

2.04.146	Gene Expression Profiling for Cutaneous Melanoma		
Original Policy Date:	July 1, 2018	Effective Date:	July 1, 2018
Section:	2.0 Medicine	Page:	Page 1 of 26

Policy Statement

Gene expression testing, including but not limited to the Pigmented Lesion Assay, in the evaluation of patients with suspicious pigmented lesions is considered **investigational**.

Gene expression testing, including but not limited to the myPath Melanoma test, in the evaluation of patients with melanocytic lesions with indeterminate histopathologic features is considered **investigational**.

Gene expression testing, including but not limited to DecisionDx-Melanoma, in the evaluation of patients with cutaneous melanoma is considered **investigational** for all indications.

Policy Guidelines

Genetic Counseling

Experts recommend formal genetic counseling for patients who are at risk for inherited disorders and who wish to undergo genetic testing. Interpreting the results of genetic tests and understanding risk factors can be difficult for some patients; genetic counseling helps individuals understand the impact of genetic testing, including the possible effects the test results could have on the individual or their family members. It should be noted that genetic counseling may alter the utilization of genetic testing substantially and may reduce inappropriate testing; further, genetic counseling should be performed by an individual with experience and expertise in genetic medicine and genetic testing methods.

Coding

There are currently no specific codes for the panel tests which are applicable to this policy. There are two genes which are included in the DermTech PLA test which are included in:

- **81401:** Molecular Pathology Procedure Level 2 (e.g., *LINC00518* and *PRAME*)

81401 does not include the remaining genes however, so this code does not represent "DermTech PLA", "MyPath Melanoma" or "DecisionDx UM Melanoma". Providers will most likely bill one of the following miscellaneous codes for these types of tests:

- **81479:** Unlisted molecular pathology procedure
- **81599:** Unlisted multianalyte assay with algorithmic analysis
- **84999:** Unlisted chemistry procedure

Description

Laboratory tests have been developed that detect the expression of different genes in pigmented lesions or melanoma tumor tissue. Test results may help providers and patients decide whether to biopsy suspicious pigmented lesions, aid in diagnosis lesions with indeterminate histopathologic lesions or determine whether to perform sentinel lymph node biopsy in patients diagnosed with stage I or II cutaneous melanoma. This report summarizes the evidence of 3 tests and is organized by indication.

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. The Pigmented Lesion Assay, myPath Melanoma, and DecisionDx-Melanoma tests are available under the auspices of the Clinical Laboratory Improvement Amendments. Laboratories that offer laboratory-developed tests must be licensed by the Clinical Laboratory Improvement Amendments for high-complexity testing. To date, the U.S. Food and Drug Administration has chosen not to require any regulatory review of this test.

Rationale

Background Cutaneous Melanoma

Cutaneous melanoma accounts for more than 90% of cases of melanoma.¹ For many decades, melanoma incidence was rapidly increasing in the United States. However, recent estimates have suggested the rise may be slowing. In 2018, more than 90,000 new cases of melanoma are expected to be diagnosed, and more than 9000 people are expected to die of melanoma.²

Risk Factors

Exposure to solar ultraviolet radiation is a major risk factor for melanoma. Most melanomas occur on the sun-exposed skin, particularly those areas most susceptible to sunburn. Likewise, features that are associated with an individual's sensitivity to sunlight, such as light skin pigmentation, red or blond hair, blue or green eyes, freckling tendency, and poor tanning ability are well-known risk factors for melanoma.^{3,4} There is also a strong association between high total body nevus counts and melanoma.⁵

Several genes appear to contribute to melanoma predisposition such as tumor suppressor gene *CDKN2A*, melanocortin-1 receptor (*MC1R*) gene, and *BAP1* variants.⁶⁻⁸ Individuals with either familial or sporadic melanoma have a 2 to 3 times increased risk of developing a subsequent primary melanoma.⁹ Several occupational exposures and lifestyle factors, such as body mass index and smoking, have been evaluated as possible risk factors for melanoma.¹⁰

Diagnosis

Primary care providers evaluate suspicious pigmented lesions to determine who should be referred to dermatology. Factors considered include both a patient's risk for melanoma as well as a visual examination of the lesion. The visual examination assesses whether the lesion has features suggestive of melanoma.

Criteria for features suggestive of melanoma have been developed. One checklist is the ABCDE checklist¹¹:

- **Asymmetry**

- Border irregularities
- Color variegation
- Diameter ≥ 6 mm
- Evolution

Another criteria commonly used is the “ugly duckling” sign.¹² An ugly duckling is a nevus that is obviously different from others in a given patient. Primary care providers generally have a low threshold for referral to dermatology.

Melanoma is difficult to diagnose based on visual examination, and the criterion standard for diagnosis is histopathology. There is a low threshold for excisional biopsy of suspicious lesions for histopathologic examination due to the procedure’s ease and low risk as well as the high probability of missing melanoma. However, the yield of biopsy is fairly low. The number of biopsies performed to yield one melanoma diagnosis has been estimated to be about 15 for U.S. dermatologists.¹³ Therefore a test that could accurately identify those lesions not needing a biopsy (i.e., a rule-out test for biopsy) could be clinically useful.

Treatment and Surveillance

Many treatments and surveillance decisions are determined by a patient’s prognostic stage group based the American Joint Committee on Cancer tumor, node, metastasis staging system.¹⁴ The prognostic groups are as follows: stage I, T1a through T2a primary melanomas without evidence of regional or distant metastases; stage II, T2b through T4b primary melanomas without evidence of lymphatic disease or distant metastases; stage III: pathologically documented involvement of regional lymph nodes or in transit or satellite metastases (N1 to N3); stage IV: distant metastases.

Patients may also undergo sentinel lymph node biopsy to gain more definitive information about the status of the regional nodes.

Wide local excision is the definitive surgical treatment of melanoma. Following surgery, patients with American Joint Committee on Cancer stage I or II (node-negative) melanoma do not generally receive adjuvant therapy. Patients with higher risk melanoma receive adjuvant immunotherapy or targeted therapy. Ipilimumab has been shown to prolong recurrence-free survival by approximately 25% compared with placebo at a median of 5.3 years in patients with resected, stage III disease.¹⁵ Nivolumab has been shown to further prolong survival compared with ipilimumab by approximately 35% at 18 months.¹⁶ For patients who are *BRAF* V600 variant-positive with stage III melanoma, the combination of dabrafenib plus trametinib has been estimated to prolong relapse-free survival by approximately 50% over 3 years.¹⁷

Patients with stage I and II disease should undergo an annual routine physical and dermatologic examination. However, follow-up strategies and intervals have not been standardized or tested, and there is no consensus. These patients typically do not receive surveillance imaging. Patients with stage III melanoma may be managed with more frequent follow-up and imaging surveillance following therapy.

Gene Expression Profiling

Gene expression profiling measures the activity of thousands genes simultaneously and creates a snapshot of cellular function. Data for gene expression profiles are generated by several molecular technologies including DNA microarrays that measures activity relative to previously identified genes and RNA-Seq that directly sequences and quantifies RNA molecules. Clinical applications of gene expression profiling include disease diagnosis, disease classification, prediction of drug response, and prognosis.

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.

Gene Expression Profiling To Guide Initial Biopsy Decisions

Clinical Context and Test Purpose

The purpose of gene expression profiling (GEP) in patients who have suspicious pigmented lesions being considered for biopsy is to inform a decision about whether to biopsy.

The question addressed in this section of the evidence review is: Does GEP improve the net health outcome in individuals with suspicious pigmented lesions?

The following PICOTS were used to select literature to inform this review.

Patients

The relevant population of interest is patients with suspicious pigmented lesions being considered for referral for biopsy, specifically those lesions meeting one or more ABCDE criteria.

Interventions

The test being considered is the DermTech Pigmented Lesion Assay (PLA). The PLA test measures expression of 6 genes (*PRAME*, *LINC00518*, *CMIP*, *B2M*, *ACTB*, *PPIA*). The *PRAME* (PReferentially expressed Antigen in Melanoma) gene encodes an antigen that is preferentially expressed in human melanomas, and that is not expressed in normal tissues (except testis).¹⁸ *LINC00518* (Long Intergenic Non-protein Coding RNA518) is a regulatory RNA molecule. The other 4 genes provide normalization values.¹⁹ The feasibility of a test like PLA was first described in Wachsman et al (2011) and Gerami et al (2014).^{20,21} and development of the specific PLA test was described in Gerami et al (2017).²²

The test is performed on skin samples of lesions at least 5 mm in diameter obtained via noninvasive, proprietary adhesive patch biopsies of a stratum corneum specimen. The test does not work on the palms of hands, soles of feet, nails, or mucous membranes, and it should not be used on bleeding or ulcerated lesions.¹⁹

The PLA test report includes 2 results. The first result is called the PLA MAGE (Melanoma Associated Gene Expression), which indicates low risk (neither *PRAME* nor *LINC00518* expression was detected), moderate risk (expression of either *PRAME* or *LINC00518* was detected), or high risk (expression of both *PRAME* and *LINC00518* was detected). The second result is as an algorithmic PLA score that ranges from 0 to 100, with higher scores indicating higher suspicion of malignant disease.¹⁹

It is not clear whether the PLA test is meant to be used as a replacement, triage, or add-on test with respect to dermoscopy. The PLA sample report states that for low-risk lesions, physicians should "consider surveillance," while for moderate- and high-risk lesions, physicians should "recommend a biopsy." It does not state whether lesions with negative results should be further evaluated with dermoscopy or other techniques to confirm the lesion should not be biopsied. Therefore, this evidence review evaluates the test as a replacement for dermoscopy. As mentioned previously, there is a low threshold for biopsy of suspicious lesions. As such, tests that can rule-out need for biopsy could be useful and thus sensitivity and negative predictive value are the performance characteristics of most interest.

Comparators

After a referral from primary care to dermatology settings, dermatologists use visual examination as well as tools such as dermoscopy to make decisions regarding biopsy of suspicious lesions. A meta-analysis of 9 studies (8487 lesions with 375 melanomas) compared dermoscopy with visual examination alone for the diagnosis of melanoma; it reported that, for clinicians with training in dermoscopy, adding dermoscopy to visual examination increased the sensitivity from 71% to 90%. The specificity numerically increased from 80% to 90%, but the difference was not statistically significant.²³ Although dermoscopy is noninvasive and may aid in decision making regarding biopsy, it is only used by approximately 50% to 80% of dermatologists in the United States due to lack of training, interest, or time required for the examination.^{24,25}

The reference standard for diagnosis of melanoma is histopathology.

Outcomes

The beneficial outcomes of a true positive test result are appropriate biopsy and diagnosis of melanoma. The beneficial outcome of a true negative test result is potentially avoiding unnecessary biopsy.

The harmful outcome of a false-positive result is having an unnecessary biopsy. The harmful outcome of a false-negative result is potential delay in diagnosis and treatment.

Timing

The timeframe of interest for calculating performance characteristics is time to biopsy result. Patients who forgo biopsy based on test results could miss or delay diagnosis of cancer. Longer follow-up would be necessary to determine the effects on overall survival.

Setting

Initial identification of potentially cancerous lesions frequently occurs in primary care but may also occur in dermatology. Patients with lesions thought to be suspicious in primary care are frequently referred to dermatology when feasible, and decisions regarding biopsy are usually made by dermatologists.

Technically Reliable

Assessment of technical reliability focuses on specific tests and operators and requires review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review, and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

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).

Determining whether a test can guide biopsy decisions is not based only on its sensitivity and specificity, but also on how the accuracy of the existing pathway for making biopsy decisions is changed by the test. Therefore, the appropriate design for evaluating performance characteristics depends on the role of the new test in the pathway for making biopsy decisions. New tests may be used as replacements for existing tests, to triage who proceeds for existing tests or add-on tests after existing tests. For replacement tests, the diagnostic accuracy of both tests should be concurrently compared, preferably in a paired design (i.e., patients receive both tests), and all patients receive the reference standard. For a triage test, a paired design is also needed, with the reference standard being performed preferably on all patients but at least for all discordant results. For an add-on test, the included patients can be limited to those who were negative after existing tests with verification of the reference standard in patients who are positive on the new test.²⁶

Study Selection Criteria

For the evaluation of clinical validity of the PLA test, studies that meet the following eligibility criteria were considered:

- Reported on a validation cohort that was independent of the development cohort
- Reported on the accuracy of the marketed version of the technology
- Included a suitable reference standard (histopathology)
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described

Studies were excluded from the evaluation of the clinical validity of the PLA test because they reported results of the development cohort,²¹ they did not use the marketed version of the test,^{20,21} did not adequately describe the patient characteristics,²⁷ or did not adequately describe patient selection criteria.²⁷

The validation cohort from the Gerami et al (2017) publication was included.²² The study characteristics are described in Table 1. The report stated that included lesions were selected by dermatologists experienced in pigmented lesion management from 28 sites in the United States, Europe, and Australia; therefore, the samples were likely not consecutive or random. Information regarding the previous testing was not provided. The flow of potential and included samples was not clear, and whether the samples were all independent or, multiple samples from the same patient were not described. Diagnosis of melanoma was based on consensus among a primary reader and 3 expert dermatopathologists. The report did not state whether the histopathologic diagnosis was blinded to the results of the PLA test but did state the diagnosis was “routinely” assessed. Interpretation of the PLA result does not depend on a reader, so it is blinded to histopathologic results. In 11% of cases originally selected, a consensus diagnosis was not reached, and these samples were not included in the training or validation cohorts. Dates of data collection were not reported. Sex and anatomic location of biopsy were reported, but other clinical characteristics (e.g., risk factors for melanoma, presenting symptoms) were not. Study results are shown in Table 2. The study training cohort included 157 samples with 80 melanomas and 77 non-melanomas. The study validation cohort included 398 samples with 87 melanomas and 311 non-melanomas. Study relevance, design, and conduct gaps are in Tables 3 and 4.

Section Summary: PLA Clinical Validity

Multiple high-quality studies are needed to establish the clinical validity of a test. The PLA test has one clinical validity study with many methodologic and reporting limitations. Therefore, performance characteristics are not well-characterized. Also, the test has not been compared with dermoscopy, another tool frequently used to make biopsy decisions.

Table 1. Clinical Validity Study Characteristics of the PLA Test for Diagnosing Melanoma

Study	Study Population	Design	Reference Standard for Dx of Melanoma	Threshold Score for PLA Test	Timing of Reference and PLA Tests	Blinding of Assessors
Gerami et al (2017) ²²	<ul style="list-style-type: none"> • Adults • Suspicious pigmented lesion ≥ 4 mm in diameter • Without obvious or suspicious nodular melanoma • 24% from extremities, 13% from head and 	<ul style="list-style-type: none"> • Retrospective • Not consecutive or random 	Histopathology; consensus diagnosis	<ul style="list-style-type: none"> • Quantitative PCR yielded an amplification curve and a measurable cycle threshold value • Either <i>LINC00518</i> or <i>PRAME</i> detected 	PLA patch before surgical biopsy; timing between patch and surgical biopsy unclear	Not clear

Study	Study Population	Design	Reference Standard for Dx of Melanoma	Threshold Score for PLA Test	Timing of Reference and PLA Tests	Blinding of Assessors
	neck, 62% from trunk <ul style="list-style-type: none"> • 55% of samples from men • Median age, 49 y (range, 19-97 y) 					

Dx: diagnosis; PCR: polymerase chain reaction.

Table 2. Clinical Validity Study Results of the PLA Test for Diagnosing Melanoma

Study	Initial N	Final N	Excluded Samples	Melanoma Prevalence	Sensitivity ^b	Specificity ^b	PPV ^b	NPV ^b
Gerami et al (2017) ²²	398 ^a	398	Before allocation to training and validation cohorts, 11% of original samples excluded due to lack of consensus diagnosis	22%	91 (83 to 96)	69 (64 to 74)	45 (38 to 53) ^c	96 (93 to 98) ^c

NPV: negative predictive value; PPV: positive predictive value.

^a 398 samples were included in the validation cohort; the number of independent patients is unclear.

^b Values are percentages with 95% confidence interval.

^c Confidence intervals provided in the report; calculated from data provided.

Table 3. Clinical Validity Study Relevance Gaps of the PLA Test

Study	Population ^a	Intervention ^b	Comparator ^c	Outcomes ^d	Duration of Follow-Up ^e
Gerami et al (2017) ²²	3. Study population characteristics not adequately described		3. No comparison to dermoscopy	3. Predictive values were not reported but were calculated based on data provided	

The evidence gaps 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. Clinical Validity Study Design and Conduct Gaps of the PLA Test

Study	Selection ^a	Blinding ^b	Delivery of Test ^c	Selective Reporting ^d	Completeness of Follow-Up ^e	Statistical ^f
Gerami et al (2017) ²²	1,2. Not clear what criteria used to select samples but it	1. Blinding of histopathology readers	1. Patch biopsy administered before	1. No registration reported		

Study	Selection ^a	Blinding ^b	Delivery of Test ^c	Selective Reporting ^d	Completeness of Follow-Up ^e	Statistical ^f
	does not appear to have been random or consecutive	not described	surgical biopsy but timing between procedures not described			

The evidence gaps 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 Follow-Up 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.

Clinically Useful

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

No direct evidence of clinical utility 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 through a chain of evidence.

A decision-impact study by Ferris et al (2017) assessed the potential impact of PLA on physicians' biopsy decisions in patients.²⁷ Forty-five dermatologists evaluated 60 clinical and dermoscopic images of atypical pigmented lesions (8 melanoma, 52 nonmelanoma). In the first round, dermatologists did not have PLA test results and, in the second round, dermatologists had access to PLA test results with the order of cases being scrambled. The dermatologists were asked whether the lesions should be biopsied after each round. Therefore, the corresponding number of biopsy decisions should be $45 \times 60 \times 2 = 5400$. Data were collected in 2014 and 2015. Results were reported for 4680 decisions with no description of the disposition of the remaining decisions. Of the 4680 reported decisions, 750 correct biopsy decisions were made without PLA results while 1331 were made with PLA results and 1590 incorrect biopsy decisions were made without PLA results while 1009 incorrect biopsy decisions were made with PLA results.

Section Summary: Clinically Useful

There is no direct evidence of clinical utility. A chain of evidence for clinical utility cannot be constructed due to lack of robust evidence of clinical validity.

GEP for Diagnosing Melanoma with Indeterminate Histopathology

Clinical Context and Test Purpose

The purpose of GEP in patients whose melanocytic lesion is indeterminate after histopathology is to aid in the diagnosis of melanoma and decisions regarding treatment and surveillance. The question addressed in this section of the evidence review is: Does GEP improve the net health outcome in individuals with indeterminate melanocytic lesions?

The following PICOTS were used to select literature to inform this review.

Patients

The relevant population of interest is patients whose melanocytic lesion is indeterminate based on clinical and histopathologic features.

Interventions

The test being considered is the Myriad myPath Melanoma test. The myPath test measures expression of 23 genes using quantitative reverse-transcription polymerase chain reaction. Fourteen genes are involved in melanoma pathogenesis and are grouped into 3 components related to cell differentiation, cell signaling, and the immune response, and 9 housekeeper genes are also included. The test is performed on 5 standard tissue sections from an existing formalin-fixed, paraffin-embedded biopsy specimen.

The myPath test report includes an algorithmic myPath score ranging from -16.7 to 11.1, with higher, positive scores indicating higher suspicion of malignant disease.²⁸ The myPath report also classifies these scores: -16.7 to -2.1 are "benign"; -2.0 to -0.1 are "indeterminate"; and 0.0 to +11.1 are "malignant". Development of the test has been described by Clarke et al (2015).²⁹ The myPath test is meant as an add-on test to standard histopathology. No recommendations for treatment or surveillance are given on the report.

Comparators

The reference standard for diagnosis of melanoma is histopathology. However, in cases of indeterminate histopathology, long-term follow-up is needed to determine the final clinical diagnosis.

Fluorescence in situ hybridization (FISH) has been evaluated as a tool to aid in the diagnosis of lesions that are indeterminate, following histopathology in 2 studies that included histologically ambiguous lesions and a clinical, long-term follow-up to establish the diagnosis. One study reported by Gaiser et al (2010) included 22 melanocytic lesions (12 indeterminate) followed for a mean of 65 months (range, 10-156 months) and reported a FISH sensitivity of 60% and a specificity of 50% for development of metastases during follow-up.³⁰ A second study, reported by Vergier et al (2011), included 90 indeterminate melanocytic lesions of which 69 had no recurrence for at least 5 years of follow-up (mean, 9 years; range, 5-19 years) and 21 lesions that exhibited metastases. The sensitivity and specificity rates of the histopathologic review combined with FISH for the clinical outcome were 76% and 90%, respectively.³¹

Outcomes

The beneficial outcomes of a true positive test result are a diagnosis of melanoma and corresponding appropriate treatment and surveillance. The beneficial outcome of a true negative test result is avoiding unnecessary surgery.

The harmful outcome of a false-positive result is having an unnecessary surgery and surveillance. The harmful outcome of a false-negative result is a delay in diagnosis and treatment.

Timing

Recurrence and metastases can occur many years after treatment of melanoma. In the 2 studies evaluating long-term outcomes of FISH (described above), the mean follow-up was approximately 5.5 and 9 years.^{30,31} In Vergier et al (2011), metastases in the FISH-negative group

occurred by 5 years.³¹ For this section of the review, at least 5 years of event-free follow-up is required to confirm negative tests.

Setting

Follow-up of melanocytic lesions that are indeterminate after histopathology is generally done in dermatology.

Technically Reliable

Assessment of technical reliability focuses on specific tests and operators and requires review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review, and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

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).

Study Selection Criteria

For the evaluation of clinical validity of the myPath test, studies that meet the following eligibility criteria were considered:

- Reported on a validation cohort that was independent of the development cohort;
- Reported on the accuracy of the marketed version of the technology;
- Included a suitable reference standard (final clinical diagnosis with at least 5 years of follow-up for negatives);
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

Studies were excluded from the evaluation of the clinical validity of the myPath test because authors did not use an appropriate reference standard,^{29,32} or did not adequately describe patient characteristics.²⁹

The Ko et al (2017) clinical validity study met selection criteria.³³ The study characteristics are described in Table 5. In Ko et al (2017), archived melanocytic neoplasms were submitted for myPath testing from university clinics in the United States and United Kingdom with additional samples acquired from Avaden BioSciences.³³ Stage I, II, and III primary cutaneous melanomas that produced distant metastases subsequent to the diagnosis and benign lesions with clinical follow-up and no evidence of recurrence of metastases were included. For benign samples, a disease-free time of at least 5 years was recommended. Information on the previous testing was not provided. It is not clear if any of the samples originally had indeterminate histopathology results. Dates of data collection were not reported. Sex, age, Breslow depth, and anatomic location were described; presenting symptoms were not reported. A total of 293 samples were submitted; of these 53 did not meet inclusion criteria and 58 (24% of those tested) failed to produce a valid test score. An additional seven samples with indeterminate results were excluded from the calculations of performance characteristics. Study results are shown in Table 6. Study relevance, design, and conduct gaps are in Tables 7 and 8.

Table 5. Clinical Validity Study Characteristics of the myPath Test for Diagnosing Melanoma

Study	Study Population	Design	Reference Standard for Diagnosis of Melanoma	Threshold Score for Positive myPath Test	Timing of Reference and myPath Tests	Blinding of Assessors
Ko et al (2017) ³³	• Primary cutaneous melanomas or benign melanocytic nevi	• Retrospective • Not consecutive or randomly selected	• For positive melanoma diagnosis: malignant lesions that produced	• Scores from 0.0 to 11.1 (i.e., "malignant")	• Final clinical diagnosis established before myPath test	Yes

Study	Study Population	Design	Reference Standard for Diagnosis of Melanoma	Threshold Score for Positive myPath Test	Timing of Reference and myPath Tests	Blinding of Assessors
	<ul style="list-style-type: none"> Mean age, 53 y 55% of samples from men 		<ul style="list-style-type: none"> distant metastases For negative melanoma diagnosis: Event-free follow-up, recommended 5 y (median, 6.2 y) 		<ul style="list-style-type: none"> Length of time between biopsy and myPath test unclear 	

Table 6. Clinical Validity Study Results of the myPath Test for Diagnosing Melanoma

Study	Initial N	Final N	Excluded Samples	Melanoma Prevalence	Sensitivity ^a	Specificity ^a	PPV ^a	NPV ^a
Ko et al (2017) ³³	240	175	<ul style="list-style-type: none"> 58 failed to produce test result 7 with indeterminate results 	54	94 (87 to 98) ^b	96 (89 to 99) ^b	97 (91 to 99) ^b	93 (85 to 97) ^b

NPV: negative predictive value; PPV: positive predictive value.

^a Values are percentages with 95% confidence interval.

^b Confidence intervals not provided in the report; calculated from data provided.

Table 7. Clinical Validity Study Relevance Gaps of the myPath Test

Study	Population ^a	Intervention ^b	Comparator ^c	Outcomes ^d	Duration of Follow-Up ^e
Ko et al (2017) ³³	4. Study population is not limited to lesions that are indeterminate following histopathology		3. No comparison to CGH or FISH		None noted

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

CGH: comparative genomic hybridization; FISH: fluorescence in situ hybridization.

^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. Clinical Validity Study Design and Conduct Gaps of the myPath Test

Study	Selection ^a	Blinding ^b	Delivery of Test ^c	Selective Reporting ^d	Completeness of Follow-Up ^e	Statistical ^f
Ko et al (2017) ³³	2. Samples not consecutive or random		1. Unclear how much time elapsed between biopsy and myPath test	1. No registration reported	2. More than 25% of samples tested did not produce results or produced indeterminate results	1. CIs for sensitivity and specificity not reported but were calculated based on data provided. NPV,

Study	Selection ^a	Blinding ^b	Delivery of Test ^c	Selective Reporting ^d	Completeness of Follow-Up ^e	Statistical ^f
						PPV were not reported

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

CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

^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 Follow-Up 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

Multiple high-quality studies are needed to establish the clinical validity of a test. The myPath test has 1 clinical validity study including long-term follow-up to establish the clinical diagnosis as the reference standard. However, it is not clear whether the study population included lesions that were indeterminate following histopathology and the study had other methodologic and reporting limitations. Therefore, performance characteristics are not well-characterized. Also, the test has not been compared with comparative genomic hybridization and FISH, other tools frequently used along with histopathology to confirm the diagnosis in challenging cases.

Clinically Useful

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

No direct evidence of clinical utility 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.

Two decision-impact studies assessed the potential impact of myPath on physicians' treatment decisions in patients with diagnostically challenging lesions.^{34,35} Given the lack of established clinical validity and no reported long-term outcomes, it is not known whether any treatment changes were clinically appropriate.

Section Summary: Clinically Useful

There is no direct evidence of clinical utility. A chain of evidence for clinical utility cannot be constructed due to lack of robust evidence of clinical validity.

GEP to Guide Decisions Regarding Sentinel Lymph Node Biopsy in Melanoma

Clinical Context and Test Purpose

The purpose of GEP in patients with melanoma is to identify low and high-risk patients classified as stage I or II according to the American Joint Committee on Cancer (AJCC) criteria. Current guidelines do not recommend adjuvant therapy or imaging surveillance for AJCC stage I or II

patients following surgery. Patients initially staged as I or II who have positive lymph nodes following sentinel lymph node biopsy (SLNB) are then eligible to be treated with adjuvant therapy as stage III patients.³⁶

The management decision to be made based on this test is not clear. The manufacturer's website indicates that physicians can use DecisionDx-Melanoma information to "consider upstaging" patients for "active systemic surveillance or referral to medical oncology for consideration of systemic drug therapy or clinical trials."³⁷ Similarly, in 1 clinical validity study (described below), the authors stated that "high-risk patients with stage I and II disease may benefit from adjuvant therapy and/or enhanced imaging protocols to allow for early detection of metastasis."³⁸ In another clinical validity study, the authors concluded that the test's "role in consideration of patients for adjuvant therapy should be examined prospectively."³⁹ This use of the test would be as a replacement for SLNB.

However, the manufacturer has also suggested in materials submitted to Evidence Street that the test can be used to select patients at low risk of being lymph node-positive who can avoid an SLNB (i.e., a triage test for SLNB). This use of the test will be the focus of the review.

The question addressed in this section of the evidence review is: Does GEP improve the net health outcome in individuals with AJCC stage I or II melanoma?

The following PICOTS were used to select literature to inform this review.

Patients

The relevant population of interest is patients with AJCC stage I or II cutaneous melanoma who are being considered for SLNB.

Interventions

The test being considered is the Castle Biosciences DecisionDx-Melanoma test. The DecisionDx test measures expression of 31 genes using quantitative reverse-transcription polymerase chain reaction. The test includes 28 prognostic gene targets and 3 endogenous control genes. The test is performed on standard tissue sections from an existing formalin-fixed, paraffin-embedded biopsy or wide local excision specimen.

Development of the test was described in Gerami et al (2015).³⁸ To develop the DecisionDx-Melanoma gene panel, Gerami et al (2015) conducted a meta-analysis of published studies that identified differential gene expression in metastatic vs nonmetastatic primary cutaneous melanoma. Of 54 identified genes, investigators selected 20 for further polymerase chain reaction analysis based on chromosomal location. Five genes from Castle Biosciences' DecisionDx-UM gene panel were added based on analysis of metastatic and nonmetastatic primary cutaneous melanoma, and 2 probes of the *BRCA1*-associated protein 1 gene, *BAP1*, which has been associated with the metastatic potential of uveal melanoma, also were added. Finally, 4 genes with minimal variation in expression level between metastatic and nonmetastatic primary cutaneous melanoma were added as controls. Patients had a minimum follow-up of 5 years unless there was a well-documented metastatic event, including positive SLNB. Information about treatments received was not provided.

The DecisionDx test report provides 2 results: a class and a probability score. The class stratifies tumors as low risk (class 1) or high risk (class 2), with subclassifications within each class (A or B) based on how close the probability score is to the threshold between class 1 and class 2. The probability score ranges from 0 to 1 and appears to be the risk of recurrence within 5 years.

DecisionDx is meant to be used as a triage test with respect to SLNB. However, the sample report makes no recommendations for SLNB, treatment, or surveillance based on test results.

Comparators

Treatment and surveillance recommendations are based on AJCC staging. SLNB may be used to get more definitive information about the status of the regional nodes compared with a physical examination. The American Society of Clinical Oncology⁴⁰ and National Comprehensive Cancer Network³⁶ have similar but not identical recommendations on which patients should undergo SLNB (patients with thickness more than approximately 1 mm or thin melanomas with other high-risk features). SLNB has a low rate of complications; in the Sunbelt Melanoma Trial, a prospective multi-institutional study of SLNB for melanoma reported by Wrightson et al (2003), less than 5% of the 2120 patients developed *major* or *minor* complications associated with SLNB.⁴¹

Online tools are available to predict prognosis based on the AJCC guidelines. The original AJCC tool was developed by Soong et al (n.d.).⁴² Callender et al (2012) incorporated SLNB results into a revised tool (<http://www.melanomacalculator.com/>).⁴³

Outcomes

For patients meeting guideline-recommended criteria for SLNB, a positive DecisionDx (class 2) test result would not change outcomes. The patients would proceed to SLNB, as they would have without the DecisionDx test, and treatment and imaging decisions would depend on SLNB results. A negative DecisionDx (class 1) test result would indicate that a patient could avoid an SLNB. Therefore, the potential beneficial outcomes of a negative result are avoidance of an SLNB. The potential harmful outcomes of a negative result are reduced time to recurrence due to not identifying node-positive patients that would be eligible for beneficial adjuvant treatment.

Timing

The risk of recurrence decreases over time but does not reach zero. In a study of 1568 patients with stage I melanoma, Dicker et al (1999) found that 80% of the recurrences occurred within the first 3 years.⁴⁴ A prospective study by Garbe et al (2003) reported that, for stage I and II patients, the risk of recurrence was low after 4.4 years.⁴⁵ Among 4731 patients treated for more than 10 years at 1 institution, Faries et al (2013) found the majority of recurrences occurred in the first 5 years.⁴⁶ However, 7% of patients experienced recurrence after 10 years (median, 16 years). Even among stage I/II patients, recurrence after 10 years occurred in 2% of patients.

Five-year recurrence-free survival (RFS) is the outcome and time-point of interest.

Setting

Follow-up of patients with stage I and II melanoma is generally done in secondary care.

Technically Reliable

Assessment of technical reliability focuses on specific tests and operators and requires review of unpublished and often proprietary information. Review of specific tests, operators, and unpublished data are outside the scope of this evidence review, and alternative sources exist. This evidence review focuses on the clinical validity and clinical utility.

Clinically Valid

Study Selection Criteria

For the evaluation of clinical validity of the DecisionDx test, studies that meet the following eligibility criteria were considered:

- Reported on a validation cohort that was independent of the development cohort;
- Reported on the accuracy of the marketed version of the technology;
- Included a suitable reference standard (5-year RFS);
- Patient/sample clinical characteristics were described
- Patient/sample selection criteria were described.

Hsueh et al (2017) was excluded from the evaluation of the clinical validity of the DecisionDx test because it did not report 5-year outcomes (median follow-up, 1.5 years).⁴⁷ Samples used in

Gerami et al (2015)⁴⁸ and Ferris et al (2017)²⁷ appear to overlap with the samples from Gerami et al (2015)³⁸ and each other and will not be considered independent validation studies for inclusion in the tables. They are described briefly following the clinical validity tables.

Two independent clinical validity studies meeting eligibility criteria have been conducted. Characteristics and results are summarized in Tables 9 and 10 and briefly in the paragraphs that follow.

Table 9. Clinical Validity Study Characteristics of the DecisionDx Test for Diagnosing Melanoma

Study	Study Population	Design	Reference Standard / Outcome Measure	Threshold Score for Positive DecisionDx Test	Timing of Reference and DecisionDx Tests	Blinding of Assessors
Gerami et al (2015) ³⁸ ; Validation subset	<ul style="list-style-type: none"> Adults Stage I-IV cutaneous melanoma (87% stage I/II) At least 5 y of FU (median, 7.0 y) 	<ul style="list-style-type: none"> Retrospective Not consecutive or randomly selected 	5-y RFS	<ul style="list-style-type: none"> Class 2 is high risk Risk threshold not provided 	<ul style="list-style-type: none"> Patient diagnosed between 1998 and 2009 Timing of DecisionDx not described 	Yes
Zager et al (2018) ³⁹	<ul style="list-style-type: none"> Stage I-III cutaneous melanoma (68% stage I/II) At least 5 y of FU (median, 7.5 y) 	<ul style="list-style-type: none"> Retrospective Not consecutive or randomly selected 	5-y RFS	<ul style="list-style-type: none"> Class 2 is high risk Class 1: probability score 0 to 0.49 Class 2: probability score 0.5 to 1 	<ul style="list-style-type: none"> Patients diagnosed between 2000 and 2014 Timing of DecisionDx not described 	Yes

DFS: disease-free survival; FU: follow-up; RFS: recurrence-free survival.

Table 10. Clinical Validity Study Results of the DecisionDx Test for Diagnosing Melanoma

Table 1. Clinical Validity Study Results of the DecisionDX Test for Diagnosing Melanoma								
Study	Initial/Final N	Excluded Samples	Events and Kaplan-Meier 5-Year RFS ^a		Sensitivity ^a	Specificity ^a	PPV ^a	NPV ^a
			Class 1	Class 2				
Gerami et al (2015) ³⁸ ; Validation subset		Samples excluded if melanoma dx not confirmed, dissectible area not acceptable						
Overall	Unclear/104		<ul style="list-style-type: none">• 4 events• RFS=97 (NR)	<ul style="list-style-type: none">• 31 events• RFS=31 (NR)• p<0.001 vs class 1	89 (73 to 97) ^b	83 (72 to 91) ^b	72 (56 to 85) ^b	93 (84 to 98) ^b
AJCC stage I and II	Unclear/78		<ul style="list-style-type: none">• 3 events• RFS=98 (NR)	<ul style="list-style-type: none">• 18 events• RFS=37 (NR)• p<0.001 vs class 1	86 (64 to 97) ^b	84 (72 to 93) ^b	67 (46 to 83) ^b	94 (84 to 99) ^b
Zager et al (2018) ³⁹		Did not meet analytic quality						

Study	Initial/Final N	Excluded Samples	Events and Kaplan-Meier 5-Year RFS ^a	Sensitivity ^a	Specificity ^a	PPV ^a	NPV ^a
		control thresholds					
Overall	601 / 523		<ul style="list-style-type: none"> 42 events RFS=88 (85 to 92) 	70 (62 to 78)	71 (67 to 76)	48 (41 to 55)	87 (82 to 90)
AJCC stage I	Unclear / 264		<ul style="list-style-type: none"> 11 events RFS=96 (94 to 99) 	35 (14 to 62) ^b	87 (82 to 91) ^b	15 (6 to 31) ^b	95 (91 to 98) ^b
AJCC stage II	Unclear / 93		<ul style="list-style-type: none"> 9 events RFS=74 (60 to 91) 	77 (61 to 89) ^b	43 (29 to 57) ^b	49 (36 to 62) ^b	72 (53 to 86) ^b

AJCC: American Joint Committee on Cancer; CI: confidence interval; Dx: diagnosis; NPV: negative predictive value; NR: not reported; PPV: positive predictive value; RFS: recurrence-free survival.

^a Values are percentages with 95% confidence interval.

^b Confidence intervals not provided in the report; calculated from data provided.

The validation cohort in Gerami et al (2015) included patients with stage 0, I, II, III, or IV disease from 6 U.S. centers (N=104).³⁸ A complete disposition of samples received from the institutions and those included in the analysis was not provided. For 78 patients in the validation cohort with AJCC stage I or II cutaneous melanoma who had either a metastatic event or had more than 5 years of follow-up without metastasis, 5-year disease-free survival was 98% (CIs not reported) for DecisionDx class I patients and 37% for DecisionDx class II patients. The positive predictive value (PPV) and negative predictive value (NPV) were 67% and 94%, respectively. CIs for performance characteristics were calculated in Table 10 based on data provided. Reclassification of patients in AJCC stages to DecisionDx classes is shown in Table 11.

Table 11. Reclassification of Patients Based on AJCC Stages to DecisionDx Classes in the Gerami Validation Cohort

AJCC Stage	DecisionDx Class		Total
	Class 1 (Low Risk), N (row %)	Class 2 (High Risk), N (row %)	
0	0	0	
Total stage I	50 (89%) ^a	6 (11%)	56
IA	37	1	
IB	10	5	
Total stage II	10 (29%)	24 (71%)	34
IIA	5	8	
IIB	5	12	
IIC	0	4	
Total stage III	1 (8%)	11 (92%)	12
Total stage IV	0 (0%)	2 (100%)	2
Total	61	43	104

Adapted from Gerami et al (2015).³⁸

AJCC: American Joint Committee on Cancer.

^a The subclass for n=3 class 1 samples are not reported.

Zager et al (2018) reported results of a second clinical validity study including AJCC stage I, II, or III primary melanoma tumors from 16 U.S. sites.³⁹ The samples were independent of the other validation studies. Of the 601 cases submitted from the institutions, 523 were included in the analysis (357 stage I/II). The excluded samples did not meet pre- and post-analytic quality control thresholds. SLN status was untested in 36% of the patients, negative in 34%, and positive in 30%. The report did not describe any adjuvant therapy that the patients received. Overall, 42 (13%) recurrence events occurred in DecisionDx class 1 patients and 100 (48%) recurrence events occurred in DecisionDx class 2 patients. The 5-year RFS estimated by Kaplan-Meier was 88% (95% CI, 85% to 92%) in class 1 and 52% (95% CI, 46% to 60%) in class 2. The reported sensitivity and specificity were 70% (95% CI, 62% to 78%) and 71% (95% CI, 67% to 76%), respectively, with a PPV of 48% (95% CI, 41% to 55%) and a NPV of 87% (95% CI, 82% to 90%). For comparison, the performance characteristics for 5-year RFS for sentinel lymph node status

among those with SLNB were: sensitivity, 66% (95% CI, 57% to 74%); specificity, 65% (95% CI, 58% to 71%); PPV, 52% (95% CI, 44% to 60%); and NPV, 76% (95% CI, 69% to 82%). Estimates stratified by AJCC stage I or II are shown in Table 10. The reclassification of patients based on SLNB status using DecisionDx classes is shown in Table 12. If DecisionDx were used as a triage test such that only class 2 received SLNB, then 159 class 1 patients would not have undergone SLNB. Of the 159 patients in class 1, 56 were SLNB-positive and were therefore eligible for adjuvant therapy. It is not clear if the SLNB-positive patients in this study received adjuvant therapy. Of the 56 patients who were DecisionDx class 1 and SLNB-positive, 22 recurrence events occurred by 5 years.

Relevance, design, and conduct gaps are summarized in Tables 13 and 14.

Table 12. Reclassification of Patients Based on SLNB Status to DecisionDx Classes

SLNB	DecisionDx Class 1 (Low Risk)			DecisionDx Class 1 (High Risk)			Total
	n (%)	Events	5-Year RFS (95% CI), %	n (%)	Events	5-Year RFS (95% CI), %	
Negative	103 (65)	15	87 (81 to 94)	77 (43)	28	67 (57 to 79)	180
Positive	56 (35)	22	61 (49 to 76)	101 (57)	60	37 (28 to 49)	157
Total	159			178			337 ^a

Adapted from Zager et al (2017).³⁹

RFS: recurrence-free survival; SLNB: sentinel lymph node biopsy.

^a 337 patients had DecisionDx results and SLNB results.

Table 13. Clinical Validity Study Relevance Gaps of the DecisionDx Test

Study	Population ^a	Intervention ^b	Comparator ^c	Outcomes ^d	Duration of Follow-Up ^e
Gerami et al (2015) ³⁸ ; Validation subset	4. Study population includes AJCC stage III/IV lesions (13%), although analysis for only stage I/II was provided	1. Risk threshold for classification into class 1 or 2 not provided.	3. Not compared to other prediction tools	2. Evidence-based treatment or surveillance pathway using the test is not described	
Zager et al (2018) ³⁹	4. Study population includes AJCC stage III lesions (32%), although analysis for only stage I/II was provided			2. Evidence-based treatment or surveillance pathway using the test is not described	

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

AJCC: American Joint Committee on Cancer.

^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 14. Clinical Validity Study Design and Conduct Gaps of the DecisionDx Test

Study	Selection ^a	Blinding ^b	Delivery of Test ^c	Selective Reporting ^d	Completeness of Follow-Up ^e	Statistical ^f
Gerami et al (2015) ³⁸ ; Validation subset	2. Not consecutive or random		1. Time between collection of biopsy and DecisionDx not described	1. No registration reported	1. No description of number of samples (if any) that failed to produce results or were indeterminate	1. CIs not reported but were calculated based on data provided
Zager et al (2018) ³⁹	2. Not consecutive or random		1. Time between collection of biopsy and DecisionDx not described	1. No registration reported	1. No description of number of samples (if any) that failed to produce results or were indeterminate	

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

CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value.

^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 Follow-Up 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.

In a subsequent analysis of patients with melanoma who had undergone SLNB, Gerami et al (2015) compared prognostic classification by DecisionDx-Melanoma with biopsy results.⁴⁸ A total of 217 patients comprised a convenience sample from a database of 406 patients previously tested with DecisionDx-Melanoma. Patients who had undergone SLNB appear to overlap with patients in Gerami et al (2015)³⁸ discussed previously. Most (73%) patients had a negative SLNB, and 27% had a positive SLNB. DecisionDx-Melanoma classified 76 (35%) tumors as low risk (class I) and 141 (65%) tumors as high risk (class II). Within the group of SLNB-negative patients, the 5-year overall survival rate was 91% in class I patients and 55% in class II patients. Within the group of SLNB-positive patients, the 5-year overall survival rate was 77% in class I patients and 57% in class II patients.

Ferris et al (2017) compared the accuracy of DecisionDx-Melanoma with the web-based AJCC Individualized Melanoma Patient Outcome Prediction Tool.²⁷ The study included 205 patients who appear to overlap with the patients in the second Gerami et al (2015) study described above. AJCC-predicted 5-year survival for each patient was categorized into low and high risk based on both a 68% predicted 5-year survival and a 79% predicted 5-year survival. The 68% and 79% cutpoints were reported to correspond to 5-year survival in patients with stage IIA and IIB, respectively, although it is unclear whether those cutpoints were prespecified, whether they were based on internal or external estimates of risk, or whether they are commonly used in practice. The prognostic sensitivity and specificity for death (median follow-up, 7 years) of the Decision-Dx Melanoma were 78% and 69%, respectively (CIs not reported). The sensitivity and specificity for the AJCC calculator with the 79% cutpoint were 60% and 74%, respectively. The combination of the DecisionDx-Melanoma and AJCC tools had a sensitivity of 82% and specificity of 62%. The cross-classification for the DecisionDx-Melanoma and AJCC tools for 5-year overall survival is shown in Table 15.

Table 15. Cross-Classification for the DecisionDx-Melanoma and AJCC Tool (79% Cutpoint) for 5-Year Overall Survival

Risk Classification (DecisionDx-Melanoma vs AJCC)	N	No. of Events	5-Year Overall Survival, %
Low/low	105	9	96
Low/high	13	2	83
High/low	30	11	71
High/high	57	28	44

Adapted from Ferris et al (2017).²⁷

AJCC: American Joint Committee on Cancer.

Section Summary: Clinically Valid

To use prognostic information for decision-making, performance characteristics should be consistent and precise. Two independent studies, using archived tumor specimens, have reported 5-year RFS in AJCC stage I or II patients. Gerami et al (2015) reported RFS rates of 98% in DecisionDx class 1 (low risk) without CIs in AJCC stage I or II patients. Zager et al (2018) reported RFS rates of 96% (95% CI, 94% to 99%) for DecisionDx class 1 in patients with AJCC stage I disease and RFS rates of 74% (95% CI, 60% to 91%) for DecisionDx class 1 in patients with AJCC stage II disease. Although CIs were not available for the first study, RFS does not appear to be well-characterized in the DecisionDx low-risk group as evidenced by the variation in estimates across studies. Zager et al (2017) also reported that 56 of 159 (35%) patients who were DecisionDx class 1 (low risk) were SLNB-positive and in those patients 22 recurrences (39%) occurred over 5 years.³⁹ If the DecisionDx test were used as a triage for SLNB, these patients would not undergo SLNB and would likely not receive adjuvant therapy, which has shown to be effective at prolonging time to recurrence in node-positive patients.

Clinical Useful

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

No direct evidence of clinical utility 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.

Four decision-impact studies have been published reporting on the impact of DecisionDx on physicians' management decisions.⁴⁹⁻⁵² Given the lack of established clinical validity and no reported long-term outcomes, it is not known whether any management changes were clinically appropriate.

For the proposed use of the test as a triage for SLNB (identify patients who can avoid SLNB), performance characteristics are not well-characterized.

For the proposed use of the test as a replacement for SLNB (identify patients who are AJCC stage I/II who should receive adjuvant therapy), an evidence-based management pathway would be needed to support the chain of evidence. The existing RCTs demonstrating that adjuvant therapy reduces recurrence included node positive patients. No evidence was identified that demonstrated that adjuvant therapy or increased surveillance improves net health outcomes in AJCC stage I or II patients who are DecisionDx class 2.

Section Summary: Clinically Useful

There is no direct evidence of clinical utility. A chain of evidence for clinical utility cannot be created due to lack of robust evidence of clinical validity and lack of evidence-based management pathway.

Summary of Evidence

For individuals with suspicious pigmented lesions (based on ABCDE and/or ugly duckling criteria) being considered for biopsy who receive gene expression profiling with the DermTech Pigmented Lesion Assay to determine which lesions should proceed to biopsy, the evidence includes observational studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, and resource utilization. The Pigmented Lesion Assay has 1 clinical validity study with many methodologic and reporting limitations. Therefore, performance characteristics are not well-characterized. Also, the test has not been compared with dermoscopy, another tool frequently used to make biopsy decisions. No direct evidence of clinical utility was identified. Given that the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility through a chain of evidence. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals who have melanocytic lesions with indeterminate histopathologic features who receive gene expression profiling with the myPath Melanoma test added to histopathology to aid in the diagnosis of melanoma, the evidence includes observational studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, change in disease status, treatment-related morbidity. The myPath test has 1 clinical validity study, which includes long-term follow-up to establish the clinical diagnosis as the reference standard. However, it is not clear if the study population included lesions that were indeterminate following histopathology and the study had other methodologic and reporting limitations. Therefore, performance characteristics are not well-characterized. No direct evidence of clinical utility was identified. Given that the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility through a chain of evidence. The evidence is insufficient to determine the effects of the technology on health outcomes.

For individuals with AJCC stage I or II cutaneous melanoma who receive gene expression profiling with the DecisionDx-Melanoma test to determine whether to perform SLNB, the evidence includes observational studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, change in disease status, resource utilization and treatment-related morbidity. The DecisionDx-Melanoma test has 2 independent clinical validity studies that have reported 5-year recurrence-free survival (RFS) in AJCC stage I or II patients. Gerami et al (2015) reported RFS rates of 98% in DecisionDx class 1 (low risk) without CIs, in AJCC stage I or II patients. Zager et al (2017) reported RFS rates of 96% (95% CI, 94% to 99%) for DecisionDx class 1 in patients with AJCC stage I disease; they also reported RFS rates of 74% (95% CI, 60% to 91%) for DecisionDx class 1 in patients with AJCC stage II disease. Although CIs were not available for the first study, RFS does not appear to be well-characterized as evidenced by the variation in estimates across studies. Zager et al (2017) also reported that in 56 patients who were DecisionDx class 1 (low risk) but SLNB-positive, 22 recurrences (39%) occurred over 5 years. If the DecisionDx test were used as a triage for SLNB, these patients would not undergo SLNB and would likely not receive adjuvant therapy, which has shown to be effective at prolonging time to recurrence in node-positive patients. No direct evidence of clinical utility was identified. Given that the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility through a chain of evidence. There is also not an explicated, evidence-based management pathway for the use of the test. The evidence is insufficient to determine the effects of the technology on health outcomes.

Supplemental Information Practice Guidelines and Position Statements

National Comprehensive Cancer Network

The National Comprehensive Cancer Network guidelines (v.2.2018) for melanoma made the following statements on use of gene expression profiling.³⁶ "While there is interest in newer prognostic molecular techniques such as gene expression profiling to differentiate primary cutaneous melanomas (before or following sentinel lymph node biopsy) is not recommended outside of a clinical study (trial)."

U.S. Preventive Services Task Force Recommendations

Not applicable.

Medicare National Coverage

There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Palmetto GBA, Wisconsin Physicians Service Insurance Corporation, and CGS Administrators, LLC have issued draft noncoverage local coverage determination for the Pigmented Lesion Assay.⁵³⁻⁵⁵

Palmetto GBA has issued a draft local coverage determination for DecisionDx-Melanoma.⁵⁶ The comment period for the draft local coverage determination closes on May 10, 2018. The draft states that the quality of the evidence is "Moderate," the strength of evidence is "Low," and weight of evidence is "Low" and that: "This contractor will cover the DecisionDx-Melanoma test for patients diagnosed with SLNB eligible T1b and T2 tumor who are being considered for SLNB. The DecisionDx-Melanoma assay should not be ordered if a patient and his/her physician do not intend to act upon the test result. Continued coverage is dependent on the publication and/or presentation of additional clinical utility data demonstrating the impact of the test's use on patient management decisions with (1) 95% or greater DMFS [distant metastasis-free survival] and MSS [melanoma-specific survival] at 3 years in patients directed to no SLNB by the test compared to standard of care, and (2) evidence of higher SLNB positivity in patients selected for this procedure by the test compared to standard of care."

Ongoing and Unpublished Clinical Trials

Some currently unpublished trials that might influence this review are listed in Table 16.

Table 16. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
NCT02355574 ^a	An Ongoing, 5-year Post Market Study to Track Clinical Application of DecisionDx-Melanoma Gene Expression Profile (GEP) Assay Results and the Impact on Patient Outcomes and Health Economics	1672	Jun 2024
NCT02355587 ^a	An Open, 5-year Registry Study to Track Clinical Application of DecisionDx-Melanoma Gene Expression Profile Assay Results and Associated Patient Outcomes	5000	Feb 2024

NCT: national clinical trial.

^a Denotes industry-sponsored or cosponsored trial.

References

1. Chang AE, Karnell LH, Menck HR. The National Cancer Data Base report on cutaneous and noncutaneous melanoma: a summary of 84,836 cases from the past decade. The American College of Surgeons Commission on Cancer and the American Cancer Society. *Cancer*. Oct 15 1998;83(8):1664-1678. PMID 9781962
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *Jan 2018*;68(1):7-30. PMID 29313949

3. Gilchrest BA, Eller MS, Geller AC, et al. The pathogenesis of melanoma induced by ultraviolet radiation. *N Engl J Med*. Apr 29 1999;340(17):1341-1348. PMID 10219070
4. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur J Cancer*. Sep 2005;41(14):2040-2059. PMID 16125929
5. Caini S, Gandini S, Sera F, et al. Meta-analysis of risk factors for cutaneous melanoma according to anatomical site and clinico-pathological variant. *Eur J Cancer*. Nov 2009;45(17):3054-3063. PMID 19545997
6. Goldstein AM, Chan M, Harland M, et al. Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. *J Med Genet*. Feb 2007;44(2):99-106. PMID 16905682
7. Wendt J, Rauscher S, Burgstaller-Muehlbacher S, et al. Human determinants and the role of melanocortin-1 receptor variants in melanoma risk independent of UV radiation exposure. *JAMA Dermatol*. Jul 1 2016;152(7):776-782. PMID 27050141
8. Wiesner T, Obenauf AC, Murali R, et al. Germline mutations in BAP1 predispose to melanocytic tumors. *Nat Genet*. Aug 28 2011;43(10):1018-1021. PMID 21874003
9. Chen T, Fallah M, Forsti A, et al. Risk of next melanoma in patients with familial and sporadic melanoma by number of previous melanomas. *JAMA Dermatol*. Jun 2015;151(6):607-615. PMID 25671687
10. Jiang AJ, Rambhatla PV, Eide MJ. Socioeconomic and lifestyle factors and melanoma: a systematic review. *Br J Dermatol*. Apr 2015;172(4):885-915. PMID 25354495
11. Abbasi NR, Shaw HM, Rigel DS, et al. Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. *Jama*. Dec 8 2004;292(22):2771-2776. PMID 15585738
12. Grob JJ, Bonerandi JJ. The 'ugly duckling' sign: identification of the common characteristics of nevi in an individual as a basis for melanoma screening. *Arch Dermatol*. Jan 1998;134(1):103-104. PMID 9449921
13. Wilson RL, Yentzer BA, Isom SP, et al. How good are US dermatologists at discriminating skin cancers? A number-needed-to-treat analysis. *J Dermatolog Treat*. Feb 2012;23(1):65-69. PMID 21756146
14. Gershenwald JES, R.A.; Hess, K.R.; et al. *Melanoma of the Skin*. Chicago, IL: American Joint Committee on Cancer; 2017.
15. Eggermont AM, Chiarion-Sileni V, Grob JJ, et al. Prolonged survival in stage III melanoma with ipilimumab adjuvant therap. *N Engl J Med*. Nov 10 2016;375(19):1845-1855. PMID 27717298
16. Weber J, Mandala M, Del Vecchio M, et al. Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *N Engl J Med*. Nov 9 2017;377(19):1824-1835. PMID 28891423
17. Long GV, Hauschild A, Santinami M, et al. Adjuvant dabrafenib plus trametinib in stage III BRAF-mutated melanomas. *N Engl J Med*. Nov 9 2017;377(19):1813-1823. PMID 28891408
18. National Center for Biotechnology Information. PRAME preferentially expressed antigen in melanoma. 2018; <https://www.ncbi.nlm.nih.gov/gene/23532>. Accessed March 30, 2018.
19. DermTech. Pigmented Lesion Assay: Non-invasive gene expression analysis of pigmented skin lesions. Performance and Development Notes. 2015; <http://dermtech.com/wp-content/uploads/2015/10/White-Paper-DermTech-Melanoma-Assay-.pdf>. Accessed March 16, 2018.
20. Wachsmann W, Morhenn V, Palmer T, et al. Noninvasive genomic detection of melanoma. *Br J Dermatol*. Apr 2011;164(4):797-806. PMID 21294715
21. Gerami P, Alsobbrook JP, 2nd, Palmer TJ, et al. Development of a novel noninvasive adhesive patch test for the evaluation of pigmented lesions of the skin. *J Am Acad Dermatol*. Aug 2014;71(2):237-244. PMID 24906614
22. Gerami P, Yao Z, Polsky D, et al. Development and validation of a noninvasive 2-gene molecular assay for cutaneous melanoma. *J Am Acad Dermatol*. Jan 2017;76(1):114-120 e112. PMID 27707590

23. Vestergaard ME, Macaskill P, Holt PE, et al. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. *Br J Dermatol*. Sep 2008;159(3):669-676. PMID 18616769
24. Murzaku EC, Hayan S, Rao BK. Methods and rates of dermoscopy usage: a cross-sectional survey of US dermatologists stratified by years in practice. *J Am Acad Dermatol*. Aug 2014;71(2):393-395. PMID 25037790
25. Engasser HC, Warshaw EM. Dermatoscopy use by US dermatologists: a cross-sectional survey. *J Am Acad Dermatol*. Sep 2010;63(3):412-419, 419.e411-412. PMID 20619490
26. Bossuyt PM, Irwig L, Craig J, et al. Comparative accuracy: assessing new tests against existing diagnostic pathways. *Bmj*. May 6 2006;332(7549):1089-1092. PMID 16675820
27. Ferris LK, Jansen B, Ho J, et al. Utility of a noninvasive 2-gene molecular assay for cutaneous melanoma and effect on the decision to biopsy. *JAMA Dermatol*. Jul 1 2017;153(7):675-680. PMID 28445578
28. Myriad. n.d. Understanding the myPath® Melanoma Results; <https://mypathmelanoma.com/about-mypath-melanoma/understanding-the-mypath-melanoma-results/>. Accessed March 30, 2018.
29. Clarke LE, Warf MB, Flake DD, 2nd, et al. Clinical validation of a gene expression signature that differentiates benign nevi from malignant melanoma. *J Cutan Pathol*. Apr 2015;42(4):244-252. PMID 25727210
30. Gaiser T, Kutzner H, Palmedo G, et al. Classifying ambiguous melanocytic lesions with FISH and correlation with clinical long-term follow up. *Mod Pathol*. Mar 2010;23(3):413-419. PMID 20081813
31. Vergier B, Prochazkova-Carlotti M, de la Fouchardiere A, et al. Fluorescence in situ hybridization, a diagnostic aid in ambiguous melanocytic tumors: European study of 113 cases. *Mod Pathol*. May 2011;24(5):613-623. PMID 21151100
32. Clarke LE, Flake DD, 2nd, Busam K, et al. An independent validation of a gene expression signature to differentiate malignant melanoma from benign melanocytic nevi. *Cancer*. Feb 15 2017;123(4):617-628. PMID 27768230
33. Ko JS, Matharoo-Ball B, Billings SD, et al. Diagnostic distinction of malignant melanoma and benign nevi by a gene expression signature and correlation to clinical outcomes. *Cancer Epidemiol Biomarkers Prev*. Jul 2017;26(7):1107-1113. PMID 28377414
34. Cockerell C, Tschen J, Billings SD, et al. The influence of a gene-expression signature on the treatment of diagnostically challenging melanocytic lesions. *Per Med*. Mar 2017;14(2):123-130. PMID 28757886
35. Cockerell CJ, Tschen J, Evans B, et al. The influence of a gene expression signature on the diagnosis and recommended treatment of melanocytic tumors by dermatopathologists. *Medicine (Baltimore)*. Oct 2016;95(40):e4887. PMID 27749545
36. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Melanoma. Version 2.2018. 2018; https://www.nccn.org/professionals/physician_gls/pdf/melanoma.pdf. Accessed March 23, 2018.
37. Castle Biosciences. Cutaneous Melanoma: DecisionDx-Melanoma Overview. n.d.; <http://castlebiosciences.com/tests/cutaneous-melanoma/>. Accessed March 30, 2018.
38. Gerami P, Cook RW, Wilkinson J, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. *Clin Cancer Res*. Jan 1 2015;21(1):175-183. PMID 25564571
39. Zager JS, Gastman BR, Leachman S, et al. Performance of a prognostic 31-gene expression profile in an independent cohort of 523 cutaneous melanoma patients. *BMC Cancer*. Feb 5 2018;18(1):130. PMID 29402264
40. Wong SL, Balch CM, Hurley P, et al. Sentinel lymph node biopsy for melanoma: American Society of Clinical Oncology and Society of Surgical Oncology joint clinical practice guideline. *J Clin Oncol*. Aug 10 2012;30(23):2912-2918. PMID 22778321
41. Wrightson WR, Wong SL, Edwards MJ, et al. Complications associated with sentinel lymph node biopsy for melanoma. *Ann Surg Oncol*. Jul 2003;10(6):676-680. PMID 12839853
42. Soong SJ, Ding S, Coit DG, et al. AJCC: Individualized melanoma patient outcome prediction tools. n.d.; <http://www.melanomaprognosis.net/>. Accessed March 21, 2018.

- Documentation for Clinical Review

- Reproduction without authorization from Blue Shield of California is prohibited

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. Inclusion or exclusion of codes does not constitute or imply member coverage or provider reimbursement.

IE

The following services may be considered investigational.

Type	Code	Description
CPT®	81401	Molecular pathology procedure level 2
	81479	Unlisted molecular pathology procedure
	81599	Unlisted multianalyte assay with algorithmic analysis
	84999	Unlisted chemistry procedure
HCPCS	None	
ICD-10 Procedure	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	Reason
07/01/2018	BCBSA Medical Policy adoption	Medical Policy Committee

Definitions of Decision Determinations

Medically Necessary: A treatment, procedure, or drug is medically necessary only when it has been established as safe and effective for the particular symptoms or diagnosis, is not investigational or experimental, is not being provided primarily for the convenience of the patient or the provider, and is provided at the most appropriate level to treat the condition.

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. Please call (800) 541-6652 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.