, Stang MT, Coyne C, Johnson JT, Stewart AF, Nikiforova MN. 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**(11):3390–3397.
7. Yip L, Wharry LI, Armstrong MJ, Silbermann A, McCoy KL, Stang MT, Ohori NP, LeBeau SO, Coyne C, Nikiforova MN, Bauman JE, Johnson JT, Tublin ME, Hodak SP, Nikiforov YE, Carty SE. A clinical algorithm for fine-needle aspiration molecular testing effectively guides the appropriate extent of initial thyroidectomy. *Ann Surg*. 2014;**260**(1):163–168.
8. Labourier E, Shifrin A, Busseniers AE, Lupo MA, Manganelli ML, Andruss B, Wylie D, Beaudenon-Huibregtse S. Molecular testing for miRNA, mRNA, and DNA on fine-needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. *J Clin Endocrinol Metab*. 2015;**100**(7):2743–2750.
9. De Napoli L, Bakkar S, Ambrosini CE, Materazzi G, Proietti A, Macerola E, Basolo F, Miccoli P. Indeterminate single thyroid nodule: synergistic impact of mutational markers and sonographic features in triaging patients to appropriate surgery. *Thyroid*. 2016; **26**(3):390–394.
10. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, Friedman L, Kloos RT, LiVolsi VA, Mandel SJ, Raab SS, Rosai J, Steward DL, Walsh PS, Wilde JI, Zeiger MA, Lanman RB, Haugen BR. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;**367**(8):705–715.
11. Al-Qurayshi Z, Deniwar A, Thethi T, Mallik T, Srivastav S, Murad F, Bhatia P, Moroz K, Sholl AB, Kandil E. Association of malignancy prevalence with test properties and performance of the gene expression classifier in indeterminate thyroid nodules. *JAMA Otolaryngol Head Neck Surg*. 2017;**143**(4):403–408.
12. Marti JL, Avadhani V, Donatelli LA, Niyogi S, Wang B, Wong RJ, Shaha AR, Ghossein RA, Lin O, Morris LG, Ho AS. Wide interinstitutional variation in performance of a molecular classifier for indeterminate thyroid nodules. *Ann Surg Oncol*. 2015;**22**(12): 3996–4001.
13. Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;**159**(3): 676–690.
14. 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**(11):E1852–E1860.
15. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, Hodak SP, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. 2014;**120**(23):3627–3634.
16. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid*. 2015;**25**(11):1217–1223.
17. Valderrabano P, Khazai L, Leon ME, Thompson ZJ, Ma Z, Chung CH, Hallanger-Johnson JE, Otto KJ, Rogers KD, Centeno BA, McIver B. Evaluation of ThyroSeq v2 performance in thyroid nodules with indeterminate cytology. *Endocr Relat Cancer*. 2017;**24**(3):127–136.
18. Taye A, Gurciullo D, Miles BA, Gupta A, Owen RP, Inabnet WB III, Beyda JN, Marti JL. Clinical performance of a next-generation sequencing assay (ThyroSeq v2) in the evaluation of indeterminate thyroid nodules. *Surgery*. 2018;**163**(1):97–103.
19. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, Gooding WE, Yip L, Ferris RL, Nikiforov YE. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer*. 2018; **124**(8):1682–1690.
20. Patel KNAT, Barbiarz J, Barth NM, Blevins TC, Duh Q-Y, Ghossein RA, Harrell RMHJ, Imtiaz U, Kennedy GC, Kim SY, Kloos RT, LiVolsi VA, Randolph GWSP, Shanik MH, Sosa JA, Traweek ST, Walsh PS, Whitney DYM, Ladenson PW. Clinical validation of an improved genomic classifier for cytologically indeterminate thyroid nodules using an NGS platform and machine learning algorithms in an independent prospective multicenter blinded cohort demonstrates improved performance. 3rd World Congress on Thyroid Cancer; July 2017; Boston, MA.
21. Nikiforov YE. Role of molecular markers in thyroid nodule management: then and now. *Endocr Pract*. 2017;**23**(8):979–988.
22. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016;**26**(1):1–133.
23. Adam MA, Thomas S, Youngwirth L, Hyslop T, Reed SD, Scheri RP, Roman SA, Sosa JA. Is there a minimum number of thyroidectomies a surgeon should perform to optimize patient outcomes? *Ann Surg*. 2017;**265**(2):402–407.
24. Lee L, How J, Tabah RJ, Mitmaker EJ. Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. *J Clin Endocrinol Metab*. 2014;**99**(8): 2674–2682.