Multibiomarker Disease Activity Blood Test for Rheumatoid Arthritis

Policy # 00442
Original Effective Date: 08/20/2014
Current Effective Date: 09/19/2018

Applies to all products administered or underwritten by Blue Cross and Blue Shield of Louisiana and its subsidiary, HMO Louisiana, Inc.(collectively referred to as the "Company"), unless otherwise provided in the applicable contract. Medical technology is constantly evolving, and we reserve the right to review and update Medical Policy periodically.

Services Are Considered Investigational
Coverage is not available for investigational medical treatments or procedures, drugs, devices or biological products.

Based on review of available data, the Company considers the use of a multi-biomarker disease activity score for rheumatoid arthritis (RA) (e.g., Vectra®‡ DA score) in all situations to be **investigational.***

Background/Overview

RHEUMATOID ARTHRITIS

Rheumatoid arthritis is characterized by chronic joint inflammation leading to painful symptoms, progressive joint destruction, and loss of function. The disorder is relatively common and associated with a high burden of morbidity for affected patients.

Treatment

Treatment of RA has undergone a shift from symptom management to a more proactive strategy of minimizing disease activity and delaying disease progression. The goal of treatment is to reduce the irreversible joint damage that occurs from ongoing joint inflammation and synovitis by keeping disease activity as low as possible. The availability of an increasing number of effective disease-modifying antirheumatic drugs has made the achievement of remission, or sustained low disease activity, a feasible goal for a large proportion of patients with RA. This treatment strategy has been called a *tight control* approach.

The concept of tight control in the management of RA has gained wide acceptance. Evidence from clinical trials has demonstrated that outcomes are improved with a tight control strategy, in which treatment targets are mainly based on measures of disease activity. In a systematic review, Schoels et al (2010) identified 7 studies that evaluated the efficacy of tight control. Four of these trials randomized patients to tight control using treatment targets or to routine management, 2 studies compared different treatment targets, and 1 study compared results from a targeted treatment with historical controls. The treatment targets were heterogeneous, including symptom-based measures, joint scores on the exam, validated treatment activity measures, lab values, or combinations of these factors. In all 4 trials that randomized patients to tight control or routine management, there was a significant decrease in the Disease Activity Score (DAS) or its 28 joints version (DAS28) and in the likelihood of achieving remission for patients in the tight control group.

Multibiomarker Disease Activity Blood Test for Rheumatoid Arthritis

Policy # 00442
Original Effective Date: 08/20/2014
Current Effective Date: 09/19/2018

According to American College of Rheumatology (ACR) guidelines, initial treatment of patients with RA is monotherapy (usually a disease-modifying antirheumatic drug). Treatment may progress to combination therapy if disease activity remains moderate or high despite monotherapy. Combination therapy may consist of additional disease-modifying antirheumatic drugs or the addition of tumor necrosis factors or non-tumor necrosis factors biologics.

Validated Disease Activity Assessment Tools
For a strategy of tight control to be successful, a reliable and valid measurement of disease activity is necessary. There are numerous disease activity measurements that can be used in clinical care.

Through a 5-stage process that included review by an expert advisory panel in RA disease activity and detailed evaluation of psychometric properties, an ACR working group determined that 6 measures were accurate reflections of disease activity: Clinical Disease Activity Index (CDAI), DAS28, Patient Activity Scale, Patient Activity Scale II, Routine Assessment of Patient Index Data 3, and the Simplified Disease Activity Index (SDAI).

Two systematic reviews were published the same year as the ACR's recommendations, one by Gaujoux-Viala et al (2012) and the other by Salaffi et al (2012), which compared disease activity measures for patients with RA. Results from the systematic reviews were consistent with the ACR working group recommendations, citing the DAS28, SDAI, and CDAI as appropriate disease activity measures for RA.

Table 1 summarizes the clinical and laboratory measurements included in each of the 6 disease activity measures recommended by ACR. The table also includes the laboratory measures included in the Vectra DA, a multi-biomarker disease activity (MBDA) test which currently does not have a recommendation from ACR.

Table 1. Clinical and Laboratory Components of Rheumatoid Arthritis Disease Activity Measurements

Recommended by ACR					No ACR Recommendation
DAS28	**CDAI and SDAI**	**PAS**	**PAS II**	**RAPID3**	**Vectra DA**
No. of swollen joints out of 28[a]	No. of swollen joints out of 28[a]	Patient describes ability to do each of 20 activities[b] as "without any difficulty," "with some difficulty," "with much difficulty," or "unable to do"	Patient describes ability to do each of 10 activities[c] as "without any difficulty," "with some difficulty," "with much difficulty," or "unable to do"	Patient describes ability to do each of 13 activities[d] as "without any difficulty," "with some difficulty," "with much difficulty," or "unable to do"	• Interleukin-6 • Tumor necrosis factor receptor type I • Vascular cell adhesion molecule 1 • Epidermal growth factor • Vascular endothelial growth factor A • YKL-40 glycoprotein • MMP-1 • MMP-3 • C-reactive protein

					• Serum amyloid A • Leptin • Resistin
No. of tender joints out of 28[a]	No. of tender joints out of 28[a]	Patient indicates need for cane, crutches, walker, wheelchair, or devices to assist with dressing or eating	Patient rates pain on scale of 0 (no pain) to 10 (severe pain)	Patient rates pain on scale of 0 (no pain) to 10 (severe pain)	
ESR (mm/h)	CRP (mg/L) (only in the SDAI, not part of CDAI calculation)	Patient indicates need for assistance in dressing, rising, eating, walking, hygiene, reaching, gripping, or chores	Patient rates how they are doing on scale of 0 (very well) to 10 (very poor)	Patient rates how they are doing on scale of 0 (very well) to 10 (very poor)	
CRP (mg/L)	Patient Global Assessment (0 [very well] to 10 [very poor])	Patient indicates if special devices needed in bathroom or kitchen			
Patient Global Assessment (0 [best] to 100 [worst])	Physician Global Assessment (0 [very well] to 10 [very poor])	Patient rates pain on scale of 0 (no pain) to 10 (severe pain)			
		Patient rates how they are doing on scale of 0 (very well) to 10 (very poor)			

Adapted by Anderson et al (2012).
ACR: American College of Rheumatology; CDAI: Clinical Disease Activity Index; CRP: C-reactive protein; DAS28: Disease Activity Score 28; ESR: erythrocyte sedimentation rate; MMP: matrix metalloproteinase; PAS: Patient Activity Scale; RAPID3: Routine Assessment of Patient Index Data 3; SDAI: Simplified Disease Activity Index.
[a] Twenty-eight joints: shoulders, elbows, wrists, metacarpophalangeal joints, proximal interphalangeal joints, and knees.
[b] Dress self; shampoo hair; stand from chair; get in and out of bed; cut meat; bring cup to mouth; open milk carton; walk outdoors on flat ground; climb 5 steps; wash and dry body; take tub bath; get on and off toilet; reach and bring down 5 pound object from above head; bend and pick up clothing from floor; open car door; open new jar; turn faucets on and off; run errands; get in and out of car; do chores (eg, vacuum or yard work).
[c] Stand from chair; walk outdoors on flat ground; get on and off toilet; reach and bring down 5 pound object from above head; open car door; do outside work such as yard work; wait in line for 15 minutes; lift heavy objects; move heavy objects; climb 2 or more flights of stairs.
[d] Dress self; get in and out of bed; bring cup to mouth; walk outdoors on flat ground; wash and dry body; bend and pick up clothing from floor; turn faucets on and off; get in and out of car; walk 2 miles; participate in recreational activities; sleep well; deal with feelings of anxiety or nervousness; deal with feelings of depression or sadness.

Vectra DA Test

The manufacturer describes Vectra DA as a complement to clinical judgment. Although not explicitly stated, it appears that the test may be used as an adjunct to other disease activity measures, to potentially identify patients at high risk of progression who would, therefore, benefit from a more aggressive treatment strategy.

The Vectra DA test scores range from 1 to 100. Categories of scores were constructed to correlate with the DAS28-CRP scale:

- 45-100: high disease activity
- 30-44: moderate disease activity
- 1-29: low disease activity.

FDA or Other Governmental Regulatory Approval

U.S. Food and Drug Administration (FDA)
Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests (LDTs) must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). The Vectra DA test (Crescendo Bioscience, South San Francisco, CA) is available under the auspices of CLIA. Laboratories that offer LDTs must be licensed by CLIA for high-complexity testing. To date, the U.S. FDA has chosen not to require any regulatory review of this test.

Centers for Medicare and Medicaid Services (CMS)
There are no Medicare National Coverage Determinations for the Vectra DA test. In July 2013, Palmetto GBA, the Medicare contractor in California, issued a positive coverage decision for the Vectra DA test. Because all Vectra DA tests are processed out of the Crescendo Bioscience laboratory in California, the test will be covered for Medicare patients in the United States.

Rationale/Source

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.

VECTRA DA TESTING FOR DISEASE ACTIVITY IN RHEUMATOID ARTHRITIS

Clinical Context and Test Purpose

The purpose of the MBDA, specifically the Vectra DA, in patients who have RA is to determine the level of disease activity (low, medium, or high) in order to inform treatment decisions.

The question addressed in this evidence review is: Does use of a MBDA (eg, Vectra DA) test, alone or as an adjunct, to predict disease activity in patients with RA, improve health outcomes compared with use of ACR–recommended measures of disease activity?

Multibiomarker Disease Activity Blood Test for Rheumatoid Arthritis

Policy # 00442
Original Effective Date: 08/20/2014
Current Effective Date: 09/19/2018

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

Patients
The relevant population of interest is individuals with RA who are being managed with a disease-modifying antirheumatic drug (DMARD).

Management of patients with RA has changed from treatment of symptoms to a tight control strategy. The objective of a tight control strategy is to minimize disease progression and joint damage by monitoring disease activity and treating aggressively if an increase in activity is predicted.

Interventions
Vectra DA provides a score indicating the level of disease activity, based on blood levels of the following 12 biomarkers: interleukin-6, tumor necrosis factor (TNF) receptor type I, vascular cell adhesion molecule 1, epidermal growth factor, vascular endothelial growth factor A, YKL-40 glycoprotein, matrix metalloproteinase 1, matrix metalloproteinase 3, C-reactive protein (CRP), serum amyloid A, leptin, and resistin.

Scores range from 1 to 100 (1-29=low disease activity; 30-44=medium disease activity; 45-100=high disease activity).

Comparators
The reference standard for disease activity is radiographic progression at a set point in time, typically 3 months to 1 year. In addition, an ACR expert panel on RA determined the following 6 disease activity measures were useful and feasible to implement in a clinical setting: Clinical Disease Activity Index (CDAI), DAS28, Patient Activity Scale, Patient Activity Scale II, Routine Assessment of Patient Index Data 3, and Simplified Disease Activity Index.

Outcomes
The goal of treating patients with RA is to improve quality of life and to prevent progression of the disease. Progression of disease causes irreversible joint damage.
If Vectra DA correctly assesses disease activity as low, the clinician may maintain medications at the same level or consider tapering the patient's medication.

If Vectra DA correctly assesses disease activity as moderate or high, the clinician may be more aggressive in disease management, by either increasing doses of current medications, switching medications, or adding medications to the treatment plan.

If Vectra DA incorrectly assesses disease activity as low, the clinician may maintain or decrease medication levels, which will allow progression of the disease and further joint damage.

If Vectra DA incorrectly assesses disease activity as moderate or high, the clinician may continue to manage the patient with higher levels of medication than is necessary to prevent disease progression, exposing the patient to unnecessary toxins. DMARDs may affect the liver, stomach, and intestines. Biologic agents may increase the risk of infection, lymphoma, and skin cancer.

Timing
The test may be run as often as a clinician needs disease activity information, typically every 3 to 6 months. A test immediately after diagnosis may serve as a baseline measurement.

For purposes of assessing Vectra DA against the reference standard of radiographic progression, 1 year is the typical time frame.

Setting
The test may be given at an outpatient rheumatology practice or an academic or community setting. Primary care may be the main source of care in some locations.

Study Selection Criteria
For the evaluation of the clinical utility of a multibiomarker disease activity test (eg, Vectra DA), studies would need to use the test as either an adjunct or a replacement to current disease activity measures to manage treatment decisions in patients with RA. Outcomes would be quality of life and measures of disease progression.

In the absence of direct evidence for the clinical utility of Vectra DA, evidence for clinical validity is evaluated, in which we can make inferences on clinical utility. For the evaluation of clinical validity, studies would need to compare Vectra DA used as an adjunct or as a replacement to ACR-recommended disease activity measures, with radiographic progression as a reference standard. Key validity outcomes of sensitivity, specificity, as well as positive (PPV) and negative (NPV) predictive values, should be reported.

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).

All evidence identified used the Vectra DA test. Eleven publications using data and serum samples from 8 studies met selection criteria and are included in this review. Table 2 summarizes study characteristics of the publications:

- Six studies (8 publications) included records and archived samples from randomized controlled trials (RCTs).
- One study (2 publications) included records and samples from a cohort study.
- One study included records and samples from a single-arm study.
- One study (2 publications) used random (weighted) samples; the remaining were convenience samples.
- All studies were retrospective analyses. Prospective protocols precluding the inference that these were nonconcurrent prospective studies could not be identified (as described by Simon et al [2009]).
- Reference standard for most studies was radiographic progression, generally at 1 year. The definition of radiographic progression varied among studies. Some studies also included moderate radiographic progression and rapid radiographic progression.
- Eight publications used thresholds of low (<30), moderate (30–44), and high (>44). One study used a threshold of remission (≤25) or no remission (>25). Two publications analyzed Vectra as a continuous score.
- Outcome assessment was blinded in 3 publications, with no report on blinding of assessors in the remaining 8 publications.

Table 3 provides clinical validity results for the studies. Below are select key points from the results:

- Samples included in the publications ranged from 52 to 524 patients.
- Only 1 publication reported sensitivity and specificity data. Hambardzumyan et al (2015) reported that the sensitivity for high Vectra risk was 98%, specificity was 17%, NPV was 97%, and PPV was 21%. The low specificity and PPV do not support the use of the test to “rule in” high-risk disease.
- Four studies reported area under the receiver operating characteristic curve (AUROC) data, and another reported positive likelihood ratios. Seven publications reported the percentage of patients progressing by Vectra class.
- In 2 publications, the AUROC for the Vectra test was numerically higher than the DAS28; however, the confidence intervals overlapped. Overlapping confidence intervals indicate uncertainty whether Vectra DA provides prognostic performance superior to DAS28.
- In 3 studies, Vectra scores were not associated with radiographic progression or rapid radiographic progression, and in another, Vectra correlated with radiographic progression at 1 time point (26 weeks), but not at baseline or 1 year.
- One study reported a higher positive likelihood ratio for Vectra compared with DAS28, again with overlapping confidence intervals.
- One study reported that significantly more patients with low Vectra scores responded to triple therapy compared with anti-TNF therapy, while significantly more patients with high Vectra scores responded to anti-TNF therapy compared with triple therapy.

Table 2. Study Characteristics of Included Studies

Study	Study Population	Design	Reference Standard (Time From Test to RP)	Threshold for Positive Index Test	Blinding of Assessors	Comment
Bakker et al (2012)[9]	CAMERA, early RA	Retrospective; convenience	RP at 2 y	NA: Continuous scores used in analysis	NR	Patients randomized to intensive or conventional treatment
Van der Helm-van Mil et al (2013)[17]	Leiden EAC cohort with symptoms <2 y; samples collected between 1995 and 2005	Retrospective; random (weighted sample)	Moderate RP and RRP at 1 y	Remission (≤25) vs not remission (>25)	Yes	Infrequent use of anti-TNFs in this population
Markusse et al (2014)[10]	BeST patients with symptoms <2 y	Retrospective; convenience	RP and RRP after 1 y	NA: Continuous scores used in analysis	NR	Same samples as Hirata et al (2013)[21]
Hambardzumyan et al (2015)[11]	SWEFOT, DMARD-naive	Retrospective; convenience	RRP at 1 y	Low (<30), moderate (30-44), and high (>44)	NR	Patients treatment with MTX until 3 mo, nonresponders randomized to MTX plus triple therapy or MTX plus infliximab
Hambardzumyan et al (2016)[12]	SWEFOT, DMARD-naive	Retrospective; convenience	RRP at 2 y	Low (<30), moderate (30-44), and high (>44)	NR	Overlapping samples with Hambardzumyan et al (2015); does not explain the differing n's
Fleischmann et al (2016)[13]	AMPLE, biologic-naive, RA ≤5 y, inadequate response to MTX	Retrospective; convenience	RP at 1 y	Low (<30), moderate (30-44), and high (>44)	Yes	Patients randomized to MTX plus abatacept or MTX plus adalimumab
Li et al (2016)[18]	Leiden EAC cohort with symptoms <2 y; samples collected between 1995 and 2005	Retrospective; random (weighted sample)	Moderate RP and RRP at 1 y	Low (<30), moderate (30-44), and high (>44)	NR	Overlap with van der Helm et al (2013)[17]
Hirata et al (2016)[14]	Patients treated with TNF inhibitor for ≥1 y at a single institution	Retrospective; convenience	Clinically relevant RP; RRP at 1 y	Low (<30), moderate (30-44), and high (>44)	Yes	Overlap with Hirata et al (2015)[22]
Bouman et al (2017)[15]	DRESS, RCT of tapering TNF inhibitors until discontinuation or flaring vs usual care	Retrospective, convenience	RP at 18 mo	Low (<30), moderate (30-44), and high (>44)	NR	Flare defined as increase in DAS28-CRP >1.2 vs baseline or increase in DAS28-CRP >0.6 vs baseline plus current DAS28 ≥3.2
Hambardzumyan et al (2017)[16]	SWEFOT, inadequate responders to MTX at 3 mo	Retrospective, convenience	EULAR criteria for response to treatment	Low (<30), moderate (30-44), and high (>44)	NR	May overlap with Hambardzumyan et al (2015) [11]
Krabbe et al (2017)[19]	Patients treated with TNF inhibitor for 1 y at a single institution	Retrospective, single arm	RP at 1 y	Remission (≤25), low (26-29), moderate (30-44), and high >44)	NR	RP defined by MRI synovitis, MRI bone marrow edema, US synovial PD score; and US GSS

AMPLE: Abatacept versus Adalimumab Comparison in Biologic-Naive RA Subjects with Background Methotrexate; BeST: Behandel Strategieën; CAMERA: Computer Assisted Management in Early Rheumatoid Arthritis; CRP: C-reactive protein; DAS28: Disease Activity Score with 28 joints; DMARD: disease-modifying antirheumatic drug; DRESS: Dose REduction Strategies of Subcutaneous TNF Inhibitors trial; EAC: Early Arthritis Clinic; EULAR: European League Against Rheumatism; GSS: grey scale synovitis; MRI: magnetic resonance imaging; MTX: methotrexate; NA: not applicable; NR: not reported; PD; power Doppler; RA: rheumatoid arthritis; RCT: randomized controlled trial; RP: radiographic progression; RRP: rapid radiographic progression; SWEFOT: Swedish Farmacotherapy; TNF: tumor necrosis factor; US: ultrasound.

Table 3. Clinical Validity Results for the Vectra DA Test

Study	Initial N	Final N	Excluded Samples	Prevalence of Condition	Clinical Validity: Risk Outcome, %			Other Reported Measures (95% CI)
					Low	Medium	High	
Bakker et al (2012)[9]	NR for CAMERA	120 samples (72 at BL, 48 at 6 mo), not clear if overlapping	Serum unavailable	NR	NR	NR	NR	Vectra: BL or 6 mo not associated with RP in multivariate analyses DAS28-CRP: AUC=0.86 (p<0.001)
Van der Helm-van Mil et al (2013)[17]	NR	163 patients (271 samples)	Not selected by sampling	Moderate RP=26%	Vectra: 7%[b] DAS28: 20%[b]		Vectra: 30%[b] DAS28: 29%[b]	Vectra PLR=4.7 (1.7 to 15.0) DAS28 PLR=1.4 (0.9 to 2.4)
Markusse et al (2014)[10]	508 in BeST	125 (91 at BL; 89 at 1 y); 84 with BL serum and 1-y radiograph	Missing or insufficient samples	RP=37%	NR	NR	NR	RP: • Vectra AUC=0.61 (0.48 to 0.73) • DAS AUC=0.37 (0.25 to 0.50) RRP: • Vectra AUC=0.77 (0.64 to 0.90) • DAS AUC=0.52 (0.36 to 0.68)
Hambardzumyan et al (2015)[11]	487 in SWEFOT	235 with complete BL demographic, serologic, radiographic, and clinical data	Incomplete data	RRP=18%	Vectra: 0% DAS28: NA	Vectra: 3% DAS28: 20%[a]	Vectra: 21% DAS28: 21%[a]	Vectra low/mod vs high: • Sensitivity: 98% • Specificity: 17% • NPV: 97% • PPV: 21%
Hambardzumyan et al (2016)[12]	487 in SWEFOT	220 patients with Vectra DA scores, CRP, ESR, and DAS28 at BL, 205 at 3 mo, 133 at 1 y	Incomplete data	RRP=30%	Vectra: • 3 mo: 9% • 1 y: 3% DAS28: • 3 mo: 24% • 1 y: 6%	Vectra: • 3 mo: 26% • 1 y: 8% DAS28: • 3 mo: 32% • 1 y: 15%	Vectra: • 3 mo: 41% • 1 y: 32% DAS28: • 3 mo: 35% • 1 y: 38%	No measures of sensitivity, specificity, or other performance characteristics reported
Fleischmann et al (2016)[13]	646 in AMPLE	524 with available data	Unavailable data	RP=30%	Vectra[a]: • ABA group: 79% • ADM group: 77% CDAI[a]: • ABA group: 64% • ADM group: 65%	Vectra[a]: • ABA group: 67% • ADM group: 55% CDAI[a]: • ABA group: 74% • ADM group: 79%	Vectra[a]: • ABA group: 63% • ADM group: 80% CDAI[a]: • ABA group: 91% • ADM group: 90%	No association between Vectra classes and RP

Li et al (2016)[18]	**NR**	**163 patients (271 samples)**	**Not selected by weighted sampling**	**Mod RP=26%; RRP=17%**	**• Vectra mod: • RP=10% • RRP=2%**	**• Vectra mod: • RP, 15%-19% • RRP, 7%-8%**	**• Vectra mod: • RP, 33%-51% • RRP, 28%-41%**	**Vectra: • AUROC mod RP=0.72 (0.64 to 0.78) • AUROC RRP=0.77 (0.69 to 0.83) DAS28: • AUROC mod RP=0.59 (0.51 to 0.67) • AUROC RRP=0.66 (0.56 to 0.75)**
Hirata et al (2016)[14]	NR	83	Patients without data at 24 wk	Clinically relevant RP=12%	• Vectra: 0% • DAS28: 0%	• Vectra: 4% • DAS28: 18%	• Vectra: 28% • DAS28: 36%	No measures of sensitivity, specificity, or other performance characteristics reported
Bouman et al (2017)[15]	171	167 (115 randomized to dose tapering, 59 randomized to usual care)	Patients without both serum samples and 18-mo radiographs	RP=26%	• RP occurred in 31% in the dose-tapering group and in 16% in the usual care group • MBDA score was not predictive of successful tapering, flare occurrence, or RP			• AUROC tapering: 0.53 (0.41 to 0.66) • AUROC flare: 0.50 (0.41 to 0.59) • AUROC RP: 0.53 (0.43 to 0.63)
Hambardzumyan et al (2017)[16]	157	157 (75 randomized to triple therapy, 82 randomized to anti-TNF therapy)	No exclusions	Responders to triple therapy 47% and responders to anti-TNF therapy was 54%	19 responded to triple therapy (88%) more than to anti-TNF therapy (18%; p=0.006)	50 responded equally to triple therapy (54%) and anti-TNF therapy (62%)	88 responded to anti-TNF therapy (58%) more than to triple therapy (35%; p=0.04)	DAS28 measure at 3 mo not associated with response to particular therapy
Krabbe et al (2017)[19]	52	• Week 10, n=46 • Week 20, n=43 • Week 30, n=42 • Week 40, n=35 • Week 50, n=35 • Week 60, n=33	Patients with missing data	RP=19%	In 10 with disease progression, 0 had low Vectra score	In 10 with disease progression, 3 had mod Vectra score	In 10 with disease progression, 7 had high Vectra score	Vectra correlated poorly with MRI/US at BL and 52 wk; Vectra correlated well with MRI/US at 26 wk

ABA: abatacept; ADM: adalimumab; AMPLE: Abatacept versus Adalimumab Comparison in Biologic-Naive RA Subjects with Background Methotrexate; AUC: area under the curve; BeST: Behandel Strategieën; BL: baseline; CAMERA: Computer Assisted Management in Early Rheumatoid Arthritis; BL: baseline; CRP: C-reactive protein; CDAI: Clinical Disease Activity Index; DAS28: 28-joint Disease Activity Score; ESR: erythrocyte sedimentation rate; MBDA: multibiomarker disease activity; Mod: moderate; MRI: magnetic resonance imaging; NA: not available; NPV: negative predictive value; NR: not reported; PLR: positive likelihood ratio; PPV: positive predictive value; RP: radiographic progression; RRP: rapid radiographic progression; SWEFOT: Swedish Farmacotherapy; TNF: tumor necrosis factor; US: ultrasound.
[a] Estimated from figure.
[b] "Low" risk is remission and "high"' risk is not remission.

Gaps Assessment for Clinical Validity

The purpose of the gaps tables (see Tables 4 and 5) is to display notable gaps identified in each study. This information is synthesized as a summary of the body of evidence following each table and provides the conclusions on the sufficiency of evidence supporting the position statement.

Table 4. Relevance Gaps for the Validity Studies

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of FU[e]
Bakker et al (2012)		3. Not consistent with current use, which is as an adjunct to other disease activity measures		3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); only AUROC reported	
Van der Helm-van Mil et al (2013)		3. Not consistent with current use, which is as an adjunct to other disease activity measures		3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); risk of remission reported	
Markusse et al (2014)		3. Not consistent with current use, which is as		3. Key clinical validity outcomes not reported	

	an adjunct to other disease activity measures	(sensitivity, specificity, predictive values); only AUROC reported
Hambardzumyan et al (2015)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	
Hambardzumyan et al (2016)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); comparisons to other biomarkers reported
Fleischmann et al (2016)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); correlations with other biomarkers reported
Li et al (2016)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); correlations with radiographic progression reported
Hirata et al (2016)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); correlations with radiographic progression reported
Bouman et al (2017)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); AUROC reported
Hambardzumyan et al (2017)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); comparisons with DAS28 reported
Krabbe et al (2017)	3. Not consistent with current use, which is as an adjunct to other disease activity measures	3. Key clinical validity outcomes not reported (sensitivity, specificity, predictive values); correlations with radiographic progression reported

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

AUROC: area under the receiver operating curve; DAS28: Disease Activity Score with 28 joints; FU: follow-up.
[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 5. Study Design and Conduct Gaps for the Validity Studies

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Bakker et al (2012)	2. Selection not random or consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Van der Helm-van Mil et al (2013)						
Markusse et al (2014)	2. Selection not random or consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Hambardzumyan et al (2015)	2. Selection not random or consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Hambardzumyan et al (2016)	2. Selection not random or consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Fleischmann et al (2016)	2. Selection not random or consecutive; convenience serum samples from an RCT					2. No statistical test reported to compare MBDA vs alternatives, only tests comparing treatment groups

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
\Li et al (2016)		1. Blinding of assessors not reported				
Hirata et al (2016)	2. Selection not random nor consecutive; convenience serum samples from an RCT					
Bouman et al (2017)	2. Selection not random nor consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Hambardzumyan et al (2017)	2. Selection not random or consecutive; convenience serum samples from an RCT	1. Blinding of assessors not reported				
Krabbe et al (2017)	2. Selection not described; single-arm study with no explanation of how patients recruited	1. Blinding of assessors not reported	4. Expertise of evaluators not described		3. High loss of follow-up: 14% at 20 wk and 30% at 50 wk	1. P values only reported for some comparisons

The evidence gaps stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
FU: follow-up; MBDA: multibiomarker disease activity; RCT: randomized controlled trial.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (ie, 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.

Subsection Summary: Gaps Assessment for Clinical Validity

Limitations in the body of evidence for clinical validity are identified in Tables 4 and 5. Relating to the relevance of the studies, only 1 of the 9 studies calculated sensitivity, specificity, and positive and negative predictive values. Most reported area under the curve measures and/or correlations with other disease activity measures. None evaluated MBDA as an adjunct to other disease activity measures, which is how the MBDA is currently being marketed.

The study design and conduct gaps table shows that 8 studies used convenience samples of serum from RCTs and one was a single-arm study with no explanation of how patients were recruited or enrolled. Eight studies did not report whether the radiographic assessors were blinded to the biomarker results.

Section Summary: Clinically Valid

Evidence for the clinical validity of the MBDA test consists of analyses of archived serum samples from RCTs as well as prospective cohort studies that have correlated MBDA with other measures of disease activity and with radiographic progression. Results from studies comparing MBDA with other disease activity measures have shown a positive correlation; however, results from studies comparing MBDA with radiographic progression are inconsistent. Only 1 study reported sensitivity and specificity, with a PPV of 21%, indicating that 4 out of 5 patients identified as positive would receive intensification of therapy unnecessarily.

Currently, MBDA is used as an adjunct to other disease activity measures. No evidence was identified that evaluated the incremental benefit of MBDA when used as an adjunct to other disease activity measures.

Overall, studies lacked reporting of sensitivity, specificity, and predictive values, which are the most informative measures for ascertaining the performance of Vectra DA in selecting high-risk patients for intensification of therapy. The evidence is insufficient to conclude the clinical validity of Vectra DA compared with ACR-recommended measures of disease activity.

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, 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 RCTs.

To demonstrate clinical utility, there should be evidence that the MBDA score is at least as good a measure of disease activity as other available measures or that the MBDA score demonstrates an incremental benefit when used as an adjunct with other disease activity measures. To demonstrate equivalence with other measures directly, an RCT comparing health outcomes of 2 groups, 1 group managed using the Vectra DA test and the other group managed by another disease activity measure is needed.

To directly demonstrate an incremental benefit when used as an adjunct, an RCT should compare health outcomes in patients receiving treatment guided by MBDA plus a disease activity measure with outcomes in patients receiving treatment guided only by the other disease activity measure.

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.

Because there is insufficient evidence that the MBDA score is clinically valid, direct evidence is needed to prove clinical utility. No trials were identified that provided direct evidence of clinical utility.

Section Summary: Clinically Useful
There are no RCTs comparing the use of the Vectra DA score with an alternative method of measuring disease activity. Additionally, there are no RCTs of Vectra DA as an adjunct to other disease activity measures compared with using the disease activity measures alone. Absent direct evidence for clinical utility, a chain of evidence could be constructed with indirect evidence proving clinical validity. However, there is insufficient evidence that MBDA is clinically valid.

SUMMARY OF EVIDENCE
For individuals who have rheumatoid arthritis who receive a MBDA (eg, Vectra DA) test as an adjunct or as a replacement of other disease activity measures, the evidence includes analyses of archived serum samples from RCTs and prospective cohort studies. Relevant outcomes are test validity, other test performance measures, symptoms, change in disease status, functional outcomes, and quality of life. Analyses comparing Vectra DA with other previously validated disease activity measures such as the DAS28 or to radiographic progression, consisted mostly of correlations, with only 1 study providing sensitivity, specificity, and positive and negative predictive values. The positive predictive value from this study was 21%. Other analyses of archived serum samples evaluated the use of Vectra DA to predict treatment response. Results from those analyses were inconsistent. The body of evidence on the Vectra DA test is insufficient to determine whether it is as good as or better than other disease activity measures. Additionally, there is no evidence evaluating Vectra DA as an adjunct to other disease activity measures. The evidence is insufficient to determine the effects of the technology on health outcomes.

References

1. Blue Cross and Blue Shield Association, Medical Policy Reference Manual, “Vectra DA Blood Test for Rheumatoid Arthritis”, 2.04.119, 6:2018.
2. Upchurch KS, Kay J. Evolution of treatment for rheumatoid arthritis. Rheumatology (Oxford). Dec 2012;51 Suppl 6:vi28-36. PMID 23221584
3. Schoels M, Knevel R, Aletaha D, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. Ann Rheum Dis. Apr 2010;69(4):638-643. PMID 20237123
4. Anderson J, Caplan L, Yazdany J, et al. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. Arthritis Care Res (Hoboken). May 2012;64(5):640-647. PMID 22473918
5. Gaujoux-Viala C, Mouterde G, Baillet A, et al. Evaluating disease activity in rheumatoid arthritis: which composite index is best? A systematic literature analysis of studies comparing the psychometric properties of the DAS, DAS28, SDAI and CDAI. Joint Bone Spine. Mar 2012;79(2):149-155. PMID 21680221
6. Salaffi F, Ciapetti A, Gasparini S, et al. The comparative responsiveness of the patient self-report questionnaires and composite disease indices for assessing rheumatoid arthritis activity in routine care. Clin Exp Rheumatol. Nov-Dec 2012;30(6):912-921. PMID 22935335

7. Curtis JR, van der Helm-van Mil AH, Knevel R, et al. Validation of a novel multibiomarker test to assess rheumatoid arthritis disease activity. Arthritis Care Res (Hoboken). Dec 2012;64(12):1794-1803. PMID 22736476
8. Eastman PS, Manning WC, Qureshi F, et al. Characterization of a multiplex, 12-biomarker test for rheumatoid arthritis. J Pharm Biomed Anal. Nov 2012;70:415-424. PMID 22749821
9. Centola M, Cavet G, Shen Y, et al. Development of a multi-biomarker disease activity test for rheumatoid arthritis. PLoS One. 2013;8(4):e60635. PMID 23585841
10. Markusse IM, Dirven L, van den Broek M, et al. A multibiomarker disease activity score for rheumatoid arthritis predicts radiographic joint damage in the BeSt study. J Rheumatol. Nov 2014;41(11):2114-2119. PMID 25128518
11. Bakker MF, Cavet G, Jacobs JW, et al. Performance of a multi-biomarker score measuring rheumatoid arthritis disease activity in the CAMERA tight control study. Ann Rheum Dis. Oct 2012;71(10):1692-1697. PMID 22596166
12. Hambardzumyan K, Bolce R, Saevarsdottir S, et al. Pretreatment multi-biomarker disease activity score and radiographic progression in early RA: results from the SWEFOT trial. Ann Rheum Dis. Jun 2015;74(6):1102-1109. PMID 24812287
13. Hambardzumyan K, Bolce RJ, Saevarsdottir S, et al. Association of a multibiomarker disease activity score at multiple time-points with radiographic progression in rheumatoid arthritis: results from the SWEFOT trial. RMD Open. 2016;2(1):e000197. PMID 26958364
14. Fleischmann R, Connolly SE, Maldonado MA, et al. Estimating disease activity using multi-biomarker disease activity scores in patients with rheumatoid arthritis treated with abatacept or adalimumab. Arthritis Rheumatol. Apr 25 2016. PMID 27111089
15. Rech J, Hueber AJ, Finzel S, et al. Prediction of disease relapses by multibiomarker disease activity and autoantibody status in patients with rheumatoid arthritis on tapering DMARD treatment. Ann Rheum Dis. Oct 19 2015. PMID 26483255
16. Li W, Sasso EH, van der Helm-van Mil AH, et al. Relationship of multi-biomarker disease activity score and other risk factors with radiographic progression in an observational study of patients with rheumatoid arthritis. Rheumatology (Oxford). Feb 2016;55(2):357-366. PMID 26385370
17. Hirata S, Li W, Defranoux N, et al. A multi-biomarker disease activity score tracks clinical response consistently in patients with rheumatoid arthritis treated with different anti-tumor necrosis factor therapies: A retrospective observational study. Mod Rheumatol. Oct 8 2014:1-6. PMID 25295918
18. Peabody JW, Strand V, Shimkhada R, et al. Impact of rheumatoid arthritis disease activity test on clinical practice. PLoS One. 2013;8(5):e63215. PMID 23667587
19. Li W, Sasso EH, Emerling D, et al. Impact of a multi-biomarker disease activity test on rheumatoid arthritis treatment decisions and therapy use. Curr Med Res Opin. Jan 2013;29(1):85-92. PMID 23176063
20. Davis JM, 3rd. Editorial: The Multi-Biomarker Disease Activity Test for rheumatoid arthritis: is it a valid measure of disease activity? Arthritis Rheumatol. Sep 2016;68(9):2061-2066. PMID 27111349
21. Fleischmann R, Connolly SE, Maldonado MA, et al. Reply. Arthritis Rheumatol. Apr 2017;69(4):867-868. PMID 27992708
22. Reiss WG, Devenport JN, Low JM, et al. Interpreting the multi-biomarker disease activity score in the context of tocilizumab treatment for patients with rheumatoid arthritis. Rheumatol Int. Feb 2016;36(2):295-300. PMID 26026604
23. Hirata S, Li W, Kubo S, et al. Association of the multi-biomarker disease activity score with joint destruction in patients with rheumatoid arthritis receiving tumor necrosis factor-alpha inhibitor treatment in clinical practice. Mod Rheumatol. Mar 30 2016:1-7. PMID 26873570
24. Hambardzumyan K, Saevarsdottir S, Forslind K, et al. A Multi-Biomarker Disease Activity Score and the choice of second-line therapy in early rheumatoid arthritis after methotrexate failure. Arthritis Rheumatol. May 2017;69(5):953-963. PMID 27992691
25. Bouman CAM, van der Maas A, van Herwaarden N, Sasso EH, van den Hoogen FHJ, den Broeder AA. A multi-biomarker score measuring disease activity in rheumatoid arthritis patients tapering adalimumab or etanercept: predictive value for clinical and radiographic outcomes. *Rheumatology (Oxford).* Jun 1 2017;56(6):973-980. PMID 28339738
26. Hambardzumyan K, Saevarsdottir S, Forslind K, et al. A Multi-Biomarker Disease Activity Score and the choice of second-line therapy in early rheumatoid arthritis after methotrexate failure. *Arthritis Rheumatol.* May 2017;69(5):953-963. PMID 27992691
27. Krabbe S, Bolce R, Brahe CH, et al. Investigation of a multi-biomarker disease activity score in rheumatoid arthritis by comparison with magnetic resonance imaging, computed tomography, ultrasonography, and radiography parameters of inflammation and damage. *Scand J Rheumatol.* Sep 2017;46(5):353-358. PMID 27682742
28. Combe B, Landewe R, Daien CI, et al. 2016 update of the EULAR recommendations for the management of early arthritis. *Ann Rheum Dis.* Jun 2017;76(6):948-959. PMID 27979873

Multibiomarker Disease Activity Blood Test for Rheumatoid Arthritis

Policy # 00442
Original Effective Date: 08/20/2014
Current Effective Date: 09/19/2018

Policy History

Original Effective Date: 08/20/2014
Current Effective Date: 09/19/2018
08/07/2014 Medical Policy Committee review
08/20/2014 Medical Policy Implementation Committee approval. New policy.
08/06/2015 Medical Policy Committee review
08/19/2015 Medical Policy Implementation Committee approval. No change to coverage.
09/08/2016 Medical Policy Committee review
09/21/2016 Medical Policy Implementation Committee approval. No change to coverage.
01/01/2017 Coding update: Removing ICD-9 Diagnosis Codes
09/07/2017 Medical Policy Committee review
09/20/2017 Medical Policy Implementation Committee approval. No change to coverage.
09/06/2018 Medical Policy Committee review
09/19/2018 Medical Policy Implementation Committee approval. No change to coverage. Title changed.
Next Scheduled Review Date: 09/2019

Coding

The five character codes included in the Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines are obtained from Current Procedural Terminology (CPT®)‡, copyright 2017 by the American Medical Association (AMA). CPT is developed by the AMA as a listing of descriptive terms and five character identifying codes and modifiers for reporting medical services and procedures performed by physician.

The responsibility for the content of Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines is with Blue Cross and Blue Shield of Louisiana and no endorsement by the AMA is intended or should be implied. The AMA disclaims responsibility for any consequences or liability attributable or related to any use, nonuse or interpretation of information contained in Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines. Fee schedules, relative value units, conversion factors and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein. Any use of CPT outside of Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines should refer to the most current Current Procedural Terminology which contains the complete and most current listing of CPT codes and descriptive terms. Applicable FARS/DFARS apply.

CPT is a registered trademark of the American Medical Association.

Codes used to identify services associated with this policy may include (but may not be limited to) the following:

Code Type	Code			
CPT	81490, 83520, 84999			
HCPCS	No codes			
ICD-10 Diagnosis	M05.00	M05.011-M05.079	M05.09-M05.10	M05.111-M05.179
	M05.19-M05.20	M05.211-M05.279	M05.29-M05.30	M05.311-M05.379
	M05.39-M05.40	M05.411-M05.479	M05.49-M05.50	M05.511-M05.579

	M05.59-M05.60	M05.611-M05.679	M05.69-M05.70	M05.711-M05.779
	M05.79-M05.80	M05.811-M05.879	M05.89-M05.9	M06.00
	M06.011-M06.079	M06.08-M06.09	M06.1	M06.20
	M06.211-M06.279	M06.28-M06.30	M06.311-M06.379	M06.38-M06.39
	M06.4	M06.80	M06.811-M06.879	M06.88-M06.89
	M06.9	M08.00	M08.011-M08.079	M08.08-M08.09
	M08.20	M08.211-M08.279	M08.28-M08.29	M08.3
	M08.40	M08.411-M08.479	M08.48	M08.80
	M08.811-M08.879	M08.88-M08.89	M08.90	M08.911-M08.979
	M08.98-M08.99	M12.00	M12.011-M12.079	M12.08-M12.09

*****Investigational – A medical treatment, procedure, drug, device, or biological product is Investigational if the effectiveness has not been clearly tested and it has not been incorporated into standard medical practice. Any determination we make that a medical treatment, procedure, drug, device, or biological product is Investigational will be based on a consideration of the following:

A. Whether the medical treatment, procedure, drug, device, or biological product can be lawfully marketed without approval of the U.S. FDA and whether such approval has been granted at the time the medical treatment, procedure, drug, device, or biological product is sought to be furnished; or
B. Whether the medical treatment, procedure, drug, device, or biological product requires further studies or clinical trials to determine its maximum tolerated dose, toxicity, safety, effectiveness, or effectiveness as compared with the standard means of treatment or diagnosis, must improve health outcomes, according to the consensus of opinion among experts as shown by reliable evidence, including:
    1. Consultation with the Blue Cross and Blue Shield Association technology assessment program (TEC) or other nonaffiliated technology evaluation center(s);
    2. Credible scientific evidence published in peer-reviewed medical literature generally recognized by the relevant medical community; or
    3. Reference to federal regulations.

‡ Indicated trademarks are the registered trademarks of their respective owners.

NOTICE: Medical Policies are scientific based opinions, provided solely for coverage and informational purposes. Medical Policies should not be construed to suggest that the Company recommends, advocates, requires, encourages, or discourages any particular treatment, procedure, or service, or any particular course of treatment, procedure, or service.