The Journal of Clinical Endocrinology & Metabolism, 2021, Vol. 106, No. 3, e1240–e1247
doi:10.1210/clinem/dgaa887
Clinical Research Article

Clinical Research Article

Outcomes of Indeterminate Thyroid Nodules Managed Nonoperatively after Molecular Testing

Catherine Y. Zhu,[1] Ines Donangelo,[2] Deepashree Gupta,[2] Dalena T. Nguyen,[1] Joana E. Ochoa,[1] Michael W. Yeh,[1] and Masha J. Livhits[1]

[1]Section of Endocrine Surgery, UCLA David Geffen School of Medicine, Los Angeles, CA 90095, USA; and [2]Division of Endocrinology, Diabetes, and Metabolism, UCLA David Geffen School of Medicine, Los Angeles, CA 90095, USA

ORCiD numbers: 0000-0001-7202-8430 (I. Donangelo); 0000-0002-8436-392X (M. J. Livhits).

Received: 3 September 2020; Editorial Decision: 20 November 2020; First Published Online: 2 December 2020; Corrected and Typeset: 4 January 2021.

Abstract

Context: Molecular testing to refine the diagnosis of cytologically indeterminate thyroid nodules has become increasingly popular, but data on long-term durability of test results and the rate of delayed operation are limited.

Objective: Determine the delayed rate of surgical resection in indeterminate nodules with benign/negative molecular testing and the risk of false-negative molecular test results.

Design: Prospective follow-up of the Gene Expression Classifier vs Targeted Next-Generation Sequencing in the Management of Indeterminate Thyroid Nodules randomized controlled trial comparing the diagnostic test performance of Afirma Gene Expression Classifier and ThyroSeq v2.

Setting: University of California, Los Angeles.

Participants: Patients who underwent thyroid biopsy with indeterminate (Bethesda III/IV) cytology (April 2016 to July 2017).

Intervention: Ultrasound surveillance.

Main Outcome Measure: False-negative rate of molecular testing.

Results: Of 95 indeterminate nodules with negative/benign molecular test results, 12 nodules underwent immediate resection (11 benign nodules, 1 noninvasive follicular thyroid neoplasm nodule with papillary-like nuclear features). Nonoperative management was pursued for 83 (87.4%) nodules. The median surveillance was 26.7 months. Ten nodules were resected during surveillance and malignancy was identified in 4 nodules (overall false-negative rate of 5.8%). In the 4 malignant nodules that underwent delayed operation, surgery was prompted by sonographic changes during surveillance.

Conclusions: The majority of indeterminate nodules with negative molecular testing have a stable clinical course over 3 years of follow-up, but our finding of a 6% false-negative

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

rate highlights the importance of continuing sonographic surveillance. Long-term studies are needed to determine the optimal length of follow-up.

Key Words: thyroid, indeterminate nodules, molecular testing, surveillance, nonoperative, thyroid cancer

The use of molecular testing for thyroid nodules with indeterminate cytology has decreased the rate of diagnostic lobectomy in the United States (1–4). Indeterminate nodules make up 15% to 30% of all thyroid fine-needle aspiration (FNA) biopsies and exhibit malignancy rates ranging from 5% to 40% (5, 6). Patients with benign/negative molecular test results have a reported < 5% risk of malignancy and are generally observed (1, 7, 8). The low false-negative rate has only been shown in limited validation studies (9–11). Most patients with benign/negative results are managed nonoperatively and do not have available histopathology for evaluation (12–15).

When the 2015 American Thyroid Association guidelines were published, the 2 most commonly used molecular tests were the Afirma Gene Expression Classifier (Afirma GEC; Veracyte, Inc, San Francisco, California) and ThyroSeq v2 (CBL Path, Inc, Rye Brook, New York) (2). Afirma GEC relies on micoarray analysis of messenger ribonucleic acid expression (9, 13), while ThyroSeq v2 uses next-generation sequencing to detect both deoxyribonucleic acid and ribonucleic acid alterations (10, 11). Our previous randomized controlled study directly compared the performance of Afirma GEC and Thyroseq v2 in 149 indeterminate thyroid nodules (12). The specificities of GEC and ThyroSeq v2 were 66% and 91%, respectively ($P < 0.01$) (12). We reported that 62% of patients with indeterminate thyroid nodules tested benign/negative with either GEC or ThyroSeq v2, and of those, 84% were observed. In all, approximately half of the indeterminate nodules were managed nonoperatively on the basis of a benign/negative molecular test.

Although prior studies have retrospectively analyzed the outcomes of indeterminate nodules selectively managed with molecular testing (16–18), we prospectively surveilled patients who reflexively underwent molecular testing for all cases of indeterminate cytology to minimize sampling bias. Both Afirma GEC and Thyroseq v2 were updated in 2017; however, we used the prior tests due to the longer duration of follow-up available and the large reservoir of patients who are currently under observation after undergoing testing with Afirma GEC or Thyroseq v2. The objective of this study was to determine the rate of delayed operation and the false-negative risk of benign/negative molecular testing.

Methods

Patients who underwent thyroid FNA within the University of California Los Angeles (UCLA) health system between May 1, 2016, and July 31, 2017, were previously enrolled in the Gene Expression Classifier vs Targeted Next-Generation Sequencing in the Management of Indeterminate Thyroid Nodules study, a randomized controlled trial comparing the diagnostic performances of Afirma GEC (Veracyte, Inc, South San Francisco, California) and ThyroSeq v2 (CBL Path, Inc, Rye Brook, New York) (12). Thyroid nodules classified as Bethesda III (atypia of undetermined significance or follicular lesion of undetermined significance) or IV (follicular neoplasm or suspicious for follicular neoplasm) by centralized cytopathology review were reflexively sent for molecular testing. Patients who were managed nonoperatively underwent annual surveillance thyroid ultrasounds through April 30, 2020. Follow-up duration was determined by time elapsed between initial FNA date and date of the most recent ultrasound or date of surgery, and nodules lost to follow-up without any surveillance imaging were excluded from analysis. Patients younger than 18 years were excluded due to the limited data on molecular testing in the pediatric patient population. Patients with a history of known thyroid malignancy and those with a concurrent thyroid nodule or lymph node FNA yielding a malignant result were excluded.

Data abstracted from imaging reports included measurements of nodule dimensions and presence of any concerning sonographic findings, such as microcalcifications, hypoechogenicity, or irregular nodule borders. Significant nodule growth was defined as > 20% growth in at least 2 nodule dimensions, with a minimum growth of 2 mm or > 50% increase in nodule volume (2). If fewer than 3 nodule dimensions were reported, the definition of nodule growth was left to the interpretation of the physician documenting the ultrasound results. Repeat FNA after a negative molecular test result was performed based on sonographic findings (increasing nodule size or development of suspicious sonographic characteristics) and/or physician preference. The cumulative incidence of sonographic changes in indeterminate nodules and the cumulative incidence of cancer diagnosis were plotted over the time of surveillance.

In patients who underwent surgical resection, data on surgical indications and histopathology results corresponding to the index nodule were recorded in a prospective database. Pathology was only considered malignant if located in the same quadrant as the originally biopsied nodule. The rate of initial false-negative molecular testing was determined based on the number of malignancies or

Non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) diagnosed on surgical resection out of all indeterminate nodules that had an initial negative molecular test result. This study was approved by the UCLA Institutional Review Board.

Results

A total of 158 patients with 165 indeterminate thyroid nodules underwent molecular testing with Afirma GEC (n = 74) or ThyroSeq v2 (n = 91) between May 1, 2016, and July 31, 2017. Ultimately, 73 nodules were resected over the 4-year study period, and the overall prevalence of cancer or NIFTP in all tested nodules was 19.4% (26/165 nodules). Among indeterminate nodules tested with Afirma GEC, 33 (44.6%) had a suspicious result; the resection rate was 81.8% (27/33), and the positive predictive value (PPV) was 36.4% (12/33). Among indeterminate nodules tested with ThyroSeq v2, 24 (26.4%) had a positive result; the resection rate was 91.7% (22/24) and the PPV was 58.3% (14/24).

A total of 100 patients underwent nonoperative management after initial thyroid FNA biopsy for 103 indeterminate nodules (44 Afirma GEC nodules and 59 ThyroSeq v2 nodules) (Table 1). The median patient age was 59 (interquartile range [IQR], 46–68) years, and 83% of patients were female. The majority of nonoperatively managed nodules had benign/negative molecular test results (n = 83 [80.6%]), while 9 (8.7%) nodules had suspicious/positive results, and 11 (10.7%) FNA samples were insufficient for molecular analysis. Surveillance ultrasound was obtained for 91 (88.3%) nodules at a median follow-up time of 26.7 (IQR, 17.4–33.2) months, while 12 (11.7%) nodules were lost to follow-up. Ultimately, 12 patients underwent delayed operation after an initial surveillance period. Reasons for undergoing surgery included repeat molecular testing newly positive after previously negative (n = 3), repeat FNA positive for malignancy (n = 1), development of compressive or hyperthyroid symptoms (n = 3), and patient preference for surgery (n = 5) (Table 2).

Of the 83 indeterminate nodules with benign/negative molecular testing that were managed nonoperatively, 74 (89.2%) had surveillance imaging performed (Fig. 1). Nodule growth was observed in 5 (6.8%) nodules, development of abnormal sonographic features (predominantly increased calcifications and/or irregular nodule borders) was observed in 6 (8.1%) nodules, and 1 (1.4%) nodule exhibited both development of abnormal features and growth; 83.7% of nodules remained stable in size and appearance. Although not in accordance with specific guidelines, repeat FNA was performed in 15 (20.3%) nodules at a median follow-up time of 20.2 (IQR, 10.3–24.4) months. Reasons for repeat FNA included sonographic nodule growth or change (n = 12), physician or patient preference (n = 2), and evaluation of new thyroid nodules resulting in concurrent repeat FNA of the index nodule (n = 1). Repeat FNA identified 3 nodules with benign cytology, 9 nodules with indeterminate cytology, 1 nodule with papillary thyroid cancer (PTC), and 2 nodules which were nondiagnostic (Table 3). Resection of the malignant nodule, which had been initially tested negative with ThyroSeq v2, revealed a papillary thyroid microcarcinoma. This nodule had a heterogenous appearance on ultrasound and only a portion of it was malignant on final pathology (0.9 cm cancer focus in a 1.4-cm nodule). Of the 9 nodules with repeat indeterminate cytology, repeat molecular testing was benign/negative in 5

Table 1. Baseline characteristics of patients with nonoperatively managed indeterminate thyroid nodules

	All (n = 103)	Benign/Negative Test Result (n = 83)	Suspicious/Positive Test Result (n = 9)	Insufficient Result (n = 11)
Median age (IQR), y	59 (46–68)	59 (45–68)	59 (50–65)	59 (54.8–61.5)
Sex, n (%)				
Female	83 (83.0%)	69 (85.2%)	9 (100.0%)	5 (50.0%)
Male	17 (17.0%)	12 (14.8%)	0 (0.0%)	5 (50.0%)
Median nodule size (IQR), cm	2.0 (1.5–2.9)	2.0 (1.6–2.9)	1.5 (1.1–2.9)	1.5 (1.0,2.7)
Bethesda category, n (%)				
III (AUS/FLUS)	97 (94.2%)	80 (96.4%)	7 (77.8%)	10 (90.9%)
IV (FN/SFN)	6 (5.8%)	3 (3.6%)	2 (22.2%)	1 (9.1%)
Initial molecular test, n (%)				
Afirma GEC	44 (42.7%)	30 (36.1%)	7 (77.8%)	7 (63.6%)
ThyroSeq v2	59 (57.3%)	53 (63.9%)	2 (22.2%)	4 (36.4%)

Abbreviations: AUS, atypia of undetermined significance; FLUS, follicular lesion of undetermined significance; FN, follicular neoplasm; GEC, gene expression classifier; IQR, interquartile range; SFN, suspicious for follicular neoplasm.

Table 2. Indications for delayed thyroidectomy after initial surveillance and surgical pathology Results

	Total, n	Surgical Pathology Result	
		Benign	Malignant
Molecular test benign/negative			
Patient preference	3	3	0
Benign symptoms[a]	2	2	0
Sonographic growth in index nodule[b]	1	1	0
Repeat FNA with positive/suspicious molecular test	3	0	3
Repeat FNA with malignant result	1	0	1
Molecular test suspicious/positive			
Patient preference	1	0	1
Molecular test insufficient			
Benign symptoms	1	0	1

Abbreviations: FNA, fine-needle aspiration; IQR, interquartile range.
[a]Benign symptoms included compressive symptoms by benign nodules and hyperthyroid symptoms.
[b]Based on physician-subjective interpretation of imaging.

nodules, suspicious/positive in 3 nodules, and 1 was not tested. In the 3 nodules with newly positive molecular test results, 2 initially tested negative with ThyroSeq v2 and 1 initially had a benign result with Afirma GEC. Subsequent testing with Afirma Genomic Sequencing Classifier (n = 1) and ThyroSeq v3 (n = 2) had suspicious/positive results. The ThyroSeq v3 positive nodules had a *BRAF V600E* mutation and combination *TP53* and *EIF1AX* mutations, respectively. The 3 patients with suspicious/positive repeat molecular test results subsequently underwent resection and were diagnosed with PTC. All 4 nodules with initial false-negative molecular testing measured less than 2 cm and underwent repeat FNA due to nodule growth and/or new suspicious sonographic features.

Nine nodules with initially suspicious/positive molecular test results were nonoperatively managed due to patient preference, of which 8 (88.9%) nodules had at least 1 surveillance ultrasound. No nodules showed significant growth. One nodule developed irregular borders and calcifications, but a repeat FNA was cytologically benign. Two other nodules had repeat FNA, with benign and indeterminate results (repeat molecular test, negative), respectively. One nodule with initial suspicious/positive molecular test results was resected after a period of surveillance and diagnosed with classic PTC. In the 11 nodules with cytology insufficient for molecular analysis and managed nonoperatively, 9 (81.8%) had at least 1 surveillance ultrasound. None of these indeterminate nodules exhibited a significant change in size or appearance. One nodule had a repeat FNA performed, which had benign cytology, and 1 patient underwent delayed surgery for compressive symptoms and was diagnosed with classic PTC.

Among the 91 indeterminate thyroid nodules initially managed nonoperatively with follow-up imaging available, 13 (14.3%) nodules grew in size or developed abnormal sonographic features, and 18 (19.8%) nodules underwent repeat FNA during surveillance (Fig. 1). In nodules that had a benign/negative initial molecular test result, the cumulative incidence of nodule change at 36 months of surveillance was 16.2% (Fig. 2). Twelve patients underwent surgery after a median of 21.0 (IQR, 12.6–31.3) months of surveillance, including 10 patients with a benign/negative initial molecular test result, 1 patient with a suspicious/positive result, and 1 patient with an insufficient result. Papillary thyroid cancer was found in 6 patients (4 in initially test-negative nodules, 1 in an initially test-positive nodule, and 1 in a nodule with an insufficient sample for molecular analysis) (Table 3). Of all the nodules that initially tested benign/negative, 5.8% (5/86 nodules) ultimately showed malignancy or NIFTP on final pathology, including 1 nodule that was resected immediately after biopsy. The false-negative rate of Afirma GEC was 6.9% (2 false-negative results of 29 nodules that initially tested benign), and the false-negative rate of ThyroSeq v2 was 5.4% (3 false-negative results of 56 nodules that initially tested negative).

Discussion

In this observational study, we prospectively followed 103 indeterminate thyroid nodules tested with either Afirma GEC or ThyroSeq v2 and that were initially managed nonoperatively. Of the 74 nodules with initial benign/negative molecular testing under surveillance, 10 nodules were ultimately resected (13.5%) and 4 nodules were PTC (5.4%). Including patients who had immediate surgery after biopsy, the overall false-negative rate was 5.8% (5/86 total nodules with negative testing and available follow-up), of which 2 were initially tested with Afirma GEC and 3 were initially tested with ThyroSeq v2. Most patients who ultimately pursued surgery did so within 2 years of initial FNA. The most common reasons for undergoing resection were patient preference, symptomatic nodules, and a suspicious/positive result on repeat molecular testing in nodules that grew and/or developed new suspicious sonographic features during surveillance.

Our findings substantiate American Thyroid Association recommendations for annual sonographic surveillance of intermediate nodules with benign/negative molecular testing (2). The 5.8% false-negative rate is slightly higher than the frequently reported false-negative rate of a benign

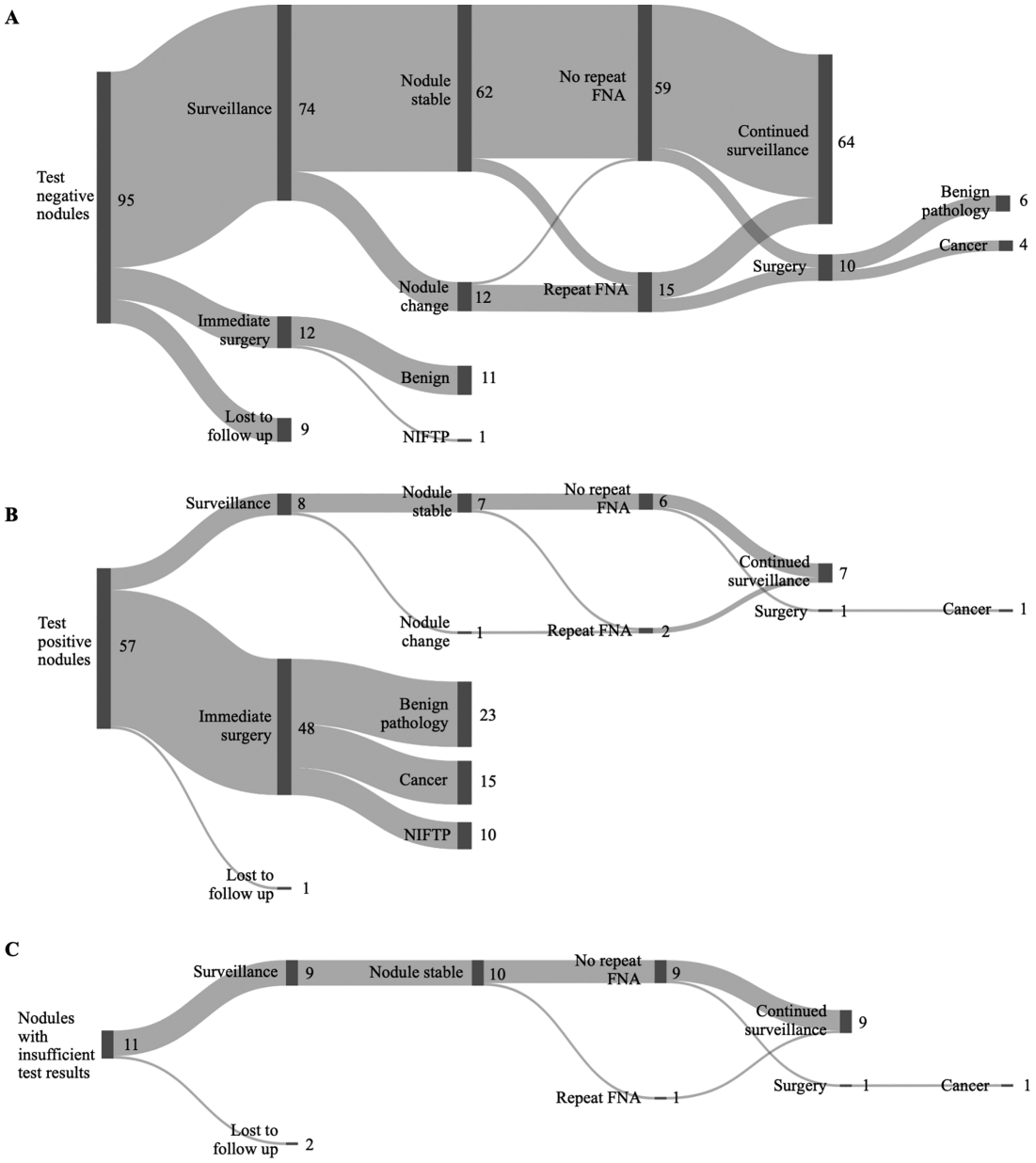

Figure 1. Sankey diagram demonstrating clinical practice patterns following molecular testing for indeterminate thyroid nodules that initially tested negative or positive. **A:** The vast majority of nodules with negative molecular testing are initially managed nonoperatively and remain sonographically stable, but a small proportion demonstrate sonographic changes, of which the majority undergo repeat FNA. **B:** The majority of nodules with positive molecular testing are initially managed with surgery, and the majority of nonoperatively managed nodules remained sonographically stable. **C:** All nodules with insufficient molecular test results remained sonographically stable; however, 1 subcentimeter nodule that was later resected due to a compressive adjacent nodule was incidentally found to harbor malignancy. Abbreviations: FNA, fine-needle aspiration; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features.

cytology result (1–3%) (19–21), highlighting the importance of ultrasound surveillance. In indeterminate nodules with initial benign/negative molecular test results, cancer was only diagnosed in nodules that exhibited a change on ultrasound (Table 3). Conversely, nodules without sonographic change that were resected due to patient preference or the development of compressive symptoms demonstrated benign histopathology. Since most indeterminate nodules with benign/negative molecular testing continued with nonoperative management, some may be harboring cancer but have not demonstrated a sonographic change. However, our results indicate that sonographic change in indeterminate nodules with benign/negative molecular test is a risk factor for cancer and has a PPV of 33.3% (4 malignancies in 12 nodules with sonographic change). Two of the false-negative patients developed sonographic changes after 3 years of follow-up, which underscores the importance of long-term follow-up in these patients.

Table 3. Management and surgical results of indeterminate nodules that had repeat FNA during follow-up

Initial Molecular Test Result	Initial Nodule Size[a] (cm)	Sonographic Findings	Repeat FNA Cytology	Repeat Molecular Test Result	Pathology Result
Negative					
GEC	2	Nodule growth	AUS	ThyroSeq v3 negative	–
GEC	2.5	Nodule growth	FLUS	ThyroSeq v2 negative	–
ThyroSeq v2	3.2	Nodule growth	AUS	GEC benign	–
ThyroSeq v2	1.4	Nodule growth	PTC	–	PTC, classic variant
ThyroSeq v2	4	Nodule change	Nondiagnostic	ThyroSeq v3 negative	–
GEC	4.6	Nodule change	Benign	ThyroSeq v3 negative	–
ThyroSeq v2	1.7	Nodule change	AUS	ThyroSeq v3 negative	–
GEC	0.9	Nodule change	AUS	ThyroSeq v3 positive (BRAF V600E)	PTC, tall cell variant
ThyroSeq v2	1.6	Nodule change	AUS	GSC suspicious	PTC, classic variant
ThyroSeq v2	1.7	Nodule growth and change	SFN	ThyroSeq v3 positive (TP53, EIF1AX)	PTC, oncocytic variant
ThyroSeq v2	2.8	Nodule stable	Nondiagnostic	–	–
GEC	1.5	Nodule stable	Benign	–	–
GEC	2.7	Nodule stable	Benign	–	–
ThyroSeq v2	1.2	Nodule stable	AUS	–	Benign
Positive					
GEC	2.8	Nodule growth	AUS	GSC benign	–
ThyroSeq v2	1.3	Nodule change	Benign	–	–
GEC	1.1	Nodule stable	Benign	ThyroSeq v3 positive (copy number alterations)	–

Abbreviations: AUS, atypia of undetermined significance; FLUS, follicular lesion of undetermined significance; FNA, fine-needle aspiration; GEC, gene expression classifier; GSC, genomic sequencing classifier; PTC, papillary thyroid carcinoma; SFN, suspicious for follicular neoplasm.
[a]In largest nodule dimension.`

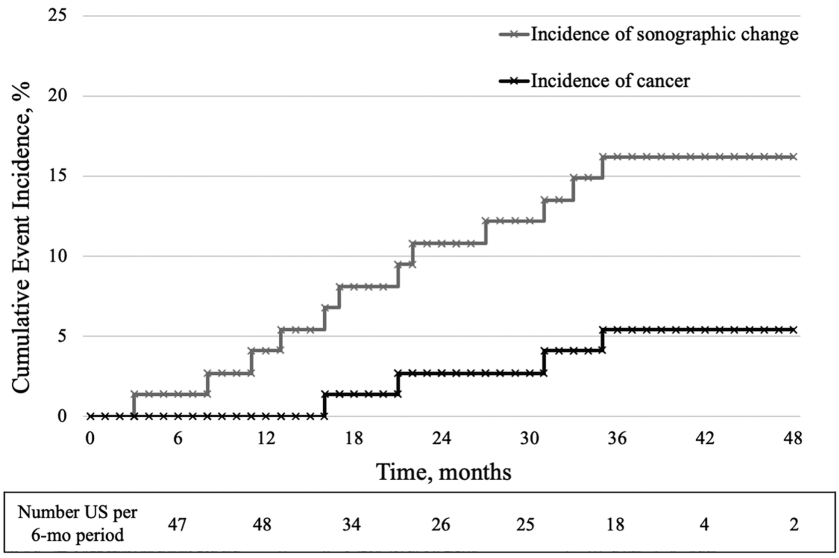

Figure 2. Cumulative event incidence during surveillance of indeterminate thyroid nodules with negative initial molecular test results. The cumulative event incidence is calculated as the proportion of all test-negative nodules that (1) developed a sonographic change, either nodule growth or development of suspicious sonographic features, such as calcifications or irregular nodule borders, at the time of surveillance ultrasound, or (2) were diagnosed with cancer after surgery was performed during the surveillance period. The plateau in incidence of change or cancer occurring after 36 months is likely a reflection of the comparatively fewer number of nodules that had surveillance imaging past 36 months. Abbreviation: US, ultrasound.

Prior studies of Afirma GEC and ThyroSeq v2 have reported high sensitivity and a negative predictive value to rule out cancer (22). One of the largest series with clinical follow-up retrospectively analyzed 405 Bethesda IV nodules tested with ThyroSeq v2 or v3 (18). Molecular testing was performed selectively, potentially introducing

bias, as nodules with clinically or sonographically suspicious features may have been resected without molecular testing. The mean follow-up was 25 months, and one-third of patients were lost to follow-up. Similar to our findings, 82% of nodules with negative molecular testing remained stable on surveillance ultrasound. The false-negative rate was 6.8% for ThyroSeq v2 and 1.5% for ThyroSeq v3, with the assumption that unresected nodules with negative molecular testing (including those lost to follow-up) were benign. Our prospective study provides clinical outcomes for Bethesda III and IV nodules tested with Afirma GEC and ThyroSeq v2, while minimizing bias by reflexive molecular testing for all nodules with indeterminate cytology and an 88% rate of surveillance with ultrasound.

In our study, surgical pathology revealed 4 patients with false-negative results. Of these, 2 nodules potentially had sampling error. One nodule initially tested benign with Afirma GEC and subsequently had a *BRAF V600E* mutation identified with ThyroSeq v3, which should have been detected by Afirma GEC (23, 24). The 2nd nodule initially tested negative with ThyroSeq v2 and subsequently had a *TP53* and *EIF1AX* mutation identified with ThyroSeq v3. These alterations were also available in ThyroSeq v2, and there is no reported change in the methodology of their detection in ThyroSeq v3 (12, 25). Since all 4 nodules were less than 2 cm in size at the time of initial FNA, the risk of sampling error should be low (26, 27). Interval malignant transformation of these nodules is unlikely due to the relatively short time between initial FNA and surgery (within 3 years in all cases) (28).

Although molecular testing platforms to improve the diagnostic accuracy of indeterminate thyroid nodules continue to evolve, the recent advances would likely not have influenced our findings. The molecular tests used in our study, Afirma GEC and ThyroSeq v2, have since been replaced by updated versions of each test: Afirma Genomic Sequencing Classifier and ThyroSeq v3. However, the shared gene panel tested by the prior and current versions of each test encompass the most commonly mutated genes in thyroid cancer (11, 29, 30). For example, in 2 of the nodules with false-negative results, repeat testing with ThyroSeq v3 identified mutations in *BRAF V600E*, *TP53*, and *EIF1AX*, all 3 of which are included in the gene panels tested by ThyroSeq v2. In our cohort, 11% of the nodules (8/74) with initially benign/negative molecular testing that were nonoperatively managed later underwent repeat testing with 1 of the updated molecular test versions, and none showed alterations in the genes added to the updated test versions. Given the large number of indeterminate nodules that have only been tested with older versions of molecular tests, long-term follow-up to assess the clinical outcomes and false-negative rate is clinically important.

There are several limitations to this study. The sample size is relatively small. By design, most nodules with benign/negative molecular testing continue to undergo nonoperative management and do not have histopathology available for evaluation. Longer follow-up is necessary to determine if additional false-negative cases will be detected at 5 or 10 years from initial FNA. The prior molecular test versions that were evaluated in the study have been updated. However, the results are likely applicable to the current test versions and are clinically relevant for the patients who were previously tested and are undergoing surveillance. Lastly, surveillance ultrasounds were performed by radiologists, endocrinologists, and endocrine surgeons. As a result, in a small number of instances, ultrasound images could not always be directly compared. Varying degrees of nodule descriptions were provided, and not all nodule sizes were documented in 3 dimensions.

Conclusions

Our study demonstrates that the majority of patients with indeterminate nodules with benign/negative molecular testing with either Afirma GEC or ThyroSeq v2 have a benign clinical course over a 3-year follow-up period. However, the importance of surveillance is validated given the 15% incidence of sonographic changes and the 6% rate of pathology-confirmed malignancy over the study period. As the false-negative rates of Afirma GEC and ThyroSeq v2 are similar, other factors such as cost and availability of specific genetic alterations that provide prognostic value may determine the choice of molecular test. Long-term follow-up studies with larger number of patients will be needed to assess the negative predictive value of molecular testing in real world practice, and to determine how long indeterminate nodules with benign/negative molecular test should be followed.

Acknowledgments

Financial Support: This study was not funded.
Clinical Trial Information: ClinicalTrials.gov (#NCT02681328).

Additional Information

Correspondence: Masha J. Livhits, MD, Section of Endocrine Surgery, 10833 Le Conte Ave, 72–228 CHS, Los Angeles, CA 90095, USA. E-mail: mlivhits@mednet.ucla.edu.

Disclosure Summary: The authors declare no competing financial interests exist.

Data Availability: Some or all datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

References

1. Jug R, Parajuli S, Ahmadi S, Jiang XS. Negative results on thyroid molecular testing decrease rates of surgery for indeterminate thyroid nodules. *Endocr Pathol*. 2019;**30**(2):134–137.
2. Haugen BR, Alexander EK, Bible KC, et al. 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.
3. Duick DS, Klopper JP, Diggans JC, et al. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Thyroid*. 2012;**22**(10):996–1001.
4. Zhu CY, Sha SW, Tseng C-H, et al. Trends in the surgical management of known or suspected differentiated thyroid cancer at a single institution, 2010–2018. *Thyroid*. 2020;**30**(11):1639–1645.
5. Baloch ZW, LiVolsi VA, Asa SL, et al. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagn Cytopathol*. 2008;**36**(6):425–437.
6. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*. 2007;**111**(6):508–516.
7. Hang JF, Westra WH, Cooper DS, Ali SZ. The impact of noninvasive follicular thyroid neoplasm with papillary-like nuclear features on the performance of the Afirma gene expression classifier. *Cancer Cytopathol*. 2017;**125**(9):683–691.
8. Samulski TD, LiVolsi VA, Wong LQ, Baloch Z. Usage trends and performance characteristics of a "gene expression classifier" in the management of thyroid nodules: an institutional experience. *Diagn Cytopathol*. 2016;**44**(11):867–873.
9. Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma gene expression classifier. *J Clin Endocrinol Metab*. 2014;**99**(1):119–125.
10. 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**(23):3627–3634.
11. 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**(11):1217–1223.
12. Livhits MJ, Kuo EJ, Leung AM, et al. Gene expression classifier vs targeted next-generation sequencing in the management of indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2018;**103**(6):2261–2268.
13. McIver B, Castro MR, Morris JC, et al. An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2014;**99**(11):4069–4077.
14. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. *Endocr Pract*. 2014;**20**(4):364–369.
15. Chaudhary S, Hou Y, Shen R, Hooda S, Li Z. Impact of the Afirma gene expression classifier result on the surgical management of thyroid nodules with category III/IV cytology and its correlation with surgical outcome. *Acta Cytol*. 2016;**60**(3):205–210.
16. Vora A, Holt S, Haque W, Lingvay I. Long-term outcomes of thyroid nodule AFIRMA GEC testing and literature review: an institutional experience. *Otolaryngol Head Neck Surg*. 2020;**162**(5):634–640.
17. Deaver KE, Haugen BR, Pozdeyev N, Marshall CB. Outcomes of Bethesda categories III and IV thyroid nodules over 5 years and performance of the Afirma gene expression classifier: a single-institution study. *Clin Endocrinol (Oxf)*. 2018;**89**(2):226–232.
18. Carty SE, Ohori NP, Hilko DA, et al. The clinical utility of molecular testing in the management of thyroid follicular neoplasms (Bethesda IV nodules). *Ann Surg*. 2020;**272**(4):621–627.
19. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference. The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol*. 2009;**132**(5):658–665.
20. Jo VY, Stelow EB, Dustin SM, Hanley KZ. Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2010;**134**(3):450–456.
21. Sarkis LM, Norlen O, Aniss A, et al. The Australian experience with the Bethesda classification system for thyroid fine needle aspiration biopsies. *Pathology*. 2014;**46**(7):592–595.
22. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012;**367**(8):705–715.
23. Kloos RT, Reynolds JD, Walsh PS, et al. Does addition of BRAF V600E mutation testing modify sensitivity or specificity of the Afirma Gene Expression Classifier in cytologically indeterminate thyroid nodules? *J Clin Endocrinol Metab*. 2013;**98**(4):E761–E768.
24. Diggans J, Kim SY, Hu Z, et al. Machine learning from concept to clinic: reliable detection of BRAF V600E DNA mutations in thyroid nodules using high-dimensional RNA expression data. *Pac Symp Biocomput*. 2015:371–382.
25. 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. *JAMA Oncol*. 2019;**5**(2):204–212.
26. Pinchot SN, Al-Wagih H, Schaefer S, Sippel R, Chen H. Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. *Arch Surg*. 2009;**144**(7):649–655.
27. McCoy KL, Jabbour N, Ogilvie JB, Ohori NP, Carty SE, Yim JH. The incidence of cancer and rate of false-negative cytology in thyroid nodules greater than or equal to 4 cm in size. *Surgery*. 2007;**142**(6):837–844; discussion 844.e831–833.
28. Durante C, Costante G, Lucisano G, et al. The natural history of benign thyroid nodules. *JAMA*. 2015;**313**(9):926–935.
29. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol*. 2011;**7**(10):569–580.
30. Harrell RM, Eyerly-Webb SA, Golding AC, Edwards CM, Bimston DN. Statistical comparison of Afirma GSC and Afirma GEC outcomes in a community endocrine surgical practice: early findings. *Endocr Pract*. 2019;**25**(2):161–164.