Performance of a Multigene Genomic Classifier in Thyroid Nodules With Indeterminate Cytology: A Prospective Blinded Multicenter Study

David L. Steward, MD; Sally E. Carty, MD; Rebecca S. Sippel, MD; Samantha Peiling Yang, MBBS, MRCP, MMed; Julie A. Sosa, MD, MA; Jennifer A. Sipos, MD; James J. Figge, MD, MBA; Susan Mandel, MD, MPH; Bryan R. Haugen, MD; Kenneth D. Burman, MD; Zubair W. Baloch, MD, PhD; Ricardo V. Lloyd, MD, PhD; Raja R. Seethala, MD; William E. Gooding, MS; Priya Kundra, MD; Kwok Seng Loh, MBBS, Carrie B. Marshall, MD; Sarah Mayson, MD; Kelly L. McCoy, MD; Min En Nga, MBBS; Kee Yuan Ngiam, MBBS, MMed; Marina N. Nikiforova, MD; Jennifer L. Poehls, MD; Matthew D. Ringel, MD; Linwah Yip, MD; Yuri E. Nikiforov, MD, PhD

IMPORTANCE Approximately 20% of fine-needle aspirations (FNA) of thyroid nodules have indeterminate cytology, most frequently Bethesda category III or IV. Diagnostic surgeries can be avoided for these patients if the nodules are reliably diagnosed as benign without surgery.

OBJECTIVE To determine the diagnostic accuracy of a multigene classifier (GC) test (ThyroSeq v3) for cytologically indeterminate thyroid nodules.

DESIGN, SETTING, AND PARTICIPANTS Prospective, blinded cohort study conducted at 10 medical centers, with 782 patients with 1013 nodules enrolled. Eligibility criteria were met in 256 patients with 286 nodules; central pathology review was performed on 274 nodules.

INTERVENTIONS A total of 286 FNA samples from thyroid nodules underwent molecular analysis using the multigene GC (ThyroSeq v3).

MAIN OUTCOMES AND MEASURES The primary outcome was diagnostic accuracy of the test for thyroid nodules with Bethesda III and IV cytology. The secondary outcome was prediction of cancer by specific genetic alterations in Bethesda III to V nodules.

RESULTS Of the 286 cytologically indeterminate nodules, 206 (72%) were benign, 69 (24%) malignant, and 11 (4%) noninvasive follicular thyroid neoplasms with papillary-like nuclei (NIFTP). A total of 257 (90%) nodules (154 Bethesda III, 93 Bethesda IV, and 10 Bethesda V) had informative GC analysis, with 61% classified as negative and 39% as positive. In Bethesda III and IV nodules combined, the test demonstrated a 94% (95% CI, 86%-98%) sensitivity and 82% (95% CI, 75%-87%) specificity. With a cancer/NIFTP prevalence of 28%, the negative predictive value (NPV) was 97% (95% CI, 93%-99%) and the positive predictive value (PPV) was 66% (95% CI, 56%-75%). The observed 3% false-negative rate was similar to that of benign cytology, and the missed cancers were all low-risk tumors. Among nodules testing positive, specific groups of genetic alterations had cancer probabilities varying from 59% to 100%.

CONCLUSIONS AND RELEVANCE In this prospective, blinded, multicenter study, the multigene GC test demonstrated a high sensitivity/NPV and reasonably high specificity/PPV, which may obviate diagnostic surgery in up to 61% of patients with Bethesda III to IV indeterminate nodules, and up to 82% of all benign nodules with indeterminate cytology. Information on specific genetic alterations obtained from FNA may help inform individualized treatment of patients with a positive test result.

Published online November 8, 2018. Corrected on December 13, 2018.

Downloaded From: https://jamanetwork.com/ on 12/05/2019
Performance of a Multigene Classifier in Thyroid Nodules With Indeterminate Cytology

Key Points

**Question** Can the diagnosis of benign disease or cancer in thyroid nodules with indeterminate cytology be established by molecular testing instead of diagnostic surgery?

**Findings** This prospective, blinded, multicenter cohort study of a multigene genomic classifier (ThyroSeq v3) test included 257 indeterminate cytology thyroid nodules with informative test results. It demonstrated a high sensitivity (94%) and reasonably high specificity (82%), with 61% of the nodules yielding a negative test result and only 3% residual cancer risk in these nodules.

**Meanings** Up to 61% of patients with indeterminate cytology thyroid nodules may avoid diagnostic surgery by undergoing multigene genomic classifier testing.

**Methods**

**Study Population**

Patients eligible for this study were aged 18 years or older, had 1 or more thyroid nodules, underwent a routine FNA procedure to collect samples for cytological examination, and agreed to provide material for molecular analysis. After FNA cytology was reported, only those patients who had at least 1 nodule that yielded a cytologic diagnosis of Bethesda III, IV, or V and underwent thyroid surgery to remove 1 or more nodules were included in the study.

**Study Design and Sample Collection**

This prospective cohort study recruited 782 patients with 1013 thyroid nodules clinically evaluated at 10 sites, 9 in the United States and 1 in Singapore, between January 2015 and December 2016. All FNA were performed using a 22g, 25g, or 27g needle depending on institutional practice. Samples were collected for molecular analysis by either (1) rinsing the residual high PPV for cancer detection but lacked sufficiently high NPV to reliably exclude malignant disease in test-negative samples. More advanced molecular tests were subsequently developed using gene expression profiling, broader panels of mutational markers, or combinations of different markers. Overall, they offered a significantly improved sensitivity and NPV. However, they suffer from either relatively low specificity and PPV, particularly for certain types of thyroid cancer, such as Hurthle cell tumors, limited clinical validation, and/or lack of reporting specific molecular information for more refined cancer risk assessment.

Recently, a new 112-gene test was developed (ThyroSeq v3 Genomic Classifier [GC]) to include a broad range of thyroid cancer-related point mutations, gene fusions, copy number alterations, and gene expression alterations with the goals of achieving both high sensitivity and specificity in detecting all types of thyroid cancer and providing detailed genomic information on the nodules sampled by FNA biopsy. This prospective, blinded, multicenter clinical validation study was undertaken to assess the diagnostic performance of this GC test in cytologically indeterminate thyroid nodules.

**Findings**

This prospective, blinded, multicenter cohort study of a multigene genomic classifier (ThyroSeq v3) test included 257 indeterminate cytology thyroid nodules with informative test results. It demonstrated a high sensitivity (94%) and reasonably high specificity (82%), with 61% of the nodules yielding a negative test result and only 3% residual cancer risk in these nodules.

**Meanings**

Up to 61% of patients with indeterminate cytology thyroid nodules may avoid diagnostic surgery by undergoing multigene genomic classifier testing.

**Methods**

**Study Population**

Patients eligible for this study were aged 18 years or older, had 1 or more thyroid nodules, underwent a routine FNA procedure to collect samples for cytological examination, and agreed to provide material for molecular analysis. After FNA cytology was reported, only those patients who had at least 1 nodule that yielded a cytologic diagnosis of Bethesda III, IV, or V and underwent thyroid surgery to remove 1 or more nodules were included in the study.

**Study Design and Sample Collection**

This prospective cohort study recruited 782 patients with 1013 thyroid nodules clinically evaluated at 10 sites, 9 in the United States and 1 in Singapore, between January 2015 and December 2016. All FNA were performed using a 22g, 25g, or 27g needle depending on institutional practice. Samples were collected for molecular analysis by either (1) rinsing the residual high PPV for cancer detection but lacked sufficiently high NPV to reliably exclude malignant disease in test-negative samples. More advanced molecular tests were subsequently developed using gene expression profiling, broader panels of mutational markers, or combinations of different markers. Overall, they offered a significantly improved sensitivity and NPV. However, they suffer from either relatively low specificity and PPV, particularly for certain types of thyroid cancer, such as Hurthle cell tumors, limited clinical validation, and/or lack of reporting specific molecular information for more refined cancer risk assessment.

Recently, a new 112-gene test was developed (ThyroSeq v3 Genomic Classifier [GC]) to include a broad range of thyroid cancer-related point mutations, gene fusions, copy number alterations, and gene expression alterations with the goals of achieving both high sensitivity and specificity in detecting all types of thyroid cancer and providing detailed genomic information on the nodules sampled by FNA biopsy. This prospective, blinded, multicenter clinical validation study was undertaken to assess the diagnostic performance of this GC test in cytologically indeterminate thyroid nodules.

**Findings**

This prospective, blinded, multicenter cohort study of a multigene genomic classifier (ThyroSeq v3) test included 257 indeterminate cytology thyroid nodules with informative test results. It demonstrated a high sensitivity (94%) and reasonably high specificity (82%), with 61% of the nodules yielding a negative test result and only 3% residual cancer risk in these nodules.

**Meanings**

Up to 61% of patients with indeterminate cytology thyroid nodules may avoid diagnostic surgery by undergoing multigene genomic classifier testing.

**Methods**

**Study Population**

Patients eligible for this study were aged 18 years or older, had 1 or more thyroid nodules, underwent a routine FNA procedure to collect samples for cytological examination, and agreed to provide material for molecular analysis. After FNA cytology was reported, only those patients who had at least 1 nodule that yielded a cytologic diagnosis of Bethesda III, IV, or V and underwent thyroid surgery to remove 1 or more nodules were included in the study.

**Study Design and Sample Collection**

This prospective cohort study recruited 782 patients with 1013 thyroid nodules clinically evaluated at 10 sites, 9 in the United States and 1 in Singapore, between January 2015 and December 2016. All FNA were performed using a 22g, 25g, or 27g needle depending on institutional practice. Samples were collected for molecular analysis by either (1) rinsing the residual high PPV for cancer detection but lacked sufficiently high NPV to reliably exclude malignant disease in test-negative samples. More advanced molecular tests were subsequently developed using gene expression profiling, broader panels of mutational markers, or combinations of different markers. Overall, they offered a significantly improved sensitivity and NPV. However, they suffer from either relatively low specificity and PPV, particularly for certain types of thyroid cancer, such as Hurthle cell tumors, limited clinical validation, and/or lack of reporting specific molecular information for more refined cancer risk assessment.

Recently, a new 112-gene test was developed (ThyroSeq v3 Genomic Classifier [GC]) to include a broad range of thyroid cancer-related point mutations, gene fusions, copy number alterations, and gene expression alterations with the goals of achieving both high sensitivity and specificity in detecting all types of thyroid cancer and providing detailed genomic information on the nodules sampled by FNA biopsy. This prospective, blinded, multicenter clinical validation study was undertaken to assess the diagnostic performance of this GC test in cytologically indeterminate thyroid nodules.
### Study Outcomes

The primary outcome was the sensitivity, specificity, NPV, and PPV of the multigene GC to predict the histopathologic diagnosis of benign nodule vs cancer/NIFTP in indeterminate thyroid nodules with Bethesda III and IV cytology. In data analysis, NIFTP was grouped together with cancer because it also represents a tumor type that requires surgery based on current practice guidelines.\(^{17,18}\) The secondary outcome was the prediction of cancer/NIFTP by specific genetic alterations in Bethesda III, IV, and V cytology nodules.

### Statistical Analysis

For the primary and secondary outcomes, the test sensitivity, specificity, PPV, and NPV with 95% Wilson confidence intervals were calculated\(^ {33}\) for individual nodules using the consensus diagnosis of central pathology as the reference standard. Using observed sensitivity and specificity, hypothetical positive and negative predictive value curves were calculated over the entire range (0%-100%) of possible disease prevalence. Among patients with nodules yielding indeterminate cytology, baseline characteristics of the included and excluded patients were compared using the Wilcoxon test and Fisher exact test. Statistical analysis was conducted with the R software package (version 3.4.2, R Foundation).\(^ {34}\) Sample size justification and the programming code used to generate results are described in eMethods 2 in the Supplement.

### Results

#### Patients and Nodules

Of the 256 eligible patients, 202 were female (79%), with a median age of 53 years (range, 18-90 years); biopsied nodules had a median size of 2.4 cm (range, 0.5-7 cm). Among the 286 eligible samples, FNA cytology diagnosis was Bethesda III in 172, Bethesda IV in 101, and Bethesda V in 13 cases. Based on the results of central pathology review, 206 nodules (72%) were classified as benign, 69 (24%) as malignant, and 11 (4%) as NIFTP. The prevalence of conditions requiring surgery, ie, cancer and NIFTP, was 28% in the entire cohort, ranging from 9% to 60% among study sites (eTable 2 in the Supplement).

---

**Figure 1. Recruitment and Exclusion of Patients and Samples in the Study**

<table>
<thead>
<tr>
<th>Potential eligible participants (1013 samples)</th>
<th>782</th>
<th>464 Excluded (663 samples)</th>
<th>464 Did not meet cytology eligibility criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>464 Excluded (663 samples)</td>
<td></td>
<td>318 Indeterminate cytology (350 samples)</td>
<td></td>
</tr>
<tr>
<td>62 Excluded (64 samples)</td>
<td></td>
<td>256 Eligible participants (286 samples)</td>
<td></td>
</tr>
<tr>
<td>54 Did not have surgery (55 samples)</td>
<td></td>
<td>24 Excluded (29 samples)</td>
<td></td>
</tr>
<tr>
<td>6 Did not have surgery (5 samples)</td>
<td></td>
<td>14 Test-negative samples</td>
<td></td>
</tr>
<tr>
<td>2 No FNA sample provided (2 samples)</td>
<td></td>
<td>152 Test-negative samples</td>
<td></td>
</tr>
<tr>
<td>1 Under 18 y old (1 sample)</td>
<td></td>
<td>Final diagnoses</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>147 Benign</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 Cancer+NIFTP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>105 Test-positive samples</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final diagnoses</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>34 Benign</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>71 Cancer+NIFTP</td>
<td></td>
</tr>
</tbody>
</table>

FNA indicates fine-needle aspiration; TNA, total nucleic acids; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features.

---

Molecular Analysis

The ThyroSeq v3 GC is a targeted next-generation sequencing test that interrogates selected regions of 112 thyroid cancer-related genes for point mutations, insertions/deletions, gene fusions, copy number alterations, or gene expression alterations.\(^ {32}\) The assay was performed at the UPMC Molecular and Genomic Pathology laboratory.\(^ {32}\) The genomic classifier was applied to assign a value to each detected genetic alteration based on the strength of association with malignancy: 0 (no association with cancer), 1 (low cancer probability), or 2 (high cancer probability). A GC score calculated for each sample is a sum of individual values of all detected alterations, with GC scores 0 and 1 accepted as test negative (score 1 commercially reported as currently negative) and scores 2 and above as test positive.\(^ {32}\)

---

Material in the aspiration needle from all passes or (2) collecting a dedicated pass into a preservative solution tube (ThyroSeqPreserve) and stored at −20°C. Samples from nodules diagnosed as Bethesda III, IV, or V with surgical follow-up were retained as eligible and shipped to the University of Pittsburgh Medical Center (UPMC) for GC testing. Application of the eligibility criteria resulted in 256 patients with thyroid nodules that yielded 286 FNA samples available for molecular analysis (Figure 1). Central pathology review was performed on 274 (96%) nodules by a panel of expert thyroid pathologists (eMethods 1 in the Supplement).

The study was double-blinded; neither cytologists nor pathologists were aware of molecular analysis results and none of the personnel involved in performing molecular analysis were aware of cytology and histopathology results. The study was approved by the institutional review boards or ethics committees of all participating study sites. Written informed consent was obtained and patients were not compensated for participation. The study protocol is available (ClinicalTrials.gov identifier: NCT02352766).
Molecular Analysis
Of 286 samples subjected to molecular analysis, 20 (7%) failed a presequencing step owing to low total nucleic acid quantity reflecting low sample cellularity, and 9 (3%) were inadequate on postsequencing analysis because the expression of thyroid cell markers was below the established acceptable level. Thus, 257 (90%) samples from 232 patients were informative for molecular analysis comprising the final study set. It included samples from 154 Bethesda III, 93 Bethesda IV, and 10 Bethesda V nodules. Molecular analysis yielded a negative test result in 152 (59%) samples and a positive result in 105 (41%) samples (eTable 4 in the Supplement). Among all 318 patients with indeterminate cytology, baseline characteristics of the included and excluded (Figure I) patients and nodules were similar (eTable 3 in the Supplement).

Overall Test Performance
The primary outcome of this study was the accurate separation of histopathological benign nodules from cancer and NIFTP in samples with Bethesda III and IV cytology. Table 1 summarizes the test sensitivity, specificity, NPV, and PPV in these cytologic groups. Overall, in Bethesda III and IV nodules, a negative or benign call rate was 61%.

Table 1. Performance of the Genomic Classifier Test in Cytologically Indeterminate Thyroid Nodules

<table>
<thead>
<tr>
<th>Performance in Bethesda III nodules (n = 154; disease prevalence 23%)</th>
<th>Performance in Bethesda IV nodules (n = 93; disease prevalence 35%)</th>
<th>Performance in Bethesda III and IV nodules (n = 247; disease prevalence 28%)</th>
<th>Performance Across the Entire Cohort (n = 257; Disease Prevalence 30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result</td>
<td>Cancer+NIFTP (n = 35)</td>
<td>Benign (n = 119)</td>
<td>Test performance, % (95% CI)</td>
</tr>
<tr>
<td>Positive</td>
<td>32</td>
<td>18</td>
<td>Sensitivity, 91 (77-97)</td>
</tr>
<tr>
<td>Negative</td>
<td>3</td>
<td>101</td>
<td>Sensitivity, 97 (85-100)</td>
</tr>
<tr>
<td>Positive</td>
<td>64</td>
<td>33</td>
<td>Sensitivity, 94 (86-98)</td>
</tr>
<tr>
<td>Negative</td>
<td>4</td>
<td>146</td>
<td>Sensitivity, 93 (86-97)</td>
</tr>
</tbody>
</table>

Table 2. Test Performance in Specific Histopathologic Types of Thyroid Lesions

<table>
<thead>
<tr>
<th>Histopathologic Diagnosis</th>
<th>Nodules, No. (%)</th>
<th>Test Correctly Classified, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>Hyperplastic follicular cell nodule</td>
<td>95 (37)</td>
</tr>
<tr>
<td></td>
<td>Hyperplastic Hürthle cell nodule</td>
<td>5 (2)</td>
</tr>
<tr>
<td></td>
<td>Follicular adenoma</td>
<td>47 (18)</td>
</tr>
<tr>
<td></td>
<td>Hürthle cell adenoma</td>
<td>34 (13)</td>
</tr>
<tr>
<td></td>
<td>NIFTP</td>
<td>11 (4)</td>
</tr>
<tr>
<td>Malignant</td>
<td>Papillary thyroid carcinoma</td>
<td>49 (19)</td>
</tr>
<tr>
<td></td>
<td>Follicular thyroid carcinoma</td>
<td>4 (2)</td>
</tr>
<tr>
<td></td>
<td>Hürthle cell carcinoma</td>
<td>10 (4)</td>
</tr>
<tr>
<td></td>
<td>Medullary thyroid carcinoma</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td></td>
<td>Metastatic carcinoma a</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Total</td>
<td>257 (100)</td>
<td>85 (80-89)</td>
</tr>
</tbody>
</table>

Abbreviations: NIFTP, Noninvasive follicular thyroid neoplasm with papillary-like nuclear features; NPV, negative predictive value; PPV, positive predictive value.

Abbreviation: NIFTP, Noninvasive follicular thyroid neoplasm with papillary-like nuclear features.

a Considering positive test result for NIFTP as correct classification.
b Metastatic renal cell carcinoma.
Among 105 cases with positive GC results, the probability of cancer/NIFTP in specific genetic alteration groups is presented in Table 3. Thirteen nodules were positive for either TERT (and HRAS) or TP53 mutations, of which 1 was a widely invasive follicular carcinoma and the other was a multifocal papillary carcinoma on surgical pathology. Thirteen nodules were positive for either BRAF V600E mutation or NTRK3, BRAF, or RET fusion. Histopathologically, these were all cancers, primarily classical papillary carcinomas. Another 60 nodules were positive for RAS, BRAF K601E, PTEN, IDH2, or Dicer1 mutation, or Pparg-Thada fusion. In this group, cancer/NIFTP was found in 37 of 60 (62%) cases and histologically benign nodules in 23 of 60 (38%); most of the cancers were follicular patterned, either follicular variant papillary or follicular carcinomas. Most common mutations (n = 45) involved RAS genes, which were associated with a diagnosis of cancer or NIFTP in 72% for HRAS, 52% for NRAS, and 40% for KRAS. Twenty-two nodules were positive for copy number alterations alone. Cancer/NIFTP was found in 13 (59%) of those, and this group was enriched in Hürthle cell carcinoma and follicular variant papillary carcinoma. Finally, 8 samples were positive for gene expression alterations alone (Table 3).

Of 34 test-positive nodules that were pathologically benign on surgery, 23 (67%) were adenomas and 11 (32%) hyperplastic nodules. Among the 34 test-positive nodules that were pathologically benign on surgery, 23 (67%) were adenomas and 11 (32%) hyperplastic nodules. Of 152 test-negative samples in the study cohort, 5 (3%) were found to be false-negative, all having a GC score of 0. They included samples from 3 Bethesda III cytology nodules, 1 Bethesda IV, and 1 Bethesda V (eTable 5 in the Supplement). Among them, there were 4 papillary carcinomas and 1 minimally invasive follicular carcinoma. These were all T1 or T2 tumors (I-4 cm), intrathyroidal and without vascular invasion or clinical evidence of nodal or distant metastasis.

**Cancer Probability in Specific Genetic Alteration Groups**

Among 105 cases with positive GC results, the probability of surgery-requiring disease, defined as cancer or NIFTP, varied depending on specific genetic alterations (Table 3). Two nodules had high-risk TERT or TP53 mutations, of which 1 was a widely invasive follicular carcinoma and the other was a multifocal papillary carcinoma on surgical pathology. Thirteen nodules were positive for either BRAF V600E mutation or NTRK3, BRAF, or RET fusion. Histopathologically, these were all cancers, primarily classical papillary carcinomas. Another 60 nodules were positive for RAS, BRAF K601E, PTEN, IDH2, or Dicer1 mutation, or Pparg-Thada fusion. In this group, cancer/NIFTP was found in 37 of 60 (62%) cases and histologically benign nodules in 23 of 60 (38%); most of the cancers were follicular patterned, either follicular variant papillary or follicular carcinomas.
ThyroSeq GC had a benign or negative call rate of 61% in GSC (eTable 7 in the Supplement). Furthermore, 91% GSC) but a specificity of 82% vs 52% in GEC and 68% overall similar sensitivity (94% ThyroSeq vs 90% GEC and comparable size validation studies, ThyroSeq GC shows an clinical studies. 10

Whereas sensitivity and specificity characterize a test independently of disease prevalence, NPV and PPV depend on the prevalence of disease in the studied population. Based on the fixed sensitivity and specificity, Bayes theorem can predict the test NPV and PPV along the spectrum of disease prevalence. 35 For the GC test, it predicts a robust NPV of 95% or higher, required to consider nonsurgical treatment by the NCCN guidelines, 36 up to a disease prevalence of 40% in Bethesda III and 60% in Bethesda IV nodules (Figure 2). This is within the range of cancer/NIFTP probability expected based on the Bethesda reporting system 12,13 and observed in most clinical studies. 10

Another commonly used molecular test for thyroid FNA samples is based on measuring expression of multiple genes either by the microarray assay (Gene Expression Classifier; GEC) 27 or RNA-Seq (Gene Sequencing Classifier; GSC). 37 In a comparable size validation studies, ThyroSeq GC shows an overall similar sensitivity (94% ThyroSeq vs 90% GEC and 91% GSC) but a specificity of 82% vs 52% in GEC and 68% in GSC (eTable 7 in the Supplement). Furthermore, ThyroSeq GC had a benign or negative call rate of 61% in indeterminate Bethesda III and IV nodules, with 82% of all histologically benign nodules yielding a negative test result. This indicates that ThyroSeq GC can prevent diagnostic surgeries for up to 61% of all of indeterminate Bethesda III to IV cytology nodules and as many as 82% of all benign nodules that yielded indeterminate cytology diagnosis. This should maximize the effect of molecular testing on the avoidance of surgery, reduction of health care costs, and improvement of patient quality of life. This is particularly important during what is widely considered as the era of thyroid cancer overdiagnosis 38 and overtreatment. 3

The multigene GC test showed robust performance in detecting all types of thyroid cancer, including Hürthle cell carcinoma. To date, the performance of existing molecular FNA tests in Hürthle cell nodules has been either not specifically reported, 28-30 not validated at all, 31 or observed to have very low specificity.39,40 In this study, all 10 Hürthle cell carcinomas were correctly classified, whereas in all types of Hürthle cell nodules the GC test negative call rate was 53%. This should allow the avoidance of diagnostic surgery in more than half of biopsied Hürthle cell nodules.

Another potential advantage of the GC test is that it provides a molecular profile of the test-positive nodules, which may help clinicians to refine the treatment of patients with Bethesda III, IV, and V nodules and a positive test. Indeed, the finding of BRAF V600E and similar alterations as well as high-risk (TERT, TPS3) mutations conferred a 100% probability of cancer in this study, in keeping with previous reports.41-43 Tumors harboring a BRAF V600E mutation are classic papillary carcinoma with a higher rate of regional lymph node metastasis. 3,19 On the contrary, RAS-like alterations were associated with a spectrum of follicular-pattern thyroid tumors, from pathologically benign adenomas to borderline NIFTP and fully invasive cancers, with a roughly 60% probability of cancer/NIFTP. These cancers are frequently encapsulated and if spread, they typically skip regional lymph nodes and metastasize hematogenously. 19 However, most thyroid cancers driven by single RAS and RAS-like mutations are minimally
invasive and low risk. The histologically benign nodules carrying these mutations are monoclonal tumors, in contrast to polyclonal hyperplastic nodules which are the most common type of benign thyroid nodules. Finally, the GC test correctly classified nodules composed of nonthyroid follicular cells, including medullary carcinoma and a metastatic tumor. This additional information on the test-positive nodules along with clinical factors may help to further individualize patient treatment.

As genetic information becomes available preoperatively, future studies are required to better understand how this information should be integrated with ultrasound and other clinical data to inform more tailored treatment of patients with thyroid nodules and cancers that have different molecular profiles. Furthermore, prospective studies will be needed to determine whether patients with the molecular signature of low-risk cancer or NIFTY can have surgery safely delayed or replaced by medical surveillance, as is currently under consideration for small thyroid cancers.44,45

Limitations
This study has several limitations. By selecting patients based on the Bethesda reporting system for thyroid cytology, the applicability of the findings is limited to practices that use this reporting system. The observed small number of samples from Bethesda V nodules did not allow meaningful test validation in this subset of nodules. By surgically removing nodules with low cancer probability genetic alterations (GC score 1) for final histological diagnosis, the long-term clinical impact of these alterations could not be established. Finally, this study was performed at moderate- to high-volume centers with established thyroid nodule imaging and clinical expertise. Thus, the results may differ for practices that have a different setting and diagnostic approaches to thyroid nodules.

Conclusions
The study documents a high sensitivity and correspondingly high NPV of the ThyroSeq GC test for Bethesda III and IV indeterminate cytology nodules, which together with high specificity may prevent diagnostic surgeries in the majority of such patients. The availability of detailed genetic information in test-positive cases may help to further inform individualized treatment for these patients after integration with imaging and other clinical information.

ARTICLE INFORMATION
Accepted for Publication: August 10, 2018.
Correction: This article was corrected on December 13, 2018, to fix typographical errors in the Methods section and Table 2 and to replace the term “neoplastic disease” with “surgery-requiring disease” in the Results section.
Published Online: November 8, 2018. doi:10.1001/jamaoncology.2018.4616
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2018 Steward DL et al. JAMA Oncology.
Author Affiliations: Department of Otolaryngology, Head and Neck Surgery, University of Cincinnati Medical Center, Cincinnati, Ohio (Steward); Division of Endocrine Surgery, University of Pittsburgh, Pittsburgh, Pennsylvania (Carty, McCoy, Yip); Division of Endocrine Surgery, University of Wisconsin, Madison (Sippel); Endocrinology Division, Department of Medicine, National University Hospital, Singapore, Singapore (Otalvaro); Division of Endocrinology, Diabetes & Metabolism, The Ohio State University School of Medicine, Columbus (Sipos, Khawaja, Ringel); Diabetes & Endocrine Care, St Peter’s Health Partners, Rensselaer, New York (Figge, Folek); Division of Endocrinology, Diabetes & Metabolism, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Mandel); Division of Endocrinology, Metabolism and Diabetes, University of Colorado School of Medicine, Aurora (Haugen, Mayson); Department of Medicine, Endocrinology Section, MedStar Washington Hospital Center, Washington, DC (Burman, Chiosea, Gomes-Lima, Kundra); Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia (Baloch); Department of Pathology and Laboratory Medicine, University of Wisconsin, Madison (Lloyd); Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania (Seethala, Nikiforova, Nikiforov); Biostatistics Facility, UPMC Hillman Cancer Center, Pittsburgh, Pennsylvania (Gooding); Departments of Otolaryngology and Immunology, UPMC Hillman Cancer Center, Pittsburgh, Pennsylvania (Ferris); Department of Otolaryngology-Head and Neck Surgery, National University Hospital, Singapore (LoH); Department of Pathology, University of Colorado School of Medicine, Denver, Colorado (Marshall); Department of Pathology, National University Hospital, Singapore (Ngam); Department of General Surgery, University Surgical Cluster, National University Hospital, Singapore (Ngiam); Division of Endocrinology, Diabetes & Metabolism, University of Wisconsin, Madison (Poehls); Department Pathology, University of Cincinnati Medical Center, Cincinnati, Ohio (H. Yang).
Author Contributions: Dr Nikiforov had full access to all the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis. Study concept and design: Steward, Haugen, Baloch, Nikiforov. Acquisition, analysis, or interpretation of data: Steward, Carty, Sipos, Yip, Sosa, Sipos, Figge, Mandel, Haugen, Burman, Lloyd, Seethala, Gooding, Chiosea, Gomes-Lima, Ferris, Folek, Khawaja, Kundra, Loh, Marshall, Mayson, McCoy. En Nga, Ngiam, Nikiforova, Poehls, Yip, Nikiforov. Drafting of the manuscript: Steward, Carty, Sipos, Yip, Sosa, Sipos, Figge, Mandel, Haugen, Baloch, Nikiforov. Administrative, technical, or material support: Sipos, Figge, Haugen, Ferris, Khawaja, Kundra, Loh, Mayson, Sosa, Ngiam, Nikiforov, Sipos, Ngiam, Sosa, Ferris, Baloch, Lloyd, Seethala, Gooding, Chiosea, Gomes-Lima, Ferris, Folek, Khawaja, Kundra, Loh, Marshall, Mayson, McCoy, En Nga, Ngiam, Nikiforov, Poehls, Yip, Nikiforov. Administrative, technical, or material support: Sipos, Figge, Haugen, Ferris, Khawaja, Kundra, Loh, Mayson, Ngiam, Sipos, Ngiam, Sosa, Ferris, Baloch, Lloyd, Seethala, Gooding, Chiosea, Gomes-Lima, Ferris, Folek, Khawaja, Kundra, Loh, Marshall, Mayson, McCoy, En Nga, Ngiam, Nikiforov, Poehls, Yip, Nikiforov. Administrative, technical, or material support: Sipos, Figge, Haugen, Ferris, Khawaja, Kundra, Loh, Mayson, Ngiam, Sipos, Ngiam, Sosa, Ferris, Baloch, Lloyd, Seethala, Gooding, Chiosea, Gomes-Lima, Ferris, Folek, Khawaja, Kundra, Loh, Marshall, Mayson, McCoy, En Nga, Ngiam, Nikiforov, Poehls, Yip, Nikiforov. Statistical analysis: Mandel, Gooding. Obtained funding: Nikiforova, Ringel, H. Yang, Nikiforov.
Conflict of Interest Disclosures: Dr Burman received Institutional Grants for Clinical Trials from AstraZeneca, Eisai, IBISA, Bayer, and Loxo. Drs Carty, Ferris, McCoy, Seethala, and Yip are employees of University of Pittsburgh Physicians, which is an affiliate of University of Pittsburgh Medical Center (UPMC). The University of Pittsburgh Medical Center has granted CBLPath Inc. a license to market ThyroSeq for commercial use. They receive no compensation, directly or indirectly, related to CBLPath Inc. Dr Chiosea receives compensation from his employer (UPMC) in connection with ThyroSeq analyses. Dr Mayson previously received research support from Rosetta Genomics. Dr Nikiforov and Dr Nikiforova have intellectual property rights related to ThyroSeq. They receive royalties associated with commercial use of ThyroSeq, Dr Sipos received an honorarium for speaking for Veracyte and Genzyme. Dr Sosa is a member of the Data Monitoring Committee of the Medullary Thyroid Cancer Registry supported by GlaxoSmithKline, Novo Nordisk, Astra Zeneca, and Eli Lilly. Dr Steward received research funding from Rosetta, Veracyte, and GenePreDX. No other disclosures were reported.
Thyroid nodules with indeterminate cytology. Ultrasound Risk Assessment of thyroid nodules: a review. JAMA. 2018;319(9):914-924.


