2.04.142 Molecular Testing in the Management of Pulmonary Nodules

<table>
<thead>
<tr>
<th>Original Policy Date:</th>
<th>July 1, 2017</th>
<th>Effective Date:</th>
<th>July 1, 2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section:</td>
<td>2.0 Medicine</td>
<td>Page:</td>
<td>Page 1 of 17</td>
</tr>
</tbody>
</table>

**Policy Statement**

Plasma-based proteomic screening, including but not limited to BDX-XL2 (Nodify XL2), in patients with undiagnosed pulmonary nodules detected by computed tomography is considered investigational.

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

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

**Policy Guidelines**

**Coding**

Effective January 1, 2021, the following CPT code may be reported for this testing:

- **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])

These tests would likely be reported with nonspecific codes such as:

- **83520:** Immunoassay for analyte other than infectious agent antibody or infectious agent antigen; quantitative, not otherwise specified
- **84999:** Unlisted chemistry procedure

The following PLA code is specific to the BDX-XL2 test by Biodesix®, Inc.:

- **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

**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.

**Related Policies**

- N/A

**Benefit Application**

Benefit determinations should be based in all cases on the applicable contract language. To the extent there are any conflicts between these guidelines and the 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 mandates (e.g., Federal Employee Program [FEP]) prohibits plans from
denying Food and Drug Administration (FDA)-approved technologies as investigational. In these
instances, plans may have to consider the coverage eligibility of FDA-approved technologies on
the basis of medical necessity alone.

**Regulatory Status**

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 (BDX-XL2 [Integrated Diagnostics
[Indi], purchased by Biodesix) and Percepta® Bronchial Genomic 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 this test.

**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.
Xpresys Lung and BDX-XL2 are plasma-based proteomic screening tests that measure the
relative abundance of proteins from multiple disease pathways associated with lung cancer
using an analytic technique called multiple reaction monitoring mass spectrometry. The role of
the tests is to aid physicians in differentiating likely benign from likely malignant nodules. If the
test yields a likely benign result, patients may choose active surveillance via serial CT scans to
monitor the pulmonary nodule. If the test yields a likely malignant result, invasive diagnostic procedures would be indicated. The test is therefore only used in the management of pulmonary nodules to rule in or out invasive diagnostic procedures and does not diagnose lung cancer.

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 of clinical samples to assess for the risk for malignancy.

Gene Expression Profiling for an Indeterminate Bronchoscopy Result
The Percepta Bronchial Genomic Classifier is a 23-gene, GEP test that analyzes genomic changes in the airways of current or former smokers to assess a patient's risk of having lung cancer, without the direct testing of a pulmonary nodule. 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 patients with undiagnosed pulmonary nodule(s) is to stratify clinical risk for malignancy and eliminate or necessitate the need for invasive diagnostic procedures.

The relevant question addressed in this evidence review is: Does plasma-based proteomic screening appropriately eliminate or necessitate the need for invasive diagnostic procedures and lead to improved net health outcomes in patients with undiagnosed pulmonary nodules detected by computed tomography (CT)?

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

Populations
The relevant population of interest is individuals with undiagnosed pulmonary nodules. 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. Therefore, additional information in the segment of patients with an indeterminate lung nodule, 8 to 30 mm in diameter, would be particularly useful.
Interventions
The test being considered is plasma-based proteomic screening. Of particular focus is the Xpresys Lung test 2 (BDX-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 posttest probability of a lung nodule being benign.

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
Several studies were identified that reported on the development and validation of a plasma-based classifier test to predict malignancy (Xpresys Lung, and Xpresys Lung 2 [BDX-XL2]). Li et al (2013) reported on an initial development that was based on a 13-protein plasma classifier.

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). 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 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. 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.

Keamey 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. 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, which includes LG3BP, C163A, and clinical risk factors.

Silvestri et al (2018) reported the validation of the Xpresys Lung version 2 (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%. 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%) was misclassified as likely benign. Of the 149 benign nodules in the study, 44% were correctly classified as likely benign. For 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. Additionally, the area under the curve of the integrated classifier was 0.76 (95% CI, 0.69-0.82), which outperformed the physician pretest probability for malignancy (0.69; 95% CI, 0.62-0.76) and the Mayo (0.69; 95% CI, 0.62-0.76), Veterans Administration (0.6; 95% CI, 0.53-0.67), and Brock (0.71; 95% CI, 0.63-0.77) models in the lower risk pretest probability (≤50%) cohort.

<table>
<thead>
<tr>
<th>Study</th>
<th>Study Population</th>
<th>Design</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vachani et al (2015)</td>
<td>141 samples associated with indeterminate pulmonary nodules</td>
<td>Retrospective analysis with existing samples</td>
<td>Selected to keep NPV of 90%</td>
<td>Yes</td>
<td>Xpresys Lung</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silvestri et al (2018)</td>
<td>178 patients with 8 to 30 mm lung nodules and</td>
<td>Prospective multicenter observational</td>
<td>Definitive diagnosis, nodule resolution, or</td>
<td>NR</td>
<td>Retrospective evaluation of performance</td>
<td>Yes</td>
<td>Xpresys Lung version 2</td>
</tr>
</tbody>
</table>

Table 1. Study Characteristics of Clinical Validity
Table 2. Summary of Diagnostic Performance Studies for Proteomic Tests to Predict Malignancy

<table>
<thead>
<tr>
<th>Study</th>
<th>Prevalence, %</th>
<th>Reference Value</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, %</th>
<th>NPV, %</th>
<th>PPV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vachani et al (2015)</td>
<td>23.1</td>
<td>0.47</td>
<td>69.5 (NR)</td>
<td>48.0 (NR)</td>
<td>84.0</td>
<td>28.6</td>
</tr>
<tr>
<td></td>
<td>23.1</td>
<td>0.39</td>
<td>83.8 (NR)</td>
<td>32.3 (NR)</td>
<td>86.9</td>
<td>27.1</td>
</tr>
<tr>
<td></td>
<td>23.1</td>
<td>0.36</td>
<td>82.1 (NR)</td>
<td>20.4 (NR)</td>
<td>89.6</td>
<td>25.8</td>
</tr>
<tr>
<td>Silvestri et al (2018)</td>
<td>16.3</td>
<td>NR</td>
<td>97 (82 to 100)</td>
<td>44 (36 to 52)</td>
<td>98 (92 to 100)</td>
<td>NR</td>
</tr>
</tbody>
</table>

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

Table 3. Study Relevance Limitations

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes</th>
<th>Duration of Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvestri et al (2018)</td>
<td>4. unclear</td>
<td>4. The enrolled patients included those who were outside of intended use.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

Table 4. Study Design and Conduct Limitations

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Blinding</th>
<th>Delivery of Test</th>
<th>Selective Reporting</th>
<th>Data Completeness</th>
<th>Statistical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvestri et al (2018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Data were collected but not...
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.
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

Clinical validation studies were identified for 2 versions of a proteomic classifier. This 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 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.

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 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 negative predictive value (NPV) in patients with undiagnosed pulmonary nodules detected by CT change clinical management (in this case, reduction of invasive procedures)? (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 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.

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 continued 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.

Gene Expression Profiling of Indeterminate Bronchoscopy Results

Clinical Context and Test Purpose

The purpose of gene expression profiling (GEP) of bronchial brushings in patients 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 relevant question addressed in this evidence review is: Does GEP of bronchial brushings reduce the need for invasive diagnostic procedures and lead to improved net health outcomes in patients with undiagnosed pulmonary nodules following indeterminate bronchoscopy results for suspected lung cancer?

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.
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. If bronchoscopy results are nondiagnostic and a suspicion of lung cancer remains, additional testing is recommended (grade 1B recommendation).

Outcomes
The potential beneficial outcomes of primary interest are 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 a 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. 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). 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%, 10-39%, 40-60%, 61-85%, 85-100%).
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. 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 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.

Table 5. Study Characteristics of Clinical Validity

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Design</th>
<th>Reference Standard</th>
<th>Threshold for Positive Index Test</th>
<th>Timing of Reference and Index Tests</th>
<th>Blinding of Assessors</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvestri et al (2015)</td>
<td>639 current or former smokers undergoing bronchoscopy for suspected lung cancer</td>
<td>Prospective, observational, cohort studies</td>
<td>Diagnosis or until 12 months after bronchoscopy</td>
<td>NR</td>
<td>Following diagnosis or 12 months</td>
<td>Yes</td>
<td>272 patients had a nondiagnostic bronchoscopy and were included in the analysis</td>
</tr>
<tr>
<td>Silvestri et al (2015)</td>
<td>AEGIS-1 (n=298)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AEGIS-2 (n=341)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Summary of Clinical Validity Studies for GEC to Predict Malignancy in Bronchial Samples

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>AUC (95% CI)</th>
<th>Sensitivity, % (95% CI)</th>
<th>Specificity, % (95% CI)</th>
<th>PPV, % (95% CI)</th>
<th>NPV, % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whitney et al (2015)</td>
<td>Training set, entire population (n=299)</td>
<td>0.78 (0.73 to 0.82)</td>
<td>93 (83 to 95)</td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training set, subset with nondiagnostic bronchoscopy (n=134)</td>
<td>0.78 (0.71 to 0.85)</td>
<td>92 (78 to 95)</td>
<td>53 (42 to 63)</td>
<td>47 (36 to 58)</td>
<td>94 (83 to 99)</td>
</tr>
<tr>
<td></td>
<td>Test set with nondiagnostic bronchoscopy (n=123)</td>
<td>0.81 (0.73 to 0.88)</td>
<td>92 (78 to 95)</td>
<td>53 (42 to 63)</td>
<td>47 (36 to 58)</td>
<td>94 (83 to 99)</td>
</tr>
<tr>
<td>Silvestri et al (2015)</td>
<td>AEGIS-1 (n=298)</td>
<td>0.78 (0.73 to 0.83)</td>
<td>88 (83 to 95)</td>
<td>47 (37 to 58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AEGIS-2 (n=341)</td>
<td>0.74 (0.68 to 0.80)</td>
<td>89 (84 to 92)</td>
<td>47 (36 to 59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subset of all patients with nondiagnostic bronchoscopy, by pretest cancer probability risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk &lt;10% (n=61)</td>
<td>7 (1 to 24)</td>
<td>100 (91 to 100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk 10%–60% (n=84)</td>
<td>40 (27 to 55)</td>
<td>91 (75 to 98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk &gt;60% (n=108)</td>
<td>84 (75 to 81)</td>
<td>38 (15 to 65)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk unknown (n=19)</td>
<td>47 (21 to 73)</td>
<td>100 (40 to 100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AUC: area under the curve; CI: confidence interval; GEC: gene expression classifier; NPV: negative predictive value; PPV: positive predictive value.
Table 7. Study Relevance Limitations

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes</th>
<th>Duration of Follow-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvestri et al (2015)</td>
<td>4. Only included patients with a history of smoking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Blinding</th>
<th>Delivery of Test</th>
<th>Selective Reporting</th>
<th>Data Completeness</th>
<th>Statistical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silvestri et al (2015)</td>
<td>11, 2. High number of excluded samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

Vachani et al (2016) reported on rates of invasive procedures from AEGIS-1 and -2. 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.

Section Summary: Clinically Valid

Two 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. For patients with an intermediate risk of lung cancer with a nondiagnostic bronchoscopic examination, the NPV was 91%. However, there has been limited replication outside of a single trial group.
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 Bronchial Genomic Classifier (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 use of the Percepta BGC in individuals with indeterminate bronchoscopy results for suspected lung cancer change clinical management (in this case, reduction of invasive procedures)? (2) Do those management changes improve outcomes?

Changes in Management
The clinical setting in which Percepta BGC 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. The results revealed that a negative (low-risk) result might reduce invasive procedure recommendations in patients diagnosed with benign disease.

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 BGC.

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 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 BGC 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 CT who receive plasma-based proteomic screening, the evidence includes prospective cohorts 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 of a proteomic classifier. This 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 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. 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. Reported receiver operating characteristic curve values ranged from 0.74 to 0.81, with a negative predictive value of 91%. Among patients with a low and intermediate pretest probability of cancer with an inconclusive bronchoscopy, 77 (85%) patients underwent invasive diagnostic procedures. However, there was a relatively high number of missed cancers. No validation of the test in other populations was identified. 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.

- **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. 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."

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

Medicare National Coverage
MolDX 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.

Ongoing and Unpublished Clinical Trials
A search of ClinicalTrials.gov in March 2021 did not identify any ongoing or unpublished trials that would likely influence this review.

References


Reproduction without authorization from Blue Shield of California is prohibited


Documentation for Clinical Review

- No records required

Coding

This Policy relates only to the services or supplies described herein. Benefits may vary according to product design; therefore, contract language should be reviewed before applying the terms of the Policy.

The following codes are included below for informational purposes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy. Policy Statements are intended to provide member coverage information and may include the use of some codes for clarity. The Policy Guidelines section may also provide additional information for how to interpret the Policy Statements and to provide coding guidance in some cases.

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT®</td>
<td>0080U</td>
<td>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</td>
</tr>
<tr>
<td></td>
<td>0092U</td>
<td>Oncology (lung), three protein biomarkers, immunoassay using magnetic nanosensor technology, plasma, algorithm reported as risk score for likelihood of malignancy</td>
</tr>
<tr>
<td></td>
<td>81554</td>
<td>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])</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>(Code effective 1/1/2021)</strong></td>
</tr>
<tr>
<td></td>
<td>83520</td>
<td>Immunoassay for analyte other than infectious agent antibody or infectious agent antigen; quantitative, not otherwise specified</td>
</tr>
<tr>
<td></td>
<td>84999</td>
<td>Unlisted chemistry procedure</td>
</tr>
</tbody>
</table>

HCPCS: None
Policy History

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

<table>
<thead>
<tr>
<th>Effective Date</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>07/01/2017</td>
<td>BCBSA Medical Policy adoption</td>
</tr>
<tr>
<td>07/01/2018</td>
<td>Policy revision without position change</td>
</tr>
<tr>
<td>02/01/2019</td>
<td>Coding update</td>
</tr>
<tr>
<td>07/01/2019</td>
<td>Policy revision without position change</td>
</tr>
<tr>
<td></td>
<td>Coding update</td>
</tr>
<tr>
<td>07/01/2020</td>
<td>Annual review. No change to policy statement. Literature review updated.</td>
</tr>
</tbody>
</table>

Definitions of Decision Determinations

Medically Necessary: 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, are: (a) consistent with Blue Shield 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 patient; 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/Experimental: A treatment, procedure, or drug is investigational when it has not been recognized as safe and effective for use in treating the particular condition in accordance with generally accepted professional medical standards. This includes services where approval by the federal or state governmental is required prior to use, but has not yet been granted.

Split Evaluation: Blue Shield of California/Blue Shield of California Life & Health Insurance Company (Blue Shield) policy review can result in a split evaluation, where a treatment, procedure, or drug will be considered to be investigational for certain indications or conditions, but will be deemed safe and effective for other indications or conditions, and therefore potentially medically necessary in those instances.

Prior Authorization Requirements (as applicable to your plan)

Within five days before the actual date of service, the provider must confirm with Blue Shield that the member's health plan coverage is still in effect. Blue Shield reserves the right to revoke an authorization prior to services being rendered based on cancellation of the member's eligibility. Final determination of benefits will be made after review of the claim for limitations or exclusions.

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: This medical policy is a guide in evaluating the medical necessity of a particular service or treatment. 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 contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.
### Appendix A

#### POLICY STATEMENT

<table>
<thead>
<tr>
<th>BEFORE</th>
<th>AFTER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Molecular Testing in the Management of Pulmonary Nodules 2.04.142</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Policy Statement:</strong> Plasma-based proteomic screening, including but not limited to BDX-XL2, in patients with undiagnosed pulmonary nodules detected by computed tomography is considered <strong>investigational</strong>.</td>
<td><strong>Policy Statement:</strong> Plasma-based proteomic screening, including but not limited to BDX-XL2 (Nodify XL2), in patients with undiagnosed pulmonary nodules detected by computed tomography is considered <strong>investigational</strong>.</td>
</tr>
<tr>
<td>Gene expression profiling on bronchial brushings, including but not limited to Percepta® Bronchial Genomic Classifier, in patients with indeterminate bronchoscopy results from undiagnosed pulmonary nodules is considered <strong>investigational</strong>.</td>
<td>Gene expression profiling on bronchial brushings, including but not limited to the Percepta® Bronchial Genomic Classifier, in patients with indeterminate bronchoscopy results from undiagnosed pulmonary nodules is considered <strong>investigational</strong>.</td>
</tr>
</tbody>
</table>