An Independent Study of a Gene Expression Classifier (Afirma™) in the Evaluation of Cytologically Indeterminate Thyroid Nodules

Bryan McIver, MB, PhD,1, M. Regina Castro, MD,2, John C. Morris, MD2, Victor Bernet, MD,3, Robert Smallridge, MD,3, Michael Henry, MD,4, Laura Kosok2, and Honey Reddi,PhD5

1Department of Head and Neck, and Endocrine Oncology, Moffitt Cancer Center, 12902 Magnolia Dr, Tampa, FL, 2Division of Endocrinology, Mayo Clinic, 200 First St SW, Rochester, MN; 3Division of Endocrinology Mayo Clinic, 4201 Belfort Rd Jacksonville, FL; 4Department of Surgical Pathology, Mayo Clinic, Rochester, MN; and 5Prevention Genetics, 3700 Downwind Dr., Marshfield, WI

Context: Molecular markers hold the promise of improved diagnostic yield in thyroid fine needle biopsy. The Afirma™ GEC, available commercially, reports a negative predictive value of 94% in the diagnosis of benign nodules following indeterminate cytology. However, there are currently no independent studies of the performance of this assay.

Objective: To assess the performance of the Afirma® Gene Expression Classifier (GEC) in an Academic Medical center.

Design: Samples for the GEC were collected according to the manufacturer’s recommended protocol, from patients undergoing thyroid FNA. We requested GEC analysis on nodules reported cytologically as follicular neoplasm orAtypia or Follicular Lesion of Undetermined Significance (AUS/FLUS), from patients willing to defer surgery.

Patients: All patients undergoing thyroid FNA during the study period, whose cytology was reported as Follicular Neoplasm or AUS/FLUS, were offered access to the test and recruited to this study.

Intervention: Patients whose GEC classifier was “benign” were offered ultrasound follow-up in lieu of surgery. Those with a “suspicious” GEC were recommended to undergo diagnostic lobectomy.

Setting: Large Academic Medical Center

Main Outcome Measure: Rate of benign and suspicious calls from the Afirma GEC and histological diagnosis following surgery.

Results: A total of 72 nucleic acid samples were sent for GEC analysis. In 12 (17%) of these samples, there was insufficient mRNA, leaving 60 Afirma results for analysis. Of these, 17 (28%) were “Benign”, while 43 (72%) were “Suspicious”. The rate of confirmed malignancy in GEC-suspicious nodules was only 16%.

Conclusion: The Afirma GEC demonstrates a lower than expected rate of benign reports in FN/HCN, and a lower than anticipated malignancy rate within GEC-suspicious nodules. These data suggest that the Positive Predictive Value of the GEC is lower than previously reported, and call into question the performance of the test when applied in the context of specialized Academic cytopathology.
Thyroid nodules are the most common endocrine tumor, with population-based screening studies identifying clinically palpable nodules in ~5% of adults, while ultrasound and autopsy studies demonstrate nodules in more than 50% of women and 20% of men over age 50 (1). Thyroid cancer, which usually presents as a nodule, is uncommon, but increasing in incidence more rapidly than any other cancer type (2). However, only a small minority of nodules, whether palpable or incidentally discovered, proves to be malignant (3). Selection of nodules for biopsy based on suspicious ultrasound features enriches the yield of malignant nodules, but the proportion that ultimately proves malignant remains a mere 10% - 15% (4).

Both the American Thyroid Association (ATA) and the American Association of Clinical Endocrinologists (AACE), working in collaboration with the Associazione Medica Endocrinologi and the European Thyroid Association, have published guidelines for the evaluation of thyroid nodules, which recommend a multistep strategy: clinical assessment, measurement of thyrotropin (TSH), ultrasound evaluation, and biopsy of nodules selected according to size and ultrasound characteristics (5, 6). For those nodules that require biopsy, Fine Needle Aspiration (FNA) cytology provides sufficient information to classify most nodules as benign (72%: range 62 - 85%), while approximately 5% (range 1 - 8%) of nodules are cytologically malignant, most often papillary thyroid carcinoma (PTC) (7).

However, 10 – 30% of biopsied nodules exhibit “indeterminate” cytology, including the subtypes of Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance (AUS/FLUS); (Suspicious for) Follicular or Hürthle Cell Neoplasm (FN/HCN); and Suspicious for Malignancy (8). Current guidelines recommend surgical resection for most of these nodules, to permit adequate pathological evaluation, though repeat biopsy is supported for AUS/FLUS, when the risk of malignancy is felt to be sufficiently low, which occurs particularly when the categorization of AUS/FLUS is driven primarily by features other than nuclear atypia (<10% vs > 50% risk of malignancy) (9).

Overall, only 15% – 35% of nodules with indeterminate cytology prove to be malignant on histological evaluation, usually either follicular variant of PTC, or follicular carcinoma (7, 8). For these cancers, a lobectomy – the diagnostic procedure of choice – is regarded as inadequate therapeutic surgery in most cases, so these patients will require a second surgical procedure, to complete the initial treatment of their malignancy. A more specific test for the preoperative diagnosis of malignancy might aid in determining the extent of initial surgery. For the 65% - 85% of indeterminate nodules that prove benign, however, a lobectomy is arguably too much surgery, required only as a diagnostic, rather than a therapeutic procedure, in the absence of local compressive symptoms or cosmetic concerns. A more sensitive preoperative test that would allow safe identification of clearly benign nodules with indeterminate cytology raises the prospect of avoiding a purely diagnostic surgery. It is in these areas that molecular markers have been developed and are being marketed to improve the preoperative evaluation of thyroid nodules.

The Afirma® Gene Expression Classifier (GEC) is a proprietary diagnostic test developed by Veracyte, Inc. (South San Francisco, Calif.) for the preoperative identification of benign thyroid nodules with indeterminate cytology. Testing is offered through a single CLIA-certified reference laboratory. The assay classifies nodules as either benign, or suspicious for malignancy, with reported post-test probability of malignancy of 5 – 6% for a “benign” result and 37% - 38% for a “suspicious” result in AUS/FLUS and FN/HCN, usually described as a negative predictive value (NPV) of 94 – 95% and positive predictive value (PPV) of 37 – 38% (10, 11). Because the risk of malignancy for a thyroid nodule with AUS/FLUS or FN/HCN and a benign GEC diagnosis is reported to be comparable to that of an operated nodule with a benign cytopathology diagnosis (11, 12) observation or ultrasound follow-up has been recommended in lieu of thyroid surgery, in a recent revision of guidelines for the management of thyroid cancer, issued by the National Cancer Cooperative Network (NCCN) (13). The data that supports this approach, however, is limited to a single confirmation study, which encompassed only 81 FN/HCN and 129 AUS/FLUS nodules, with wide confidence intervals for sensitivity (68% - 99%) and specificity (36% - 63%), and significant consequent uncertainty in both the NPV (79% - 99%) and PPV (23% - 52%) of the test, even as applied within this trial setting, with a known pretest probability of malignancy (11). Two subsequent studies have been reported, demonstrating higher PPV (54% and 57% respectively) and implying a lower NPV (estimated at 92% and 90%), though neither of these studies was able directly to assess NPV because most patients with benign GEC results have not undergone surgery (14, 15). To increase the reported experience of the GEC, we therefore set out to assess the performance of the test as applied in a large Academic, multispecialty clinic setting, using both highly specialized cytopathology services and a clinical assessment to determine which samples should be reflexed to the GEC assay.

Materials and Methods

Mayo Clinic was the first center that was not part of the industry-sponsored confirmatory study (11) to receive the Veracyte des-
ignation of “Enabled Center”. This designation allowed our group, and subsequently a limited number of other Academic centers, to send appropriately collected nucleic acid specimens directly for Afirma GEC analysis, based on our in-house specialized cytopathology services, rather than being required to use Thyroid Cytopathology Partners (TCP, Austin, Texas), the cytopathology group chosen by Veracyte to act as gatekeepers to the commercial assay. Starting in May 2011, we began to offer access to the GEC to adult patients over age 21 years, with nodules > 1.0cm, whose cytology was reported as FN/HCN, or AUS/FLUS. After obtaining written informed consent, an ultrasound guided fine needle aspiration biopsy (US-FNA) was conducted in the usual way, using 4–6 needle passes with a 27-gauge needle, placed into the index nodule under direct ultrasound control. Either capillary action or gentle aspiration was used to obtain cytological material in a minimally traumatic fashion and the aspirated material was spread onto labeled glass slides and fixed in ethanol preservative for cytological evaluation. At the end of the procedure, either 1 or 2 additional needle passes were collected from the index nodule and washed into the GEC collection tube, containing a proprietary nucleic acid preservative, following the protocol recommended by Veracyte.

Patients were offered access to the test if they attended Mayo Clinic for assessment of a thyroid nodule and underwent thyroid FNA within the Endocrine Clinic in Rochester, MN between May 2011 and December 2012; in Jacksonville, FL, between October 2011 and December 2012; or in the Radiology department of Mayo Clinic, Rochester, between December 2011 and December 2012. Afirma GEC samples were labeled and stored at 4 degrees Centigrade, pending receipt of the cytology report, typically within 2–4 hours. Following a further discussion with the patient, the samples were shipped to Veracyte for GEC analysis, from nodules that were reported either as FN/HCN or AUS/FLUS, for patients who were willing to defer surgery and who were deemed not to be at particularly high risk for malignancy (no history of head or neck irradiation; negative family history of thyroid cancer; no prior history of thyroid cancer; no worrisome imaging characteristics). Patients with symptomatic nodules, with worrisome imaging features, or with significant risk factors for thyroid cancer, in whom a decision had already been made to take the patient to surgery, were not recommended for the GEC. Their nucleic acid samples were disposed of, along with those from patients with cytology that was unsatisfactory, definitively benign, malignant, or suspicious for malignancy.

Samples were shipped on ice to Veracyte’s CLIA laboratory in South San Francisco, by overnight courier, using the cold-packs and shipping containers provided by Veracyte. Samples that would arrive on Saturday morning were held over the weekend at 4 degrees Centigrade, and shipped on Monday morning for a Tuesday morning delivery. Afirma-GEC reports were received, typically within 7–10 days and scanned into the patients’ electronic medical record.

We maintained a prospective register of patients undergoing the GEC assay, including demographic, ultrasonographic, cytological, surgical, histological and follow-up findings. Patients with AUS/FLUS or FN/HCN and a GEC classifier result of “suspicious” were recommended to undergo diagnostic lobectomy, intraoperative frozen section and immediate completion thyroidectomy, if malignancy was diagnosed. Patients whose GEC classifier was reported to be “benign” were offered the option of continued ultrasound follow-up, with plans to intervene surgically only if the nodule enlarged or changed over time. A structured follow-up plan was implemented to ensure these patients were not lost to follow-up, with repeat ultrasound recommended after 3–6 months, and then annually for at least 5 years.

Statistical analysis was performed using the SPSS Statistical Package. Comparisons were made using binomial distribution statistics and confidence intervals were estimated using the normal distribution approximations.

Results

As illustrated in Figure 1, we performed FNA biopsies on 1207 nodules from 984 patients within the “enabled system”, which permitted collection of a nucleic acid sample. Of these, 12 (1.0%) were categorized as AUS/FLUS on cytology and 93 (7.7%) were FN/HCN, rates that are in keeping with previously published data from this Institution, though lower than those reported from several other Academic and community centers (14). The necessary nucleic acid sample was not collected in 15 of these patients, at the discretion of the Physician performing the biopsy, leaving a total of 90 GEC’s available for analysis from cytologically indeterminate nodules in eligible patients (10 with a cytologic diagnosis of AUS/FLUS, 17 HCN and 63 FN).

Of these 90 patients, 18 (20%) chose to move ahead with a diagnostic or therapeutic lobectomy, rather than consider use of the GEC. In 6 of these cases, the decision was prompted by concerns about a highly suspicious appearance on ultrasound (n = 3), the synchronous discovery of a contralateral malignancy (n = 2), or (in 1 case) the finding that the nodule was PET-positive, leading to increased concern about a malignant condition. In an additional 4 cases, the nodule was large and/or causing compressive symptoms. The remaining 8 patients underwent surgery because either the patient or their treating Physician was sufficiently concerned about the nodule that a decision not to operate would not have been an acceptable choice. Consequently, the GEC was not performed on any of these 18 patients. Thyroid lobectomy was performed for diagnostic purposes on all of these cases, except the two in whom a contralateral malignancy was identified on preoperative cytology.

A total of 72 nucleic acid samples were sent to Veracyte for GEC analysis. In 12 (17%) of these samples, there was insufficient quantity or quality of mRNA to obtain a result, leaving a total of 60 Afirma results available for analysis. Of these, 17 (28%) were reported as “Benign”, while 43 (72%) were reported as “Suspicious”. Patients with benign GEC results were offered the opportunity to defer diagnostic surgery and undergo ultrasound-based follow-up as an alternative. Fifteen of the 17 patients with benign GEC reports chose this conservative management,
while two patients chose surgery, a rate of conservative
management that closely correlates with the findings of
our previous clinical utility study (12). Two patients with
a GEC-benign classifier (one HCN, one FN) underwent
later surgery, in one case because of the growth of the
index nodule after 11 months of follow up, in the other
case because of patient’s preference 6 months later, despite
lack of growth. Of these four GEC-benign nodules that
were resected, three proved to be histologically benign at
surgery, while the fourth (which demonstrated no growth
on follow up) proved to be a 3.2 cm follicular carcinoma,
exhibiting focal capsular and vascular invasion. Twelve
patients with benign-GEC results remain under active fol-
low-up after a median of 9.5 months.

Of the 44 patients with suspicious GEC results, 32
(73%) have undergone surgery to date. Of the remaining
12 patients, 4 have contraindications to surgery, and 8
patients have chosen to avoid surgery despite medical rec-
ommendations to proceed. A total of 5 cancers have been
identified in the 32 operated thyroids, giving a post-test
probability of cancer of 15.6%. These malignancies en-
compassed 2 classic PTC, 1 follicular variant PTC and 1
Follicular thyroid cancer. The remaining lesion was de-
scribed as either an “atypical adenoma with Hürthle Cell

![Flow diagram of patient disposition and outcomes in the study.](image-url)
features”, or a minimally invasive Hürthle Cell cancer. For the purpose of this analysis, we have considered this nodule to be a malignancy. A total of 31 patients with cytology suspicious for neoplasm (18 FN and 13 HCN) underwent surgery. Table 1 summarizes the results according to GEC category and final histopathology. Five patients with cytology read as AUS/FLUS had samples submitted for GEC. All of these were read as GEC suspicious and underwent surgery. Of these, 1 lesion proved to be malignant (microPTC arising within a larger adenoma), the others showed: 1 follicular adenoma, 1 adenomatous nodule and 2 hyperplastic nodules. Because most of our patients with a benign GEC report have not undergone surgery, we are unable to directly assess the probability of cancer in GEC-benign nodules (the NPV). These patients remain under careful long-term follow-up.

Discussion

Remarkable advances have been made over the last 2 decades in our understanding of the genetic and molecular changes that drive thyroid neoplasia (16). Chromosomal rearrangements involving the RET protooncogene, or the V600E point mutation in the BRAF gene, underpin most PTC, while mutations of RAS and rearrangements of the PPARγ gene have been implicated in a significant proportion of follicular thyroid cancer (FTC). Nikiforov and col-

Table 1. PERFORMANCE OF GENE EXPRESSION CLASSIFIER ACCORDING TO FINAL HISTOPATHOLOGY DIAGNOSIS FOR CYTOLOGICALLY INDETERMINATE (SUSPICIOUS) NODULES THAT UNDERWENT SURGERY

<table>
<thead>
<tr>
<th>GEC RESULTS</th>
<th>HISTOPATHOLOGY</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2†</td>
<td>13**</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>BENIGN</td>
<td>1†</td>
<td>2 #</td>
<td>3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3</td>
<td>15</td>
<td>18</td>
</tr>
</tbody>
</table>

*1 Papillary Ca (PC), 1 Follicular Ca (FC); ** 9 Follicular adenomas (FA), 1 Nodular hyperplasia, 3 Hurthle Cell adenomas (HCA); † 1 Follicular Ca; # 2 Follicular adenomas

SUSPICIOUS FOR HURTHLE CELL NEOPLASM (SHCN) n = 13

<table>
<thead>
<tr>
<th>GEC RESULTS</th>
<th>HISTOPATHOLOGY</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2*</td>
<td>10**</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>BENIGN</td>
<td>0</td>
<td>1 #</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2</td>
<td>11</td>
<td>13</td>
</tr>
</tbody>
</table>

*1 PC, 1 Hürthle Cell Ca (HCC); ** 1 Benign (per report), 1 FA with HC change, 7 HCA, 1 Hashimoto’s thyroiditis; # 1 HCA

SUSPICIOUS FOR NEOPLASM (SFN + SHCN) n = 31

<table>
<thead>
<tr>
<th>GEC RESULTS</th>
<th>HISTOPATHOLOGY</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>BENIGN</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5</td>
<td>26</td>
<td>31</td>
</tr>
</tbody>
</table>

Sensitivity (4/5) 80%; Specificity (3/26) 12%; PPV (4/27) 15%; NPV(3/4) 75%; Prevalence of malignancy = 16%

ATYPIA/FLUS n = 5

<table>
<thead>
<tr>
<th>GEC RESULTS</th>
<th>HISTOPATHOLOGY</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1*</td>
<td>4**</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>BENIGN</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

*1 microPTC, arising in larger adenoma; ** 1 FA; 1 adenomatous nodule, 2 hyperplastic nodules

Performance across the Entire Data set of indeterminate samples (SFN + SHCN + Atypia/FLUS)

<table>
<thead>
<tr>
<th>GEC RESULTS</th>
<th>HISTOPATHOLOGY</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>BENIGN</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6</td>
<td>30</td>
<td>36</td>
</tr>
</tbody>
</table>

Sensitivity (5/6) 83%; Specificity(3/30) 10%; PPV (5/32) 16%; NPV (3/4) 75%; Prevalence of malignancy = 17%.
leagues have been instrumental in developing the concept of using a panel of oncogenes to more accurately diagnose thyroid cancers in the preoperative phase (17, 18). The high specificity of these oncogenes in predicting cancer—particularly BRAF and RET/PTC, with specificities close to 100%—allows the surgeon to plan an appropriate cancer surgery when a mutation is present, even when cytology alone is indeterminate. However, our understanding of oncogenic triggers is incomplete, particularly for Hürthle Cell cancer and Follicular Variant PTC, in which only a minority exhibits RAS mutations (16, 19). Consequently, the currently available oncogene panels lack sensitivity in the detection of these malignant subtypes, resulting in a number of false-negative results.

An alternative approach has been developed by Veracyte Inc. (South San Francisco, CA), marketed under the trade name Afirma™. Based on a commercially available gene expression chip (Affymetrix Inc, Santa Clara, CA), Veracyte has developed a “gene expression classifier” (GEC), using a proprietary algorithm to distinguish “benign” from “suspicious” nodules, based on the expression pattern of mRNA extracted from one or two dedicated FNA needle passes. The algorithm utilizes a screening “cassette” of 25 genes designed to identify medullary thyroid cancer (MTC) and certain metastatic malignancies. Thereafter, the expression levels of 142 genes are processed through a “support vector machine” (SVM) that classifies the expression pattern as either “benign” or “suspicious” (10). The SVM is a supervised machine-learning algorithm that identifies patterns of gene expression in a recursive learning process, which was trained on a set of mRNA samples derived from nodules with known histology. In the only clinical validation study yet published, the performance of the GEC was assessed in a set of 256 FNA samples with indeterminate cytology, including 81 FN/HCN and 129 AUS/FLUS (11). Within the AUS/FLUS group, sensitivity of the GEC was 90%, specificity 53% and NPV 95%, while the figures for FN/HCN were 81 FN/HCN and 129 AUS/FLUS (11). Within the AUS/FLUS group, sensitivity of the GEC was 90%, specificity 53% and NPV 95%, while the figures for FN/HCN were 81%. Within the AUS/FLUS group, sensitivity of the GEC was 90%, specificity 53% and NPV 95%, while the figures for FN/HCN were 81%.

However, a community practice based study, utilizing Veracyte’s chosen cytopathology group, TCP, reported that only 34% of indeterminate samples proved benign on the GEC, closer to our data, with only 28% of samples receiving a benign GEC result. Li et al, in a theoretical cost effectiveness modeling analysis, assumed that surgery could be avoided in “almost three fourths of currently performed surgeries in patients with benign nodules”, concluding that the test would be cost-effective, saving money for the health care system as a whole, despite the high direct cost of the assay (20). The much lower rate of surgery avoidance that could have been achieved by Alexander et al (11) (41% of calls in AUS/FLUS and FN/HCN categories were benign), by Harrell et al (15) (34% benign GEC reports) and in our experience (28% benign) call into question that analysis and suggest that the costs of widespread implementation of this molecular test may be substantially higher than initially reported, because of the greater number of tests needed to be run in order to avoid one surgery. In our study, surgery might have been avoided in one patient for every 4 tests run, whereas one surgery was avoided for every 2 tests run in a study based on TCP cytopathology (12), in which ~50% of tested samples were reported benign on the Afirma GEC (P < .05). Furthermore, because of the uncertainty about the true NPV achieved by the GEC, we have implemented a structured follow-up plan to ensure these patients were not lost to follow-up, with repeat ultrasound after 3 – 6 months, and then annually for at least 5 years. This more intensive follow-up is, we believe, clinically appropriate, but may add to the overall costs of the assay, compared to the analysis by Li et al (20).

In our study, only 15.6% of the nodules that were GEC suspicious, for which final histology is known, have proven malignant, a PPV for malignancy (PPV) substantially and significantly lower than the 38% reported by Alexander et al (P < .05), which was itself rather lower than the 48% PPV in early reports (10, 11) and dramatically lower than the 57% PPV reported recently by Harell et al (15).

The most obvious explanation for this discrepancy might be a significantly lower than reported specificity of the Afirma GEC, when applied to samples from nodules reported as AUS/FLUS or FN/HCN by the specialized thyroid cytopathologists at Mayo Clinic, compared to either Veracyte’s cytopathology partners, TCP, or to the “local cytopathologists” who provided the data for the Veracyte confirmatory study (11). Because we have elected not to operate on all patients with an Afirma-GEC benign result, we are unable to directly calculate the precise sensitivity and specificity of the assay in our hands, a limitation of all the subsequent reports (14, 15). If we assume that the
single false negative result we have identified (one GEC-
benign patient had a Follicular carcinoma), is the only
such error, the assay would have a sensitivity in our hands
of 83% (Table 2), well within the 95% confidence interval
(CI) of 68% - 99% reported by Alexander et al (11). How-
ever, the lower incidence of cancer we have identified in
GEC-suspicious nodules yields a specificity of Afirma in
our study of only 36%, significantly lower than the 49%
and 53% specificity reported by Alexander et al (P < .05)
in FN/HCN and AUS/FLUS, respectively (11). The second
contributor to the lower PPV for Afirma in our experience
is likely to be the low pretest probability of malignancy in
the group of nodules we have selected for GEC analysis.
Only 5 cancers have been identified in the 32 nodules that
have gone to surgery so far, for a histological malignancy
rate of 15.6%. Including the 1 false negative within the
GEC-benign group, the overall malignancy rate in our
study is only ~12.5%. Although low, this probability of
malignancy is close to the 14% risk of malignancy re-
ported for FN/HCN at Mayo Clinic in previous studies
(21) that were completed without use of genetic markers,
and is similar to the rates of malignancy reported by sev-
eral other large Academic medical centers, including MD
Anderson Cancer Center (16%) and Northwestern Uni-
versity (15%), although it stands in contrast to some other
centers, including Yale University (48%) and Washington
Hospital Medical Center (49%): see Wang et al for a de-
tailed meta-analysis (22). The pretest risk of malignancy in
our study may have been further lowered (from 14% to
~12.5%) by our decision not to send for GEC analysis
samples from patients we deemed to be at high risk for
cancer. Patients with compressive symptoms, large nod-
ules, highly suspicious ultrasound features, or other risk
factors for cancer were recommended to undergo surgery
and the GEC was not requested on these patients (Figure
1), a decision that may have depleted the study pool of
some cancers, but which, in our view, was medically ap-
propriate. The rate of malignancy among this group of
patients was, indeed, slightly higher than among the pa-
tients for whom the GEC was sent. Four of the 18 (22%)
patients for whom the GEC was drawn, but not analyzed,
had malignancy diagnosed at the time of surgery, of which
3 (17%) were cancers within the index nodule. However,
because the decision to operate had already been made in
these patients, we believe that the GEC would have rep-
resented an unnecessary medical expenditure, which
would not have altered the management of the patient.

The consequence of this low pretest probability of ma-
lignancy is to lower the post-test probability of malign-
ancy, both for benign and suspicious results. The rela-
tionship between the pretest probability of malignancy
and the achieved NPV and PPV of the Afirma GEC is
shown in Figure 2A, assuming a sensitivity of 90% and a
specificity of 49% for the test, as reported by Alexander et
al (11). As the pretest probability of malignancy increases,
the achieved NPV falls, while the PPV rises. Consequently,
the PPV and NPV achieved by the Afirma GEC assay de-
pend critically on the pretest probability of malignancy,
and indeed can be defined only when that risk of malign-
ancy is known. An NPV of 94% and PPV of 37%, as
reported by Alexander et al in FN/HCN, are achieved with
a pretest risk of malignancy of 25%. However, based on
the same test sensitivity and specificity, with a pretest risk
of malignancy of 12%, the GEC is predicted to offer a PPV
of 17%, close to the PPV we have demonstrated in this
study (Figure 2B). Under these circumstances, we would
predict that the NPV in our study should be higher, at
98%, but caution that this assumes that the point esti-
mates for sensitivity and specificity reported by Alexander
et al, are accurate, an assumption that awaits independent
confirmation and which is not entirely supported by our
study. As a consequence of these uncertainties, our pa-
tients with benign GEC results who have not yet under-
gone surgery remain under close surveillance, pending the
publication of additional confirmatory studies.

In summary, our experience of the Afirma GEC has

Table 2. Estimates for sensitivity, specificity, negative predictive value (NPV) and positive predictive value (PPV) of
Afirma GEC. For the purpose of this analysis we have included all of the GEC-benign nodules (n = 16), but only
those GEC-suspicious nodules that have undergone surgery (n = 32). One of 16 GEC-benign nodules has proven to
be malignant at surgery; this analysis is based on the presumption that this was the only false negative result in this
group.

<table>
<thead>
<tr>
<th>AFIRMA GEC RESULTS</th>
<th>MALIGNANT</th>
<th>BENIGN</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUSPICIOUS</td>
<td>5</td>
<td>27</td>
<td>32</td>
</tr>
<tr>
<td>BENIGN</td>
<td>1</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>TOTAL</td>
<td>6</td>
<td>42</td>
<td>48</td>
</tr>
</tbody>
</table>

Sensitivity = 83%; specificity = 36%; prevalence = 12.5%

Negative predictive value of benign GEC = 94% (assumes no additional false negatives)
Positive predictive value of suspicious GEC = 15.6%.
demonstrated a lower than expected rate of benign Afirma GEC reports in AUS/FLUS and FN/HCN, increasing the number of tests needed to avoid one surgery from two to four and raising questions about the costs of widespread application of this assay. In addition, we found the PPV of a “suspicious” classifier result to be lower than previously reported (16% vs 38%), so that more than 80% of GEC-suspicious nodules proved to be benign at surgery. This disappointing result, however, is consistent with the performance of the classifier as reported by Alexander et al, when applied to a group of patients at low risk for malignancy, and is a reminder that the performance of this or any other molecular test, depends critically on the input cytopathology. Unless pretest probability of malignancy is known, claims of reproducible NPV and PPV are spurious and should be treated with caution. Additional confirmatory studies are necessary to assess the performance characteristics of the Afirma GEC, before widespread adoption of this technology can be recommended.

Acknowledgments

The authors are grateful to Ms. Desirae Howe-Clayton and Mr. Russell Ward for their valuable contributions to sample collection and handling, and management of the clinical database. We are also grateful to Ms. Heidi Gautier, of Veracyte Inc, who developed the “Enabled Model” and assisted in writing the Standard Operating Procedures for use of the Afirma GEC.

Address all correspondence and requests for reprints to: Bryan McIver MBPhD, Chair, Department of Head, Neck and Endocrine Oncology, Leader, Endocrine Tumor Program, Moffitt Cancer Center, 12902 Magnolia Drive USF, Tampa, FL 33612, E-mail: Bryan.McIver@moffitt.org.

This work was supported by .

Conflicts or Dualities of Interest: BM has received speaker honoraria and scientific consulting fees from Veracyte Inc. and Asuragen Inc. No other author has any relevant financial conflict of interest.

References


Legend to Figure 1 Continued. . .

38%, specificity of 83% and 12% pretest probability of malignancy yields the observed PPV of 16% and an NPV of 94%.


