Received: 2 April 2021 | Revised: 21 April 2021 | Accepted: 29 April 2021

DOI: 10.1002/dc.24765

ORIGINAL ARTICLE

WILEY

Performance of Afirma genomic sequencing classifier vs gene expression classifier in Bethesda category III thyroid nodules: An institutional experience

Lin Zhang MD PhD | Brian Smola BS | Madelyn Lew MD | Judy Pang MD | Richard Cantley MD | Liron Pantanowitz MD | Amer Heider MD | Xin Jing MD

Department of Pathology and Clinical Laboratories, University of Michigan, Ann Arbor, Michigan

Correspondence
Xin Jing MD Faculty Suite Rm. 36-1221-38, 2800 Plymouth Rd, Building 35, Department of Pathology and Clinical Laboratories, University of Michigan, Ann Arbor, MI 48109.
Email: xinjing@med.umich.edu

Abstract

Background: Afirma gene expression classifier (GEC) is an adjunct to thyroid fine needle aspiration shown to improve pre-operative risk assessment and reduce unnecessary surgery of indeterminate thyroid nodules. Genomic sequencing classifier (GSC) is a newer version aiming to improve specificity and positive predictive value (PPV) of Afirma testing. There are limited studies comparing GSC vs GEC. This study was undertaken to compare these classifiers in terms of diagnostic performance and effect on clinical management of indeterminate thyroid nodules.

Methods: The study cohort consisted of patients with thyroid nodules that had a recurrent cytologic diagnosis of atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS) and were tested by either GEC or GSC. Patient demographics, nodule size, and clinical follow-up were recorded. Benign call rate (BCR) of Afirma testing, rate of subsequent surgery (RSS), rate of histology-confirmed malignancy (RHM), as well as diagnostic sensitivity, specificity, PPV, negative predicative value (NPV), and accuracy were calculated and compared between GSC and GEC cohorts.

Results: Among 264 AUS/FLUS thyroid nodules, 127 and 137 were tested with GEC and GSC, respectively. Compared to GEC, GSC demonstrated increased BCR (77.3% vs 52%), decreased RSS (31.4% vs 51.2%), greater RHM (29% vs 9.8%) associated with a suspicious Afirma result, as well as improved specificity (82.8% vs 54.5%), PPV (29% vs 9.8%), and diagnostic accuracy (83.9% vs 56.7%), while maintaining high sensitivity and NPV.

Conclusion: Afirma GSC substantially improved BCR, RSS, RHM, and diagnostic performance, enhancing appropriate triage and thereby helped avoid unnecessary surgery in AUS/FLUS thyroid nodules.

KEYWORDS

Afirma testing, atypia of undetermined significance/follicular lesion of undetermined significance, gene expression classifier, genomic sequencing classifier, indeterminate thyroid nodules

1 | INTRODUCTION

Fine needle aspiration (FNA) is considered the principle diagnostic test for evaluation of thyroid nodules, aiming to separate benign, non-neoplastic thyroid nodules which may be managed conservatively from neoplastic or malignant nodules in which surgical intervention is necessary. The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) has standardized reporting of thyroid FNAs into six diagnostic categories, including (I) non-diagnostic, (II) benign, (III) atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS), (IV) follicular neoplasm/suspicious for follicular neoplasm (FN/SFN), (V) suspicious for malignancy (SFM), and (VI) malignant. Each of these categories has an implied risk of malignancy linked to recommended clinical management.[1] Although wide implementation of TBSRTC has significantly reduced unnecessary surgical interventions, challenges still remain. In this regard, 15.8 to 38.2% of thyroid FNAs assessed in different institutions fall into the indeterminate categories of TBSRTC, namely AUS/FLUS, FN/SFN, and SFM.[2] Our previous data showed that subsequent hemithyroidectomy or total thyroidectomy specimens revealed non-neoplasia (ie, benign nodular hyperplasia or lymphocytic thyroiditis) in nearly 50% of patients with indeterminate thyroid nodules categorized as AUS/FLUS.[3,4]

Today, molecular testing has become an important adjunct to FNA in further stratification and management of indeterminate thyroid nodules, in terms of improving the diagnostic accuracy and malignant risk assessment, as well as reducing unnecessary surgeries.[5] Among several commercially available tests, Afirma gene expression classifier (GEC) became available in 2012, aiming to improve preoperative risk assessment of indeterminate thyroid nodules. It tested mRNA expression of 167 genes and further stratified indeterminate thyroid nodules into either benign or suspicious categories. Nodules with a benign GEC result carry a similar malignant risk as implied in benign nodules diagnosed by cytology (BSRTC category II) and could accordingly be managed by clinical and/or ultrasound imaging follow-up; whereas those with a suspicious GEC result may require surgical intervention. As a rule-out testing method, GEC was designed to identify benign nodules with a sensitivity of 92% and a negative predictive value (NPV) of 85 to 95%. This test, however, was associated with a specificity of only 52% and a positive predictive value (PPV) of 37%. In addition, Hürthle cell predominant lesions were often classified as suspicious by GEC but were proven to be benign on subsequent resections.[6] An updated version of Afirma testing, called the Gene Sequencing Classifier (GSC) was introduced in 2017 to address the aforementioned shortcomings of GEC. GSC uses next-generation sequencing, incorporating an ensemble model composed of 12 independent classifiers (10 196 genes with 1115 core genes) and seven other components (parathyroid, medullary thyroid carcinoma, BRAFV600E, RET/PTC and RET/PTC3 detection modules, Hürthle cell index and Hürthle neoplasm index). GSC has demonstrated improved specificity (68%) and PPV (47%) while maintaining high sensitivity and NPV in the same blinded, multicenter cohort used to validate GEC. GSC also showed improved specificity with Hürthle cell predominant lesions.[7] However, there have been limited studies comparing the diagnostic performance of GSC with that of GEC in real world practice settings.[8,9]

In our institution, FNA samples were collected for Afirma GEC (from January 2013 to July 2017) or GSC (since July 2017) testing for thyroid nodules with a recurrent AUS/FLUS diagnosis. The current study was undertaken to report our own institutional experience of utilizing GSC vs GEC as an adjunct to FNA in the assessment and management of thyroid nodules categorized as AUS/FLUS (TBSRTC category III).

2 | MATERIAL AND METHODS

This retrospective study was approved by the institutional review board (IRB) at the University of Michigan in Ann Arbor, Michigan. The study cohorts included consecutive thyroid nodules that underwent FNA and had a recurrent cytologic diagnosis of AUS/FLUS (TBSRTC category III) and were tested by either GEC (January 2013-July 2017) or GSC (July 2017-June 2020). All the nodules were followed by either surgical intervention or at least 6 months of clinical and/or ultrasound follow-up monitoring. Nodules with an Afirma result of "non-diagnostic" (due to inadequate sample) and nodules lacking both surgical follow-up and appropriate clinical and/or ultrasound monitoring were excluded from the study.

Ultrasound-guided thyroid FNAs were performed by radiologists and/or endocrinologists with cytology-assisted rapid on-site adequacy assessment. Two dedicated passes were simultaneously collected into the Afirma-provided fixative vial for thyroid nodules that had a previous diagnosis of AUS/FLUS. FNA specimens were then assessed by subspecialty board certified cytopathologists and diagnoses were reported using TBSRTC system. When a recurrent diagnosis of AUS/FLUS was rendered, the aforementioned pre-collected samples were sent to Veracyte's CLIA laboratory (South San Francisco, CA) for Afirma GEC or GSC testing.

The following information from individual patients were collected and recorded: age, gender, size of thyroid nodule, subsequent surgical interventions if available, and the corresponding histologic diagnosis (if surgically treated), as well as stability (lack of change in nodule size and characteristics) of non-surgically removed nodules during the period of at least 6 months of clinical and/or ultrasound monitoring. The nodules with stable and benign characteristics were considered benign (clinical benign diagnosis).

Benign call rate (BCR) of Afirma testing, rate of subsequent surgery (RSS), rate of histology-confirmed malignancy (RHM), as well as the diagnostic parameters including sensitivity, specificity, PPV, NPV, and diagnostic accuracy, were calculated for each cohort as follows:

BCR = number of nodules with benign Afirma result/total number of nodules with Afirma testing.

RSS = number of nodules surgically resected/total number of nodules with Afirma testing.

RHM = number of histology-confirmed malignant nodules/number of nodules with suspicious Afirma result.

Sensitivity = number of nodules with Afirma suspicious result and histology-proven malignancy (True positive)/number of all histology-proven malignant nodules (True positive + False negative).

Specificity = Number of nodules with Afirma benign result and a subsequent surgical and/or clinical benign diagnosis (True negative)/numbers of all benign nodules (True negative + False positive).

PPV = True positive/all nodules with Afirma suspicious result (True positive + False positive).

NPV = True negative/all nodules with Afirma benign result (True negative + false negative).

Diagnostic accuracy = (True positive + True negative)/total number of nodules.

The above parameters were compared between GEC and GSC cohorts using Social Science Statistics (https://www.socscistatistics.com/tests/). Pearson's chi-square or Fisher exact test for categorical variables and student T-test for continuous variables were performed. Statistical significance was defined as a two-tailed *P*-value of <.05 for all analysis.

3 | RESULTS

3.1 | Study cohorts

A total of 264 thyroid nodules fulfilled the aforementioned inclusion criteria of the study. Among which, 127 and 137 nodules were tested by GEC and GSC, respectively. Table 1 summarizes the patients demographic and size distribution of the thyroid nodules in the GEC cohort and GSC cohort. The patients in the GEC cohort (n = 120) and GSC cohort (n = 125) had a similar age range and female predominance (73.3% vs 73.6%). In both GEC and GSC cohorts, nodules measuring from >2 cm to 4 cm represented the majority (47.2% vs 52.6%) of cases followed by nodules measuring from >1 cm to 2 cm (37.8% vs 40.9%). The rest of the nodules were either >4 cm (10.2% vs 4.4%) or 1 cm (4.7% vs 2.2%). No significant difference was seen in these parameters between the two cohorts.

TABLE 1 Characteristics of patients and Bethesda category III thyroid nodules

	Afirma GEC	Afirma GSC
Number of patients	120	125
Age range (median)	18-90 (55)	19-83 (47)
Sex (F/M)	88 (73.3%)/ 32 (26.7%)	92 (73.6%)/ 33 (26.4%)
Number of nodules	127	137
Nodule size		
1 cm	6 (4.7%)	3 (2.2%)
1–2 cm	48 (37.8%)	56 (40.9%)
2–4 cm	60 (47.2%)	72 (52.5%)
>4 cm	13 (10.2%)	6 (4.4%)

3.2 | Benign call rate

As can be seen in Figure 1, GEC interpreted 66 out of 127 thyroid nodules as benign with a BCR of 52.0% and the remaining 61 nodules (48.0%) were reported as suspicious. Among the 137 thyroid nodules tested with GSC, 106 had a benign result with a BCR of 77.4% vs the remaining 31 (22.6%) that were suspicious. The BCR was significantly greater in the GSC cohort than that of GEC cohort (Chi square = 18.7, $P < .01$).

3.3 | Rate of subsequent surgery

Among the 127 thyroid nodules tested with GEC, 14 benign and 51 suspicious were surgically removed with a RSS of 51.2%. Among the 137 thyroid nodules tested with GSC, 14 benign and 29 suspicious were resected with a RSS of 31.4%, which was significantly lower than that of the GEC cohort (Chi square = 10.68, $P < .01$). At least 6 months of clinical and/or ultrasound monitoring was performed for the remaining 62 (48.8%) and 94 (68.6%) nodules in the GEC and GSC cohorts, respectively (Figure 1).

3.4 | Follow up of thyroid nodules with a benign Afirma test result

Among the thyroid nodules with a benign Afirma result in both the GEC cohort (n = 66) and GSC cohort (n = 106), no malignancy was

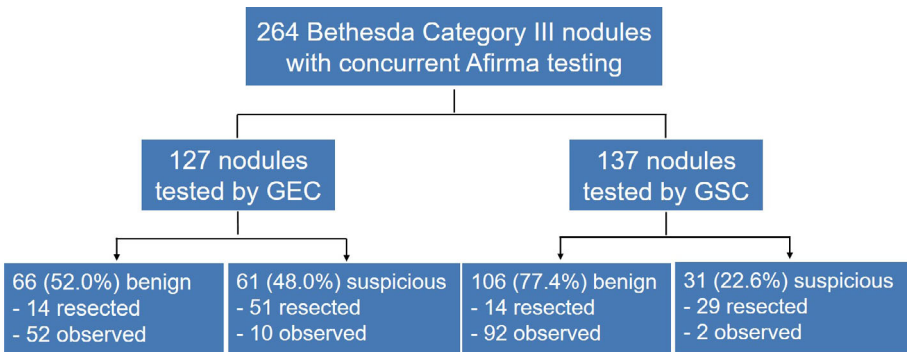

FIGURE 1 Study flowchat [Color figure can be viewed at wileyonlinelibrary.com]

identified in surgically removed nodules in either GEC (n = 14, 21.2%) or GSC (n = 14, 13.2%) cohorts. Clinical and/or ultrasound monitoring during the subsequent period of at least 6 months revealed stable and benign characteristics in the remaining 52 (78.8%) and 92 (86.8%) nodules included in the GEC cohort and GSC cohort, respectively (Table 2).

TABLE 2 Surgical/clinical follow-up of nodules with benign Afirma result

	GEC (66)	GSC (106)
Surgical resection	14 (21.2%)	14 (13.2%)
Non-neoplastic		
Nodular hyperplasia	12	9
Lymphocytic thyroiditis	0	2
Neoplastic		
Follicular adenoma	2	3
Malignant	0	0
Clinical stable/benign	52 (78.8%)	92 (86.8%)

TABLE 3 Surgical/clinical follow-up of nodules with suscipious Afirma result

	GEC (61)	GSC (31)
Surgical resection	51 (83.6%)	29 (93.5%)
Non-neoplastic		
Nodular hyperplasia	25	13
Lymphocytic thyroiditis	1	0
Neoplastic		
Follicular adenoma	18	6
NFITP	1	1
Malignant		
PTC, Classic	2	5
PTC, Follicular	3	1
PTC, Tall cell	1	0
Follicular carcinoma	0	1
Hurthle cell carcinoma	0	1
Medullary carcinoma	0	1
Clinical stable/benign	10 (16.4%)	2 (6.5%)

3.5 | Follow up of thyroid nodules with a suspicious Afirma test result

Among the thyroid nodules with a suspicious Afirma result in the GEC cohort (n = 61), 51 were surgically removed. Of which, six showed malignancy with a RHM of 9.8%, including two classic, three follicular variants and one tall cell variant of papillary thyroid carcinoma. No malignancy was identified upon histologic examination of the remaining 45 resected nodules. With regard to the suspicious nodules in the GSC cohort (n = 31), 29 were surgically removed and nine proved to be malignant. RHM reached 29.0%, which was significantly greater than that of GEC (Chi square = 5.55 and $P < .05$). The nine malignant nodules included six papillary thyroid carcinomas (five classic and one follicular variant), one follicular carcinoma, one Hürthle cell carcinoma, and one medullary thyroid carcinoma. No malignancy was identified upon histologic examination of the remaining 20 resected nodules. Furthermore, clinical and/or ultrasound monitoring during the period of at least 6 months revealed stable and benign characteristics in the remaining 10 (16.4%) and 2 (6.5%) nodules included in the GEC cohort and GSC cohort, respectively (Table 3).

3.6 | Afirma testing in thyroid nodules with predominant Hürthle cells

Among the 12 nodules with predominant Hürthle cells in the GEC cohort, four (33.3%) were reported as benign and 8 (66.7%) were reported as suspicious. However, no malignancy was identified in any of these nodules with either surgical (n = 8) or at least 6 months of clinical/ultrasound monitoring (n = 4). The GSC cohort contained 19 nodules and among these 17 (89.5%) were reported as benign vs 2 (10.5%) as suspicious. Histology-confirmed malignancy was evident in one suspicious nodule while the remaining 18 nodules were considered stable/benign with either surgical follow up (n = 4) or at least 6 months of clinical/ultrasound monitoring (n = 14) (Table 4).

3.7 | Diagnostic performance of GEC vs GSC

Both testing methods offered 100% sensitivity and 100% NPV. A greater specificity, PPV and diagnostic accuracy were achieved with GSC compared to GEC. The differences in PPV and diagnostic accuracy were significant with and without including clinical stable/benign

TABLE 4 Surgical/clinical follow-up of nodules with predominant Hürthle cells

	GEC	GSC		
	Benign (4)	Suspicious (8)	Benign (17)	Suspicious (2)
Surgical resection				
Hürthle cell nodule	2	1	2	1
Hürthle cell adenoma		5	1	
Hürthle cell carcinoma				1
Clinical stable/Benign	2	2	14	

TABLE 5 Comparison of the diagnostic performance of Afirma GEC vs GSC

		GEC	GSC	Statistics
All cases	Sensitivity	100%	100%	
	Specificity	54.5%	82.8%	Chi square = 16.8, $P < .01$
	PPV	9.8%	29.0%	Chi square = 5.5, $P < .05$
	NPV	100%	100%	
	Accuracy	56.7%	83.9%	Chi square = 23.7, $P < .01$
Resected cases	Sensitivity	100%	100%	
	Specificity	23.7%	41.2%	Chi square = 3.1, $P = .08$
	PPV	11.8%	31.0%	Chi square = 4.5, $P < .05$
	NPV	100%	100%	
	Accuracy	30.8%	53.5%	Chi square = 5.6, $P < .05$

nodules. The difference in specificity was significant when including clinical stable/benign nodules (Table 5).

4 | DISCUSSION

The data from the current study indicates that compared to GEC testing, GSC has an increased BCR (77.3% vs 52.0%), decreased RSS (31.4% vs 51.2%), and a greater RHM (29.0% vs 9.8%) associated with a suspicious Afirma test result. While maintaining diagnostic sensitivity of 100% and NPV of 100%, GSC has improved specificity (all cases - 82.8% vs 54.5%; resected cases - 41.2% vs 23.7%) and PPV (all cases - 29.0% vs 9.8%; resected cases - 31.0% vs 11.8%), as well as diagnostic accuracy (all cases - 83.9% vs 56.7%; resected cases - 53.5% vs 30.8%). A similar trend was identified in the recently published meta-analysis of seven studies and a single institutional study that consisted of thyroid nodules categorized as AUS/FLUS (TBSRTC category III) and FN/SFN (TBSRTC category IV).[8,9]

Age distribution and female to male ratio of the patient cohorts, as well as size range and proportion of the targeted thyroid nodules in the current study are similar to that of the previously published multicenter validation study for GSC, in which all nodules had surgical follow up.[7] Like the prior validation study, the current study demonstrated excellent diagnostic sensitivity and NPV for both GSC and GEC. However, the specificity (41.2%) and PPV (31.0%) of GSC in surgically treated AUS/FLUS nodules of the current study differed from what have been previously demonstrated in AUS/FLUS nodules of the validation study in which specificity of 60 to 80% and PPV of 37 to 65% were achieved.

There are a few studies comparing the performance of GSC vs GEC in real world practice settings, all of which included thyroid nodules categorized as TBSRTC category III (AUS/FLUS) and/or category IV (FN/SFN).[8,9] When focusing on the AUS/FLUS category, an improved BCR following utilization of GSC was evident in these studies, ranging from 61.0 to 80.6%..[9-13] GSC-generated BCR of 77% demonstrated by the current study falls into the previously reported range. It is unclear whether the altered BCRs may be partially related to potential changes in the percentage of AUS/FLUS diagnoses among the different time frames. In our institution, AUS/FLUS represented 17.3% of all thyroid FNA interpretations during the period of 54 months when GEC was utilized. The percentage of AUS/FLUS diagnoses among all FNA interpretations reached 24.4% within the 34 months of using GSC. Regardless, no malignancy was identified after at least 6 months of clinical /ultrasound monitoring or surgical follow up in any of the thyroid nodules that had a GSC or GEC "benign" testing result in the current study. One of the published studies assessed individual rates of surgical treatment separately for AUS/FLUS and FN/SFN. Accordingly, a reduced surgical rate for AUS/FLUS nodules following GSC (14.9%) vs GEC (51.3%) were documented in this prior study.[11] The current study also demonstrated a significant reduction in surgical rate following implementation of GSC, from 51.2 to 31.4%.

Similar to previous studies, the current study demonstrated that while maintaining a high level of sensitivity and NPV, GSC improved specificity and PPV. When including only AUS/FLUS nodules with surgical follow up, PPV of GSC in our study (31.0%) appears to be lower than that of two other single institutional studies (57 and 52%).[9,11] The current study also demonstrated a similar PPV of GSC (29%) while including all AUS/FLUS cases with surgical follow up or clinical/ultrasound monitoring (assuming all were benign), which was lower than what has been previously reported (57 and 80%).[11,13] Unlike the current study which did not classify Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features (NIFTP) as malignant, the two aforementioned studies counted NIFTP as "malignant" due to the current recommendations for hemithyroidectomy.[9,11] These facts may contribute to the variation in PPV of GSC among different institutions.

Overinterpretation of benign Hürthle cell predominant nodules followed by unnecessary surgeries was a major concern for GEC.[14] The validated study and studies from real world practice demonstrated that GSC prompted a higher BCR in Hürthle cell predominant nodules, ranging from 60% to 80%, thereby potentially avoiding unnecessary surgical interventions.[7,11-13,15] In spite of a limit number of AUS/FLUS nodules with predominant Hürthle cells, the current study revealed a BCR of 90% for GSC vs 34.4% for GEC. Further, all GSC-identified "benign" nodules were confirmed to be benign or stable upon surgical follow up or clinical/ultrasound monitoring.

In our institution, Afirma testing is routinely applied to those thyroid nodules which are categorized into TBSRTC category III (AUS/FLUS) a second time. Thus, the current study focuses on our institutional experience in AUS/FLUS nodules only while intuitional data on FN/SFN category is absent. Similar to previous studies from other institutions, limitations of the current study were small case cohort and short follow up periods for the GSC cohort. It is noteworthy to mention that many factors may attribute to the findings among different institutions, which include but are not limited to variations in diagnostic thresholds for AUS/FLUS, interobserver variability among pathologists, and selection bias in surgical cases.

In summary, compared to GEC, GSC substantially improved BCR, RSS and diagnostic performance for AUS/FLUS thyroid nodules, enhancing appropriate triage and treatment, and avoiding unnecessary surgeries in AUS/FLUS thyroid nodules.

CONFLICT OF INTEREST

All authors have no conflict of interest.

AUTHOR CONTRIBUTION

Lin Zhang: acquired/analyzed/interpreted data and resource materials, drafted the manuscript and contributed significant revisions on subsequent drafts, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Brian Smola: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Madelyn Lew: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Judy Pang: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Richard Cantley: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Liron Pantanowitz: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Amer Heider: contributed revisions for intellectual content, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

Xin Jing: designed concept of article, acquired/analyzed/interpreted data and resource materials, drafted the manuscript and contributed significant revisions on subsequent drafts, contributed to final approval of version to be published, and agrees to be accountable for all aspects of the work in ensuring questions related to accuracy or integrity of any part of the work are appropriate investigated and resolved.

DATA AVAILABILITY STATEMENT

Data available on request from the authors

ORCID

Lin Zhang https://orcid.org/0000-0003-4831-6736
Brian Smola https://orcid.org/0000-0003-2049-1970
Madelyn Lew https://orcid.org/0000-0002-6279-0330
Judy Pang https://orcid.org/0000-0003-0589-8262
Richard Cantley https://orcid.org/0000-0002-6564-3889
Liron Pantanowitz https://orcid.org/0000-0001-8182-5503
Amer Heider https://orcid.org/0000-0001-6522-2454
Xin Jing https://orcid.org/0000-0002-1768-3472

REFERENCES

1. Cibas ES, Ali SZ. The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid*. 2017;27(11):1341-1346.
2. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda system for reporting thyroid cytopathology: a meta-analysis. *Acta Cytol*. 2012;56(4):333-339.
3. Lee B, Smola B, Roh MH, Hughes DT, Miller BS, Jing X. The impact of using the Bethesda system for reporting thyroid cytology diagnostic criteria on the follicular lesion of undetermined significance category. *J Am Soc Cytopathol*. 2014;3(3):131-136.
4. Kantola S, Virani N, Haus C, Hipp J, Zhao L, Jing X. Prospective evaluation of impact of using the Bethesda system for reporting thyroid cytopathology: an institutional experience. *J Am Soc Cytopathol*. 2015;4(1):25-29.
5. Nishino M. Molecular cytopathology for thyroid nodules: a review of methodology and test performance. *Cancer Cytopathol*. 2016;124(1):14-27.
6. 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.
7. Patel KN, Angell TE, Babiarz J, et al. Performance of a genomic sequencing classifier for the preoperative diagnosis of Cytologically indeterminate thyroid nodules. *JAMA Surg*. 2018;153(9):817-824.
8. Vuong HG, Nguyen TPX, Hassell LA, Jung CK. Diagnostic performances of the Afirma gene sequencing classifier in comparison with the gene expression classifier: a meta-analysis. *Cancer Cytopathol*. 2021;129(3):182-189.
9. Geng Y, Aguilar-Jakthong JS, Moatamed NA. Comparison of Afirma gene expression classifier with gene sequencing classifier in indeterminate thyroid nodules: a single-institutional experience. *Cytopathology*. 2021;32(2):187-191.

10. Angell TE, Heller HT, Cibas ES, et al. Independent comparison of the Afirma genomic sequencing classifier and gene expression classifier for Cytologically indeterminate thyroid nodules. *Thyroid*. 2019;29(5):650-656.
11. Endo M, Nabhan F, Porter K, et al. Afirma gene sequencing classifier compared with gene expression classifier in indeterminate thyroid nodules. *Thyroid*. 2019;29(8):1115-1124.
12. Wei S, Veloski C, Sharda P, Ehya H. Performance of the Afirma genomic sequencing classifier versus gene expression classifier: an institutional experience. *Cancer Cytopathol*. 2019;127(11):720-724.
13. San Martin VT, Lawrence L, Bena J, et al. Real-world comparison of Afirma GEC and GSC for the assessment of Cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2020;105(3):e428-e435.
14. Brauner E, Holmes BJ, Krane JF, et al. Performance of the Afirma gene expression classifier in Hurthle cell thyroid nodules differs from other indeterminate thyroid nodules. *Thyroid*. 2015;25(7):789-796.
15. 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.

How to cite this article: Zhang L, Smola B, Lew M, et al. Performance of Afirma genomic sequencing classifier vs gene expression classifier in Bethesda category III thyroid nodules: An institutional experience. *Diagnostic Cytopathology*. 2021; 49:921–927. https://doi.org/10.1002/dc.24765