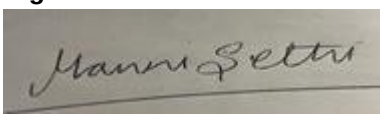
Prior Authorization Review Panel
MCO Policy Submission

A separate copy of this form must accompany each policy submitted for review.
Policies submitted without this form will not be considered for review.

Plan: Keystone First	**Submission Date**: 6/1/24
Policy Number: ccp.1099	**Effective Date:** 9/2014 **Revision Date:** May 1, 2024
Policy Name: Afirma® thyroid FNA analysis for indeterminate thyroid nodules	
Type of Submission – Check all that apply: **New Policy** **X Revised Policy*** **Annual Review – No Revisions** **Statewide PDL**	
***All revisions to the policy must be highlighted using track changes throughout the document.** **Please provide any clarifying information for the policy below:** See tracked changes below.	
Name of Authorized Individual (Please type or print): Manni Sethi, MD, MBA, CHCQM	**Signature of Authorized Individual:** 

Keystone First
Family of Health Plans

Afirma® thyroid FNA analysis for indeterminate thyroid nodules

Clinical Policy ID: CCP.1099

Recent review date: 5/2024

Next review date: 9/2025

Policy contains: Molecular testing; thyroid neoplasm; thyroid nodule.

Keystone First has developed clinical policies to assist with making coverage determinations. Keystone First's clinical policies are based on guidelines from established industry sources, such as the Centers for Medicare & Medicaid Services (CMS), state regulatory agencies, the American Medical Association (AMA), medical specialty professional societies, and peer-reviewed professional literature. These clinical policies along with other sources, such as plan benefits and state and federal laws and regulatory requirements, including any state- or plan-specific definition of "medically necessary," and the specific facts of the particular situation are considered by Keystone First, on a case by case basis, when making coverage determinations. In the event of conflict between this clinical policy and plan benefits and/or state or federal laws and/or regulatory requirements, the plan benefits and/or state and federal laws and/or regulatory requirements shall control. Keystone First's clinical policies are for informational purposes only and not intended as medical advice or to direct treatment. Physicians and other health care providers are solely responsible for the treatment decisions for their patients. Keystone First's clinical policies are reflective of evidence-based medicine at the time of review. As medical science evolves, Keystone First will update its clinical policies as necessary. Keystone First's clinical policies are not guarantees of payment.

Coverage policy

Afirma® Thyroid Analysis using the Genomic Sequencing Classifier (Veracyte, Inc., South San Francisco, California) is clinically proven and, therefore, may be medically necessary to rule out thyroid neoplasm when all of the following criteria are met (Haugen, 2016, 2017; National Comprehensive Cancer Network, 2024):

- Member is at least 18 years old.
- Thyroid nodule is at least one centimeter on ultrasonography.
- Presence of thyroid nodules with one or more prior non-diagnostic fine needle aspirates, described as either:
    - Atypia (or follicular lesion) of undetermined significance (Bethesda System for Reporting Thyroid Cytopathology Category III, Cibas, 2017).
    - Follicular neoplasm or suspicious for follicular neoplasm, including Hűrthle cell (Bethesda System for Reporting Thyroid Cytopathology Category IV, Cibas, 2017).
- Results are expected to influence clinical management (Gharib, 2016).
- Member is not undergoing thyroid surgery for diagnostic confirmation.

The Afirma Xpression Atlas test is investigational/not clinically proven and, therefore, not medically necessary.

Limitations

Frequency of testing using the Afirma Genomic Sequencing Classifier is limited to once per lifetime per member. In the event of a second unrelated thyroid nodule, additional testing may be medically necessary if the same criteria as the initial thyroid nodule are met.

Alternative covered services

- Thyroid stimulating hormone levels.
- Conventional imaging (e.g., ultrasonography, elastosonography).
- Functional and molecular imaging (e.g., ^{99m}Tc-MIBI scintigraphy, positron emission tomography).
- Open or closed thyroid biopsy.

Background

An estimated 43,720 new cases of thyroid cancer and 2,120 deaths from thyroid cancer will occur in 2023 in the United States. Thyroid cancer affects females more often than males and usually occurs in people ages 25 to 65 years (National Cancer Institute, undated).

Levels of thyroid stimulating hormone, clinical features, and high-resolution ultrasound findings of the thyroid gland and neck determine risk of malignancy. Cytologic examination of the thyroid with ultrasound-guided, fine-needle aspiration is used to rule out malignancy in suspicious nodules, but it produces indeterminate results in up to 30% of cases (Bose, 2019). Indeterminate aspirates comprise three Bethesda System for Reporting Thyroid Cytopathology categories, each with differing associated risks of malignancy and corresponding treatment (de Koster, 2018; Krane, 2015):

- Category III — Atypia of undetermined significance/follicular lesion of undetermined significance is considered a low-risk indeterminate category with an estimated 5% to 15% of malignancy. Repeat fine-needle aspiration is often recommended.
- Category IV — Follicular neoplasm/suspicious for a follicular neoplasm has a 15% to 30% risk of malignancy. Lobectomy is recommended in most cases.
- Category V — Suspicious for malignancy has a 60% to 75% risk of malignancy (typically papillary carcinoma with less distinct nuclear features). Thyroid surgery is recommended.

The significance of Hürthle cells found in follicular thyroid pathology can be challenging to interpret, as they can present in both benign and malignant neoplastic thyroid disease. Carcinomas of the Hürthle cell type account for about 10% to 15% of all follicular carcinomas. Histologically, the key determinant for classifying follicular carcinomas of the Hürthle cell type is vascular and/or capsular invasion. No definitive cytological criteria exist for determining the benign or malignant nature of Hürthle cells before or during thyroidectomy, and prognostic indicators remain controversial (Auger, 2014).

The prevalence of malignancy in indeterminate nodules increases when signs and symptoms are present. In ambiguous cases, thyroid resection is performed for a definitive diagnosis. Molecular diagnostic tests are proposed as a triage tool to either “rule in” or “rule out” malignancy with greater accuracy and defer unnecessary surgeries for benign thyroid nodules. Laboratory-developed testing for genetic mutations associated with thyroid cancer employs reverse transcriptase polymerase chain reaction assay microarrays and next-generation ribonucleic acid (RNA) sequencing.

Afirma has developed three molecular diagnostic tests for thyroid analysis: the first-generation Afirma Gene Expression Classifier; second-generation Afirma Genomic Sequencing Classifier; and Afirma Xpression Atlas. The Afirma Gene Expression Classifier measured the activity or “expression” levels of 167 genes in a thyroid

nodule sample. The Afirma Genomic Sequencing Classifier, introduced in 2017, and the Afirma Xpression Atlas, introduced in 2018, employ next-generation RNA genomic sequencing to measure hundreds more genetic mutations linked to thyroid cancer (Veracyte, Inc, 2024).

The Genomic Sequencing Classifier has replaced the Gene Expression Classifier test. Both are considered "rule out" tests with high negative predictive values for determining the probability that an indeterminate thyroid nodule is truly benign, thereby identifying individuals with greater certainty who can safely undergo periodic surveillance and avoid overdiagnosis and unnecessary thyroidectomies. Test results are reported as either "benign" or "suspicious for malignancy." A "benign" result indicates very little chance (≤ 5%) of cancer. A "suspicious for malignancy" result implies the test could not rule out the presence of cancer with complete certainty (i.e., the risk of cancer has increased) (Krane, 2020).

The Xpression Atlas was developed as an adjunct to the Genomic Sequencing Classifier for cytologically indeterminate, Bethesda V/VI (suspicious for malignancy or malignant), and Genomic Sequencing Classifier-suspicious thyroid nodules to further refine the risk of cancer. Although Xpression Atlas is not a cancer rule-out test, it has high positive predictive value for characterizing the presence of tumor oncogenic drivers that may inform decisions regarding systemic targeted therapy (Krane, 2020).

Findings

Guideline recommendations for the Afirma Gene Expression Classifier are equivocal, reflecting uncertainty in its clinical role. The American College of Clinical Endocrinologists recommends neither for nor against the use of this test for cytologically indeterminate nodules, stressing a complementary role to cytopathology when the results are expected to influence clinical management (Gharib, 2016).

The National Comprehensive Cancer Network (2024) suggests molecular diagnostic testing for evaluating Bethesda III or IV cytopathology results, taking into consideration the clinical, radiographic, and cytologic features of each patient. Molecular diagnostic testing, such as multigene assays or individual mutational analysis, may inform decisions about targeted therapy options for advanced disease.

The American Thyroid Association strongly recommended (based on weak evidence) counseling patients regarding the potential benefits and limitations of molecular testing and uncertainties in the therapeutic and long-term clinical implications of results (Haugen, 2016, 2017). A guideline for pediatric thyroid nodules states no studies exist to determine the usefulness of multigene expression classifier in children (Francis, 2015).

The evidence supporting use of Afirma for testing thyroid nodules is primarily taken from five large systematic reviews/meta-analyses:

- A review of seven studies determined that the Afirma gene expression classifier test for indeterminate thyroid nodules had a 95.7% sensitivity ($P = .09$), and 30.5% specificity ($P < .01$). Authors termed the test "an excellent tool to rule out malignancy" (Santhanam, 2016).
- An analysis of 12 studies identified the most common methodologic flaw with Afirma testing occurred when un-excised benign intermediate thyroid nodules were excluded from analyses, making estimates of specificity and negative predictive value unreliable. Authors suggest including benign nodules that do not develop malignant signs or symptoms during a follow-up in any accuracy calculations (Duh, 2017).
- A meta-analysis of 18 studies (n = 5,290) found sensitivity of the gene expression classifier test was 95.5% ($P < .001$) and specificity of 22.1% ($P < .001$). High sensitivity makes ruling out malignant nodules likely, but over half of nodules with suspicious results require more evaluation (Liu, 2019).
- An assessment of seven studies found the Afirma gene sequencing classifier (versus gene expression classifier) had a higher benign call rate (65.3% versus 43.8%, $P < .001$), a lower resection rate (26.8%

versus 50.1%, *P* < .001), and a higher risk of malignancy (60.1% versus 37.6%, *P* < .001). Authors state these improvements could reduce unnecessary surgeries and tailor clinical decisions (Vuong, 2021).

- A review of 40 studies (n = 7,565) found the best performance in diagnostic testing of indeterminate thyroid nodules was by Thyroseq v3 (Area under the curve 0.95); however, there are a paucity of studies for Thyroseq v3. The more frequently used Afirma gene sequencing classifier (0.90) and Thyroseq v2 (0.88) achieved slightly lower rates (Silaghi, 2021).

In addition, a number of single studies, each with hundreds of subjects, provide supportive evidence for use of Afirma testing of thyroid nodules:

- Use of the Afirma test (n = 3,789) on 85 patients classified as malignant by fine needle aspiration correctly identified 78 of the 85 as suspicious (92% sensitivity) (Alexander, 2012). A follow-up showed Afirma gene expression classifier significantly altered care recommendations, as 4 of 175 benign were recommended for surgery versus 141 of 149 that were suspicious (*P* < .01) (Alexander, 2014).
- A study (n = 563) found the benign cell rate was 47.9% for Afirma gene expression classifier versus 65.8% for the genomic sequencing classifier (*P* = .0006). The higher rate, occurring predominantly among nodules with Hürthle-cell cytology, will lead to fewer surgeries (Angell, 2019).
- Fine needle aspiration (n = 519) identified 58 as indeterminate, and Afirma gene expression classifier identified 36 (62%) of these as suspicious (62%), and 20 (34%) as benign. Of the 36 suspicious cases, 21 had malignant final pathology, versus 2 of the 20 benign cases (Harrell, 2014).
- Afirma gene expression classifier testing in 475 patients who had undergone thyroid surgery showed surgery was clinically indicated for other reasons in 22% of patients with benign results. The false-negative percentage (surgically proven false negatives divided by the total Afirma benign patients) was 7.3% (Harrell, 2018).
- Afirma gene sequencing classifier (n = 481) identified fewer indeterminate cytology nodules as suspicious, compared with gene expression classifier (38.8% versus 58.4%). Authors conclude the gene sequencing classifier further reduces surgery in indeterminate thyroid nodules by improving specificity of Afirma results without compromising sensitivity (Harrell, 2019).

In 2023, we added two systematic reviews/meta-analyses (DiGennaro, 2022; Lee, 2022), one cost effectiveness analysis, and a new cohort study (Hu, 2022). All included trials of the Afirma gene expression classifier and Afirma genomic sequencing classifier for assessing indeterminate thyroid nodules. No policy changes are warranted.

DiGennaro (2022) found the following diagnostic tests showed high diagnostic accuracy for assessing malignancy risk in cytologically indeterminant thyroid nodules, although there were methodological limitations in the studies:

- Afirma gene expression classifier (38 studies): Sensitivity = 92%, specificity = 26%, negative likelihood ratio = 0.32, positive likelihood ratio = 1.24, and area under the receiver operating curve = 0.83.
- Afirma genomic sequencing classifier (10 studies): Sensitivity = 94%, specificity = 38%, negative likelihood ratio = 0.18, positive likelihood ratio = 1.52, area under the receiver operating curve = 0.91.

Another systematic review/meta-analysis found both second-generation molecular tests—Afirma genomic sequencing classifier (seven studies) and Thyroseq v3 (six studies)—showed high sensitivity and high negative predictive value for ruling out malignancy in thyroid nodules with indeterminate cytology results, including Bethesda III and IV categories. There were no statistically significant differences in the diagnostic performance of the two tests. Further long-term outcome data are needed (Lee, 2022).

A cost effectiveness analysis, taken from the Canadian perspective, reported a strategy using the Afirma Gene Expression Classifier for examining a cytologically-indeterminate solitary thyroid nodule with no additional high-risk features was cost-effective for avoiding unnecessary thyroid surgery. The incremental cost effectiveness ratio was $4,234.22 per surgery avoided, which varied based on the cost of the molecular test and the willingness-to-pay threshold to avoid unnecessary thyroid surgery (Dharampal, 2022).

Results of a prospective cohort study of 343 patients with 375 indeterminate thyroid nodules found both the Afirma Genomic Sequencing Classifier and ThyroSeq v3 tests performed well in ruling out malignancy in sonographically low/intermediate-suspicion thyroid nodules but had limited diagnostic value in sonographically high-suspicion nodules. The impact of these test results on clinical decision making is the subject of an ongoing clinical trial (Hu, 2022: clinicaltrials.gov identifier NCT02681328).

In 2024, we updated the references, added a new meta-analysis, and removed the Gene Expression classifier from coverage, as the Genomic Sequencing Classifier test has replaced it. The Afirma Xpression Atlas test remains investigational.

In a meta-analysis of 13 post validation studies (n = 1,976 participants), compared to the results from the original validation study performed under controlled conditions, use of the Genomic Sequencing Classifier in real world conditions resulted in a higher benign call rate (67% versus 54%, $P < .05$), which is the percentage of cytomorphologically indeterminate cases with subsequent benign or negative test results. The negative predictive value (reported with 95% confidence intervals) (0.995 [0.98 to 0.999] versus 0.961 [0.904 to 0.989], $P = 0.018$) was higher for ruling out thyroid cancer in participants with indeterminate thyroid nodules and benign Genomic Sequencing Classifier results (Nasr, 2023). No other policy changes are warranted.

References

On March 18, 2024, we searched PubMed and the databases of the Cochrane Library, the U.K. National Health Services Centre for Reviews and Dissemination, the Agency for Healthcare Research and Quality, and the Centers for Medicare & Medicaid Services. Search terms were “Thyroid Nodule” (MeSH), “Gene Expression Profiling” (MeSH), “thyroid,” and “Afirma.” We included the best available evidence according to established evidence hierarchies (typically systematic reviews, meta-analyses, and full economic analyses, where available) and professional guidelines based on such evidence and clinical expertise.

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. Doi: 10.1056/NEJMoa1203208.

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. Doi: 10.1210/jc.2013-2482.

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. Doi: 10.1089/thy.2018.0726.

Auger M. Hürthle cells in fine-needle aspirates of the thyroid: A review of their diagnostic criteria and significance. *Cancer cytopathology*. 2014;122(4):241-9. Doi: 10.1002/cncy.21391.

Bose S, Sacks W, Walts AE. Update on molecular testing for cytologically indeterminate thyroid nodules. *Adv Anat Pathol.* 2019;26(2):114-123. Doi: 10.1097/pap.0000000000000211.

Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid.* 2017;27(11):1341-1346. Doi: 10.1089/thy.2017.0500.

de Koster EJ, de Geus-Oei L-F, Dekkers OM, et al. Diagnostic utility of molecular and imaging biomarkers in cytological indeterminate thyroid nodules. *Endocr Rev*. 2018;39(2):154-191. Doi: 10.1210/er.2017-00133.

Dharampal N, Smith K, Harvey A, Paschke R, Rudmik L, Chandarana S. Cost-effectiveness analysis of molecular testing for cytologically indeterminate thyroid nodules. *J Otolaryngol Head Neck Surg*. 2022;51(1):46. Doi: 10.1186/s40463-022-00604-7.

DiGennaro C, Vahdatzad V, Jalali MS, et al. Assessing bias and limitations of clinical validation studies of molecular diagnostic tests for indeterminate thyroid nodules: Systematic review and meta-analysis. *Thyroid*. 2022;32(10):1144-1157. Doi: 10.1089/thy.2022.0269.

Duh QY, Busaidy NL, Rahilly-Tierney C, Gharib H, Randolph G. A systematic review of the methods of diagnostic accuracy studies of the Afirma gene expression classifier. *Thyroid.* 2017;27(10):1215-1222. Doi: 10.1089/thy.2016.0656.

Francis GL, Waguespack SG, Bauer AJ, et al. Management guidelines for children with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2015;25(7):716-59. Doi: 10.1089/thy.2014.0460.

Gharib H, Papini E, Garber JR, et al. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules – 2016 update. *Endocr Pract.* 2016;22(5):622-639. Doi: 10.4158/ep161208.gl.

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. Doi: 10.4158/ep13330.or.

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. Doi: 10.4158/ep-2018-0395.

Harrell RM, Eyerly-Webb SA, Pinnar NE, et al. Community endocrine surgical experience with false-negative Afirma GEC® results: 2011-2017. *Endocr Pract.* 2018;24(7):622-627. Doi: 10.4158/ep-2017-0263.

Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association Guidelines Task Force on thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2016;26(1):1-133. Doi: 10.1089/thy.2015.0020.

Haugen BR, Sawka AM, Alexander EK, et al. American Thyroid Association guidelines on the management of thyroid nodules and differentiated thyroid cancer task force review and recommendation on the proposed renaming of encapsulated follicular variant papillary thyroid carcinoma without invasion to noninvasive follicular thyroid neoplasm with papillary-like nuclear features. *Thyroid.* 2017;27(4):481-483. Doi: 10.1089/thy.2016.0628.

Hu TX, Nguyen DT, Patel M, et al. The effect modification of ultrasound risk classification on molecular testing in predicting the risk of malignancy in cytologically indeterminate thyroid nodules. *Thyroid*. 2022;32(8):905-916. Doi: 10.1089/thy.2021.0659.

Krane JF, Cibas ES, Alexander EK, Paschke R, Eszlinger M. Molecular analysis of residual ThinPrep material from thyroid FNAs increase diagnostic sensitivity. *Cancer Cytopathol*. 2015;123(6):356-61. Doi: 10.1002/cncy.21546.

Krane JF, Cibas ES, Endo M, et al. The Afirma Xpression Atlas for thyroid nodules and thyroid cancer metastases: Insights to inform clinical decision-making from a fine-needle aspiration sample. *Cancer cytopathology*. 2020;128(7):452-459. Doi: 10.1002/cncy.22300.

Lee E, Terhaar S, McDaniel L, et al. Diagnostic performance of the second-generation molecular tests in the assessment of indeterminate thyroid nodules: A systematic review and meta-analysis. *Am J Otolaryngol*. 2022;43(3):103394. Doi: 10.1016/j.amjoto.2022.103394.

Liu Y, Pan B, Xu L, Fang D, Ma X, Lu H. The diagnostic performance of Afirma gene expression classifier for the indeterminate thyroid nodules: A meta-analysis. *Biomed Res Int*. 2019;2019:7150527. Doi: 10.1155/2019/7150527.

Nasr CE, Andrioli M, Endo M, et al. Real-world performance of the Afirma Genomic Sequencing Classifier (GSC) — a meta-analysis. *J Clin Endocrinol Metab*. 2023;108(6):1526-1532. Doi: 10.1210/clinem/dgac688.

National Cancer Institute. SEER cancer stat facts: Thyroid cancer. National Cancer Institute. Bethesda, MD. https://seer.cancer.gov/statfacts/html/thyro.html. Undated.

National Comprehensive Cancer Network. NCCN guidelines. Thyroid carcinoma. Version 2.2024. www.nccn.org. Published March 12, 2024.

Santhanam P, Khthir R, Gress T, et al. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: A meta-analysis. *Med Oncol.* 2016;33(2):14. Doi: 10.1007/s12032-015-0727-3.

Silaghi CA, Vozovanu V, Georgescu CE, et al. Thyroseq v3, Afirma GSC, and microRNA panels versus previous molecular tests in the preoperative diagnosis of indeterminate thyroid nodules: A systematic review and meta-analysis. *Front Endocrinol (Lausanne)*. 2021;12:649522. Doi: 10.3389/fendo.2021.649522.

Veracyte, Inc. Afirma Genomic Sequencing Classifier. Exceptional thyroid cancer diagnostics. https://www.veracyte.com/diagnostics/thyroid-cancer. Published 2024.

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. Doi: 10.1002/cncy.22332.

Policy updates

3/2014: initial review date and clinical policy effective date: 9/2014

3/2016: Policy references updated.

3/2018: Policy references updated.

3/2019: Policy references updated, coverage modified, and policy ID changed.

3/2020: Policy references updated and coverage expanded.

3/2021: Policy references updated.

3/2022: Policy references updated.

5/2023: Policy references updated.

5/2024: Policy references updated. Coverage modified.