European experience with the Afirma Gene Expression Classifier for indeterminate thyroid nodules

Lončar, Ivona; van Velsen, Evert F.S.; Massolt, Elske T.; van Kemenade, Folkert J.; van Engen-van Grunsven, Adriana C.H.; van Hemel, Bettien M.; van Nederveen, Francien H.; Netea-Maier, Romana; Links, Thera P.; Peeters, Robin P.

Published in:
Head and Neck

DOI:
10.1002/hed.27472

Publication date:
2023

Received: 19 December 2022 | Revised: 30 April 2023 | Accepted: 11 July 2023

DOI: 10.1002/hed.27472

ORIGINAL ARTICLE

WILEY

European experience with the Afirma Gene Expression Classifier for indeterminate thyroid nodules: A clinical utility study in the Netherlands

Ivona Lončar PhD[1] | **Evert F. S. van Velsen MD, PhD**[2] | **Elske T. Massolt MD, PhD**[3] | **Folkert J. van Kemenade MD, PhD**[4] | **Adriana C. H. van Engen-van Grunsven MD, PhD**[5] | **Bettien M. van Hemel MD, PhD**[6] | **Francien H. van Nederveen MD, PhD**[7] | **Romana Netea-Maier MD, PhD**[8] | **Thera P. Links MD, PhD**[9] | **Robin P. Peeters MD, PhD**[2] | **Tessa M. van Ginhoven MD, PhD**[1]

[1]Department of Surgical Oncology and Gastrointestinal Surgery, Erasmus MC Cancer Institute, Rotterdam, the Netherlands

[2]Department of Internal Medicine, Erasmus MC Academic Center for Thyroid Diseases, Rotterdam, the Netherlands

[3]Department of Internal Medicine, Albert Schweitzer Hospital, Dordrecht, the Netherlands

[4]Department of Pathology, Erasmus MC, Rotterdam, the Netherlands

[5]Department of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands

[6]Department of Pathology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

[7]Laboratory for Pathology, Dordrecht, the Netherlands

[8]Division of Endocrinology, Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands

[9]Department of Endocrinology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

Correspondence
Ivona Lončar, Department of Surgical Oncology and Gastrointestinal Surgery, Erasmus MC Cancer Institute, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands.
Email: i.loncar@erasmusmc.nl

Funding information
Veracyte

Abstract

Background: The Gene Expression Classifier (GEC) and Genomic Sequencing Classifier (GSC) were developed to improve risk stratification of indeterminate nodules. Our aim was to assess the clinical utility in a European population with restrictive diagnostic workup.

Methods: Clinical utility of the GEC was assessed in a prospective multicenter cohort of 68 indeterminate nodules. Diagnostic surgical rates for Bethesda III and IV nodules were compared to a historical cohort of 171 indeterminate nodules. Samples were post hoc tested with the GSC.

Results: The GEC classified 26% as benign. Surgical rates between the prospective and historical cohort did not differ (72.1% vs. 76.6%). The GSC classified 59% as benign, but misclassified six malignant lesions as benign.

Conclusion: Implementation of GEC in management of indeterminate nodules in a European country with restrictive diagnostic workup is currently not

supported, especially in oncocytic nodules. Prospective studies with the GSC in European countries are needed to determine the clinical utility.

KEYWORDS

fine-needle aspiration, gene expression classifier, genomic sequencing classifier, indeterminate thyroid nodule

1 | INTRODUCTION

Thyroid nodules are identified in approximately 5% of the adult population by clinical examination and even up to 50% on ultrasound or autopsy examination.[1,2] However, only 7%–15% of thyroid nodules harbors malignancy.[3] Current workup of thyroid nodules consists of a neck ultrasound with, if indicated, cytological assessment after ultrasound-guided fine-needle aspiration (FNA).[4] After FNA approximately 15%–30% of the nodules are reported as cytologically indeterminate, including the Bethesda III (B3, Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance), Bethesda IV (B4, Follicular Neoplasm or Suspicious for a Follicular Neoplasm), and Bethesda V (B5, Suspicious for malignancy) categories.[5,6] Following diagnostic surgical resection, only one in three indeterminate nodules proves to be malignant, suggesting a large potential to avoid diagnostic surgeries.[7] To reduce the number of diagnostic surgeries several molecular tests have been developed in the past decade, using different approaches including microRNA classifiers (Rosetta GX Reveal and ThyraMIR) and next-generation sequencing-based assays (Thyroseq v2 and v3).[8–11] In 2011 Veracyte's Afirma Gene Expression Classifier (GEC) became commercially available for patients with Bethesda III or IV thyroid nodules. This assay uses messenger RNA expression patterns of 167 genes to reclassify Bethesda III and IV nodules as "benign" or "suspicious."[12] A GEC benign result in patients with Bethesda III or IV nodules leads to a malignancy risk of <5%, which is similar to the benign Bethesda II category. These data justify a wait-and-see policy based on a benign GEC result,[13,14] whereas in patients with a GEC suspicious result, the estimated malignancy rate of 37%–38% often results in diagnostic surgery.[12] In 2017, Veracyte updated the GEC to the Genomic Sequencing Classifier (GSC) and added the Xpression Atlas (XA) assay in 2019 showing improved specificity and positive predictive value.[15,16] The GSC uses whole transcriptome RNA sequencing and machine learning algorithms to reclassify nodules, combined with the XA which uses RNA sequencing to detect genomic variants and fusions. Most clinical validation and utility studies of the GEC and GSC have been performed in the United States, with only one publication from Europe so far.[17]

Historically, the Netherlands has a conservative approach towards patients with thyroid nodules compared to many other countries, with a high threshold for neck ultrasound leading to relatively low number of incidental findings. Since 2007, the Dutch national thyroid cancer guidelines advocate to only pursue neck ultrasound in patients with a palpable thyroid nodule and to refrain from further diagnostics in incidentally found thyroid nodules by other imaging such as MRI and CT-scans, except for FDG-PET/CT avid thyroid nodules.[18] These restrictive diagnostic workup protocols result in a selected population with less incidentally discovered low-grade papillary thyroid cancer and a lower burden of microcarcinomas compared to other countries.[19] The utility of the GEC in a population with long-standing restrictive workup protocols has not been assessed before and provides insights for future implementations as de-escalating trends in diagnostic workups are upcoming.[20] This study therefore assesses the clinical utility of the GEC in a European country with restrictive diagnostic workup by describing the impact on surgical treatment and concordancy between GEC and GSC results and definitive histopathology.

2 | METHODS

2.1 | Patient selection and study design

This prospective multicenter cohort study included patients with an indeterminate Bethesda III or IV thyroid nodule undergoing Afirma GEC testing within three university hospitals collaborating in the Dutch Thyroid Cancer Group (Erasmus MC, Radboud UMC, UMC Groningen) between May 2016 and January 2018. Institutional Review Board approval was obtained for this study and all patients provided written informed consent (MEC-2015-637). Patients were eligible for inclusion if 18 years or older, had a cytologically proven Bethesda III or IV nodule without prior molecular testing on the index nodule. Exclusion criteria were the presence of Bethesda V or VI nodules, or pathological lymph nodes, indicative

of thyroid malignancy. Patients were excluded if surgery was not applicable or was already likely or planned regardless of the FNA result (i.e., due to cosmetic reasons, rapid nodule growth, mechanical symptoms, patients request, or physician's recommendation). In the Netherlands thyroid nodules are not primarily referred to a university hospital. The participating university hospitals were depended on referrals from peripheral non-university hospitals. Therefore, patients could be enrolled in two ways (Figure 1). Patients could be enrolled if they were already diagnosed with an indeterminate nodule in a non-university hospital (route A), but could also be referred directly to clinically evaluate a thyroid nodule without a cytopathological diagnosis yet (route B). In case an FNA was performed in the referring hospital, revision of this FNA was performed by the pathologists from the university hospitals. All nodules were given a second ultrasound-guided FNA in the participating university hospitals for standard cytopathological diagnosis. Two extra needle passes were collected into the GEC collection tube containing the nucleic acid preservative. GEC samples were stored and shipped according to Veracyte CLIA Laboratory procedures, after the local pathologist (re)confirmed the Bethesda III or IV diagnosis.[12] Applying eligibility criteria after FNA resulted in a cohort of 68 nodules in 66 patients (Figure 1). The decision to perform diagnostic surgery was made during local tumor board meetings, based on all available clinical data including the GEC result. Surgery was performed in the university or referring hospital. Final histopathology was obtained by the local pathologist. In general, thyroid histopathology and specifically thyroid tumors with a follicular growth pattern can pose difficulties for definitive histological diagnosis.[21] Therefore, all postoperative specimens have been blindly reviewed by a panel of four experienced pathologists (Folkert J. van Kemenade, Francien H. van Nederveen, Adriana C. H. van Engen-van Grunsven, and Bettien M. van Hemel) with a high level of agreement. In case of discrepancies a consensus meeting was held. Consensus was reached if three out of four pathologists agreed. Patients with a malignant histopathological outcome were treated according to the current guidelines. If the decision to wait-and-see was made, a follow-up ultrasound was recommended after 6 months and thereafter annually for at least 5 years in the university hospitals to reduce loss of follow-up. If the nodule increased or changed over time, patients were offered the possibility for new FNA and/or surgery. During the inclusion period of this study, the Afirma Gene Sequencing Classifier (GSC) was released. Therefore, the ad hoc decision was made to additionally analyze all previously collected samples with the GSC and the XA version 1. However, GSC-testing was performed retrospectively and no clinical decisions were made based on these results. Data regarding patients' characteristics, GEC and GSC results, and surgical, histopathological, and follow-up findings were collected. To evaluate the surgical rate in patients with a cytologically indeterminate thyroid nodule, without additional GEC testing, a retrospective chart review from January 2012 to December 2015 in the same university hospitals was performed. Same inclusion and exclusion criteria were applied. Patients from the

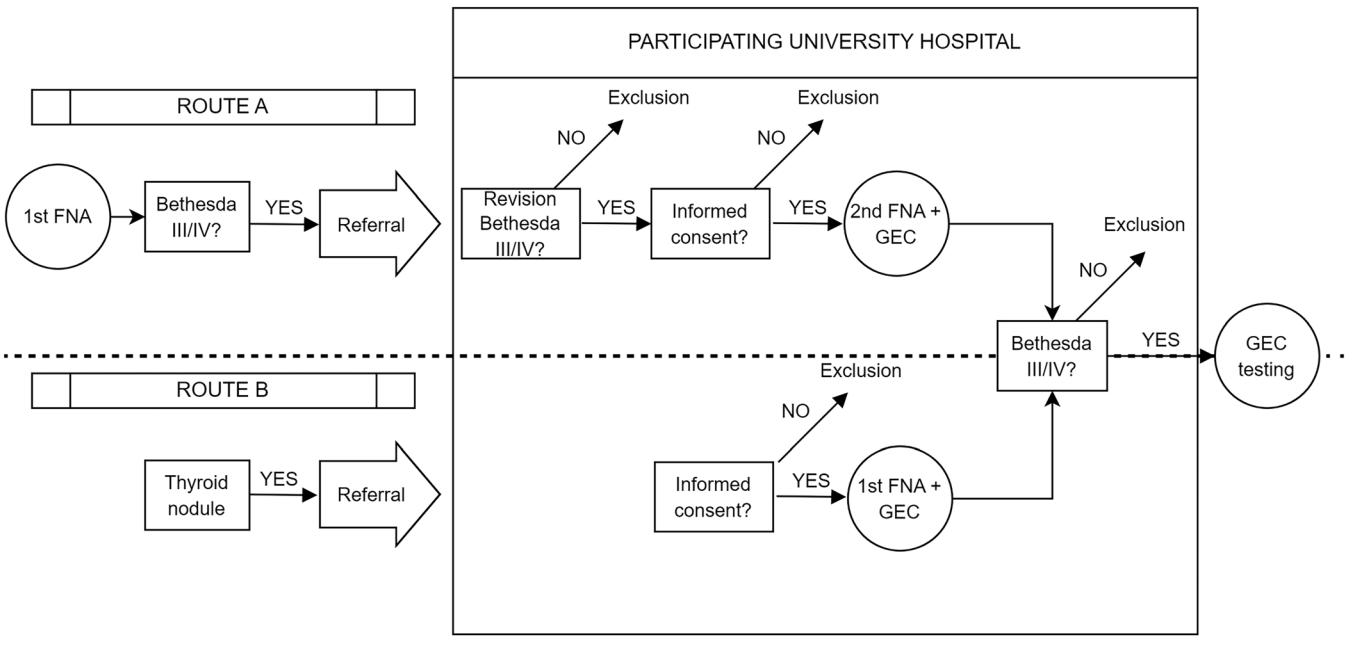


FIGURE 1 Study design

historical cohort were adult patients with Bethesda III or IV cytopathological diagnosis without molecular testing on the index thyroid nodule. The same exclusion criteria were applied.

2.2 | Outcomes

Primary outcome is the surgical rate within the first year after diagnosis of an indeterminate nodule. We compared the historical cohort with the prospective GEC cohort to assess the clinical utility of the GEC on surgical rates. Thereafter, we retrospectively assessed the possible clinical utility of the GSC.

2.3 | Statistical analysis

Data were analyzed using descriptive statistics. Categorical variables are displayed as count (n) and percentage (%). Continuous variables are displayed by mean ± standard deviation (SD). Differences between groups were evaluated with independent Student's t test or Mann–Whitney U test and chi-square or Fisher's exact test, where appropriate. A p-value <0.05 was considered statistically significant and analyses were performed using IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY).

3 | RESULTS

3.1 | Characteristics of the prospective and historical cohort

In the prospective GEC tested cohort, all patients were enrolled by route A (Figure 1). Median time between first and second FNA was 48.5 (range 1–319) days. In total 68 nodules in 66 patients were included, out of 109 patients assessed for eligibility (Figure 2). Of these 29 patients were excluded after reclassification of their FNA. Mean age was 52.5 ± 12.2 years, and 47 (71.2%) were women (Table 1). Cytology diagnosis was Bethesda III in 27 (39.7%) and Bethesda IV in 41 (60.3%) nodules. Oncocytic-dominant cytology was present in 38 (55.9%) nodules. Mean nodule size was 30.3 ± 12.5 mm. In the GEC cohort 49 (72.1%) patients underwent surgery, of which 16 (32.7%) were malignant on histopathology (Table 2).

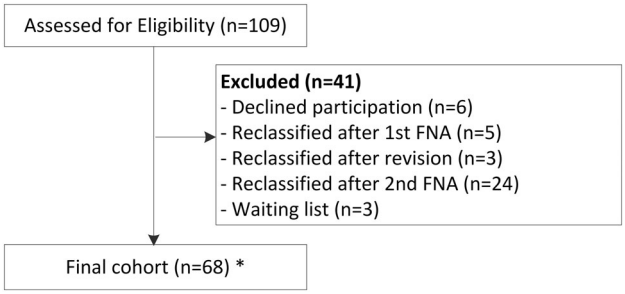


FIGURE 2 Inclusion flowchart (*2 patients had 2 nodules)

A total of 171 patients were included in the historical cohort. Mean age was 54.4 ± 14.8 years, and 117 patients (68.4%) were women. Cytology diagnosis was Bethesda III in 81 (47.4%) and Bethesda IV in 90 (52.6%) nodules. Oncocytic-dominant cytology was present in 52 (30.4%) nodules. Mean nodule size was 29.7 ± 15.8 mm. In this cohort 131 (76.6%) patients underwent surgery, of which 39 (29.8%) had a malignant and 92 (70.2%) a benign lesion. Age, sex, and cytology was similar in both cohorts. However, nodules in the prospective cohort more often showed oncocytic features on cytology (52 [30.4%] vs. 38 [55.9%]; $p < 0.001$). The surgical rate was not statistically different between the two cohorts (72.1% vs. 76.6%; $p = 0.355$). When excluding nodules with oncocytic features on cytology in both cohorts, the surgical rate in the prospective cohort is lower than the historical cohort (50.0% vs. 71.4%; $p = 0.026$).

3.2 | GEC results

A total of 68 nodules underwent GEC analysis (Figure 3). Of these, 18 (26%) were classified as GEC benign, 49 (72%) as GEC suspicious, and 1 (2%) as GEC parathyroid. In one patient, GEC was repeated, because it resulted in a nondiagnostic outcome in the first attempt. Patients with a GEC benign and GEC parathyroid result were offered a wait-and-see policy with ultrasound follow-up for 5 years. However, two (11%) patients with a GEC benign nodule were operated within the first year after surgery, and both were benign at histopathology. One patient had undergone surgery out of personal preference and one patient underwent surgery because of another GEC suspicious nodule in the same thyroid lobe. The GEC classified 18 nodules (26%) as benign and during follow-up ($n = 16$) or after surgery ($n = 2$) no malignancies were found. Forty-nine (72%) nodules were classified as GEC suspicious, of which 2 were scheduled for follow-up and 47 (96%) were operated upon of which 16 (34%) had malignant disease. The GEC thus correctly identified all 16 malignancies as suspicious (Table 3).

3.3 | Follow-up

Two patients were operated more than 1 year after diagnosis. One patient had a GEC suspicious nodule and was

TABLE 1 Characteristics of prospective GEC tested and historical cohort

	Prospective ($n = 68$)	Historical ($n = 171$)	p-value
Sex[a]			0.744
Female	47 (71.2%)	117 (68.4%)	
Male	19 (28.8%)	54 (31.6%)	
Age (year)	52.5 ± 12.2	53.4 ± 14.8	0.567
Bethesda category			0.283
III	27 (39.7%)	81 (47.4%)	
IV	41 (60.3%)	90 (52.6%)	
Oncocytic-dominant cytology	38 (55.9%)	52 (30.4%)	<0.001
Nodule size (mm)	30.3 ± 12.5	29.7 ± 15.8	0.809
Surgical treatment	49 (72.1%)	131 (76.6%)	0.355

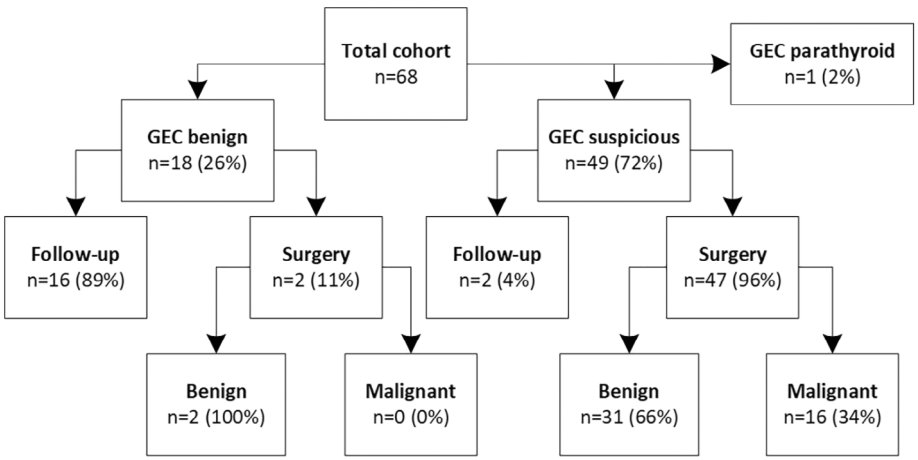
[a]Sixty-eight nodules in 66 patients.

TABLE 2 Histopathology results in operated prospective GEC tested and historical cohort

	Prospective ($n = 49$)	Historical ($n = 131$)	p-value
Malignant	16 (32.7%)	39 (29.8%)	0.647
PTC	5 (10.2%)	26 (19.8%)	0.140
FTC	7 (14.3%)	7 (5.3%)	0.041
OC	4 (8.2%)	5 (3.8%)	0.252
MTC	0 (0.0%)	1 (0.8%)	0.544
Benign	33 (67.3%)	92 (70.2%)	0.647
OA	11 (22.4%)	14 (10.7%)	0.037
FA	12 (24.5%)	55 (42.0%)	0.038
HN	10 (20.4%)	23 (17.6%)	0.854

Abbreviations: FA, follicular adenoma; FTC, follicular thyroid carcinoma; HN, hyperplastic nodule; MTC, medullary thyroid carcinoma; OA, oncocytic adenoma; OC, oncocytic carcinoma; PTC, papillary thyroid carcinoma.

FIGURE 3 Flowchart GEC results prospective cohort



reluctant to undergo surgery because of fear for surgical complications, but eventually chose to undergo surgery in the second year of follow-up (Table S1, Supporting Information; number 59). Another patient with a GEC benign nodule (Table S1; number 12) chose surgery over follow-up in the second year after diagnosis, because of fear of cancer in the context of thyroid cancer diagnosis in a first degree family member. Both nodules revealed a follicular adenoma at histopathology. Median follow-up time of all 17 unoperated patients was 65.0 (62.5–68.0)

TABLE 3 Afirma GEC and GSC results for Bethesda III and IV nodules

	Total	Malignant	Benign
GEC	65		
Suspicious		16	31
Benign		0	18
GSC	64		
Suspicious		10	14
Benign		6	34

Note: Unoperated nodules were considered benign. Unoperated patients with suspicious results ($n = 2$) and patients with parathyroid results ($n = 1$) were excluded from this table. One GSC sample was not sufficient for analysis.

months, no patients were loss to follow-up and no clinical suspicion of malignancy was encountered.

3.4 | GSC and XA results

At a later stage, all 68 samples were additionally analyzed with the GSC (GSC results were not used in clinical decision making due to the timing). Of these, 40 (59%) were GSC benign, 26 (38%) GSC suspicious, 1 (1.5%) GSC parathyroid, and 1 (1.5%) was not sufficient for analysis. The benign call rate (BCR) was 59% and significantly higher than the GEC BCR (59% vs. 26%; p-value <0.001). Unoperated nodules with suspicious results ($n = 2$) and parathyroid results ($n = 1$) were excluded from Table 3, leading to 64 samples with a diagnosis of benign or malignant. The GSC failed to classify 6 out of 16 (37.5%) malignant nodules as suspicious and classified them benign instead. These six malignancies were correctly classified as suspicious with the GEC. On cytopathology, five had oncocytic-dominant morphology. Histopathology after resection showed three minimally invasive oncocytic carcinomas, two minimally invasive follicular thyroid carcinomas, and one widely invasive follicular thyroid carcinoma. In total 9 of 67 samples showed specific genomic variants and 5 fusions (Table S1).

4 | DISCUSSION

This study assessed the clinical utility of the GEC for the first time in a European country with restrictive diagnostic work-up protocols for thyroid nodules. Incorporating the GEC in diagnostic work-up did not influence the overall surgical rate of indeterminate thyroid nodules compared to a historical series from the same centers, as 72% was classified as GEC suspicious warranting further diagnostic interventions. The updated GSC increased the BCR reducing the surgical rate, but would have misclassified six malignant lesions as benign.

4.1 | GEC low BCR

The first GEC clinical validation study achieved 52% specificity.[12] This was confirmed in a multicenter clinical utility study reporting a GEC BCR of 51%, meaning that half of all indeterminate thyroid nodules were reclassified as benign and thereby potentially reducing the number of diagnostic surgeries.[22] Later, a meta-analysis of 19 studies reported a pooled GEC BCR of 45% in 2568 indeterminate thyroid nodules, ranging between 26% and 61%.[23] In the current Dutch study population, the GEC BCR was among the lowest with 26%. As a consequence, the 72.1% surgical rate did not differ from the historical cohort and literature.[7] We are not the first to report that institutional rates of surgery did not change after GEC implementation.[24] The lower BCR in the Dutch population might be explained by several factors. First, the restrictive diagnostic workup protocol for thyroid nodules could have resulted in different patient populations by selection, thereby affecting the ratio of different histological types of thyroid nodules between populations.[18,19,25] This is also reflected by the EUROCARE-5 population-based study, which shows that 20% of the thyroid cancers in the Netherlands is an FTCs.[26] This is higher compared to other European countries and double the number of all FTCs in the United States according to the SEER database (10.8%),[27] which is most likely due to the more restrictive protocols in the Netherlands for thyroid nodule evaluation. Another consequence of the restrictive workup is a cohort with larger nodules than reported in other series, as FNA is only performed in palpable and therefore larger thyroid nodules.[25] This underlines the need for more clinical data from different populations. To date only one European study testing the GEC has been reported.[17] This Italian single center study used the GEC in 25 patients and found a 64% BCR, which is higher than previously reported studies. However, the Italian consensus for thyroid cytology uses subclasses for indeterminate thyroid nodules that are different from the Bethesda categories and makes it difficult to compare results. Also, Italy is one of the countries with the highest burden of thyroid cancer incidence and over diagnosis worldwide,[28] which is reflected by the fact that more than 50% of all thyroid cancers are papillary thyroid microcarcinomas.[29] Compared to Italy, the burden of microcarcinomas in the Netherlands is strikingly lower, as 15.7% of the total amount of thyroid cancers between 2005 and 2015 was a microcarcinoma.[19] Second, multiple studies

have demonstrated that the GEC categorizes nodules with oncocytic-dominant morphology as GEC suspicious more often, which leads to higher surgical rates in these populations.[30–34] In this Dutch cohort 55.9% of the patients had a nodule with oncocytic-dominant cytology, of which almost 90% had a suspicious GEC result and only 31% proved to be malignant after resection. Current findings underline the suggestion by Brauner et al. that oncocytic neoplasms constitute a distinct molecular class of tumors, in which genetic differentiation of benign from malignant is not well described and remains a challenge in molecular profiling.[30] A sensitivity analysis showed that the exclusion of oncocytic-dominant nodules in both cohorts decreased the surgical rate in the GEC cohort to 50.0%, which is significantly lower than the historical surgical rate. Altogether, these results in general increase the concern of the usefulness of the Afirma GEC in oncocytic lesions. The predominance of oncocytic aspirates in the prospective cohort could be explained by the restrictive evaluation of thyroid nodules in the Netherlands, but might also be explained by our study design. Some degree of bias could have occurred as most GEC tests were performed on repeat FNA unlike the first validation study which performed the GEC on first FNA passes.[12] Repeat FNAs could have led to the reclassification and exclusion of more non-oncocytic and Bethesda III lesions and therefore the concentration of oncocytic-dominant cytology and histologically follicular thyroid cancers and oncocytic adenomas (Table 2) in the prospective cohort. Performing repeat FNAs in other than Bethesda III nodules is not common practice in the Netherlands, which explains the difference in the occurrence of oncocytic-dominant cytology in the prospective and retrospective cohort. Time between first and second FNA varied from 1 to 319 days. It could be speculated that recurrent FNA's lead to an inflammatory response in the target nodule influencing the GEC analyses of following biopsies.

4.2 | GSC higher BCR, missing malignancies

In 2017, the GSC was released with higher BCRs, especially in nodules with oncocytic-dominant cytology. The BCR of the GSC was higher than the BCR of the GEC in nodules with oncocytic-dominant cytology (63.2% vs. 22%) resulting in a decreased surgical rate (47.8% vs. 34.7%) in two American studies.[15,35] Our results confirm this increase in BCR from 26% to 59%. If only GSC suspicious nodules would undergo surgery, the surgical rate would have been reduced from 72.1% to 38.2% and primary endpoint of the study would have been met (38.2% vs. 76.6%, p-value <0.001). However, the GSC still would have classified six malignant thyroid nodules as benign in our series. All six cases were either oncocytic carcinomas or follicular thyroid carcinomas on histopathological examination, of which the widely invasive carcinoma was a shocking finding. Although confirmed by a panel of four experienced pathologists, we note that only 3 of 6 (50%) of the GSC false-negative cases had 100% histological concordance before the consensus meeting (Table S1) compared to 33 of 42 (79%) for the other operated cases. This trend may indicate more histological challenges among these false-negative cases. The study by San Martin et al. also reported false-negative results with the GSC (1 oncocytic carcinoma and 2 follicular variants PTCs) among 82 GSC benign cases and recommended to interpret the results with caution, as the higher BCR could imply an increased risk in false-negative results.[35] Still, our GSC findings of a high false-negative rate overall or among oncocytic-dominant histologies are unique among currently published experiences.[36]

4.3 | Clinical perspective

The diagnostic workup of thyroid nodules is a challenge in itself, which is confirmed by the fact that 29 of our 109 eligible patients were excluded because of reclassification of their FNA. Over the last years, other methods have been developed to reduce diagnostic surgeries. Other molecular profiling panels such as the ThyroSeq v3 are designed and being validated with a published BCR of 61% in indeterminate thyroid nodules.[9] However, a recent clinical trial randomizing indeterminate thyroid nodules to evaluation with ThyroSeq v3 or Afirma GSC found a low positive predicitive value of 20% with both test in oncocytic aspirates.[37] These results together with our results highlight the need for further evaluation and improvement of diagnostic performance of molecular tests in oncocytic nodules. Also, other diagnostics modalities are explored to improve the risk stratification of indeterminate thyroid nodules. The EfFECTS trial showed that an FDG-PET/CT-driven diagnostic workup of indeterminate thyroid nodules reduces the number of diagnostic surgeries by 40%.[38] Unfortunately, FDG-PET/CT visual assessment once again did not contribute to reduction of diagnostic surgeries in patients with oncocytic nodules. Another study also showed that FDG-PET/CT could be used to stratify the cancer risk of thyroid nodules with an intermediate ultrasound assessment.[39] Nodules classified as EU-TIRADS 4 without FDG-PET/CT uptake could be refrained from further investigation. Third, the American College of Radiology and the European Thyroid Association have developed Thyroid

Imaging, Reporting and Data System classifications (ACR-TIRADS and EU-TIRADS).[40,41] The TIRADS classification is designed to sonographically assess the need for FNA based of an estimated risk of malignancy. Over the last years this system has been widely implemented in order to improve the sensitivity of the diagnostic work-up of thyroid nodules.

4.4 | Limitations

First, clinical decisions in this study were made using the GEC test only. During the inclusion period of this study, the GSC test was released and samples were post-hoc tested using the GSC. Therefore the GSC results should be interpreted with caution. Second, the low BCR in this study is driven by the high rate of oncocytic aspirates, as a consequence of performing repeat FNA, which is a limitation of this study and data should be interpreted with respect to this limitation. Third, the before after design rather than a randomized study has limitations and cannot correct for (hidden) confounders. Furthermore, historical bias could have played a role, when comparing the prospective and retrospective cohorts. However, a historical cohort design was chosen to obtain enough controls and to compare the surgical rate from the GEC cohort to the surgical rate from a period without GEC. A control cohort from the same time period could have introduced selection bias. This study is limited by verification bias because to date only three of the GEC benign patients were treated surgically. However, patients are monitored closely and to date no patient in follow-up had an indication for thyroid surgery because of nodule growth or malignant FNA cytology. Long-term follow-up studies should confirm the true nature of GEC benign thyroid nodules. That said, the question remains how long the follow-up should be. It is hypothesized that a 2- to 3-year follow-up is sufficient based on two studies that investigated the malignancy detection rate of initially benign thyroid nodules on the long term.[42,43] In comparison, the median follow-up time in the current study is 68 months and we therefore consider the GEC benign thyroid nodules as clinically benign. Lastly, this is a descriptive study of the clinical implementation of a new diagnostic test. This study is not designed and therefore underpowered to calculated specific test parameters.

5 | CONCLUSION

The results of this study currently do not support the implementation of the Afirma GEC in the management of indeterminate thyroid nodules in a European country with restrictive diagnostic workup, especially in nodules with oncocytic morphology. Non-oncocytic aspirates might benefit from molecular testing, but further studies are needed to assess the clinical utility in these nodules. Although post hoc analysis does not support the use of the GSC test due to the false-negative results, further prospective studies are needed to determine the clinical utility of this test.

ACKNOWLEDGMENT

Logistics of this investigator initiated research project was supported by an unrestricted research grant from Veracyte.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

ORCID

Ivona Lončar https://orcid.org/0000-0001-7300-5580

REFERENCES

1. Hegedus L. Clinical practice. The thyroid nodule. *N Engl J Med*. 2004;351(17):1764-1771.
2. Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med*. 1993;328(8):553-559.
3. Alexander EK, Cibas ES. Diagnosis of thyroid nodules. *Lancet Diabetes Endocrinol*. 2022;10(7):533-539.
4. Sakorafas GH. Thyroid nodules; interpretation and importance of fine-needle aspiration (FNA) for the clinician—practical considerations. *Surg Oncol*. 2010;19(4):e130-e139.
5. American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer, Cooper DS, Doherty GM, Haugen BR, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009;19(11):1167-1214.
6. Cibas ES, Ali SZ, NCI Thyroid FNA State of the Science Conference. The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol*. 2009;132(5):658-665.
7. Wang CC, Friedman L, Kennedy GC, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid*. 2011;21(3):243-251.
8. Nikiforov YE, Carty SE, Chiosea SI, et al. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. 2014;120(23):3627-3634.
9. Steward DL, Carty SE, Sippel RS, et al. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol*. 2019;5(2):204-212.

10. Lupo MA, Walts AE, Sistrunk JW, et al. Multiplatform molecular test performance in indeterminate thyroid nodules. *Diagn Cytopathol*. 2020;48(12):1254-1264.
11. Benjamin H, Schnitzer-Perlman T, Shtabsky A, et al. Analytical validity of a microRNA-based assay for diagnosing indeterminate thyroid FNA smears from routinely prepared cytology slides. *Cancer Cytopathol*. 2016;124(10):711-721.
12. 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.
13. Chudova D, Wilde JI, Wang ET, et al. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metab*. 2010;95(12):5296-5304.
14. Duick DS, Klopper JP, Diggans JC, et al. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. *Thyroid*. 2012;22(10):996-1001.
15. 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.
16. Angell TE, Wirth LJ, Cabanillas ME, et al. Analytical and clinical validation of expressed variants and fusions from the whole transcriptome of thyroid FNA samples. *Front Endocrinol (Lausanne)*. 2019;10:612.
17. Andrioli M, Carocci S, Alessandrini S, et al. Testing for Afirma in thyroid nodules with high-risk indeterminate cytology (TIR3B): first Italian experience. *Endocr Pathol*. 2020;31(1):46-51.
18. Links T. Richtlijn voor de diagnostiek, behandeling en follow-up van patiënten met gedifferentieerd (niet-medullair) schildkliercarcinoom 2015. https://richtlijnendatabase.nl/richtlijn/schildkliercarcinoom/algemeen.html. Updated February 16, 2015
19. Lončar I, van Dijk SPJ, Metman MJH, et al. Active surveillance for papillary thyroid microcarcinoma in a population with restrictive diagnostic workup strategies. *Thyroid*. 2021;31(8):1219-1225.
20. Kim BW, Yousman W, Wong WX, Cheng C, McAninch EA. Less is more: comparing the 2015 and 2009 American Thyroid Association guidelines for thyroid nodules and cancer. *Thyroid*. 2016;26(6):759-764.
21. Suster S. Thyroid tumors with a follicular growth pattern: problems in differential diagnosis. *Arch Pathol Lab Med*. 2006;130(7):984-988.
22. Alexander EK, Schorr M, Klopper J, et al. Multicenter clinical experience with the Afirma Gene Expression Classifier. *J Clin Endocrinol Metab*. 2014;99(1):119-125.
23. Valderrabano P, Hallanger-Johnson JE, Thapa R, Wang X, McIver B. Comparison of postmarketing findings vs the initial clinical validation findings of a thyroid nodule gene expression classifier: a systematic review and meta-analysis. *JAMA Otolaryngol Head Neck Surg*. 2019;145(9):783-792.
24. Sacks WL, Bose S, Zumsteg ZS, et al. Impact of Afirma Gene Expression Classifier on cytopathology diagnosis and rate of thyroidectomy. *Cancer Cytopathol*. 2016;124(10):722-728.
25. Metman MJH, Lončar I, Kruijff S, Engelsman AF, van Ginhoven TM. Is less always more in a national well-differentiated thyroid cancer population? *Eur J Surg Oncol*. 2020;46:709-711.
26. Dal Maso L, Tavilla A, Pacini F, et al. Survival of 86,690 patients with thyroid cancer: a population-based study in 29 European countries from EUROCARE-5. *Eur J Cancer*. 2017;77:140-152.
27. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA*. 2017;317(13):1338-1348.
28. Li M, Dal Maso L, Vaccarella S. Global trends in thyroid cancer incidence and the impact of overdiagnosis. *Lancet Diabetes Endocrinol*. 2020;8(6):468-470.
29. Malandrino P, Pellegriti G, Attard M, et al. Papillary thyroid microcarcinomas: a comparative study of the characteristics and risk factors at presentation in two cancer registries. *J Clin Endocrinol Metab*. 2013;98(4):1427-1434.
30. 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.
31. Parajuli S, Jug R, Ahmadi S, Jiang XS. Hurthle cell predominance impacts results of Afirma Gene Expression Classifier and ThyroSeq molecular panel performance in indeterminate thyroid nodules. *Diagn Cytopathol*. 2019;47(11):1177-1183.
32. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. *Endocr Pract*. 2014;20(4):364-369.
33. Lastra RR, Pramick MR, Crammer CJ, LiVolsi VA, Baloch ZW. Implications of a suspicious Afirma test result in thyroid fine-needle aspiration cytology: an institutional experience. *Cancer Cytopathol*. 2014;122(10):737-744.
34. McIver B, Castro MR, Morris JC, et al. An independent study of a Gene Expression Classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metab*. 2014;99(11):4069-4077.
35. 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.
36. 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.
37. Livhits MJ, Zhu CY, Kuo EJ, et al. Effectiveness of molecular testing techniques for diagnosis of indeterminate thyroid nodules: a randomized clinical trial. *JAMA Oncol*. 2021;7(1):70-77.
38. de Koster EJ, de Geus-Oei LF, Brouwers AH, et al. [(18)F]FDG-PET/CT to prevent futile surgery in indeterminate thyroid nodules: a blinded, randomised controlled multicentre trial. *Eur J Nucl Med Mol Imaging*. 2022;49:1970-1984.
39. Trimboli P, Piccardo A, Alevizaki M, et al. Dedicated neck (18) F-FDG PET/CT: an additional tool for risk assessment in thyroid nodules at ultrasound intermediate risk. *Clin Endocrinol (Oxf)*. 2019;90(5):737-743.
40. Grant EG, Tessler FN, Hoang JK, et al. Thyroid ultrasound reporting lexicon: White paper of the ACR Thyroid Imaging,

Reporting and Data System (TIRADS) Committee. *J Am Coll Radiol*. 2015;12:1272-1279.
41. Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L. European thyroid association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *Eur Thyroid J*. 2017;6(5):225-237.
42. Lee S, Skelton TS, Zheng F, et al. The biopsy-proven benign thyroid nodule: is long-term follow-up necessary? *J Am Coll Surg*. 2013;217(1):81-88; discussion 8–9.
43. Nou E, Kwong N, Alexander LK, Cibas ES, Marqusee E, Alexander EK. Determination of the optimal time interval for repeat evaluation after a benign thyroid nodule aspiration. *J Clin Endocrinol Metab*. 2014;99(2):510-516.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lončar I, van Velsen EFS, Massolt ET, et al. European experience with the Afirma Gene Expression Classifier for indeterminate thyroid nodules: A clinical utility study in the Netherlands. *Head & Neck*. 2023;45(9):2227-2236. doi:10.1002/hed.27472