| 2.04.142 | Molecular Testing in the Management of Pulmonary Nodules | | |
|---|---|---|---|
| **Original Policy Date:** | July 1, 2017 | **Effective Date:** | November 1, 2025 |
| **Section:** | 2.0 Medicine | **Page:** | Page 1 of 25 |

## Policy Statement

I. Plasma-based proteomic screening, including but not limited to Nodify XL2® (BDX-XL2), Nodify CDT®, and REVEAL Lung Nodule Characterization (MagArray), in individuals with undiagnosed pulmonary nodules detected by computed tomography is considered **investigational**.

II. Gene expression profiling on bronchial brushings, including but not limited to the Percepta® Genomic Sequencing Classifier, in individuals with indeterminate bronchoscopy results from undiagnosed pulmonary nodules is considered **investigational**.

Note: For individuals enrolled in health plans subject to the Biomarker Testing Law (Health & Safety Code Section 1367.667 and the Insurance Code Section 10123.209), Centers for Medicare & Medicaid Services (CMS) Local Coverage Determination (LCD) may also apply. Please refer to the Medicare National and Local Coverage section of this policy and MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy for reference.

**NOTE**: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

## Policy Guidelines

Plans may need to alter local coverage medical policy to conform to state law regarding coverage of biomarker testing.

### Coding

See the Codes table for details.

## Description

Plasma-based proteomic screening and gene expression profiling of bronchial brushing are molecular tests available in the diagnostic workup of pulmonary nodules. To rule out malignancy, invasive diagnostic procedures such as computed tomography-guided biopsies, bronchoscopies, or video-assisted thoracoscopic procedures are often required, but each carry procedure-related complications ranging from postprocedure pain to pneumothorax. Molecular diagnostic tests have been proposed to aid in risk-stratifying patients to eliminate or necessitate the need for subsequent invasive diagnostic procedures.

### Summary of Evidence

For individuals with undiagnosed pulmonary nodules detected by computed tomography who receive plasma-based proteomic screening, the evidence includes prospective cohorts, retrospective studies, and prospective-retrospective studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, morbid events, hospitalizations, and resource utilization. Clinical validation studies were identified for 2 versions (Xpresys Lung, and Xpresys Lung version 2 [now Nodify XL2]) of a proteomic classifier and another lung nodule characterization test (REVEAL).

The Nodify XL2 classifier has undergone substantial evolution, from a 13-protein assay to a 2-protein assay integrated with clinical factors. Because of this evolution, the most relevant studies are with the most recent version (Xpresys Lung version 2 [now Nodify XL2]). One validation study on version 2 has been identified. The classifier has been designed to have high specificity for malignant pulmonary nodules, and the validation study showed a specificity of 97% for patients with a low-to-moderate pretest probability (≤50%) of a malignant pulmonary nodule. The primary limitation of this study is that a high number of patients were excluded from the study due to incomplete clinical data or because they were subsequently determined to be outside of the intended use population. It is unclear if the intended use population was determined a priori. Validation in an independent sample in the intended use population is needed. No recent clinical validation studies were identified for the Nodify CDT test or the Nodify Lung testing strategy. The REVEAL validation study was a retrospective study that demonstrated use as a rule-out test in conjunction with the Veteran's Affairs (VA) Clinical Factors Model when the samples were considered inconclusive or intermediate risk by the VA model. The REVEAL model subsequently correctly identified 65% of intermediate-risk samples as either low or high risk. The negative predictive value and sensitivity were both 94%. Limitations included a small sample size and use in conjunction with just 1 type of testing model. Validation in an independent sample in the intended use population with additional probability models is needed. Indirect evidence suggests that a proteomic classifier with a high negative predictive value has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. However, long-term follow-up data would be required to determine the survival outcomes in patients with a missed diagnosis of lung cancer at earlier, more treatable stages. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with undiagnosed pulmonary nodules following indeterminate bronchoscopy results for suspected lung cancer who receive gene expression profiling of bronchial brushings, the evidence includes multicenter prospective studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, morbid events, hospitalizations, and resource utilization. A 3-cohort, prospective, multicenter study validated the second generation Percepta Genomic Sequencing Classifier (GSC) test in an independent sample set, showing high sensitivity for the rule-out portion of the classifier and high specificity for the rule-in portion of the classifier. For intermediate pretest risk patients with an inconclusive bronchoscopy, Percepta GSC can down-classify the risk of primary lung cancer to low with a 91% negative predictive value, or up-classify the risk to high with a 65% positive predictive value. Further assessment of clinical utility is warranted. Also, where the test would fall in the clinical pathway (i.e., other than indeterminate bronchoscopy) is uncertain. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

**Additional Information**
Not applicable.

## Related Policies

- N/A

## Benefit Application

Benefit determinations should be based in all cases on the applicable member health services contract language. To the extent there are conflicts between this Medical Policy and the member health services contract language, the contract language will control. Please refer to the member's contract benefits in effect at the time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal law may prohibit health plans from denying FDA-approved Healthcare Services as investigational or experimental. In these instances, Blue Shield of California may be obligated to determine if these FDA-approved Healthcare Services are Medically Necessary.

## Regulatory Status

**Cal. Health & Safety Code §1367.667, Insurance Code Section 10123.209, and Welfare and Institutions Code 14132.09**
California laws that requires insurers to cover biomarker testing for the diagnosis, treatment, appropriate management, or ongoing monitoring of an enrollee's disease or condition to guide treatment decisions, as prescribed.

**Clinical Laboratory Improvement Amendments (CLIA) and FDA Regulatory Overview**
Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). Xpresys Lung 2, now Nodify XL2 (BDX-XL2; Integrated Diagnostics [Indi], purchased by Biodesix); Nodify CDT (Biodesix); REVEAL Lung Nodule Characterization (MagArray); and Percepta Genomic Sequencing Classifier (Veracyte) are available under the auspices of the CLIA. Laboratories that offer laboratory-developed tests must be licensed by the CLIA for high-complexity testing. To date, the U.S. Food and Drug Administration (FDA) has chosen not to require any regulatory review of these tests.

## Rationale

**Background**
**Pulmonary Nodules**
Pulmonary nodules are a common clinical problem that may be found incidentally on a chest x-ray or computed tomography (CT) scan or during lung cancer screening studies of smokers. The primary question after the detection of a pulmonary nodule is the probability of malignancy, with subsequent management of the nodule based on various factors such as the radiographic characteristics of the nodules (e.g., size, shape, density) and patient factors (e.g., age, smoking history, previous cancer history, family history, environmental/occupational exposures). The key challenge in the diagnostic workup for pulmonary nodules is appropriately ruling in patients for invasive diagnostic procedures and ruling out patients who should forego invasive diagnostic procedures. However, due to the low positive predictive value of pulmonary nodules detected radiographically, many unnecessary invasive diagnostic procedures and/or surgeries are performed to confirm or eliminate the diagnosis of lung cancer.

**Proteomics**
Proteomics is the study of the structure and function of proteins. The study of the concentration, structure, and other characteristics of proteins in various bodily tissues, fluids, and other materials has been proposed as a method to identify and manage various diseases, including cancer. In proteomics, multiple test methods are used to study proteins. Immunoassays use antibodies to detect the concentration and/or structure of proteins. Mass spectrometry is an analytic technique that ionizes proteins into smaller fragments and determines mass and composition to identify and characterize them.

**Plasma-Based Proteomic Screening for Pulmonary Nodules**
Plasma-based proteomic screening has been investigated to risk-stratify pulmonary nodules as likely benign to increase the number of patients who undergo serial CT scans of their nodules (active surveillance), instead of invasive procedures such as CT-guided biopsy or surgery. Additionally, proteomic testing may also determine a likely malignancy in clinically low-risk or intermediate-risk pulmonary nodules, thereby permitting earlier detection in a subset of patients.

Nodify XL2 (BDX-XL2) is a plasma-based proteomic screening test that measures the relative abundance of proteins from multiple disease pathways associated with lung cancer using an analytic technique called multiple reaction monitoring mass spectroscopy. The test helps physicians identify lung nodules that are likely benign or at lower risk of cancer. If the test yields a "likely benign" or "reduced risk" result, patients may choose active surveillance via serial CT scans to monitor the pulmonary nodule. Earlier generations of the Nodify XL2 test include Xpresys Lung® and Xpresys Lung 2®.

Nodify CDT® is a proteomic test that uses multi-analyte immunoassay technology to measure autoantibodies associated with tumor antigens. The test helps physicians identify lung nodules that are likely malignant or at higher risk of cancer. Patients with a "high level" Nodify CDT test result have a higher risk of malignancy than predicted by clinical factors alone; invasive diagnostic procedures would be indicated in these cases.

The Nodify XL2 and Nodify CDT tests are therefore only used in the management of pulmonary nodules to rule out or rule in invasive diagnostic procedures; they do not diagnose lung cancer. These tests are offered together as Biodesix's Nodify Lung® testing strategy, but physicians may also choose to order each test independently.

REVEAL Lung Nodule Characterization (MagArray) is a plasma-protein biomarker test that may aid clinicians in characterizing indeterminate pulmonary nodules (4 to 30 mm) in current smokers 25 years of age and older. The test is based on a multianalyte assay with a proprietary algorithmic analysis using immunoassay, microarray, and magnetic nanoparticle detection techniques to obtain laboratory data for calculation of the risk score for lung cancer. The REVEAL Lung Nodule Characterization is presented on a scale from 0 to 100 with a single cut point at 50. The score is based on the measurement of 3 clinical factors (age, sex, and nodule diameter) and 3 proteins (epidermal growth factor receptor, prosurfactant protein B, and tissue inhibitor of metalloproteinases 1) associated with the presence of lung cancer. It may aid a clinician in the decision to perform a biopsy or to consider routine monitoring. It is not intended as a screening or stand-alone diagnostic assay.

### Gene Expression Profiling

Gene expression profiling (GEP) is the measurement of the activity of genes within cells. Messenger RNA serves as the bridge between DNA and functional proteins. Multiple molecular techniques such as Northern blots, ribonuclease protection assay, in situ hybridization, spotted complementary DNA arrays, oligonucleotide arrays, reverse transcriptase polymerase chain reaction, and transcriptome sequencing are used in GEP. An important role of GEP in molecular diagnostics is to detect cancer-associated gene expression in clinical samples to assess the risk for malignancy.

### Gene Expression Profiling for an Indeterminate Bronchoscopy Result

The first generation Percepta® Bronchial Genomic Classifier was a 23-gene, GEP test that analyzed genomic changes in the airways of current or former smokers to assess a patient's risk of having lung cancer, without direct testing of a pulmonary nodule. This classifier was designed to be a "rule-out" test for intermediate-risk patients. The second generation Percepta Genomic Sequencing Classifier was developed to serve as both a "rule-in" test and a "rule-out" test, thereby increasing its potential utility in improving risk stratification. The test is indicated for current and former smokers following an indeterminate bronchoscopy result to determine the subsequent management of pulmonary nodules (e.g., active surveillance or invasive diagnostic procedures) and does not diagnose lung cancer.

### Literature Review

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance

of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test. The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

## Plasma-Based Proteomic Screening of Pulmonary Nodules

### Clinical Context and Test Purpose

The purpose of plasma-based proteomic screening in individuals with undiagnosed pulmonary nodule(s) is to stratify clinical risk for malignancy and eliminate or necessitate the need for invasive diagnostic procedures.

The following PICO was used to select literature to inform this review.

#### *Populations*

The relevant population of interest is individuals with undiagnosed pulmonary nodules detected by computed tomography (CT). In particular, as outlined in the evidence-based American College of Chest Physicians guidelines (2013) on the diagnosis and management of lung cancer, decision-making about a single indeterminate lung nodule 8 to 30 mm in diameter on a CT scan is complicated, requiring input about the patient's pretest probability of lung cancer, the characteristics of the lung nodule on CT, and shared decision-making between the patient and physician about follow-up.[1,] Therefore, additional information in the segment of individuals with an indeterminate lung nodule, 8 to 30 mm in diameter, would be particularly useful.

#### *Interventions*

The tests being considered are plasma-based proteomic screening, the Nodify XL2 (BDX-XL2; formerly Xpresys Lung 2), Nodify CDT, and REVEAL Lung Nodule Characterization tests. Nodify XL2 (BDX-XL2) measures the abundance of 2 plasma proteins (LG3BP and C163A) and combines the results with 5 clinical risk factors (age, smoking status, nodule diameter, edge characteristics, and location) to provide a post-test probability of a lung nodule being benign. Nodify CDT measures 7 autoantibodies associated with tumor antigens to provide a post-test probability of a lung nodule being malignant. These 2 tests are offered alone, or in conjunction with each other as the Nodify Lung. REVEAL Lung Nodule Characterization (MagArray) measures 3 plasma proteins (epidermal growth factor receptor, prosurfactant protein B, and tissue inhibitor of metalloproteinases 1) associated with the presence of lung cancer and combines the results with 3 clinical factors (age, sex, and nodule diameter) to provide algorithmic scoring to quantify the likelihood of lung cancer as a risk assessment tool.

#### *Comparators*

The following practice is currently being used: standard diagnostic workup using clinical and radiographic risk factors.

#### *Outcomes*

The potential beneficial outcomes of primary interest are avoiding an unneeded invasive biopsy of a nodule that would be negative for lung cancer or initiating a biopsy for a nodule that would otherwise have been followed with serial CTs.

Potential harmful outcomes are those resulting from false-positive or false-negative test results. False-positive test results can lead to unnecessary invasive diagnostic procedures and procedure-related complications. False-negative test results can lead to lack of pulmonary nodule surveillance or lack of appropriate invasive diagnostic procedures to diagnose a malignancy.

The time frame for evaluating test performance varies from the initial CT scan to an invasive diagnostic procedure up to 2 years later, which would be the typical follow-up needed for some lung nodules.

### Study Selection Criteria

For the evaluation of clinical validity of the plasma-based proteomic screening test, studies that met the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores)
- Included a suitable reference standard
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described

### Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

### Review of Evidence

***Nodify XL2 (BDX-XL2; previously Xpresys Lung and Xpresys Lung 2)***

Several studies were identified that reported on the development and validation of Xpresys Lung, and Xpresys Lung 2/Nodify XL2 (BDX-XL2).

Li et al (2013) reported on an initial development that was based on a 13-protein plasma classifier.[2,]

Vachani et al (2015) reported the validation of Xpresys Lung, which was an 11-protein plasma classifier designed to identify likely benign lung nodules (Tables 1 and 2).[3,] This retrospective, blinded analysis evaluated existing samples (N=141) associated with indeterminate pulmonary nodules 8 to 30 mm in diameter. The performance of the classifier in identifying benign nodules was tested at predefined reference values. For example, using a population-based non-small-cell lung cancer prevalence estimate of 23% for indeterminate pulmonary nodules 8 to 30 mm in diameter, the classifier identified likely benign lung nodules with a 90% negative predictive value (NPV) and a 26% positive predictive value, at 92% sensitivity and 20% specificity, with the lower bound of the classifier's performance at 70% sensitivity and 48% specificity. Additional sample diagnostic characteristics, selected to keep the study's target negative predictive value of 90%, are shown in Table 2. Classifier scores for the overall cohort were statistically independent of patient age, tobacco use, nodule size, and chronic obstructive pulmonary disease diagnosis. The classifier also demonstrated incremental diagnostic performance in combination with a 4-parameter clinical model.

Vachani et al (2015) reported on a multicenter prospective-retrospective study of patients with indeterminate pulmonary nodules.[4,] A plasma protein classifier was used on 475 patients with nodules 8 to 30 mm in diameter who had an invasive procedure to confirm the diagnosis. Using the classifier, 32.0% (95% confidence interval [CI], 19.5 to 46.7) of surgeries and 31.8% (95% CI, 20.9 to 44.4) of invasive procedures (biopsy and/or surgery) on benign nodules could have been avoided, while 24.0% (95% CI, 19.2 to 29.4) of patients with malignancy would have been triaged to CT surveillance. By comparison, 24.5% (95% CI, 16.2 to 34.4) of patients with malignancy were routed to CT surveillance using clinical parameters alone.

Kearney et al (2017) conducted an exploratory study that combined the 11-protein plasma classifier (Xpresys Lung) with clinical risk factors using 222 samples associated with a lung nodule of 8 to 20 mm in diameter from the reclassification study by Vachani et al (2015) described above.[5,] The study determined that the ratio of LG3BP to a normalizer protein C163A was the diagnostic and normalizer protein pair with the highest area under the curve (60%). At a sensitivity of 90% and specificity of

33%, the ratio of the proteomic marker was more accurate than clinical risk factors, and the combination of the clinical risk factors with the proteomic markers was more accurate than either alone. This study led to the development of the Xpresys Lung version 2/Nodify XL2, which includes LG3BP, C163A, and clinical risk factors.

Silvestri et al (2018) reported the validation of the Xpresys Lung version 2/Nodify XL2 (BDX-XL2) in a prospective multicenter observational study (Pulmonary Nodule Plasma Proteomic Classifier [PANOPTIC]) that enrolled 685 patients with lung nodules of 8 to 30 mm and a low pretest probability of malignancy ≤50%.[6] After exclusions for missing clinical data or a pretest probability of >50%, 178 patients remained in the intended use population. Of these, 66 were classified as likely benign, 65 of which had a benign nodule, while 1 of 29 malignant nodules (3%) were misclassified as likely benign. Of the 149 benign nodules in the study, 44% were correctly classified as likely benign. Of the 71 patients who had invasive procedures, 42 had benign nodules. Use of the integrated proteomic classifier would have reduced the number of patients undergoing an invasive procedure to 27, a 36% relative risk reduction, with 1 malignant nodule misclassified as benign.

In an extended analysis and 2-year follow-up of the PANOPTIC trial, Tanner et al (2021) found that all nodules designated as benign at year 1 remained benign by imaging at year 2 with no change in pathologic diagnoses or nodule size by CT.[7] Additionally, the area under the curve of the integrated classifier was 0.76 (95% CI, 0.69 to 0.82), which outperformed the physician pretest probability for malignancy (0.69; 95% CI, 0.62 to 0.76) and the Mayo (0.69; 95% CI, 0.62 to 0.76), Veterans Administration (0.6; 95% CI, 0.53 to 0.67), and Brock (0.71; 95% CI, 0.63 to 0.77) models in the lower risk pretest probability (≤50%) cohort.

**Table 1. Study Characteristics of Clinical Validity**

| Study | Study Population | Design | Reference Standard | Threshold for Positive Index Test | Timing of Reference and Index Tests | Blinding of Assessors | Comment |
|---|---|---|---|---|---|---|---|
| **Vachani et al (2015)[3]** | 141 samples associated with indeterminate pulmonary nodules | Retrospective analysis with existing samples | | Selected to keep NPV of 90% | | Yes | Xpresys Lung |
| **Silvestri et al (2018)[6] PANOPTIC** | 178 patients with 8 to 30 mm lung nodules and low pretest probability | Prospective multicenter observational | Definitive diagnosis, nodule resolution, or 1 year of radiographic stability | NR | Retrospective evaluation of performance | Yes | Xpresys Lung version 2 |

NPV: negative predictive value; NR: not reported; PANOPTIC: Pulmonary Nodule Plasma Proteomic Classifier.

**Table 2. Summary of Diagnostic Performance Studies for Proteomic Tests to Predict Malignancy**

| Study | Prevalence, % | Reference Value | Sensitivity, % (95% CI) | Specificity, % | NPV, % | PPV, % |
|---|---|---|---|---|---|---|
| **Vachani et al (2015)[8,3]** | 23.1 | 0.47 | 69.5 (NR) | 48.0 (NR) | 84.0 | 28.6 |
| | 23.1 | 0.39 | 83.8 (NR) | 32.3 (NR) | 86.9 | 27.1 |
| | 23.1 | 0.36 | 82.1 (NR) | 20.4 (NR) | 89.6 | 25.8 |
| **Silvestri et al (2018)[6] PANOPTIC** | 16.3 | NR | 97 (82 to 100) | 44 (36 to 52) | 98 (92 to 100) | NR |

CI: confidence interval; NPV: negative predictive value; NR: not reported; PANOPTIC: Pulmonary Nodule Plasma Proteomic Classifier; PPV: positive predictive value.

Limitations of the 2 validation studies are described in Tables 3 and 4. The primary limitation of the study by Vachani et al (2015) is that the technology is very different from the current marketed version. The primary limitation of the study by Silvestri et al (2018) is that a high number of patients were excluded from the study due to incomplete clinical data or because they were subsequently determined to be outside of the intended use population. It is unclear if the intended use population was determined a priori.

### Table 3. Study Relevance Limitations

| Study | Population[a] | Intervention[b] | Comparator[c] | Outcomes[d] | Duration of Follow-Up[e] |
|---|---|---|---|---|---|
| Vachani et al (2015)[3], | | 3. Not the current version of the test. | | | |
| Silvestri et al (2018)[6], PANOPTIC | 4. The enrolled patients included those who were outside of intended use. | | | | |

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

### Table 4. Study Design and Conduct Limitations

| Study | Selection[a] | Blinding[b] | Delivery of Test[c] | Selective Reporting[d] | Data Completeness[e] | Statistical[f] |
|---|---|---|---|---|---|---|
| Vachani et al (2015)[3], | | | | | | |
| Silverstri et al (2018)[6],PANOPTIC | | | | 2. Data were collected but not reported for the 214 patients with a pretest probability >50%. | 2. A high number of patients (n=234) were excluded. | |

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

### *Nodify CDT and Nodify Lung*

No recent literature was identified for Nodify CDT or Nodify Lung (performing the Nodify XL2 and Nodify CDT tests in conjunction) that meets the evidence requirements of this review.

### *REVEAL Lung Nodule Characterization*

Trivedi et al (2018) reported on a clinical validation study for the REVEAL Lung Nodule Characterization test using retrospective human plasma samples and associated clinical data from current smokers 25 to 85 years of age with indeterminate lung nodules measuring 4 to 30 mm in diameter.[9,] Plasma samples from patients with metastatic disease or previously diagnosed lung cancer were excluded. The REVEAL test was used in conjunction with the Veteran's Affairs (VA) Clinical Factors Model, with the objective to add discriminatory information when the VA model classified samples as inconclusive or intermediate risk. Ninety-seven samples were included in the validation study. Of the 97 samples, 68 were grouped as having intermediate risk by the VA model. The REVEAL model correctly identified 44 (65%) of these intermediate-risk samples as low (n=16) or high (n=28) risk. The REVEAL assay NPV was 94% and its sensitivity was 94%, suggesting potential application as a rule-out test to increase the confidence of providers to avoid aggressive interventions for patients for whom the VA model result is inconclusive or intermediate risk.

### Section Summary: Clinically Valid

Clinical validation studies were identified for 2 versions (Xpresys Lung, and Xpresys Lung 2 [now Nodify XL2]) of a proteomic classifier and another lung nodule characterization test (REVEAL). The Nodify XL2 classifier has undergone substantial evolution, from a 13-protein assay to a 2-protein assay integrated with clinical factors. Because of this evolution, the most relevant studies are with the most recent version 2. One validation study on version 2 (Xpresys Lung 2 [now Nodify XL2]) has been identified. The classifier has been designed to have high specificity for malignant pulmonary nodules, and the validation study showed a specificity of 97% for patients with a low to moderate pretest probability (≤50%) of a malignant pulmonary nodule. The primary limitation of this study is that a high number of patients were excluded from the study due to incomplete clinical data or because they were subsequently determined to be outside of the intended use population. It is unclear if the intended use population was determined a priori. Validation in an independent sample in the intended use population is needed. No relevant recent studies were identified for Nodify CDT or Nodify Lung. The REVEAL validation study was a retrospective study that demonstrated use as a rule-out test in conjunction with the VA Clinical Factors Model when the samples were considered inconclusive or intermediate risk by the VA model. The REVEAL model subsequently correctly identified 65% intermediate-risk samples as either low or high risk. The NPV and sensitivity were both 94%. Limitations included a small sample size and use in conjunction with just 1 type of testing model. Validation in an independent sample in the intended use population with additional probability models is needed.

### Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, more effective therapy, or avoid unnecessary therapy or testing.

### Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials (RCTs).

No evidence directly demonstrating improved outcomes in patients managed with the Xpresys Lung, Xpresys Lung 2/Nodify XL2 (BDX-XL2), or Nodify CDT tests, or the Nodify Lung testing strategy was identified.

## Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

A chain of evidence was developed, which addresses 2 key questions: (1) Does the use of a proteomic classifier with a high NPV in patients with undiagnosed pulmonary nodules detected by CT change clinical management (in this case, reduction of invasive procedures)?; and (2) Do those management changes improve outcomes relative to a clinical classifier?

### *Changes in Management*

The patient population for which a proteomic classifier with a high NPV is used is individuals with undiagnosed pulmonary nodules detected by CT.
Indirect evidence regarding Xpresys Lung version 2 suggests that 36% of invasive procedures (biopsy and/or surgery) on benign nodules could have been avoided, if the test is used in patients with a low to moderate (≤50%) pretest probability of malignancy. Three percent of malignant lesions may be missed, although these patients would be followed by CT to verify lack of progression.
One decision impact study reporting on clinical management changes, but not on outcomes after decisions for invasive procedures were made, has suggested that, in at least some cases, decisions for invasive procedures may be changed. Pritchett et al (2023) reported on the impact of the Nodify XL2 test on physician decision-making for recommending invasive procedures among patients with undiagnosed pulmonary nodules detected by CT in the ORACLE study.[10,] This propensity score matching cohort study compared patients with a low to moderate (≤50%) pretest probability of malignancy in the ORACLE prospective, multicenter, observational registry (classifier arm) to retrospective chart review of control patients treated with typical care. The results revealed that classifier testing result might reduce invasive procedure recommendations in patients diagnosed with benign disease. Of the 197 patients tested in the classifier group, 162 (82%) were benign and 35 (18%) were malignant. Patients with a benign nodule in the classifier arm were 74% less likely to undergo an invasive procedure as compared to patients in the control group (absolute difference, -14%; 95% CI, -19.5% to -7.9%; p<.001). There was 1 invasive procedure per 20 patients in the benign nodule classifier group compared to 1 invasive procedure per 5 patients in the control group (odds ratio, 0.23; 95% CI, 0.09 to 0.53; p<.001). In other words, for every 7 benign nodules tested with the Nodify XL2 test, 1 unnecessary invasive procedure was avoided. The rate of patients in the classifier group with a malignant nodule was not statistically different than the control group.

### *Improved Outcomes*

Indirect evidence suggests that use of a proteomic classifier with a high NPV has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. Compared with the standard care plan, some patients without cancer will have avoided an unnecessary invasive procedure, which is weighed against the increase in missed cancers in patients who had lung cancer but tested as negative on the proteomic classifier with a high NPV test. Whether the tradeoff between avoiding unneeded surgeries and the potential for missed cancer is worthwhile depends, in part, on patient and physician preferences. Missed malignancies would likely continue to be followed by active surveillance using low-dose CT imaging. In the context of lung cancers, overall survival depends on detection of lung cancer at early, more treatable stages. Avoiding invasive procedures in situations where patients are at a very low likelihood of having lung cancer is likely beneficial, given the known complications (e.g., pneumothorax). However, reductions in unnecessary invasive procedures must be weighed against outcomes and harms associated with a missed diagnosis of lung cancer at earlier, more treatable stages.

## Section Summary: Clinically Useful

Indirect evidence suggests that a proteomic classifier with a high NPV has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. However, stronger clinical validity data would be needed to rely on indirect evidence for

clinical utility, or long-term follow-up data would be required to determine the survival outcomes in patients with a missed diagnosis of lung cancer at earlier, more treatable stages.

### Gene Expression Profiling of Indeterminate Bronchoscopy Results

#### Clinical Context and Test Purpose

The purpose of gene expression profiling (GEP) of bronchial brushings in individuals who undergo bronchoscopy for the diagnosis of suspected lung cancer but who have an indeterminate cytology result is to stratify the clinical risk for malignancy and eliminate the need for invasive diagnostic procedures.

The following PICO was used to select literature to inform this review.

*Populations*

The relevant population of interest is individuals with undiagnosed pulmonary nodules following indeterminate bronchoscopy results for suspected lung cancer.

*Interventions*

The test being considered is GEP of bronchial brushings: Percepta Genomic Sequencing Classifier (GSC), previously Percepta Bronchial Genomic Classifier (BGC).

*Comparators*

The following practice is currently being used: standard diagnostic workup. The management of patients with suspected lung cancer who have an indeterminate bronchoscopy result is not entirely standardized. However, it is likely that, in standard practice, many patients would have a surgical biopsy, transthoracic needle aspiration, or another test, depending on the location of the nodule. In 2013, the American College of Chest Physicians recommended bronchoscopy to confirm diagnosis in patients who have suspected lung cancer with a central lesion.[11,] If bronchoscopy results are nondiagnostic and suspicion of lung cancer remains, additional testing is recommended (grade 1B recommendation).

*Outcomes*

The potential beneficial outcome of primary interest is avoiding an unneeded invasive biopsy of a nodule that would be negative for lung cancer.

Potential harmful outcomes are those resulting from false-positive or false-negative test results. False-positive test results can lead to unnecessary invasive diagnostic procedures and procedure-related complications. False-negative test results can lead to lack of pulmonary nodule surveillance or lack of appropriate invasive diagnostic procedures to diagnose malignancy.

The time frame for outcome measures varies from the short-term development of invasive diagnostic procedure-related complications to long-term procedure-related complications, development of malignancy, or overall survival.

#### Study Selection Criteria

For the evaluation of clinical validity of the plasma-based proteomic screening test, studies that met the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores).
- Included a suitable reference standard.
- Patient/sample clinical characteristics were described.
- Patient/sample selection criteria were described.

### Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

### Review of Evidence

Whitney et al (2015) reported on the development and initial validation of an RNA-based gene expression classifier from airway epithelial cells designed to be predictive of cancer in current and former smokers undergoing bronchoscopy for suspected lung cancer.[12,] Samples were from patients in the Airway Epithelium Gene Expression In the DiagnosiS of Lung Cancer (AEGIS) trials, which were 2 prospective, observational, cohort studies (AEGIS-1, AEGIS-2), for current or former smokers undergoing bronchoscopy for suspected lung cancer. Cohort details are described in Silvestri et al (2015) below. A total of 299 samples from AEGIS-1 (223 cancer-positive and 76 cancer-free subjects) were used to derive the classifier. Data from 123 patients in a prior study with a nondiagnostic bronchoscopy were used as an independent test set. In the final model, the classifier included 17 genes, patient age, and gene expression correlates and was reported as a dichotomous score (≥0.65 as cancer-positive, <0.65 as cancer-negative). The performance characteristics of the classifier in the training and test set are shown in Table 6.

Silvestri et al (2015) reported on the diagnostic performance of the gene expression classifier developed in Whitney et al (2015), in a sample of 639 patients enrolled in 2 multicenter prospective studies (AEGIS-1, N=298 patients; AEGIS-2, N=341 patients).[13,] Study characteristics are summarized in Table 5. The study enrolled patients who were undergoing clinically indicated bronchoscopy for a diagnosis of possible lung cancer and had a history of smoking. Before the bronchoscopy, the treating physician assessed each patient's probability of having cancer with a 5-level scale (<10%, 10% to 39%, 40% to 60%, 61% to 85%, >85%). Patients were followed until a diagnosis was established (either at the time of bronchoscopy or subsequently by another biopsy means) or until 12 months after bronchoscopy.

A total of 855 patients in AEGIS-1 and 502 patients in AEGIS-2 met enrollment criteria.[13,] After exclusions due to sample quality issues, loss to follow-up, lack of final diagnosis, or nonprimary lung cancer, 341 subjects were available in the validation set for AEGIS-2. For AEGIS-1, patients were randomized to the development (described above) or validation (n=298) sets. Of the 639 patients in the validation study who underwent bronchoscopy, 272 (43%; 95% CI, 39 to 46) had a nondiagnostic examination. The prevalence of lung cancer was 74% and 78% in AEGIS-1 and AEGIS-2, respectively. The overall test characteristics in AEGIS-1 and AEGIS-2 are summarized in Table 6. The classifier improved the prediction of cancer compared with bronchoscopy alone but comparisons with a clinical predictor were not reported. For the subset of 272 patients with a nondiagnostic bronchoscopy, the classifier performance was presented by the pretest physician-predicted risk of cancer. For most subpopulations, there was a very high NPV. However, there were 13 false negatives, 10 of which were considered at high risk (>60%) of cancer pre-bronchoscopy. Study limitations are summarized in Tables 7 and 8.

Vachani et al (2016) reported on rates of invasive procedures from AEGIS-1 and -2.[14,] Of 222 patients, 188 (85%) had an inconclusive bronchoscopy and follow-up procedure data available for analysis. Seventy-seven (41%) patients underwent an additional 99 invasive procedures, which included surgical lung biopsy in 40 (52%) patients. Benign and malignant diseases were ultimately diagnosed in 62 (81%) and 15 (19%) patients, respectively. Among those undergoing surgical biopsy, 20 (50%) were performed in patients with benign disease. If the classifier had been used to guide decision-making, procedures could have been avoided in 21 (50%) of 42 patients who had additional invasive testing. Further, among 35 patients with an inconclusive index bronchoscopy who were diagnosed with lung cancer, the sensitivity of the classifier was 89%, with 4 (11%) patients having a false-negative classifier result. Invasive procedures after an inconclusive bronchoscopy occur frequently, and most are performed in patients ultimately diagnosed with benign disease.

Mazzone et al (2022) conducted a prospective, multicenter, blinded, clinical validation study on individuals (N=412) who currently or formerly were smokers and were undergoing bronchoscopy for suspected lung cancer from the AEGIS-1/AEGIS-2 cohorts and the Percepta Registry.[15,] The sensitivity, specificity, and predictive values were calculated using predefined thresholds. Study characteristics and results are summarized in Tables 5 and 6, respectively. Investigators noted that Percepta GSC performance was similar between the AEGIS-1 and -2 cohorts and the Percepta Registry with an overall area under the curve of 0.73 (95% CI, 68.3 to 78.4), demonstrating the robustness of the classifier performance across different patient cohorts. Investigators also estimated the potential utility of Percepta GSC in decreasing invasive procedure utilization, had the classifier result been available to manage these lesions. It was determined that, if the classifier results were used in nodule management, 50% of patients with benign lesions and 29% of patients with malignant lesions undergoing additional invasive procedures could have avoided these procedures. Study limitations are summarized in Table 7 and Table 8.

**Table 5. Study Characteristics of Clinical Validity**

| Study | Study Population | Design | Reference Standard | Threshold for Positive Index Test | Timing of Reference and Index Tests | Blinding of Assessors | Comments |
|---|---|---|---|---|---|---|---|
| **Silvestri et al (2015)[13,]** | 639 current or former smokers undergoing bronchoscopy for suspected lung cancer (White, 76% to 78%; Black, 18% to 19%; Other, 1% to 5%) | Prospective, observational, cohort studies | Diagnosis or until 12 months after bronchoscopy | NR | Following diagnosis or 12 months | Yes | Percepta BGC<br><br>272 patients had a nondiagnostic bronchoscopy and were included in the analysis |
| **Mazzone et al (2022)[15,]** | 412 current or former smokers undergoing bronchoscopy for suspected lung cancer | Prospective, multicenter study | Diagnosis or until 12 months after bronchoscopy | NR | Following diagnosis or 12 months | Yes | Percepta GSC |

BGC: bronchial genomic classifier; GSC: genomic sequencing classifier; NR: not reported.

**Table 6. Summary of Clinical Validity Studies for Gene Expression Classifier to Predict Malignancy in Bronchial Samples**

| Study | Population | AUC (95% CI) | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) | Percepta GSC Result | Post-test NPV or PPV, % (95% CI) | % Reclassified Risk of Malignancy |
|---|---|---|---|---|---|---|---|---|---|
| **Whitney et al (2015)[12,]** | Training set, entire population (n=299) | 0.78 (0.73 to 0.82) | 93 | 57 | | | | | |
| | Training set, subset with nondiagnostic | 0.78 (0.71 to 0.85) | | | | | | | |

| Study | Population | AUC (95% CI) | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) | Percepta GSC Result | Post-test NPV or PPV, % (95% CI) | % Reclassified Risk of Malignancy |
|---|---|---|---|---|---|---|---|---|---|
| | bronchoscopy (n=134) | | | | | | | | |
| | Test set with nondiagnostic bronchoscopy (n=123) | 0.81 (0.73 to 0.88) | 92 (78 to 98) | 53 (42 to 63) | 47 (36 to 58) | 94 (83 to 99) | | | |
| **Silvestri et al (2015)[13],** | AEGIS-1 (n=298) | 0.78 (0.73 to 0.83) | 88 (83 to 95) | 47 (37 to 58) | | | | | |
| | AEGIS-2 (n=341) | 0.74 (0.68 to 0.80) | 89 (84 to 92) | 47 (36 to 59) | | | | | |
| | Subset of all patients with nondiagnostic bronchoscopy, by pretest cancer probability risk | | | | | | | | |
| | Risk <10% (n=61) | | | | 7 (1 to 24) | 100 (89 to 100) | | | |
| | Risk 10%-60% (n=84) | | | | 40 (27 to 55) | 91 (75 to 98) | | | |
| | Risk >60% (n=108) | | | | 84 (75 to 81) | 38 (15 to 65) | | | |
| | Risk unknown (n=19) | | | | 47 (21 to 73) | 100 (40 to 100) | | | |
| **Mazzone et al (2022)[15],** | Low pre-test risk of malignancy (n=80 [4 malignant, 68 benign, 8 clinical benign]); cancer prevalence 5.0% | | 57.4 (44.8 to 69.3)[a] | 100 (39.8 to 100)[b] | | | Very low | NPV: 100 (91.0 to 100) | 54.5 |
| | Intermediate pre-test risk of malignancy (n=188 [53 malignant, 102 benign, 33 clinical benign]); cancer prevalence 28.2% | | 37.3 (27.9 to 47.4)[a] | 90.6 (79.3 to 96.9)[b] | | | Low | NPV: 91.0 (80.8 to 96.0) | 29.4 |
| | | | 94.1 (87.6 to 97.8)[a] | 28.3 (16.8 to 42.3)[b] | | | High | PPV: 65.4 (43.8 to 82.1) | 12.2 |
| | High pre-test risk of malignancy (n=144 [106 malignant, 34 benign, 4 clinical | | 91.2 (76.3 to 98.1)[a] | 34.0 (25.0 to 43.8)[b] | | | Very high | PPV: 91.5 (77.9 to 97.0) | 27.3 |

| Study | Population | AUC (95% CI) | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) | Percepta GSC Result | Post-test NPV or PPV, % (95% CI) | % Reclassified Risk of Malignancy |
|---|---|---|---|---|---|---|---|---|---|
| | benign]); cancer prevalence 73.6% | | | | | | | | |

AEGIS: Airway Epithelium Gene Expression In the Diagnosis of Lung Cancer; AUC: area under the curve; CI: confidence interval; GSC: genomic sequencing classifier; NPV: negative predictive value; PPV: positive predictive value.
[a] Sensitivity is calculated on malignant patients only.
[b] Specificity is calculated on benign patients only, excluding clinical benign.

## Table 7. Study Relevance Limitations

| Study | Population[a] | Intervention[b] | Comparator[c] | Outcomes[d] | Duration of Follow-Up[e] |
|---|---|---|---|---|---|
| **Silvestri et al (2015)[13,]** | 4. Only included patients with a history of smoking | | | | |
| **Mazzone et al (2022)[15,]** | 4. Only included patients with a history of smoking | | | | 1. Follow-up only required to be 12 months to determine benign status, thus a few indolent lung cancers could have been present |

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true positives, true negatives, false positives, false negatives cannot be determined).

**Table 8. Study Design and Conduct Limitations**

| Study | Selection[a] | Blinding[b] | Delivery of Test[c] | Selective Reporting[d] | Data Completeness[e] | Statistical[f] |
|---|---|---|---|---|---|---|
| Silvestri et al (2015)[13], | | | | | 2. High number of excluded samples | |

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison to other tests not reported.

### Section Summary: Clinically Valid

Three multicenter prospective studies have provided evidence of the clinical validity of a bronchial genomic classifier in current or former cigarette smokers undergoing bronchoscopy for suspicion of lung cancer. The most recent study was a 3-cohort study that validated the second generation Percepta GSC test in an independent sample set. High sensitivity with modest specificity for the rule-out portion of the classifier, and high specificity with modest sensitivity for the rule-in portion was confirmed.

### Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, more effective therapy, or avoid unnecessary therapy or testing.

### Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No evidence directly demonstrating improved outcomes in patients managed with the Percepta GSC or BGC was identified.

### Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.
A chain of evidence was developed, which addresses 2 key questions: (1) Does the use of the Percepta GSC in individuals with indeterminate bronchoscopy results for suspected lung cancer change clinical management (in this case, reduction of invasive procedures)?; and (2) Do those management changes improve outcomes?

#### *Changes in Management*

The clinical setting in which Percepta GSC is meant to be used is not well-defined: individuals who are suspected to have cancer but who have a nondiagnostic bronchoscopy.

One decision impact study reporting on clinical management changes, but not on outcomes after decisions for invasive procedures were made, has suggested that, in at least some cases, decisions for invasive procedures may be changed. Ferguson et al (2016) reported on the impact of the Percepta BGC on physician decision-making for recommending invasive procedures among patients

with an inconclusive bronchoscopy.[16,] The results revealed that a negative (low-risk) result might reduce invasive procedure recommendations in patients diagnosed with benign disease.

Lee et al (2021) provided additional data on the effect of Percepta BCG on clinical management decisions among patients (N=283) with low or intermediate-risk lung nodules who had at least 1 year of follow-up.[17,] The availability of Percepta results led to 34.3% of patients having their risk of malignancy downgraded. Two-thirds of these patients switched from a planned invasive procedure to surveillance.

*Improved Outcomes*

Indirect evidence suggests that use of the Percepta BGC has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. Compared with the standard care plan, some patients without cancer will have avoided an unnecessary invasive procedure, which is weighed against the small increase in missed cancers in patients who had cancer but tested as negative (low-risk) on the Percepta GSC.

Whether the tradeoff between avoiding unneeded surgeries and the potential for missed cancer is worthwhile depends, in part, on patient and physician preferences. Missed malignancies would likely be continued to be followed by active surveillance by low-dose CT imaging. In the context of lung cancers, overall survival depends on the detection of lung cancer at early, more treatable stages. Avoiding invasive procedures in situations where patients are at very low likelihood of having lung cancer is likely beneficial, given the known complications (e.g., pneumothorax). However, reductions in unnecessary invasive procedures must be weighed against outcomes and harms associated with a missed diagnosis of lung cancer at earlier, more treatable stages.

**Section Summary: Clinically Useful**

Direct evidence of the clinical utility for GEP of bronchial brushings is lacking. Indirect evidence suggests that Percepta GSC has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. However, long-term follow-up data would be required to determine the survival outcomes in patients with a missed diagnosis of lung cancer at earlier, more treatable stages.

**Summary of Evidence**

For individuals with undiagnosed pulmonary nodules detected by computed tomography who receive plasma-based proteomic screening, the evidence includes prospective cohorts, retrospective studies, and prospective-retrospective studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, morbid events, hospitalizations, and resource utilization. Clinical validation studies were identified for 2 versions (Xpresys Lung, and Xpresys Lung version 2 [now Nodify XL2]) of a proteomic classifier and another lung nodule characterization test (REVEAL). The Nodify XL2 classifier has undergone substantial evolution, from a 13-protein assay to a 2-protein assay integrated with clinical factors. Because of this evolution, the most relevant studies are with the most recent version (Xpresys Lung version 2 [now Nodify XL2]). One validation study on version 2 has been identified. The classifier has been designed to have high specificity for malignant pulmonary nodules, and the validation study showed a specificity of 97% for patients with a low-to-moderate pretest probability (≤50%) of a malignant pulmonary nodule. The primary limitation of this study is that a high number of patients were excluded from the study due to incomplete clinical data or because they were subsequently determined to be outside of the intended use population. It is unclear if the intended use population was determined a priori. Validation in an independent sample in the intended use population is needed. No recent clinical validation studies were identified for the Nodify CDT test or the Nodify Lung testing strategy. The REVEAL validation study was a retrospective study that demonstrated use as a rule-out test in conjunction with the Veteran's Affairs (VA) Clinical Factors Model when the samples were considered inconclusive or intermediate risk by the VA model. The REVEAL model subsequently correctly identified 65% of intermediate-risk samples as either low or high risk. The negative predictive value and sensitivity were both 94%. Limitations

included a small sample size and use in conjunction with just 1 type of testing model. Validation in an independent sample in the intended use population with additional probability models is needed. Indirect evidence suggests that a proteomic classifier with a high negative predictive value has the potential to reduce the number of unnecessary invasive procedures to definitively diagnose benign disease versus malignancy. However, long-term follow-up data would be required to determine the survival outcomes in patients with a missed diagnosis of lung cancer at earlier, more treatable stages. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with undiagnosed pulmonary nodules following indeterminate bronchoscopy results for suspected lung cancer who receive gene expression profiling of bronchial brushings, the evidence includes multicenter prospective studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, morbid events, hospitalizations, and resource utilization. A 3-cohort, prospective, multicenter study validated the second generation Percepta Genomic Sequencing Classifier (GSC) test in an independent sample set, showing high sensitivity for the rule-out portion of the classifier and high specificity for the rule-in portion of the classifier. For intermediate pretest risk patients with an inconclusive bronchoscopy, Percepta GSC can down-classify the risk of primary lung cancer to low with a 91% negative predictive value, or up-classify the risk to high with a 65% positive predictive value. Further assessment of clinical utility is warranted. Also, where the test would fall in the clinical pathway (i.e., other than indeterminate bronchoscopy) is uncertain. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

## Supplemental Information

The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

### Practice Guidelines and Position Statements

Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

### American College of Chest Physicians

In 2013, the American College of Chest Physicians published evidence-based clinical practice guidelines on the diagnosis and management of lung cancer, including pulmonary nodules, which is discussed in the patient population parameters in the 'Plasma-Based Proteomic Screening of Pulmonary Nodules' section.[18,]

### American Thoracic Society

In 2017, the American Thoracic Society published a position statement on the evaluation of molecular biomarkers for the early detection of lung cancer.[19,] The Society states that "a clinically useful molecular biomarker applied to the evaluation of lung nodules may lead to expedited therapy for early lung cancer and/or fewer aggressive interventions in patients with benign lung nodules." To be considered clinically useful, a molecular diagnosis "must lead to earlier diagnosis of malignant nodules without substantially increasing the number of procedures performed on patients with benign nodules" or "fewer procedures for patients with benign nodules without substantially delaying the diagnosis of cancer in patients with malignant nodules."

### National Comprehensive Cancer Network

The National Comprehensive Cancer Network (NCCN) guidelines for non-small cell lung cancer, small cell lung cancer, or lung cancer screening do not mention plasma-based proteomic screening testing or gene expression profiling as a potential diagnostic or screening tool.[20,21,20,]

## U.S. Preventive Services Task Force Recommendations
Not applicable.

## Medicare National and Local Coverage
Some plans will provide limited coverage for the BDX-XL2 test (Biodesix) for the management of a lung nodule between 8 and 30 mm in diameter, in patients at least 40 years of age and with a pre-test cancer risk of 50% or less, as assessed by the Mayo Clinic Model for Solitary Pulmonary Nodules. Per Biodesix, both the Nodify XL2 and Nodify CDT tests are $0 out of pocket for covered Medicare beneficiaries.[22,]

Some plans will provide limited coverage for the PERCEPTA Bronchial Genomic Classifier (Veracyte) to identify patients with clinical low- or intermediate-risk of malignancy, after a non-diagnostic bronchoscopy, who may be followed with computed tomography surveillance in lieu of further invasive biopsies or surgery. A patient's clinical risk of malignancy may be ascertained by the McWilliams or Gould risk assessment models. Coverage does not include clinical high-risk patients or patients with known lung cancer. Per Veracyte, the PERCEPTA Genomic Sequencing Classifier test is covered by Medicare.[23,]

Local coverage guidance for California is provided by the Molecular Diagnostic Services Program (MolDX®) in the document MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy and associated Billing and Coding: MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy along with information available in the DEX® Diagnostics Exchange Registry.

MolDx will provide limited coverage for molecular tests to aid in the diagnosis or exclusion of lung cancer in a patient with an indeterminate pulmonary nodule (IPN) following a non-diagnostic bronchoscopy when ALL of the following conditions are met:

1. The beneficiary has undergone bronchoscopy for an indeterminate pulmonary nodule AND
    a. The bronchoscopy has failed to provide a specific histopathological diagnosis such that further diagnostic procedures would otherwise be considered necessary to pursue a specific diagnosis (non-diagnostic bronchoscopy); AND
    b. Test results will be used to meaningfully inform patient management within the framework of nationally recognized consensus guidelines.
    c. The nodule cannot or will not be evaluated by an alternate methodology (EBUS, FNA, etc.) for a specific diagnosis prior to receipt of molecular test results.
2. The beneficiary does NOT have any of the following:
    a. Personal history of lung cancer
    b. Current diagnosis of cancer or high clinical suspicion for cancer
    c. An overall low risk for pulmonary malignancy such that test results would not meaningfully alter patient management and significantly improve patient outcomes.
    d. An overall high risk for pulmonary malignancy such that test results would not meaningfully alter patient management and significantly improve patient outcomes.
3. The beneficiary has not been tested with the same or similar assay for the same clinical indication.
4. The beneficiary is within the population and has the indication for which the test was developed and is covered. The lab providing the test is responsible for clearly indicating to treating clinicians the population and indication for test use.
5. The test has demonstrated clinical validity and utility, establishing a clear and significant biological/molecular basis for stratifying patients and subsequently selecting (either positively or negatively) a clinical management decision in a clearly defined population.
6. Clinical validity of any analytes (or expression profiles) measured must be established through a study published in the peer-reviewed literature for the intended use of the test in the intended population.

7. Rule-out tests should have a high sensitivity and negative predictive value (NPV) such that patients can be safely selected for a less aggressive management strategy without delay to diagnosis due to false negative results.
8. Rule-in tests should have a high specificity and positive predictive value (PPV) such that patients can be safely selected for more aggressive management without significantly increasing procedures in patients without cancer due to false positive results.
9. The test demonstrates analytical validity including both analytical and clinical validations. If the test relies on an algorithm (which may range in complexity from a threshold determination of a single numeric value to a complex mathematical or computational function), the algorithm must be validated in a cohort that is not a development cohort for the algorithm.
10. Tests utilizing a similar methodology or evaluating a similar molecular analyte to a test for which there is a generally accepted testing standard or for which existing coverage exists must demonstrate equivalent or superior test performance (i.e., sensitivity and/or specificity) when used for the same indication in the same intended-use population. New tests that become available with significantly improved performance may render older tests no longer compliant with this policy.
11. The test successfully completes a Molecular Diagnostic Services Program (MolDX®) technical assessment that ensures the test is reasonable and necessary as described above.

According the DEX® Diagnostics Exchange registry, MolDx has made a positive coverage determination the following tests tests to aid in the diagnosis or exclusion of lung cancer in a patient with an indeterminate pulmonary nodule (IPN) following a non-diagnostic bronchoscopy:

- Percepta® Bronchial Genomic Classifier (Veracyte Inc)
- BDX-XL2 or Nodify XL2 (Biodesix, Inc.) [0080U]
- Nodify CDT (Biodesix, Inc.) [0360U]

The Reveal Lung Nodule Characterization (MagArray, Inc.) [0092U] is listed as non-covered by MolDx on the DEX® Diagnostics Exchange.

### Ongoing and Unpublished Clinical Trials

Some currently ongoing trials that might influence this review are listed in Table 9.

**Table 9. Summary of Key Trials**

| NCT No. | Trial Name | Planned Enrollment | Completion Date |
|---|---|---|---|
| **Ongoing** | | | |
| **NCT04171492[a]** | A Multicenter, Randomized Controlled Trial, Prospectively Evaluating the Clinical Utility of the Nodify XL2 Proteomic Classifier in Incidentally Discovered Low to Moderate Risk Lung Nodules | 2000 | Dec 2026 |
| **Unpublished** | | | |
| **NCT03766958[a]** | An Observational Registry Study to Evaluate the Performance of the BDX-XL2 Test | 842 | May 2024 |

NCT: national clinical trial.
[a] Denotes industry-sponsored or cosponsored trial.

## References

1. Gould MK, Donington J, Lynch WR, et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. May 2013; 143(5 Suppl): e93S-e120S. PMID 23649456

2. Li XJ, Hayward C, Fong PY, et al. A blood-based proteomic classifier for the molecular characterization of pulmonary nodules. Sci Transl Med. Oct 16 2013; 5(207): 207ra142. PMID 24132637
3. Vachani A, Pass HI, Rom WN, et al. Validation of a multiprotein plasma classifier to identify benign lung nodules. J Thorac Oncol. Apr 2015; 10(4): 629-37. PMID 25590604
4. Vachani A, Hammoud Z, Springmeyer S, et al. Clinical Utility of a Plasma Protein Classifier for Indeterminate Lung Nodules. Lung. Dec 2015; 193(6): 1023-7. PMID 26376647
5. Kearney P, Hunsucker SW, Li XJ, et al. An integrated risk predictor for pulmonary nodules. PLoS One. 2017; 12(5): e0177635. PMID 28545097
6. Silvestri GA, Tanner NT, Kearney P, et al. Assessment of Plasma Proteomics Biomarker's Ability to Distinguish Benign from Malignant Lung Nodules: Results of the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) Trial. Chest. Sep 2018; 154(3): 491-500. PMID 29496499
7. Tanner NT, Springmeyer SC, Porter A, et al. Assessment of Integrated Classifier's Ability to Distinguish Benign from Malignant Lung Nodules: Extended Analyses and 2-Year Follow-Up Results of the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) Trial. Chest. Mar 2021; 159(3): 1283-1287. PMID 33171158
8. Vachani A, Pass HI, Rom Wn, et al. Validation of a multiprotein plasma classifier to identify benign lung nodules. Supplement. J Thorac Oncol. April 2015;10(4):629-637. https://cdn-links.lww.com/permalink/jto/a/jto_10_4_2015_01_07_massion_jto-d-14-00912_sdc1.pdf. Accessed March 27, 2025.
9. Trivedi NN, Arjomandi M, Brown JK, et al. Risk assessment for indeterminate pulmonary nodules using a novel, plasma-protein based biomarker assay. Biomed Res Clin Pract. Dec 2018; 3(4). PMID 32913898
10. Pritchett MA, Sigal B, Bowling MR, et al. Assessing a biomarker's ability to reduce invasive procedures in patients with benign lung nodules: Results from the ORACLE study. PLoS One. 2023; 18(7): e0287409. PMID 37432960
11. Rivera MP, Mehta AC, Wahidi MM. Establishing the diagnosis of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. May 2013; 143(5 Suppl): e142S-e165S. PMID 23649436
12. Whitney DH, Elashoff MR, Porta-Smith K, et al. Derivation of a bronchial genomic classifier for lung cancer in a prospective study of patients undergoing diagnostic bronchoscopy. BMC Med Genomics. May 06 2015; 8: 18. PMID 25944280
13. Silvestri GA, Vachani A, Whitney D, et al. A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer. N Engl J Med. Jul 16 2015; 373(3): 243-51. PMID 25981554
14. Vachani A, Whitney DH, Parsons EC, et al. Clinical Utility of a Bronchial Genomic Classifier in Patients with Suspected Lung Cancer. Chest. Jul 2016; 150(1): 210-8. PMID 26896702
15. Mazzone P, Dotson T, Wahidi MM, et al. Clinical validation and utility of Percepta GSC for the evaluation of lung cancer. PLoS One. 2022; 17(7): e0268567. PMID 35830375
16. Ferguson JS, Van Wert R, Choi Y, et al. Impact of a bronchial genomic classifier on clinical decision making in patients undergoing diagnostic evaluation for lung cancer. BMC Pulm Med. May 17 2016; 16(1): 66. PMID 27184093
17. Lee HJ, Mazzone P, Feller-Kopman D, et al. Impact of the Percepta Genomic Classifier on Clinical Management Decisions in a Multicenter Prospective Study. Chest. Jan 2021; 159(1): 401-412. PMID 32758562
18. Detterbeck FC, Lewis SZ, Diekemper R, et al. Executive Summary: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. May 2013; 143(5 Suppl): 7S-37S. PMID 23649434
19. Mazzone PJ, Sears CR, Arenberg DA, et al. Evaluating Molecular Biomarkers for the Early Detection of Lung Cancer: When Is a Biomarker Ready for Clinical Use? An Official American Thoracic Society Policy Statement. Am J Respir Crit Care Med. Oct 01 2017; 196(7): e15-e29. PMID 28960111

20. National Comprehensive Cancer Network. NCCN Guidelines Version 3.2025: Non-Small Cell Lung Cancer. 2024; https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed March 26, 2025.
21. National Comprehensive Cancer Network. NCCN Guidelines Version 4.2025: Small Cell Lung Cancer. 2023; https://www.nccn.org/professionals/physician_gls/pdf/sclc.pdf. Accessed March 25, 2025.
22. Biodesix. Nodify Lung: Lung Nodule Management. 2024; https://www.biodesix.com/our-tests/nodify-lung. Accessed March 27, 2025.
23. Veracyte. Percepta Genomic Sequencing Classifier for your patients. 2024; https://lung.veracyte.com/percepta-gsc/for-your-patients/. Accessed March 27, 2025.
24. Centers for Medicare & Medicaid Services. (2024, September 22). Local Coverage Determination (LCD): MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy (L39680). U.S. Department of Health & Human Services. https://www.cms.gov/medicare-coverage-database/view/lcd.aspx?lcdid=39680&ver=4. Accessed October 15, 2025.
25. Centers for Medicare & Medicaid Services. (2024, September 22). Local Coverage Determination (LCD) Reference Article: Billing and Coding: MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy (A59507). U.S. Department of Health & Human Services. https://www.cms.gov/medicare-coverage-database/view/article.aspx?articleId=59507&ver=4. Accessed October 15, 2025.
26. Palmetto GBA. (n.d.). Welcome to DEX® Diagnostics Exchange Registry. Login page. https://app.dexzcodes.com/login#!newSplashPage

## Documentation for Clinical Review

- No records required

## Coding

*The list of codes in this Medical Policy is intended as a general reference and may not cover all codes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy.*

| Type | Code | Description |
|---|---|---|
| CPT® | 0080U | Oncology (lung), mass spectrometric analysis of galectin-3-binding protein and scavenger receptor cysteine-rich type 1 protein M130, with five clinical risk factors (age, smoking status, nodule diameter, nodule-spiculation status and nodule location), utilizing plasma, algorithm reported as a categorical probability of malignancy |
| | 0092U | Oncology (lung), three protein biomarkers, immunoassay using magnetic nanosensor technology, plasma, algorithm reported as risk score for likelihood of malignancy |
| | 0360U | Oncology (lung), enzyme-linked immunosorbent assay (ELISA) of 7 autoantibodies (p53, NY-ESO-1, CAGE, GBU4-5, SOX2, MAGE A4, and HuD), plasma, algorithm reported as a categorical result for risk of malignancy |
| | 81554 | Pulmonary disease (idiopathic pulmonary fibrosis [IPF]), mRNA, gene expression analysis of 190 genes, utilizing transbronchial biopsies, diagnostic algorithm reported as categorical result (e.g., positive or negative for high probability of usual interstitial pneumonia [UIP]) |
| | 83520 | Immunoassay for analyte other than infectious agent antibody or infectious agent antigen; quantitative, not otherwise specified |
| | 84999 | Unlisted chemistry procedure |

| Type | Code | Description |
|---|---|---|
| HCPCS | None | |

## Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

| Effective Date | Action |
|---|---|
| 07/01/2017 | BCBSA Medical Policy adoption |
| 07/01/2018 | Policy revision without position change |
| 02/01/2019 | Coding update |
| 07/01/2019 | Policy revision without position change<br>Coding update |
| 07/01/2020 | Annual review. No change to policy statement. Literature review updated. |
| 07/01/2021 | Annual review. Literature review updated. Policy statement, policy guidelines and literature updated. Coding update. |
| 07/01/2022 | Annual review. Policy statement, guidelines and literature updated. |
| 11/01/2025 | Policy reactivated. Policy archived from 08/01/2023 to 10/31/2025. |

## Definitions of Decision Determinations

**Healthcare Services**: For the purpose of this Medical Policy, Healthcare Services means procedures, treatments, supplies, devices, and equipment.

**Medically Necessary:** Healthcare Services that are Medically Necessary include only those which have been established as safe and effective, are furnished under generally accepted professional standards to treat illness, injury or medical condition, and which, as determined by Blue Shield of California, are: (a) consistent with Blue Shield of California medical policy; (b) consistent with the symptoms or diagnosis; (c) not furnished primarily for the convenience of the patient, the attending Physician or other provider; (d) furnished at the most appropriate level which can be provided safely and effectively to the member; and (e) not more costly than an alternative service or sequence of services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of the member's illness, injury, or disease.

**Investigational or Experimental:** Healthcare Services which do not meet ALL of the following five (5) elements are considered investigational or experimental:

A. The technology must have final approval from the appropriate government regulatory bodies.
   - This criterion applies to drugs, biological products, devices and any other product or procedure that must have final approval to market from the U.S. Food and Drug Administration ("FDA") or any other federal governmental body with authority to regulate the use of the technology.
   - Any approval that is granted as an interim step in the FDA's or any other federal governmental body's regulatory process is not sufficient.
   - The indications for which the technology is approved need not be the same as those which Blue Shield of California is evaluating.

B. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.
   - The evidence should consist of well-designed and well-conducted investigations published in peer-reviewed journals. The quality of the body of studies and the consistency of the results are considered in evaluating the evidence.

- The evidence should demonstrate that the technology can measure or alter the physiological changes related to a disease, injury, illness, or condition. In addition, there should be evidence, or a convincing argument based on established medical facts that such measurement or alteration affects health outcomes.

C. The technology must improve the net health outcome.
   - The technology's beneficial effects on health outcomes should outweigh any harmful effects on health outcomes.

D. The technology must be as beneficial as any established alternatives.
   - The technology should improve the net health outcome as much as, or more than, established alternatives.

E. The improvement must be attainable outside the investigational setting.
   - When used under the usual conditions of medical practice, the technology should be reasonably expected to satisfy Criteria C and D.

## Feedback

Blue Shield of California is interested in receiving feedback relative to developing, adopting, and reviewing criteria for medical policy. Any licensed practitioner who is contracted with Blue Shield of California or Blue Shield of California Promise Health Plan is welcome to provide comments, suggestions, or concerns. Our internal policy committees will receive and take your comments into consideration. Our medical policies are available to view or download at www.blueshieldca.com/provider.

For medical policy feedback, please send comments to: MedPolicy@blueshieldca.com

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department at (800) 541-6652, or the Transplant Case Management Department at (800) 637-2066 ext. 3507708 or visit the provider portal at www.blueshieldca.com/provider.

*Disclaimer: Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as member health services contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member health services contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.*

## Appendix A

| POLICY STATEMENT | |
|---|---|
| **BEFORE** | **AFTER**<br>**Blue font: Verbiage Changes/Additions** |
| **Reactivated Policy**<br><br>**Policy Statement:**<br>N/A | **Molecular Testing in the Management of Pulmonary Nodules 2.04.142**<br><br>**Policy Statement:**<br>I. Plasma-based proteomic screening, including but not limited to Nodify XL2® (BDX-XL2), Nodify CDT®, and REVEAL Lung Nodule Characterization (MagArray), in individuals with undiagnosed pulmonary nodules detected by computed tomography is considered **investigational**.<br><br>II. Gene expression profiling on bronchial brushings, including but not limited to the Percepta® Genomic Sequencing Classifier, in individuals with indeterminate bronchoscopy results from undiagnosed pulmonary nodules is considered **investigational**.<br><br>Note: For individuals enrolled in health plans subject to the Biomarker Testing Law (Health & Safety Code Section 1367.667 and the Insurance Code Section 10123.209), Centers for Medicare & Medicaid Services (CMS) Local Coverage Determination (LCD) may also apply. Please refer to the Medicare National and Local Coverage section of this policy and MolDX: Molecular Biomarkers for Risk Stratification of Indeterminate Pulmonary Nodules Following Bronchoscopy for reference. |