

Medical Policy



Title: Laboratory Tests Post Transplant and for Heart Failure

| Professional | Institutional |
|---|---|
| Original Effective Date: January 6, 2005 | Original Effective Date: March 12, 2013 |
| Revision Date(s): March 12, 2013; October 16, 2013; January 1, 2016; July 1, 2018; September 12, 2018; December 20, 2018; July 1, 2019; October 1, 2019; May 18, 2020; February 4, 2022 | Revision Date(s): October 16, 2013; January 1, 2016; July 1, 2018; September 12, 2018; December 20, 2018; July 1, 2019; October 1, 2019; May 18, 2020; February 4, 2022 |
| Current Effective Date: February 4, 2022 | Current Effective Date: February 4, 2022 |

State and Federal mandates and health plan member contract language, including specific provisions/exclusions, take precedence over Medical Policy and must be considered first in determining eligibility for coverage. To verify a member's benefits, contact [Blue Cross and Blue Shield of Kansas Customer Service](#).

The BCBSKS Medical Policies contained herein are for informational purposes and apply only to members who have health insurance through BCBSKS or who are covered by a self-insured group plan administered by BCBSKS. Medical Policy for FEP members is subject to FEP medical policy which may differ from BCBSKS Medical Policy.

The medical policies do not constitute medical advice or medical care. Treating health care providers are independent contractors and are neither employees nor agents of Blue Cross and Blue Shield of Kansas and are solely responsible for diagnosis, treatment and medical advice.

If your patient is covered under a different Blue Cross and Blue Shield plan, please refer to the Medical Policies of that plan.

| Populations | Interventions | Comparators | Outcomes |
|--|---|--|---|
| Individuals: • With chronic heart failure | Interventions of interest are: • Soluble suppression of tumorigenicity-2 assay to determine prognosis and/or to guide management | Comparators of interest are: • Standard prognostic markers, including B-type natriuretic peptide levels | Relevant outcomes include: • Overall survival • Quality of life • Hospitalizations |

| Populations | Interventions | Comparators | Outcomes |
|---|---|---|--|
| Individuals: • With heart transplant | Interventions of interest are: • Soluble suppression of tumorigenicity-2 assay to determine prognosis and/or to predict acute cellular rejection | Comparators of interest are: • Routine endomyocardial biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |
| Individuals: • With heart transplant | Interventions of interest are: • my TAI _{HEART} assay to determine prognosis and/or to predict acute cellular rejection | Comparators of interest are: • Routine endomyocardial biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events Hospitalizations |
| Individuals: • With heart transplant | Interventions of interest are: • Measurement of volatile organic compounds to assess cardiac allograft rejection | Comparators of interest are: • Routine endomyocardial biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |
| Individuals: • With heart transplant | Interventions of interest are: • Gene expression profiling to assess cardiac allograft rejection | Comparators of interest are: • Routine endomyocardial biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |
| Individuals: • With heart transplant | Interventions of interest are: • Gene expression with testing of donor-derived cell-free DNA in blood to assess cardiac allograft rejection | Comparators of interest are: • Routine endomyocardial biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |
| Individuals: • With renal transplant who are undergoing surveillance or have clinical suspicion of allograft rejection | Interventions of interest are: • Testing donor-derived cell-free DNA in blood to assess renal allograft rejection | Comparators of interest are: • Renal biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |
| Individuals: • With lung transplant who are undergoing surveillance allograft rejection | Interventions of interest are: • Testing donor-derived cell-free DNA in blood to assess lung allograft rejection | Comparators of interest are: • Bronchoscopy with biopsy | Relevant outcomes include: • Overall survival • Test validity • Morbid events • Hospitalizations |

DESCRIPTION

Clinical assessment and noninvasive imaging of chronic heart failure can be limited in accurately diagnosing patients with heart failure because symptoms and signs can poorly correlate with objective methods of assessing cardiac dysfunction. For management of heart failure, clinical signs and symptoms (e.g., shortness of breath) are relatively crude markers of decompensation and occur late in the course of an exacerbation. Thus, circulating biomarkers have potential benefit in heart failure diagnosis and management. why is there an issue locally and not when I use my personal computer.

In transplant recipients, despite the progress in immunosuppressant therapy, the risk of rejection remains. Diagnosis of allograft rejection continues to rely on clinical monitoring and histologic confirmation by tissue biopsy. However, due to limitations of tissue biopsy, including a high degree of interobserver variability in the grading of results and its potential complications, less invasive alternatives have been investigated. Several laboratory-tested biomarkers of transplant rejection have been evaluated and are commercially available for use. The laboratory tests for heart transplant rejection currently evaluated in this policy include the Presage® ST2 Assay kit, which measures the soluble suppression of tumorigenicity-2 (sST2) protein biomarker; the myTAI_{HEART} assay, which uses cell-free DNA to measure a panel of single nucleotide polymorphisms; the Heartsbreath test, which measures breath markers of oxidative stress; the AlloMap test, which uses gene expression profiling (GEP); and the HeartCare test, which combines AlloMap GEP testing with the AlloSure Heart test for donor-derived cell-free DNA (dd-cfDNA). Also included in this policy are the AlloSure dd-cfDNA tests for assessment of renal and lung transplant rejection.

OBJECTIVE

The objective of this evidence review is to determine whether the measurement of various selected biomarkers improves the detection of allograft rejection in transplant patients or in the diagnosis and management of heart failure, thus improving net health outcomes.

BACKGROUND

Heart Failure

Heart failure is a major cause of morbidity and mortality worldwide. The term *heart failure* refers to a complex clinical syndrome that impairs the heart's ability to move blood through the circulatory system.¹ The prevalence of heart failure in the U.S. between 2013 and 2016 was an estimated 6.2 million for Americans ≥20 years old, up from 5.7 million between 2009 and 2012.^{2,3} Heart failure is the leading cause of hospitalization among people older than age 65 years, with direct and indirect costs estimated at \$37 billion annually in the U.S.² Although survival has improved with treatment advances, absolute mortality rates of heart failure remain near 50% within 5 years of diagnosis.

Physiology

Heart failure can be caused by disorders of the pericardium, myocardium, endocardium, heart valves or great vessels, or metabolic abnormalities. Individuals with heart failure may present with a wide range of left ventricular (LV) anatomy and function. Some have normal LV size and preserved ejection fraction; others have severe LV dilatation and depressed ejection fraction. However, most patients present with key signs and symptoms secondary to congestion in the lungs from impaired LV myocardial function.¹ They include dyspnea, orthopnea, and paroxysmal dyspnea. Other symptoms include weight gain due to fluid retention, fatigue, weakness, and exercise intolerance secondary to diminished cardiac output.

Diagnosis

Initial evaluation of a patient with suspected heart failure is typically based on clinical history, physical examination, and chest radiograph. Because people with heart failure may present with nonspecific signs and symptoms (e.g., dyspnea), accurate diagnosis can be challenging. Therefore, noninvasive imaging procedures (e.g., echocardiography, radionuclide angiography) are used to quantify pump function of the heart, thus identifying or excluding heart failure in patients with characteristic signs and symptoms. These tests can also be used to assess prognosis by determining the severity of the underlying cardiac dysfunction.¹ However, clinical assessment and noninvasive imaging can be limited in accurately evaluating patients with heart failure because symptoms and signs can poorly correlate with objective methods of assessing cardiac dysfunction.^{4,5,6} Thus, invasive procedures (e.g., cardiac angiography, catheterization) are used in select patients with presumed heart failure symptoms to determine the etiology (i.e., ischemic vs. nonischemic) and physiologic characteristics of the condition.

Treatment

Patients with heart failure may be treated using a number of interventions. Lifestyle factors such as the restriction of salt and fluid intake, monitoring for increased weight, and structured exercise programs are beneficial components of self-management. A variety of medications are available to treat heart failure. They include diuretics (e.g., furosemide, hydrochlorothiazide, spironolactone), angiotensin-converting enzyme inhibitors (e.g., captopril, enalapril, lisinopril), angiotensin receptor blockers (e.g., losartan, valsartan, candesartan), b-blockers (e.g., carvedilol, metoprolol succinate), and vasodilators (e.g., hydralazine, isosorbide dinitrate). Numerous device-based therapies are also available. Implantable cardioverter defibrillators reduce mortality in patients with an increased risk of sudden cardiac death. Cardiac resynchronization therapy improves symptoms and reduces mortality for patients who have disordered LV conduction evidenced by a wide QRS complex on electrocardiogram. Ventricular assist devices are indicated for patients with end-stage heart failure who have failed all other therapies and are also used as a bridge to cardiac transplantation in select patients.¹

Heart Failure Biomarkers

Because of limitations inherent in standard clinical assessments of patients with heart failure, a number of objective disease biomarkers have been investigated to diagnose and assess heart

failure patient prognosis, with the additional goal of using biomarkers to guide therapy.⁷ They include a number of proteins, peptides, or other small molecules whose production and release into circulation reflect the activation of remodeling and neurohormonal pathways that lead to LV impairment. Examples include B-type natriuretic peptide (BNP), its analogue N-terminal pro B-type natriuretic peptide (NT-proBNP), troponin T and I, renin, angiotensin, arginine vasopressin, C-reactive protein, and norepinephrine.^{1,7}

BNP and NT-proBNP are considered the reference standards for biomarkers in assessing heart failure patients. They have had a substantial impact on the standard of care for diagnosis of heart failure and are included in the recommendations of all major medical societies, including the American College of Cardiology Foundation and American Heart Association,¹ European Society of Cardiology,⁸ and the Heart Failure Society of America.⁹ Although natriuretic peptide levels are not 100% specific for the clinical diagnosis of heart failure, elevated BNP or NT-proBNP levels in the presence of clinical signs and symptoms reliably identify the presence of structural heart disease due to remodeling and heightened risk for adverse events. Natriuretic peptides also can help in determining the prognosis of heart failure patients, with elevated blood levels portending a poorer prognosis.

In addition to diagnosing and assessing the prognosis of heart failure patients, blood levels of BNP or NT-proBNP have been proposed as an aid for managing patients diagnosed with chronic heart failure.^{1,10,11} Levels of either biomarker rise in response to myocardial damage and LV remodeling, whereas they tend to fall as drug therapy ameliorates symptoms of heart failure. Evidence from a large number of randomized controlled trials (RCTs) that have compared BNP- or NT-proBNP-guided therapy with clinically guided adjustment of pharmacologic treatment of patients who had chronic heart failure has been assessed in recent systematic reviews and meta-analyses. However, these analyses have not consistently reported a benefit for BNP-guided management. Savarese et al (2013) published the largest meta-analysis to date—a patient-level meta-analysis that evaluated 2686 patients from 12 RCTs.¹⁰ This meta-analysis showed that NT-proBNP-guided management was associated with significant reductions in all-cause mortality and heart failure-related hospitalization compared with clinically guided treatment. Although BNP-guided management in this meta-analysis was not associated with significant reductions in these parameters, differences in patient numbers and characteristics may explain the discrepancy. Troughton et al (2014) conducted a second patient-level meta-analysis that included 11 RCTs with 2000 patients randomized to natriuretic peptide-guided pharmacologic therapy or usual care.¹¹ The results showed that, among patients 75 years of age or younger with chronic heart failure, most of whom had impaired left ventricular ejection fraction, natriuretic peptide-guided therapy was associated with significant reductions in all-cause mortality compared with clinically guided therapy. Natriuretic-guided therapy also was associated with significant reductions in hospitalization due to heart failure or cardiovascular disease.

Suppression of Tumorigenicity-2 Protein Biomarker

A protein biomarker, ST2, has elicited interest as a potential aid to predict prognosis and manage therapy of heart failure.^{12,13,14,15,16,17,18} This protein is a member of the interleukin-1 (IL-1) receptor

family. It is found as a transmembrane isoform (ST2L) and a soluble isoform (sST2), both of which have circulating IL-33 as their primary ligand. ST2 is a unique biomarker that has pluripotent effects in vivo. Thus, binding between IL-33 and ST2L is believed to have an immunomodulatory function via T-helper type 2 lymphocytes and was initially described in the context of cell proliferation, inflammatory states, and autoimmune diseases.¹⁹ However, the IL-33/ST2L signaling cascade is also strongly induced through the mechanical strain of cardiac fibroblasts or cardiomyocytes. The net result is mitigation of adverse cardiac remodeling and myocardial fibrosis, which are key processes in the development of heart failure.²⁰ The soluble isoform of ST2 is produced by lung epithelial cells and cardiomyocytes and is secreted into circulation in response to exogenous stimuli, mechanical stress, and cellular stretch. This form of ST2 binds to circulating IL-33, acting as a "decoy," thus inhibiting the IL-33-associated antiremodeling effects of the IL-33/ST2L signaling pathway. Thus, on a biologic level, IL-33/ST2L signaling plays a role in modulating the balance of inflammation and neurohormonal activation and is viewed as pivotal for protection from myocardial remodeling, whereas sST2 is viewed as attenuating this protection. In the clinic, blood concentrations of sST2 appear to correlate closely with adverse cardiac structure and functional changes consistent with remodeling in patients with heart failure, including abnormalities in filling pressures, chamber size, and systolic and diastolic function.^{7,14,16}

An enzyme-linked immunosorbent-based assay is commercially available for determining sST2 blood levels (Presage ST2 Assay).¹⁷ The manufacturer claims a limit of detection of 1.8 ng/mL for sST2, and a limit of quantification of 2.4 ng/mL, as determined according to Clinical and Laboratory Standards Institute guideline EP-17-A. Mueller and Dieplinger (2013) reported a limit of detection of 2.0 ng/mL for sST2 in their study.¹⁷ In the same study, the assay had a within-run coefficient of variation of 2.5% and a total coefficient of variation less than 4.0%, demonstrated linearity within the dynamic range of the assay calibration curve, and exhibited no relevant interference or cross-reactivity.

The ST2 biomarker is not intended to diagnosis heart failure because it is a relatively nonspecific marker that is increased in many other disparate conditions that may be associated with acute or chronic manifestations of heart failure.^{16,17} Although the natriuretic peptides (BNP, NT-proBNP) reflect different physiologic aspects of heart failure compared with sST2, they are considered the reference standard biomarkers when used with clinical findings to diagnose, prognosticate, and manage heart failure and as such are the comparator to sST2.

Heart Transplant Rejection

Most cardiac transplant recipients experience at least a single episode of rejection in the first year after transplantation. The International Society for Heart and Lung Transplantation (2005) modified its grading scheme for categorizing cardiac allograft rejection.²¹ The revised (R) categories are listed in Table 1.

Table 1. Revised Grading Schema for Cardiac Allograft Rejection

| New Grade | Definition | Old Grade |
|-----------|--------------|-----------|
| 0R | No rejection | |

| New Grade | Definition | Old Grade |
|------------------|--------------------|------------------|
| 1R | Mild rejection | 1A, 1B, and 2 |
| 2R | Moderate rejection | 3A |
| 3R | Severe rejection | 3B and 4 |

Acute cellular rejection is most likely to occur in the first 6 months after transplantation, with a significant decline in the incidence of rejection after this time. Although immunosuppressants are required on a life-long basis, dosing is adjusted based on graft function and the grade of acute cellular rejection determined by histopathology. Endomyocardial biopsies are typically taken from the right ventricle via the jugular vein periodically during the first 6 to 12 months post transplant. The interval between biopsies varies among clinical centers. A typical schedule is weekly for the first month, once or twice monthly for the following 6 months, and several times (monthly to quarterly) between 6 months and 1-year post transplant. Surveillance biopsies may also be performed after the first postoperative year (e.g., on a quarterly or semiannual basis). This practice, although common, has not been demonstrated to improve transplant outcomes. Some centers no longer routinely perform endomyocardial biopsies after 1 year in patients who are clinically stable.

While the endomyocardial biopsy is the criterion standard for assessing heart transplant rejection, it is limited by a high degree of interobserver variability in the grading of results and potential morbidity that can occur with the biopsy procedure. Also, the severity of rejection may not always coincide with the grading of the rejection by biopsy. Finally, a biopsy cannot be used to identify patients at risk of rejection, limiting the ability to initiate therapy to interrupt the development of rejection. For these reasons, an endomyocardial biopsy is considered a flawed criterion standard by many. Therefore, noninvasive methods of detecting cellular rejection have been explored. It is hoped that noninvasive tests will assist in determining appropriate patient management and avoid overuse or underuse of treatment with steroids and other immunosuppressants that can occur with false-negative and false-positive biopsy reports. Two techniques are commercially available for the detection of heart transplant rejection.

NONINVASIVE HEART TRANSPLANT REJECTION TESTS

Heartsbreath Test

The Heartsbreath test, a noninvasive test that measures breath markers of oxidative stress, has been developed to assist in the detection of heart transplant rejection. In heart transplant recipients, oxidative stress appears to accompany allograft rejection, which degrades membrane polyunsaturated fatty acids and evolving alkanes and methylalkanes that are, in turn, excreted as volatile organic compounds in breath. The Heartsbreath test analyzes the breath methylated alkane contour, which is derived from the abundance of C4 to C20 alkanes and monomethylalkanes and has been identified as a marker to detect grade 3 (clinically significant) heart transplant rejection.

AlloMap

Another approach has focused on patterns of gene expression of immunomodulatory cells, as detected in the peripheral blood. For example, microarray technology permits the analysis of the expression of thousands of genes, including those with functions known or unknown. Patterns of gene expression can then be correlated with known clinical conditions, permitting a selection of a finite number of genes to compose a custom multigene test panel, which then can be evaluated using polymerase chain reaction techniques. AlloMap (CareDx) is a commercially available molecular expression test that has been developed to detect acute heart transplant rejection or the development of graft dysfunction. The test involves polymerase chain reaction-expression measurement of a panel of genes derived from peripheral blood cells and applies an algorithm to the results. The proprietary algorithm produces a single score that considers the contribution of each gene in the panel. The score ranges from 0 to 40. The AlloMap website states that a lower score indicates a lower risk of graft rejection; the website does not cite a specific cutoff for a positive test.²² All AlloMap testing is performed at the CareDx reference laboratory in California.

HeartCare

Cell-free DNA (cfDNA), released by damaged cells, is normally present in healthy individuals.²³ In patients who have received transplants, donor-derived cell-free DNA (dd-cfDNA) may be also present. It is proposed that allograft rejection, which is associated with damage to transplanted cells, may result in an increase in dd-cfDNA. HeartCare (CareDx) is a commercially-available test that combines AlloMap gene expression profiling with a next-generation sequencing assay that quantifies the fraction of dd-cfDNA in cardiac transplant recipients relative to total cfDNA. The AlloMap score, AlloMap score variability, and AlloSure % dd-cfDNA are reported.

Presage ST2 Assay

In addition to its use as a potential aid to predict prognosis and manage therapy of heart failure, elevated serum ST2 levels have also been associated with an increased risk of antibody-mediated rejection following a heart transplant. For this reason, ST2 has also been proposed as a prognostic marker post heart transplantation and as a test to predict acute cellular rejection (graft-versus-host disease). The Presage ST2 Assay, described above, is a commercially available sST2 test that has been investigated as a biomarker of heart transplant rejection.

myTAI_{HEART} Biomarker

Using proprietary myTAI_{HEART} software, the myTAI_{HEART} test uses multiplexed, high-fidelity amplification followed by allele-specific qPCR of a panel of 94 highly informative bi-allelic single nucleotide polymorphisms (SNPs) and two controls to quantitatively genotype cell free DNA in the patient's plasma after cardiac transplant, and accurately distinguish "donor specific" cell free DNA originating from the engrafted heart from "self-specific" cell free DNA originating from the recipient's native cells.²⁴ The ratio of donor specific cell free DNA to total cell free DNA is reported as the donor fraction (%) and categorizes the patient as at low or increased risk of moderate (grade 2R) to severe (grade 3R) acute cellular rejection: low donor fractions indicate less damage to the transplanted heart and a lower risk for rejection, while increased donor fractions indicate more damage to the transplanted heart and an increased risk for rejection. Testing with

myTAI_{HEART} does not require a donor specimen. The test is indicated for use in heart transplant recipients who are 2 months of age or older and ≥ 8 days post-transplant, restricted to use in single organ post-heart transplant patients, and is contraindicated in patients who:

- are pregnant
- currently have or in the past have had another transplanted organ (solid organ or allogeneic bone marrow)
- have post-transplant lymphoproliferative disease
- have cancer or have had cancer within the previous 2 years
- are on mechanical circulatory support
- are closely related to the transplant donor

Other laboratory-tested biomarkers of heart transplant rejection have been evaluated. They include brain natriuretic peptide, troponin, and soluble inflammatory cytokines. Most have had low accuracy in diagnosing rejection. Preliminary studies have evaluated the association between heart transplant rejection and micro-RNAs or high-sensitivity cardiac troponin in cross-sectional analyses but the clinical use has not been evaluated.^{25,26,}

Renal Transplant Rejection

Allograft dysfunction is typically asymptomatic and has a broad differential, including graft rejection. Diagnosis and rapid treatment are recommended to preserve graft function and prevent loss of the transplanted organ. For a primary kidney transplant, graft survival at 1 year is 94.7%; at 5 years, graft survival is 78.6%.^{27,}

Surveillance of transplant kidney function relies on routine monitoring of serum creatinine, urine protein levels, and urinalysis.^{28,} Allograft dysfunction may also be demonstrated by a drop in urine output or, rarely, as pain over the transplant site. With clinical suspicion of allograft dysfunction, additional noninvasive workup including ultrasonography or radionuclide imaging may be used. A renal biopsy allows a definitive assessment of graft dysfunction and is typically a percutaneous procedure performed with ultrasonography or computed tomography guidance. Biopsy of a transplanted kidney is associated with fewer complications than biopsy of a native kidney because the allograft is typically transplanted more superficially than a native kidney. Renal biopsy is a low-risk invasive procedure that may result in bleeding complications; loss of a renal transplant, as a complication of renal biopsy, is rare.^{29,}

Kidney biopsies allow for diagnosis of acute and chronic graft rejection, which may be graded using the Banff Classification.^{30,31,} Pathologic assessment of biopsies demonstrating acute rejection allows clinicians to further distinguish between acute cellular rejection and antibody-mediated rejection, which are treated differently.

NONINVASIVE RENAL TRANSPLANT REJECTION TESTS

AlloSure

AlloSure Kidney (CareDx) is a commercially available, next-generation sequencing assay that quantifies the fraction of dd-cfDNA in renal transplant recipients relative to total cfDNA by measuring 266 single nucleotide variants. Separate genotyping of the donor or recipient is not required but patients who receive a kidney transplant from a monozygotic (identical) twin are not eligible for this test. The fraction of dd-cfDNA relative to total cfDNA present in the peripheral blood sample is cited in the report. For patients undergoing surveillance, a routine testing schedule is recommended for longitudinal monitoring.

Prospera

Prospera Kidney (Natera) is a commercially available assay that uses massively multiplexed PCR (mmPCR) followed by next-generation sequencing (NGS) to quantify the fraction of dd-cfDNA in renal transplant recipients. Donor versus recipient cfDNA is differentiated via an advanced bioinformatics analysis of >13,000 single-nucleotide polymorphisms (SNPs) without the need for prior recipient or donor genotyping or computational adjustments for related donors. The manufacturer recommends use of the test when there is clinical suspicion of active rejection and for regular surveillance of subclinical rejection.³² In a surveillance scenario, regular testing is recommended at 1, 2, 3, 4, 6, 9 and 12 months after renal transplant or most recent rejection.³³ Thereafter, the test should be repeated quarterly. The proportion of dd-cfDNA relative to total cfDNA is reported, with detection of $\geq 1\%$ dd-cfDNA indicating increased risk for active rejection. The percent dd-cfDNA change between tests is also reported.

Lung Transplant Rejection

Despite advances in induction and maintenance immunosuppressive regimens, lung transplant recipients have a median overall survival of 6 years, with more than a third of patients receiving treatment for acute rejection in the first year after transplant.^{34,35} Acute cellular rejection, lymphocytic bronchiolitis, and antibody-mediated rejection are all risk factors for subsequent development of chronic lung allograft dysfunction (CLAD). Pathologic grading of acute cellular rejection is based on the histological assessment of perivascular and interstitial mononuclear cell infiltrates. Antibody-mediated rejection may be clinical (symptomatic or asymptomatic allograft dysfunction) or subclinical (normal allograft function). Key diagnostic criteria established via consensus by the International Society for Heart and Lung Transplantation include the presence of antibodies directed toward donor human leukocyte antigens and characteristic lung histology with or without evidence of complement 4d within the graft.³⁶ The most common phenotype of CLAD is a persistent obstructive decline in lung function known as bronchiolitis obliterans syndrome (BOS), which is graded based on the degree of decrease in FEV₁. Approximately 50% of patients develop BOS within 5 years post-transplant. Median survival following a diagnosis of BOS is 3-5 years. Acute rejection may present with non-specific physical symptoms or be asymptomatic. However, the role of surveillance bronchoscopy for screening asymptomatic patients for acute rejection is controversial, and performance of surveillance bronchoscopies varies across transplant centers.

NONINVASIVE LUNG TRANSPLANT REJECTION TESTS

AlloSure

AlloSure Lung (CareDx) is a commercially available, NGS assay that quantifies the fraction of dd-cfDNA in lung transplant patients relative to total cfDNA by measuring single nucleotide polymorphisms. The test is intended to provide a direct, noninvasive measure of organ injury in lung transplant patients who are undergoing surveillance. Suggested thresholds for severe injury, injury, and quiescence are 1%, 0.85%, and <0.5%, respectively.³⁷

REGULATORY STATUS

Regulatory Status

The U.S. Food and Drug Administration (FDA) has cleared multiple biomarker tests for the detection of heart and renal allograft rejection. Table 2 provides a summary of the biomarker tests currently included in this policy that have FDA clearance.

Table 2. Select Biomarker Tests for Detection of Heart or Renal Allograft Rejection Cleared by the U.S. Food and Drug Administration

| Test | Manufacturer | FDA Clearance Type, Product Number | FDA Clearance Date | Indicated Use |
|---------------------------------------|----------------------|--|--------------------|---|
| Heartsbreath™ | Menssana Research | Humanitarian device exemption, H030004 | 2004 | To aid in diagnosing grade 3 heart transplant rejection in patients who have received heart transplants within the preceding year. The device is intended as an adjunct to, and not as a substitute for, endomyocardial biopsy and is also limited to patients who have had endomyocardial biopsy within the previous month. |
| AlloMap® Molecular Expression Testing | CareDx, formerly XDx | 510(k), k073482 | 2008 | The test is to be used in conjunction with clinical assessment, for aiding in the identification of heart transplant recipients with stable allograft function and a low probability of moderate-to-severe transplant rejection. It is intended for patients at least 15 years old who are at least 2 months post transplant. |
| Presage® ST2 Assay kit | Critical Diagnostics | 510(k), k093758 | 2011 | For use with clinical evaluation as an aid in assessing the prognosis of |

| Test | Manufacturer | FDA Clearance Type, Product Number | FDA Clearance Date | Indicated Use |
|------|--------------|------------------------------------|--------------------|---|
| | | | | patients diagnosed with chronic heart failure |

FDA: Food and Drug Administration.

Laboratory Developed Tests

There are also commercially available laboratory-developed biomarker tests for the detection of heart and renal allograft rejection. 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. To-date, AlloSure (CareDx) renal and lung and Prospera (Natera) renal dd-cfDNA tests are regulated under the Clinical Laboratory Improvement Amendments standards.

myTAI_{HEART} is also a laboratory developed test (LDT) developed for clinical diagnostic performance exclusively in the College of American Pathologists (CAP) and Clinical Laboratory Improvement Amendment (CLIA) accredited TAI Diagnostics Clinical Reference Laboratory.²⁴ This test was developed and its performance characteristics were determined by TAI Diagnostics.

These LDTs have not been cleared or approved by the FDA.

Other Tests

Other commercially available LDTs without FDA clearance or approval for use have been excluded from this evidence review when studies reporting on the clinical validity of the marketed version of the test could not be identified and/or where the test is marketed for research use only. Excluded tests and their descriptions are summarized for reference purposes in Table 3.

Table 3. Biomarker Tests Excluded from Review

| Test | Manufacturer | Technology | Indications for Use |
|---------------|--------------|------------------|--|
| KidneyCare® | CareDx | dd-cfDNA and GEP | Available as a research tool through the OKRA Registry. |
| AlloSeq® HCT | CareDx | NGS | To aid in the assessment of engraftment following HCT via NGS analysis of 202 biallelic SNPs. The fraction of recipient and donor genomic DNA is reported. The test is marketed for research use only. |
| AlloSeq® Tx17 | CareDx | NGS | An NGS test utilizing Hybrid Capture Technology conducted pre-transplant to identify optimal transplant matches. The test sequences full HLA genes and other transplant-associated genes (KIR, MICA/B, C4, HPA, ABO). This test is marketed for research use only. |

| Test | Manufacturer | Technology | Indications for Use |
|---------------|-----------------------------|-------------------------|--|
| Viracor TRAC® | Eurofins | dd-cfDNA | To aid in the diagnosis of solid organ transplant rejection via NGS analysis. The fraction of dd-cfDNA is reported. ¹ |
| MMDx® Heart | Kashi Clinical Laboratories | Tissue-based microarray | Tissue-based microarray mRNA gene expression test of 1283 genes post-transplant to provide a probability score of rejection as a complement to conventional biopsy processing. The test is not marketed to provide information for the diagnosis, prevention, or treatment of disease or to aid in the clinical decision-making process. |
| MMDx® Kidney | Kashi Clinical Laboratories | Tissue-based microarray | Tissue-based microarray mRNA gene expression test of 1494 genes post-transplant to provide a probability score of rejection as a complement to conventional biopsy processing. The test is not marketed to provide information for the diagnosis, prevention, or treatment of disease or to aid in the clinical decision-making process. |

dd-cfDNA: donor-derived cell-free DNA; GEP: gene expression profiling; HCT: hematopoietic cell transplantation; HLA: human leukocyte antigen; MMDx: molecular microscope diagnostic system; NGS: next-generation sequencing; OKRA: Outcomes in KidneyCare in Renal Allografts; SNP: single-nucleotide polymorphism; TRAC: transplant rejection allograft check.

¹ Published studies reporting on the clinical validity of the marketed version of the test were not identified.

POLICY

- A. The use of the Presage ST2 Assay to evaluate the prognosis of patients diagnosed with chronic heart failure is considered **experimental / investigational**.
- B. The use of the Presage ST2 Assay to guide management (e.g., pharmacologic, device-based, exercise) of patients diagnosed with chronic heart failure is considered **experimental / investigational**.
- C. The use of the Presage ST2 Assay in the post cardiac transplantation period, including but not limited to predicting prognosis and predicting acute cellular rejection, is considered **experimental / investigational**.
- D. The use of the myTAI_{HEART} assay in the post cardiac transplantation period, including but not limited to predicting prognosis and predicting acute cellular rejection, is considered **experimental / investigational**.
- E. The measurement of volatile organic compounds to assist in the detection of moderate grade 2R (formerly grade 3) heart transplant rejection is considered **experimental / investigational**.
- F. The use of peripheral blood gene expression profile tests or peripheral blood measurement of donor-derived cell-free DNA (dd-cfDNA), alone or in combination in the management of patients after heart transplantation, including but not limited to the detection of acute heart transplant rejection or heart transplant graft dysfunction, is considered **experimental / investigational**.
- G. The use of peripheral blood measurement of donor-derived cell-free DNA in the management of patients after renal transplantation, including but not limited to the detection of acute renal transplant rejection or renal transplant graft dysfunction, is considered **experimental / investigational**.
- H. The use of peripheral blood measurement of dd-cfDNA in the management of patients after lung transplantation, including but not limited to the detection of acute lung transplant rejection or lung transplant graft dysfunction, is considered **experimental / investigational**.

POLICY GUIDELINES

The U.S. Food and Drug Administration has indicated that the Heartsbreath (Menssana Research) test is only for use as an aid in the diagnosis of grade 3 (now known as grade 2R) heart transplant rejection in patients who have received heart transplants within the preceding year and who have had endomyocardial biopsy within the previous month.

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.

RATIONALE

This evidence review has been updated regularly with searches of the PubMed database. The most recent literature update was performed through October 20, 2021.

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.

USE OF SOLUBLE SUPPRESSION OF TUMORIGENICITY-2 LEVELS IN CHRONIC HEART FAILURE PATIENTS

Clinical Context and Test Purpose

The purpose of the Soluble Suppression of Tumorigenicity-2 (sST2) assay is to determine prognosis and/or to guide management in patients with chronic heart failure as an alternative to or an improvement on existing tests and clinical assessment.

The question addressed in this evidence review is: Do sST2 assays determine prognosis and/or guide treatment in patients with chronic heart failure and improve net health outcomes?

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

Populations

The relevant population of interest is individuals with chronic heart failure.

Interventions

The test being considered is sST2 assay to determine prognosis and/or to guide management. Elevated sST2 levels are purported to predict a higher risk of poor outcomes.

Comparators

Comparators of interest include standard prognostic markers, including B-type natriuretic peptide levels and clinical assessment.

Outcomes

The general outcomes of interest are overall survival (OS), quality of life, and hospitalizations. Follow-up of 6-12 months would be appropriate to assess quality of life outcomes.

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 sST2 testing, methodologically credible studies were selected using the following principles:

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

REVIEW OF EVIDENCE**Subanalyses from Randomized Controlled Trials**

A number of clinical studies in which sST2 blood levels were determined using the Presage ST2 Assay have reported that there is an association between ST2 levels and adverse outcomes in patients diagnosed with chronic heart failure. A substantial body of biomarker evidence has been reported retrospectively from subsets of patients enrolled in randomized controlled trials (RCTs) of heart failure interventions. These RCTs include the Valsartan Heart Failure Trial (Val-HeFT)³⁸; Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training (HF-ACTION)³⁹; Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA)⁴⁰; ProBNP Outpatient Tailored Chronic Heart Failure study (PROTECT).⁴¹ Although patients in these RCTs were well-characterized and generally well-matched between study arms, the trials were neither intended nor designed specifically to evaluate biomarkers as risk predictors. At present, no prospectively gathered evidence is available from an RCT in which sST2 levels were compared with levels of a B-type natriuretic peptide (BNP or N-terminal pro B-type natriuretic peptide [NT-proBNP]) to predict risk for adverse outcomes among well-defined cohorts of patients with diagnosed chronic heart failure. Key results of larger individual studies are summarized in Table 4.

Table 4. Summary of Selected Clinical Studies of sST2 to Predict Outcomes in Chronic Heart Failure Patients

| Study | Population | Mean Age, y | Study Description and Biomarkers | Primary Endpoints | Mean FU | Synopsis of Findings |
|--|---|-------------|---|--|---------|---|
| Ky et al (2011) ⁴² , | Ambulatory CHF (N = 1,141, 75% of Penn HF Study population) | 56 | Retrospective analysis of sST2 and NT-proBNP levels and their incremental usefulness over clinical SHFM | Mortality or cardiac transplant | 2.8 y | <ul style="list-style-type: none"> Elevated sST2 levels associated with increased risk (adjusted p =.002) sST2 in plus NT-proBNP levels showed moderate improvement over SHFM in predicting outcomes (p =.017) |
| Bayes-Genis et al (2012) ⁴³ , | Ambulatory decompensated HF (N = 891) | 70 | Retrospective analysis of sST2 and NT-proBNP levels from consecutive series | Mortality | 2.8 y | <ul style="list-style-type: none"> Elevated sST2 and NT-proBNP levels provided independent and additive prognostic information for elevated risk of mortality (p <.001) |
| Broch et al (2012) ⁴⁴ , | Ischemic CHF (N = 1,149, 30% of CORONA RCT) | 72 | Retrospective analysis of sST2, NT-proBNP, and CRP levels | CV mortality, nonfatal myocardial infarction or stroke | 2.6 y | <ul style="list-style-type: none"> Elevated sST2 levels were independently associated with increased risk for mortality, hospitalization due to HF, or any CV hospitalization (p <.001) sST2 did not provide |

| Study | Population | Mean Age, y | Study Description and Biomarkers | Primary Endpoints | Mean FU | Synopsis of Findings |
|-------------------------------------|---|-------------|---|--|---------|---|
| | | | | | | additive prognostic information vs. NT-proBNP |
| Felker et al (2013) ⁴⁵ , | Ambulatory HF (N = 910, 39% of HF-ACTION RCT) | 59 | Retrospective analysis of sST2 and NT-proBNP levels | Mortality, hospitalization, functional capacity | 2.5 y | <ul style="list-style-type: none"> Elevated sST2 levels were independently associated with increased risk for mortality, hospitalization due to HF, or any CV hospitalization (p <.000) sST2 and NT-proBNP provided independent prognostic information sST2 did not provide additive prognostic information vs. NT-proBNP |
| Gaggin et al (2013) ⁴⁶ , | Recently decompensated CHF (n=151, 100% of PROTECT RCT) | 63 | Retrospective analysis of sST2 and NT-proBNP levels | Composite outcome (worsening HF, hospitalization for HF, clinically significant CV events) | 0.8 y | <ul style="list-style-type: none"> Elevated sST2 levels associated with increased risk for adverse CV outcome (p <.001) sST2 and NT-proBNP did not provide independent prognostic information |

| Study | Population | Mean Age, y | Study Description and Biomarkers | Primary Endpoints | Mean FU | Synopsis of Findings |
|------------------------------------|--|-------------|---|---|--------------------|---|
| Anand et al (2014) ⁴⁷ , | CHF (n=1,650, 33% of Val-HeFT RCT) | 63 | Retrospective analysis of sST2, NT-proBNP, and other biomarker levels | All-cause mortality and composite outcome (mortality, SCD with resuscitation, hospitalization for HF, or administration of IV inotropic or vasodilator drug for ≥4 h without hospitalization) | | <ul style="list-style-type: none"> Elevated sST2 levels were independently associated with increased risk of poor outcomes (p <.001) Baseline sST2 levels did not provide substantial prognostic information when added to a clinical model that included NT-proBNP levels |
| Zhang et al (2015) ⁴⁸ , | De novo HF or decompensated CHF (N = 1161) | 58 | Prospective analysis of sST2 in a hospitalized sample at 1 center in China | All-cause mortality | 1 y | <ul style="list-style-type: none"> Elevated sST2 levels were independently associated with increased risk of all-cause mortality (p <.001) after adjustment for clinical risk factors and NT-proBNP levels |
| Dupuy et al (2016) ⁴⁹ , | HF for ≥6 mo (N = 178) | 75 | Prospective analysis of sST2, NT-proBNP, and other biomarker levels in a sample from 1 center in France | All-cause mortality and CV mortality | 42 mo ^a | <ul style="list-style-type: none"> Elevated sST2 levels were independently associated with increased risk for all-cause mortality and CV mortality (p <.001) |

| Study | Population | Mean Age, y | Study Description and Biomarkers | Primary Endpoints | Mean FU | Synopsis of Findings |
|-------|------------|-------------|----------------------------------|-------------------|---------|---|
| | | | | | | <ul style="list-style-type: none"> In multivariate analysis, sST2 and CRP significantly associated with all-cause mortality and CV mortality |

CHF: chronic heart failure; CRP: C-reactive protein; CV: cardiovascular; FU: follow-up; HF: heart failure; HF-ACTION: Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training; IV: intravenous; NT-proBNP: N-terminal pro B-type natriuretic peptide; PROTECT: ProBNP Outpatient Tailored Chronic Heart Failure study; RCT: randomized controlled trial; SCD: sudden cardiac death; SHFM: Seattle Heart Failure Model; sST2: soluble suppression of tumorigenicity-2; Val-HeFT: Valsartan Heart Failure Trial.

^a Median.

Meta-Analyses

Aimo et al (2017) pooled findings of studies on the prognostic value of sST2 for chronic heart failure in a meta-analysis.⁵⁰ The meta-analysis selected 7 studies, including post hoc analyses of RCTs, and calculated the association between the Presage ST2 Assay and health outcomes. A pooled analysis of 7 studies found that sST2 was a statistically significant predictor of overall mortality (hazard ratio [HR] = 1.75; 95% confidence interval [CI], 1.37 to 2.22). Moreover, a pooled analysis of 5 studies found that sST2 was a significant predictor of cardiovascular mortality (HR = 1.79; 95% CI, 1.22 to 2.63).

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.

No evidence is available from randomized or nonrandomized controlled studies in which outcomes from groups of well-matched patients managed using serial changes in sST2 blood levels were compared with those managed using the reference standard of BNP or NT-proBNP levels.

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.

No inferences can be drawn about the clinical utility of sST2 levels for chronic heart failure.

Section Summary: Use of Soluble Suppression of Tumorigenicity-2 in Chronic Heart Failure Patients

Several analyses, mostly retrospective, have evaluated whether sST2 levels are associated with disease prognosis, especially mortality outcomes. Studies mainly found that elevated sST2 levels were statistically associated with an elevated risk of mortality. A pooled analysis of study results found that sST2 levels significantly predicted overall mortality and cardiovascular mortality. Several studies, however, found that sST2 test results did not provide additional prognostic information compared with BNP or NT-proBNP levels. In general, it appears that elevated sST2 levels predict higher risk of poor outcomes better than lower levels. The available evidence is limited by interstudy inconsistency and differences in patient characteristics, particularly the severity of heart failure, its etiology, duration, and treatment. Furthermore, most of the evidence was obtained from retrospective analyses of sST2 levels in subsets of larger patient cohorts within RCTs, potentially biasing the findings. The evidence primarily shows associations between elevated sST2 levels and poor outcomes but does not go beyond that in demonstrating a clinical connection among biomarker status, treatment received, and clinical outcomes.

USE OF SOLUBLE ST2 SUPPRESSION OF TUMORIGENICITY-2 IN POST-HEART TRANSPLANTATION PATIENTS

Clinical Context and Test Purpose

The purpose of sST2 assay is to determine prognosis and/or to predict acute cellular rejection in patients with heart transplantation an alternative to or an improvement on existing tests. The question addressed in this evidence review is: Does the use of the sST2 assay determine prognosis and/or predict acute cellular rejection in patients undergoing heart transplantation and improve net health outcomes?

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

Populations

The relevant population of interest is individuals with heart transplantation.

Interventions

The test being considered is sST2 assay to determine prognosis and/or to predict acute cellular rejection.

Comparators

Comparators of interest include endomyocardial biopsy for predicting acute cellular rejection.

Outcomes

The general outcomes of interest are OS, quality of life, and hospitalizations.

Table 5. Significant Outcomes for Post–Heart Transplantation Patients.

| Outcomes | Details | Timing |
|------------------|--|------------------------------|
| Morbid events | Short-term and long-term events, such as acute cellular rejection, myocardial infarction, and stroke | 30 days, 6 months, 1-5 years |
| Hospitalizations | Inpatient hospital admissions | 30 days, 6 months, 1-5 years |

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 sST2 testing, methodologically credible studies were selected using the following principles:

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

Observational Studies

Serum ST2 levels have been proposed as a prognostic marker post heart transplantation and as a test to predict acute cellular rejection (graft-versus-host disease). There is very little evidence available for these indications. Januzzi et al (2013) retrospectively assessed sST2 levels in 241 patients post–heart transplant.⁵¹ Over a follow-up out to 7 years, sST2 levels were predictive of total mortality (HR = 2.01; 95% CI, 1.15 to 3.51; p =.01). Soluble ST2 levels were also associated with risk of acute cellular rejection, with a significant difference between the top and bottom quartiles of sST2 levels in the risk of rejection (p =.003).

Pascual-Figal et al (2011) reported on 26 patients post cardiac transplantation with and an acute rejection episode.⁵² Soluble ST2 levels were measured during the acute rejection episode and compared with levels measured when acute rejection was not present. Soluble ST2 levels were higher during the acute rejection episode (130 ng/mL) than during the nonrejection period (50 ng/mL; p =.002). Elevated sST2 levels greater than 68 ng/mL had a positive predictive value (PPV) of 53% and a negative predictive value (NPV) of 83% for the presence of acute cellular rejection. The addition of sST2 levels to serum BNP resulted in incremental improvement in identifying rejection episodes.

Table 6. Summary of Key Nonrandomized Clinical Validity Study Characteristics

| Study | Study Type | Country | Dates | Participants | Treatment | Follow-Up |
|-------------------------------------|---------------|---------------|-----------|---|--------------------------------|------------------|
| Januzzi (2013) ^{51,} | Retrospective | United States | NR | Post-cardiac transplantation | sST2 levels assessment (n=241) | Median 7.1 years |
| Pascual-Figal (2011) ^{52,} | Retrospective | Spain | 2002-2007 | Post-cardiac transplantation with acute rejection | sST2 levels assessment (n=26) | Median 3 months |

NR: not reported, sST2: soluble suppression of tumorigenicity-2.

Table 7. Summary of Key Nonrandomized Clinical Validity Study Results

| Study | Total Mortality | ST2 Levels | PPV | NPV |
|-------------------------------------|------------------|--------------------------------|-----|-----|
| Januzzi (2013) ^{51,} | | | NR | NR |
| HR (95% CI) | 2.02 (1.16-3.52) | ≥ 30 ng/mL at 7-year follow-up | NA | NA |
| <i>P-value</i> | .01 | NR | NA | NA |
| Pascual-Figal (2011) ^{52,} | | | 53% | 83% |
| Rejection Episode | NR | 130 ng/mL (IQR 60-238 ng/mL) | NA | NA |
| Nonrejection Period | NR | 50 ng/mL (IQR 28-80 ng/mL) | NA | NA |
| <i>P-value</i> | NR | .002 | NA | NA |

HR: hazard ratio; IQR: interquartile range; NA: not applicable; NPV: negative predictive value; NR: not reported, PPV: positive predictive value.

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.

No RCTs were identified using sST2 levels that directed patient management in heart transplantation patients and which assessed patient outcomes.

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.

No inferences can be drawn about the clinical utility of sST2 levels for patients with heart transplantation.

Section Summary: Use of Soluble Suppression of Tumorigenicity-2 sST2 in Post-Heart Transplantation Patients

Few studies are available, and they are observational and retrospective. No prospective studies were identified that provide high-quality evidence on the ability of sST2 levels to predict transplant outcomes. One retrospective study (N = 241) found that sST2 levels were associated with acute cellular rejection and mortality; another study (N = 26) found that sST2 levels were higher during an acute rejection episode than before rejection.

USE OF MYTAI_{HEART} IN POST-HEART TRANSPLANTATION PATIENTS**Clinical Context and Test Purpose**

The purpose of myTAI_{HEART} is to determine prognosis and/or to predict acute cellular rejection in patients with heart transplantation as an alternative to or an improvement on existing tests.

The question addressed in this evidence review is: Does the use of the myTAI_{HEART} assay determine prognosis and/or predict acute cellular rejection in patients undergoing heart transplantation and improve net health outcomes?

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

Populations

The relevant population of interest is individuals with heart transplantation.

Interventions

The test being considered is myTAI_{HEART} assay to determine prognosis and/or to predict acute cellular rejection.

Comparators

Comparators of interest include endomyocardial biopsy for predicting acute cellular rejection.

Outcomes

The general outcomes of interest are OS, quality of life, and hospitalizations.

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 myTAI_{HEART} testing, methodologically credible studies were selected using the following principles:

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

Observational Studies

Serum donor fraction (%), defined as the ratio of donor specific cell free DNA to total cell free DNA, has been proposed as a test to predict acute cellular rejection. In a study funded by TAI Diagnostics, Inc., North et al (2020) performed a blinded clinical validation study on 158 matched pairs of endomyocardial biopsy-plasma samples collected from 76 volunteer adult and pediatric heart transplant recipients (ages 2 months or older, and 8 days or more post-transplant) between June of 2010 and Aug 2016 from 2 Milwaukee transplant centers.²⁴ Based on acute cellular rejection grade as defined by the 2004 International Society for Heart and Lung Transplantation (ISHLT) classification, Receiver Operating Characteristic (ROC) analysis was performed to evaluate diagnostic accuracy across all possible cutoffs. To maximize diagnostic accuracy, Youden's Index was used to select the optimal cutoff, found to correspond to a donor fraction value of 0.32%. Using this cutoff, clinical performance characteristics of the assay included a NPV of 100.00% for grade 2R or higher acute cellular rejection, with 100.00% sensitivity and 75.48% specificity; Area under the Curve (AUC) for this analysis was 0.842, indicative of robust ability of the donor fraction assay to rule out 2R or greater acute cellular rejection for donor fraction values less than 0.32%. There was no statistically significant correlation of donor fraction with age. Donor fraction elevation can also be caused by other forms of injury to the donor heart such as acute cellular rejection 1R, acute antibody-mediated rejection (AMR), and presence of coronary artery vasculopathy (CAV), thereby requiring correlation of myTAI_{HEART} results with other clinical indicators.

In study funded by a grant from the National Institutes of Health and TAI Diagnostics, Inc., Richmond et al (2019) assessed 174 post-cardiac transplant patients from 7 centers (ages 2.4 months-73.4 years) with myTAI_{HEART} testing (before transplant; 1, 4, and 7 days following transplant; and at discharge from transplant hospitalization) using blinded analysis of biopsy-paired samples.⁵³ All the patients were followed for at least 1 year. Donor fraction, defined as the ratio of cell free DNA specific to the transplanted organ to the total amount of cell free DNA present in a blood sample was higher in acute cellular rejection 1R/2R (n = 15) than acute cellular rejection 0R (healthy) (n = 42) (p = .02); an optimal donor fraction threshold (0.3%) was determined by the use of ROC analysis, revealing an AUC of 0.814 with a sensitivity of 0.65, specificity of 0.93, and an NPV of 81.8% for the absence of any allograft rejection.

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.

No studies were identified using myTAI_{HEART} levels that directed patient management in heart transplantation patients and which assessed patient outcomes.

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.

No inferences can be drawn about the clinical utility of myTAI_{HEART} levels for patients with heart transplantation.

Section Summary: Use of myTAI_{HEART} in Post–Heart Transplantation Patients

Few studies are available, and they are observational and of small sample size. No high-quality evidence was found supporting the ability of myTAI_{HEART} levels to predict transplant outcomes. A clinical validation study (N = 76) reports donor fraction cutoff value of 0.32% cutoff provides a NPV of 100.00% for grade 2R or higher acute cellular rejection, with 100.00% sensitivity and 75.48% specificity; an additional prospective study (N = 174) found that myTAI_{HEART} levels were associated with acute cellular rejection with an optimal donor fraction threshold of 0.3% to rule out the presence of either acute cellular rejection or antibody-mediated rejection.

MEASUREMENT OF VOLATILE ORGANIC COMPOUNDS FOR HEART TRANSPLANT**Clinical Context and Test Purpose**

The purpose of measuring volatile organic compounds in patients with a heart transplant is to assess for heart allograft rejection in a noninvasive manner.

The question addressed in this evidence review is: Does the measurement of volatile organic compounds improve the diagnostic assessment of allograft rejection in heart transplant patients?

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

Populations

The relevant population of interest are individuals with a heart transplant.

Interventions

The test being considered measures volatile organic compounds to assess for allograft rejection.

Comparators

The following test is currently being used to diagnose heart allograft rejection: routine endomyocardial biopsy.

Outcomes

The general outcomes of interest are OS, test validity, morbid events, and hospitalizations. Follow-up over months to years is necessary to monitor for signs of allograft rejection.

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 the clinical validity of measuring volatile organic compounds, studies that met the following eligibility criteria were considered:

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

Observational Studies

The U.S. Food and Drug Administration approval of the Heartsbreath test was based on the results of the Heart Allograft Rejection: Detection with Breath Alkanes in Low Levels (HARDBALL) study sponsored by the National Heart, Lung, and Blood Institute.⁵⁴ The HARDBALL study was a 3-year, multicenter study of 1061 breath samples in 539 heart transplant patients. Before the scheduled endomyocardial biopsy, patient breath was analyzed by gas chromatography and mass spectroscopy for volatile organic compounds. The amount of C4 to C20 alkanes and monomethylalkanes was used to derive the marker for rejection, known as the breath methylated alkane contour. The breath methylated alkane contour results were compared with subsequent biopsy results, as interpreted by 2 readers using the International Society for Heart and Lung Transplantation biopsy grading system as the criterion standard for rejection.²¹

The authors of the HARDBALL study reported that the abundance of breath markers that measured oxidative stress was significantly greater in grade 0, 1, or 2 rejection than in healthy normal persons. In contrast, in grade 3 rejection, the abundance of breath markers that measure oxidative stress was found to be reduced, most likely due to accelerated catabolism of alkanes and methylalkanes that make up the breath methylated alkane contour. The authors also reported that in identifying grade 3 rejection, the NPV of the breath test (97.2%) was similar to endomyocardial biopsy (96.7%) and that the breath test could potentially reduce the total number of biopsies performed to assess for rejection in patients at low-risk for grade 3 rejection. The sensitivity of the breath test was 78.6% vs 42.4% with biopsy. However, the breath test had

a lower specificity (62.4%) and a lower PPV (5.6%) in assessing grade 3 rejection than a biopsy (specificity, 97%; PPV = 45.2%). In addition, the breath test was not evaluated in grade 4 rejection.

Findings from the HARDBALL study were published by Phillips et al (2004). No subsequent studies evaluating the use of the Heartsbreath test to assess for graft rejection were identified in literature updates.

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.

No RCTs assessing the measurement of volatile organic compounds to diagnose cardiac allograft rejection were 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.

Because the clinical validity of measuring volatile organic compounds to assess for cardiac allograft rejection has not been established, a chain of evidence to support clinical utility cannot be constructed.

Section Summary: Measurement of Volatile Organic Compounds for Heart Transplant

A published study found that for identifying grade 3 (now grade 2R) rejection, the NPV of the breath test the study evaluated (97.2%) was similar to endomyocardial biopsy (96.7%), and the sensitivity of the breath test (78.6%) was better than that for biopsy (42.4%). However, the breath test had a lower specificity (62.4%) and a lower PPV (5.6%) in assessing grade 3 rejection than a biopsy (specificity, 97%; PPV = 45.2%). The breath test was also not evaluated for grade 4 rejection. At present, no studies evaluating the clinical utility for the measurement of volatile organic compound testing for heart transplant have been identified.

GENE EXPRESSION PROFILING AND DONOR-DERIVED CELL-FREE DNA TESTING FOR HEART TRANSPLANT

Clinical Context and Test Purpose

The purpose of gene expression profiling (GEP) and donor-derived cell-free DNA (dd-cfDNA) testing in patients with a heart transplant is to assess for allograft rejection.

The question addressed in this evidence review is: Does the use of GEP testing, alone or in combination with dd-cfDNA testing, improve the diagnostic assessment of allograft rejection in heart transplant patients?

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

Populations

The relevant population of interest are individuals with heart transplants.

Interventions

The test being considered is GEP to assess for allograft rejection (i.e., AlloMap), used alone or in combination with AlloSure Heart dd-cfDNA testing. The combination of these tests is commercially marketed as HeartCare (CareDx).

AlloMap test results are reported on a scale from 0 to 40, with a proposed high-risk cutoff of ≥ 30 for patients < 6 months post-transplant and ≥ 34 for patients ≥ 6 months post-transplant. The AlloSure test reports the fraction of dd-cfDNA, with a proposed high-risk cutoff of $\geq 0.15\%$. Clinical interpretation of alternate thresholds for quiescence ($< 0.12\%$), injury (0.20%) and severe injury (0.35%) have also been proposed. The HeartCare report provides the AlloMap score, AlloMap score variability, and AlloSure percent dd-cfDNA. Direct guidance for the combined interpretation of results is not provided in the HeartCare report, but potential clinical implications of concordant and discordant test scenarios have been proposed.⁵⁵

Comparators

The following test is currently being used to diagnose cardiac allograft rejection: routine endomyocardial biopsy.

Outcomes

The general outcomes of interest are OS, test validity, morbid events, and hospitalizations. Follow-up over months to years is needed to monitor for signs of allograft rejection. Beneficial outcomes resulting from a true-negative test result are avoiding unnecessary subsequent biopsy. Harmful outcomes resulting from a false-positive result may include unnecessary biopsy or unnecessary treatment. Harmful outcomes from a false-negative result are increased risk of adverse transplant outcomes.

In a triage scenario, the test would need to identify precisely a group of patients that could safely forgo biopsy; therefore, the sensitivity, NPV, and negative likelihood ratio are key test performance characteristics.

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 the clinical validity of GEP testing, studies that met the following eligibility criteria were considered:

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

AlloMap

Systematic Reviews

A TEC Assessment (2011) reviewed the evidence on the use of GEP using the AlloMap test.⁵⁶ The Assessment concluded that the evidence was insufficient to permit conclusions about the effect of the AlloMap test on health outcomes. Key evidence in the TEC Assessment is described below.

Nonrandomized Studies

Patterns of gene expression for the development of the AlloMap test were studied in the Cardiac Allograft Rejection Gene Expression Observation (CARGO) study, which included 8 U.S. cardiac transplant centers enrolling 629 cardiac transplant recipients.⁵⁷ The study included the discovery and validation phases. In the discovery phase, patient blood samples were obtained during the endomyocardial biopsy, and the expression levels of more than 7000 genes involved in immune responses were assayed and compared with the biopsy results. A subset of 252 candidate genes was identified, from which a panel of 11 genes was selected for evaluation. A proprietary algorithm was applied to the results, producing a single score that considers the contribution of each gene in the panel.

The validation phase of the CARGO study, published by Deng et al (2006), was prospective, blinded, and enrolled 270 patients.⁵⁷ Primary validation was conducted using samples from 63 patients independent from discovery phases of the study and enriched for biopsy-proven evidence of rejection. A prospectively defined test cutoff value of 20 resulted in a sensitivity of 84% of patients with moderate/severe rejection but a specificity of 38%. Of note, in the “training set” used in the study, these rates were 80% and 59%, respectively. The authors evaluated the 11-gene expression profile on 281 samples collected at 1 year or more from 166 patients who were representative of the expected distribution of rejection in the target population (and not involved in discovery or validation phases of the study). When a test cutoff of 30 was used, the NPV (no moderate/severe rejection) was 99.6%; however, only 3.2% of specimens had grade 3 or higher rejection. In this population, grade 1B scores were found to be significantly higher than grade 0, 1A, and 2 scores but were similar to grade 3 scores.

A second prospective multicenter study evaluating the clinical validity of GEP with the AlloMap test (CARGO II) was published by Crespo-Leiro et al (2016).⁵⁸ The study enrolled 499 heart transplant recipients undergoing surveillance for allograft rejection. The reference standard for rejection status was histologic grade from an endomyocardial biopsy performed on the same day

as blood samples were collected. Blood samples need to be collected 55 days or more post transplant, more than 30 days after blood transfusion, more than 21 days after administration of prednisone 20 mg/day or more, and more than 60 days after treatment for a prior rejection. Patients had a total of 1579 eligible blood samples for which paired GEP scores and endomyocardial biopsy rejection grades were available.

As in the original CARGO study, the proportion of cases of rejection was small. The prevalence of moderate-to-severe rejection (grade 2R/>3A) reported by local pathologists was 3.2%, which was reduced to 2.0% when confirmation from 1 or more other independent pathologists was required. At a GEP cutoff of 34, for patients who were at least 2 to 6 months post transplant, the sensitivity of GEP for detecting grade 2R/>3A was 25.0%, and the specificity was 88.7%. The PPV and NPV were 4.0% and 98.4%, respectively. Using the same cutoff of 34, for patients more than 6 months post transplant, the sensitivity of GEP was 25.0%, the specificity was 88.8%, the PPV was 4.3%, and the NPV was 98.3%. The number of true-positives used in the above calculations was 5 (9.1%) of 55 for patients at least 2 to 6 months post transplant and 6 (10.2%) of 59 for patients more than 6 months post transplant.

Kanwar et al (2021) published data from the Outcomes AlloMap Registry (OAR) indicating that asymptomatic or active cytomegalovirus infection is associated with significantly higher AlloMap scores among heart transplant recipients compared to those without infection, even in the absence of acute rejection, potentially resulting in unnecessary biopsies among surveillance patients.⁵⁹ Donor-derived cell-free DNA levels measured by the AlloSure Heart test available for a small subset of samples (5.3%) were not significantly different between groups. The authors conclude that further assessment of the combined use of AlloMap and AlloSure scores is required to determine if this will improve differentiating infection-related from rejection-related immune activation. The combined use of these tests, commercially available as HeartCare (CareDx), is addressed in the following section.

HeartCare

The commercially available HeartCare (CareDx) test combines AlloMap GEP testing with AlloSure Heart measurement of percent dd-cfDNA. The combined use of GEP and dd-cfDNA testing for surveillance of acute rejection was assessed in a single-center, retrospective study conducted by Gondi et al (2021) between February 2019 and March 2020.⁶⁰ Patients (N=153) were required to be \geq 55 days post transplant, hemodynamically stable, \geq 15 years of age, and single-organ recipients. The majority of patients were male (74.5%) and white (78.4%) with an average age of 54.5 years. Patients were assessed once monthly between 2 and 12 months, every 3 months between 12 and 24 months, and every 6 months between 24 and 36 months posttransplant. Pre-specified thresholds for GEP scores were \geq 30 for patients < 6 months post transplant and \geq 34 for patients \geq 6 months post-transplant. The pre-specified threshold for percent dd-cfDNA was \geq 0.20% based on a prior study of the AlloSure test by Khush et al (2019),⁶¹ described in the following section. In patients < 6 months post-transplant, endomyocardial biopsy was performed regardless of test results. For patients \geq 6 months post-transplant who received both GEP and dd-cfDNA testing, endomyocardial biopsy was canceled in patients with dd-cfDNA < 0.20% regardless of AlloMap score. In patients with positive AlloMap scores but negative dd-cfDNA, endomyocardial biopsy could be performed or deferred in favor of repeat dd-cfDNA testing.

Among 495 samples, overall test result distributions were 59.6% for patients negative on both tests, 12.3% for patients positive by dd-cfDNA only, 22.6% for patients positive by GEP only, and 5.5% positive by both GEP and dd-cfDNA. The combined testing approach resulted in a 12.7% reduction (48 biopsies) in endomyocardial biopsy volume compared to GEP testing alone. Among the 172 biopsies performed, 2 patients with cell-mediated rejection were identified, with corresponding dual-positive tests. Two patients with antibody-mediated rejection were identified, with corresponding tests that were only positive by dd-cfDNA. The study is limited by its retrospective design, incomplete evaluation of performance characteristics, and lack of reporting on health outcomes.

AlloSure

Khush et al (2019) published performance characteristics for the AlloSure Heart dd-cfDNA test as assessed in the D-OAR prospective, multicenter registry study.⁶¹ Patients already undergoing AlloMap testing for surveillance were eligible for inclusion; however following a protocol amendment, dd-cfDNA specimens were only obtained in patients with clinical suspicion of rejection and a planned for-cause biopsy after 2016 through 2018. The majority of dd-cfDNA samples (81%) were drawn in the first-year post-transplant. The D-OAR cohort included 841 biopsy-paired dd-cfDNA results, of which 587 were performed for routine surveillance of rejection. Overall, cell-mediated rejection and antibody-mediated rejection were biopsy-confirmed in 17 and 18 cases, respectively. The AUC for detecting acute rejection was 0.64 (95% CI, 0.52 to 0.75). At a 0.2% cutoff for dd-cfDNA, the sensitivity, specificity, PPV, and NPV for detection of acute rejection was 80%, 44%, 8.9%, and 97.1% respectively. For the subgroup of patients undergoing surveillance, the sensitivity, specificity, PPV, and NPV were 38.1%, 84.0%, 8.1%, and 97.3%, with a corresponding AUC of 0.61 (95% CI, 0.46 to 0.74). Among for-cause samples, the sensitivity, specificity, PPV, and NPV were 53.8%, 76.1%, 11.6%, and 96.6%, respectively. The study is limited by the protocol changes designed to increase the number of observed rejection events overall and low availability of concurrent dd-cfDNA results with respect to biopsy specimens (58%).

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.

AlloMap

Randomized Controlled Trials

Kobashigawa et al (2015) published the results of a pilot RCT evaluating the use of the AlloMap test in patients who were 55 days to 6 months post transplant.⁶² The trial design was similar to that of the Invasive Monitoring Attenuation through Gene Expression (IMAGE) RCT, discussed

next. Sixty subjects were randomized to rejection monitoring with AlloMap or with endomyocardial biopsy at prespecified intervals of 55 days and 3, 4, 5, 6, 8, 10, and 12 months post transplant. The threshold for a positive AlloMap test was set at 30 for patients 2 to 6 months post transplant and 34 for patients after 6 months post transplant, based on data from the CARGO study. Endomyocardial biopsy outside of the scheduled visits was obtained in either group if there was clinical or echocardiographic evidence of graft dysfunction and for the AlloMap group if the score was above the specified threshold. The incidence of the primary outcome at 18 months post transplant (a composite outcome of the first occurrence of any of the following: death or retransplant, rejection with hemodynamic compromise, or allograft dysfunction due to other causes) did not differ significantly between the AlloMap and biopsy groups (10% vs 17%; $p = .44$). The number of biopsy-proven rejection episodes (International Society for Heart and Lung Transplantation grading system $\geq 2R$) within the first 18 months did not differ significantly between groups (3 in the AlloMap group vs 1 in the biopsy group; $p = .31$). Of the rejections in the AlloMap group, 1 was detected after an elevated routine AlloMap test, while 2 were detected after patients presenting with hemodynamic compromise. As in the IMAGE study, a high proportion of rejection episodes were detected by clinical signs or symptoms (however, this study had only 3 rejection episodes in the AlloMap group).

In 2010, the results of the IMAGE study were published.^{63,64} This was an industry-sponsored, nonblinded, noninferiority RCT that compared outcomes in 602 patients managed with the AlloMap test ($n=297$) or with routine endomyocardial biopsies ($n=305$). The trial included adults from 13 centers who underwent cardiac transplantation between 1 and 5 years prior to participating, were clinically stable and had a left ventricular ejection fraction of at least 45%. To increase enrollment, the trial protocol was later amended to include patients who had undergone transplantation between 6 months and 1 year prior to participating; this subgroup ultimately comprised only 15% of the final sample ($n=87$). Each transplant center used its own protocol for determining the intervals for routine testing. At all sites, patients in both groups underwent clinical and echocardiographic assessments in addition to the assigned surveillance strategy. According to the study protocol, patients underwent biopsy if they had signs or symptoms of rejection or allograft dysfunction at clinic visits (or between visits) or if the echocardiogram showed a left ventricular ejection fraction decrease of at least 25% compared with the initial visit. Additionally, patients in the AlloMap group underwent biopsy if their test score was above a specified threshold; however, if they had 2 elevated scores with no evidence of rejection found on 2 previous biopsies, no additional biopsies were required. The AlloMap test score varied from 0 to 40, with higher scores indicating a higher risk of transplant rejection. The investigators initially used 30 as the cutoff for a positive score; the protocol was amended to use a cutoff of 34 to minimize the number of biopsies needed. Fifteen patients in the AlloMap group and 26 in the biopsy group did not complete the trial.

The primary outcome was a composite variable: (1) the first occurrence of rejection with hemodynamic compromise; (2) graft dysfunction due to other causes; (3) death; or (4) retransplantation. Use of the AlloMap test was considered noninferior to the biopsy strategy if the 1-sided upper boundary of the 95% CI for the hazard ratio comparing the 2 strategies was less than the prespecified margin of 2.054. The margin was derived using the estimate of a 5% event

rate per year in the biopsy group, taken from published observational studies, and allowing for an event rate of up to 10% per year in the AlloMap group.

According to Kaplan-Meier analysis, the 2-year event rate was 14.5% in the AlloMap group and 15.3% in the biopsy group. The corresponding hazard ratio was 1.04 (95% CI, 0.67 to 1.68). The upper boundary of the CI of the hazard ratio (1.68) fell within the prespecified noninferiority margin (2.054); thus, GEP was considered noninferior to endomyocardial biopsy. Death from all causes, a secondary outcome, did not differ significantly between groups. There were 13 (6.3%) deaths in the AlloMap group and 12 (5.5%) in the biopsy group ($p = .82$). During follow-up, there were 34 treated episodes of graft rejection in the AlloMap group. Only 6 (18%) of the 34 patients with graft rejection presented solely with elevated AlloMap scores. Twenty (59%) patients presented with clinical signs/symptoms and/or graft dysfunction on echocardiogram and 7 patients had an elevated AlloMap score plus clinical signs/symptoms with or without graft dysfunction on echocardiogram. In the biopsy group, 22 patients were detected solely due to an abnormal biopsy.

A total of 409 biopsies were performed in the AlloMap group and 1249 in the biopsy group. Most biopsies in the AlloMap group (67%) were performed because of elevated gene profiling scores. Another 17% were performed due to clinical or echocardiographic manifestations of graft dysfunction, and 13% were performed as part of routine follow-up after treatment for rejection. There was 1 (0.3%) adverse event associated with biopsy in the AlloMap group and 4 (1.4%) in the biopsy group. In terms of quality of life, the physical health and mental health summary scores of the 12-Item Short-Form Health Survey were similar in the 2 groups at baseline and did not differ significantly between groups at 2 years.

A limitation of the trial was that the threshold for a positive AlloMap test was changed partway through the study; thus, the optimal test cutoff remains unclear. Moreover, the trial was not blinded, which could have affected treatment decisions based on clinical findings, such as whether to recommend a biopsy. In addition, the study did not include a group that only received clinical and echocardiographic assessment, so the value of AlloMap testing beyond that of clinical management alone cannot be determined. The uncertain incremental benefit of the AlloMap test is highlighted by the finding that only 6 of the 34 treated episodes of graft rejection detected during follow-up in the AlloMap group were initially identified solely due to an elevated GEP score. Since 22 episodes of asymptomatic rejection were detected in the biopsy group, the AlloMap test does not appear to be a sensitive test, possibly missing more than half of the episodes of asymptomatic rejection. Because clinical outcomes were similar in the 2 groups, there are at least 2 possible explanations: the clinical outcome of the study may not be sensitive to missed episodes of rejection, or it is not necessary to treat asymptomatic rejection. In addition, the trial was only statistically powered to rule out more than a doubling of the rate of the clinical outcome, which some may believe is an insufficient margin of noninferiority. Finally, only 15% of the final study sample had undergone transplantation less than 1 year before study participation; therefore, findings might not be generalizable to the population of patients 6 to 12 months post transplant.

HeartCare

Direct evidence of clinical utility was not identified for the HeartCare test.

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 the clinical validity of GEP testing, alone or in combination with dd-cfDNA testing, to assess for cardiac allograft rejection has not been established, a chain of evidence to support clinical utility cannot be constructed.

SUBSECTION SUMMARY: GENE EXPRESSION PROFILING FOR HEART TRANSPLANT**AlloMap**

The 2 studies (CARGO, CARGO II) examining the diagnostic performance of GEP using the AlloMap test for detecting moderate or severe rejection were flawed by lack of a consistent threshold (i.e., 20, 30, or 34) for determining positivity and by a small number of positive cases. In the available studies, although the NPVs were relatively high (i.e., at least 88%), the performance characteristics were calculated based on the detection of 10 or fewer cases of rejection each. Moreover, the PPV in the CARGO II study was only 4.0% for patients who were at least 2 to 6 months post transplant and 4.3% for patients more than 6 months post transplant. The ability of the AlloMap test to differentiate between infection-related and rejection-related graft injury has also been called into question.

The most direct evidence on the clinical utility of GEP using the AlloMap test comes from a large RCT comparing a GEP-directed strategy with an endomyocardial biopsy-directed strategy for detecting rejection; it found that the GEP-directed strategy was noninferior. However, given the high proportion of rejection episodes in the GEP-directed strategy group detected by clinical signs/symptoms, the evidence is insufficient to determine that health outcomes are improved because of the uncertain incremental benefit of GEP. In addition, a minority of subjects assessed were in the first year post transplant. Results from a pilot RCT would suggest that GEP may have a role in evaluating for heart transplant rejection beginning at 55 days post transplant, but the trial was insufficiently powered to permit firm conclusions about the noninferiority of early GEP use.

SUBSECTION SUMMARY: GENE EXPRESSION PROFILING WITH DONOR-DERIVED CELL-FREE DNA TESTING FOR HEART TRANSPLANT**HeartCare**

One retrospective study assessing the combined use of GEP testing with AlloMap and dd-cfDNA testing with AlloSure Heart reported a 12.7% reduction in endomyocardial biopsy volume when combined testing was used compared to AlloMap alone. However, this observation is limited by a lack of reporting on long-term health outcomes and incomplete diagnostic performance assessment for combined testing. The independent diagnostic performance of the AlloSure dd-

cfDNA test was evaluated in 1 registry study, which was flawed by protocol changes and low availability of concurrent dd-cfDNA results. While NPVs were generally high (>96.6%), PPVs were low overall (8.9%), in surveillance patients (8.1%), and in patients with clinical suspicion of rejection (11.6%).

DONOR-DERIVED CELL-FREE DNA TESTING FOR RENAL TRANSPLANT

Clinical Context and Test Purpose

The purpose of dd-cfDNA testing in patients with renal transplant who are undergoing surveillance or have clinical suspicion of allograft rejection is to detect allograft rejection.

The question addressed in this evidence review is: Does testing for dd-cfDNA improve outcomes in renal transplant patients who are undergoing surveillance or have clinical suspicion of allograft rejection?

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

Populations

The relevant population of interest are individuals with renal transplants who are undergoing surveillance or who have a clinical suspicion of allograft rejection.

Clinical suspicion of allograft rejection may be indicated by clinical symptoms (e.g., pain) or dynamic changes in laboratory parameters.

Allograft dysfunction is typically asymptomatic and has a broad differential, including graft rejection. Diagnosis and rapid treatment are recommended to preserve graft function and prevent loss of the transplanted organ.

Interventions

The test being considered is dd-cfDNA testing to assess for renal allograft rejection (i.e., AlloSure or Prospera).

Various clinical pathways have been proposed for these tests. Use of the Prospera test is recommended when there is clinical suspicion of active rejection and for regular surveillance of subclinical rejection.³² In a surveillance scenario, regular testing is recommended at 1, 2, 3, 4, 6, 9, and 12 months after renal transplant or most recent rejection. Thereafter, the test should be repeated quarterly. The proportion of dd-cfDNA relative to total cfDNA is reported, with detection of $\geq 1\%$ dd-cfDNA indicating increased risk for active rejection. The percent dd-cfDNA change between tests is also reported.³³ In the surveillance scenario, patients with a negative result may avoid biopsy and it is recommended that a positive test result is incorporated with clinical findings to determine whether a biopsy is indicated. When there is clinical suspicion of rejection, testing is recommended as an adjunct to biopsy for treatment response monitoring, or as a rule-out test for biopsy.

For the AlloSure test, various dd-cfDNA thresholds are suggested depending on the clinical scenario and include the detection of antibody-mediated rejection (ABMR) in patients with donor-specific antibodies (DSA), the detection of "likely" active rejection, the prediction of adverse outcomes as an adjunct to biopsy-confirmed T cell-mediated (TCMR) 1A/borderline rejections, and for the exclusion of active rejection.⁶⁵ A routine testing schedule is also recommended, and details regarding its clinical rationale have been published.⁶⁶

Comparators

The following test is currently being used to confirm a clinical suspicion of allograft rejection: renal biopsy. The adoption of protocol (i.e., surveillance) biopsies varies across transplant centers and its use is not standardized.

Clinical suspicion of allograft rejection may be indicated by physical symptoms and/or dynamic changes in laboratory parameters (e.g., serum creatinine, estimated glomerular filtration rate [eGFR], DSA).

Outcomes

The general outcomes of interest are OS, test validity, morbid events, and hospitalizations. Follow-up over months to years is needed to monitor for signs of allograft rejection.

For a primary kidney transplant, graft survival at 1 year is 94.7%; at 5 years, graft survival is 78.6%.²⁷

Beneficial outcomes resulting from a true-negative test result are avoiding unnecessary subsequent biopsy. Harmful outcomes resulting from a false-positive result may include an unnecessary biopsy or unnecessary treatment. Harmful outcomes from a false-negative result are increased risk of adverse transplant outcomes.

In a triage scenario, the test would need to identify precisely a group of patients that could safely forgo biopsy; therefore, the sensitivity, NPV, and negative likelihood ratio are key test performance characteristics.

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 the clinical validity of dd-cfDNA testing, studies that met the following eligibility criteria were considered:

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

Observational Studies

Major study results are summarized in Table 8.

AlloSure

Development of the AlloSure test was conducted in the multicenter prospective Circulating Donor-Derived Cell-Free DNA in Blood for Diagnosing Acute Rejection in Kidney Transplant Recipients (DART) study by Bloom et al (2017), which both recruited patients who were less than 3 months after renal transplant (n=245) and recruited renal transplant patients requiring a biopsy for suspicion of graft rejection (n=139).⁶⁷ For the primary analysis, an active rejection was defined as the combined categories of T cell-mediated rejection, acute/active antibody-mediated rejection, and chronic/active antibody-mediated rejection as defined by the Banff working groups. Only patients undergoing biopsy were considered; further exclusion of biopsies that were not for cause or had an inadequate or incomplete collection of biopsies or corresponding blood samples or had prior allograft in situ. These exclusions resulted in the main study cohort of 102 patients (107 biopsies). Within this population, acute rejection was noted in 27 patients (27 biopsies). After statistical analysis accounting for multiple biopsies from the same patient, the threshold dd-cfDNA fraction corresponding to acute rejection was set to 1.0% or higher. In the main study group, this resulted in a sensitivity of 59% (95% CI, 44% to 74%) and specificity of 85% (95% CI, 79% to 81%) for detecting active rejection versus no rejection. Using the original data set including all biopsies performed for clinical suspicion of rejection, 58 cases of acute rejection were diagnosed in 204 biopsies (170 patients). This PPV was 61% and the NPV 84%. Biopsies performed for surveillance (n=34 biopsies) were excluded from analysis in this study, as only 1 biopsy for surveillance demonstrated acute rejection. Study limitations included the absence of a validation data set.

Huang et al (2019) conducted a smaller single center that recruited 63 renal transplant patients with suspicion of rejection that had AlloSure assessment of dd-cfDNA within 30 days of an allograft biopsy.⁶⁸ Median years from transplant to dd-cfDNA measurement was 2.0 (interquartile range, 0.3 to 6.5). Within this population, biopsy found acute rejection in 34 (54%) of patients; 10 (15.9%) were cell-mediated only, 22 (25.4%) were antibody-mediated only, and 2 (3.2%) were mixed cell-mediated and antibody-mediated. In contrast to the study by Bloom et al (2017), the optimal threshold for a positive dd-cfDNA result was identified as $\geq 0.74\%$. For the outcome of any rejection (i.e., cell-mediated, antibody-mediated, or mixed), use of this threshold was associated with an overall sensitivity of 79.4%, specificity of 72.4%, PPV of 77.1%, and NPV of 75.0%. Discrimination of rejection differed by biopsy findings, however. For the subgroup of patients with antibody-mediated rejection, the sensitivity was 100%, specificity was 71.8%, PPV was 68.6%, and NPV was 100%. The dd-cfDNA test did not discriminate rejection in patients with cell-mediated rejection, as evidenced by an AUC of 0.43 (95% CI, 0.17 to 0.66). The major limitations of this study are its small sample size and single-center setting.

Stites et al (2020) assessed clinical outcomes in 79 patients diagnosed with TCMR 1A/borderline rejection with simultaneous AlloSure assessment of dd-cfDNA across 11 centers between June 2017 and May 2019.⁶⁹ Timing of testing with respect to the date of transplantation was not reported. Elevated levels of dd-cfDNA ($\geq 0.5\%$) were detected in 42 (53.2%) patients. No statistically significant differences between dd-cfDNA distributions when stratified by protocol

versus for-cause biopsies was detected ($p = .7307$). Elevated levels of dd-cfDNA were associated with adverse clinical outcomes compared to patients with low levels ($< 0.5\%$), including decline in eGFR (8.5% versus 0%; $p = .004$), de novo DSA formation (40% versus 2.7%; $p < .0001$), and future or persistent rejection (21.4% versus 0%; $p = .003$). The authors hypothesize that the use of dd-cfDNA may complement histological evaluation and risk stratify patients with TCMR 1A or borderline rejection identified on biopsy and propose the use of reference ranges as opposed to absolute dd-cfDNA cutoff thresholds.

Additional analyses of the DART study have reported on associations between first-year AlloSure dd-cfDNA fraction or serial variability and subsequent eGFR decline (Sawinski et al [2021]),⁷⁰ and combined use of dd-cfDNA and DSA testing to diagnose active antibody-mediated rejection (Jordan et al [2018], Mayer et al [2021]).^{71,72}

Puliyanda et al (2021) conducted a prospective study of 67 pediatric renal transplant recipients enrolled across 2 medical centers between 2017 and 2019.⁷³ Patients had a median age of 11 years (interquartile range [IQR], 4 to 13) and median time post-transplant to first AlloSure dd-cfDNA measurement was 55.6 months. Nineteen patients (28.4%) received dd-cfDNA testing in the absence of clinical suspicion of rejection. Median dd-cfDNA scores in the surveillance versus for-cause cohorts were 0.37% (IQR, 0.19% to 1.10%) and 0.47% (IQR, 0.24% to 2.15%), respectively. Among patients undergoing surveillance, 26.3% (5/19 patients) had a dd-cfDNA score $>1\%$ with biopsies indicating 4 cases of antibody-mediated rejection and 1 case of mixed rejection. Among patients with clinical suspicion of rejection, 43.8% (21/48 patients) had dd-cfDNA scores $>1\%$. All for-cause biopsies showed evidence of rejection, including 11 cases of antibody-mediated rejection, 2 cases of T cell-mediated rejection, and 8 cases of mixed rejection. An additional 7 patients with clinical suspicion of rejection underwent biopsy despite dd-cfDNA scores $< 1\%$, revealing 4 cases without rejection, 1 case with antibody-mediated rejection, 1 case with cell-mediated rejection, and 1 case of mixed rejection. Among all patients with biopsy-matched results (33/67), dd-cfDNA $>1\%$ was associated with a sensitivity of 86% and specificity of 100%, with a corresponding AUC of 0.996 ($p = .002$). No significant difference in serum creatinine change from baseline to testing was identified for those with rejection compared to those without. The study is limited by the small sample size and lack of biopsy-matched data for a complete assessment of false negatives. The authors also note that the 1% dd-cfDNA cutoff threshold was used based on prior studies in adults and it is unclear if this is appropriate for the pediatric population. Additionally, the authors suggest that relative increases in dd-cfDNA, as opposed to absolute values, may be more valuable in the pediatric population, given that appropriate cutoff thresholds may depend on child age and size.

Prospera

Sigdel et al (2019) evaluated the diagnostic accuracy of the Prospera dd-cfDNA test in a retrospective analysis of 300 biorepository plasma samples from kidney transplant recipients at a single academic medical center.⁷⁴ Of the 300 samples (193 patients), 217 were biopsy-matched with 38 cases of active rejection, 72 cases of borderline rejection, 82 with stable allografts, and 15 cases of other kidney injuries. The sample cohort was demographically diverse, including women (42.5%), Hispanic and Latino patients (34.6%), Black or African American patients (14%), and pediatric patients (20%). Indication for renal transplantation was unknown in 45.6%

of samples. The majority of samples (72.3%) were drawn on the day of surveillance (n = 114 [52.5%] patients) or clinically indicated biopsy (n=103 [47.5%] patients). Timing of tests with respect to the date of transplantation was not reported. Biopsies were evaluated by a single pathologist according to 2017 Banff criteria and classified as active rejection or non-rejection (i.e., borderline rejection, other injury, or stable allograft status). Median dd-cfDNA levels were significantly higher in biopsy-proven active rejection (2.32%) versus non-rejection subgroups (0.47%; p <.0001). All subtypes of active rejection could be detected, and median dd-cfDNA did not differ significantly between antibody-mediated (2.2%), T cell-mediated (2.7%), and combined subtypes (2.6%).

Sigdel et al (2019) also assessed the performance characteristics of eGFR, which was calculated as a function of serum creatinine with adjustments for age, sex, and race based on the Modification of Diet in Renal Disease (MDRD) Study equation.⁷⁴ At a cutoff threshold of < 60, the sensitivity and specificity for eGFR were lower compared to dd-cfDNA, at 67.8% (95% CI, 51.3% to 84.2%) and 65.3% (95% CI, 57.6% and 73.0%), respectively, with a corresponding AUC of 0.74 (95% CI, 0.66 to 0.83). However, the relevance of absolute eGFR measurements is limited as dynamic changes in laboratory parameters (e.g., serum creatinine elevation, eGFR decline) are used to flag impaired kidney function in clinical practice in the transplant population. Separate eGFR estimates in the for-cause subgroup were not reported. Major limitations of this study include its retrospective design and single-center setting. While the dd-cfDNA cutoff was prespecified, it was based on prior studies of the AlloSure test and may not be optimized for Prospera.

Bunnapradist et al (2021) noted that while % dd-cfDNA is a promising noninvasive biomarker for detecting renal allograft rejection, levels can be artificially depressed by high levels of circulating cfDNA, which may be observed in patients who are obese, have recently undergone surgery, have medical complications, or receive certain medications, potentially leading to false-negative results.⁷⁵ The authors suggest that a combination of dd-cfDNA fraction and absolute quantity thresholds may improve the sensitivity of allograft rejection while maintaining high specificity.

Table 8. Diagnostic Performance Characteristics of dd-cfDNA Tests for Detection of Active Kidney Allograft Rejection

| Study; dd-cfDNA threshold | Biopsy-Matched Samples | Prevalence, n (%) ^a | Sensitivity, % (95% CI) | Specificity, % (95% CI) | AUC (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) |
|--|------------------------|--------------------------------|-------------------------|-------------------------|---------------------|-----------------|-----------------|
| AlloSure | | | | | | | |
| Bloom et al (2017) (≥1%)⁶⁷ | | | | | | | |
| For-cause, dd-cfDNA | 107 | 27 (25.2) | 59 (44 to 74) | 85 (79 to 81) | 0.74 (0.61 to 0.86) | 61 (NR) | 84 (NR) |

| Study; dd-cfDNA threshold | Biopsy-Matched Samples | Prevalence, n (%)^a | Sensitivity, % (95% CI) | Specificity, % (95% CI) | AUC (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) |
|--|-------------------------------|--------------------------------------|----------------------------------|----------------------------------|---------------------|--|--|
| For-cause, SCr | 204 | 58 (28.4) | NR | NR | 0.54 (0.43 to 0.66) | NR | NR |
| Huang et al (2019) ($\geq 0.74\%$)^{68,} | | | | | | | |
| For-cause, any rejection | 63 (patients) | 34 (54) | 79.4 (NR) | 72.4 (NR) | 0.71 (0.58 to 0.85) | 77.1 (NR) | 75 (NR) |
| For-cause, CMR | 63 (patients) | 10 (16) | NR | NR | 0.42 (0.17 to 0.66) | NR | NR |
| Prospera | | | | | | | |
| Sigdel et al (2019) ($\geq 1\%$)^{74,} | | | | | | | |
| Overall, dd-cfDNA | 217 | 33 (17.5) | 88.7 (77.7 to 99.8) | 72.6 (65.4 to 79.8) | 0.87 (0.80 to 0.95) | 52.0 (44.7 to 59.2) ^c | 95.1 (90.5 to 99.7) ^c |
| Overall, eGFR | 217 | 33 (17.5) | 67.8 (51.3 to 84.2) | 65.3 (57.6 to 73.0) | 0.74 (0.66 to 0.83) | 39.4 (31.6 to 47.3) ^c | 85.9 (75.9 to 92.2) ^c |
| Surveillance, dd-cfDNA | 114 | 12 (11.4) | 92.3 (64.0 to 99.8) | 75.2 (65.7 to 83.3) | 0.89 (0.79 to 0.99) | 55.4 (46.2 to 64.7) ^c 32.4 (24.8 to 41.1) ^d | 96.7 (90.6 to 99.9) ^c 98.7 (92.0 to 99.8) ^d |
| For-cause, dd-cfDNA | 103 | 25 (24.3) | 84.0 (63.9 to 95.5) ^b | 68.0 (56.4 to 78.1) ^b | NR | 45.7 (36.8 to 54.8) ^d | 93.0 (84.2 to 97.1) ^d |

AUC: area under the receiver-operating curve; CI: confidence interval; CMR: cell-mediated rejection; dd-cfDNA: donor-derived cell-free DNA; eGFR: estimated glomerular filtration rate; NPV: negative predictive value; NR: not reported; PPV: positive predictive value; SCr: serum creatinine.

^a Study disease prevalence.

^b Calculated based on reported case numbers.

^c Projected value as reported based on assumed disease prevalence of 25% in an at-risk population.

^d Calculated value based on study disease prevalence.

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.

No RCTs assessing the clinical utility of dd-cfDNA (i.e., AlloSure, Prospera) testing to diagnose renal allograft rejection were 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.

Because the clinical validity of dd-cfDNA (i.e., AlloSure, Prospera) testing to assess for renal allograft rejection has not been established, a chain of evidence to support clinical utility cannot be constructed.

Section Summary: Donor-Derived Cell-Free DNA Testing for Renal Transplant

A discovery phase prospective study using the AlloSure test has been performed in a multicenter setting. A subsequent smaller single-center study that explored variation in clinical validity based on different rejection mechanisms found the strongest performance characteristics for AlloSure with antibody-mediated rejection. A single retrospective study of the Prospera test reported a PPV and NPV of 52% and 95%, respectively, for detection of active rejection among a combined cohort of patients undergoing surveillance or for-cause biopsies using the 1% dd-cfDNA threshold previously proposed for AlloSure. Larger prospective studies validating dd-cfDNA thresholds for active rejection are needed to develop conclusions for each test. At present, no studies evaluating the clinical utility for AlloSure or Prospera dd-cfDNA testing were identified.

DONOR-DERIVED CELL-FREE DNA TESTING FOR LUNG TRANSPLANT

Clinical Context and Test Purpose

The purpose of dd-cfDNA testing in patients with lung transplant who are undergoing surveillance is to detect allograft rejection.

The question addressed in this evidence review is: Does testing for dd-cfDNA improve outcomes in lung transplant patients who are undergoing surveillance for allograft rejection?

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

Populations

The relevant population of interest are individuals with lung transplants who are undergoing surveillance for allograft rejection.

Interventions

The test being considered is dd-cfDNA testing to assess for lung allograft rejection (i.e., AlloSure).

A regular testing schedule is recommended for patients undergoing surveillance, with monthly testing in the first-year post-transplant and quarterly in the years 2-3. The proportion of dd-cfDNA relative to total cfDNA is reported. The report also notes that a threshold of >0.85% dd-cfDNA is associated with a higher probability of acute cellular rejection, CLAD, and antibody-mediated rejection and that the NPV is maximized at a % dd-cfDNA cutoff of 0.20%.

Comparators

The following test is currently being used to confirm a clinical suspicion of allograft rejection: bronchoscopy with transbronchial biopsy.

Outcomes

The general outcomes of interest are OS, test validity, morbid events, and hospitalizations. Follow-up over months to years is needed to monitor for signs of allograft rejection.

Beneficial outcomes resulting from a true-negative test result are avoiding unnecessary subsequent biopsy. Harmful outcomes resulting from a false-positive result may include an unnecessary biopsy or unnecessary treatment. Harmful outcomes from a false-negative result are increased risk of adverse transplant outcomes.

In a triage scenario, the test would need to identify precisely a group of patients that could safely forgo biopsy; therefore, the sensitivity, NPV, and negative likelihood ratio are key test performance characteristics.

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 the clinical validity of dd-cfDNA testing, studies that met the following eligibility criteria were considered:

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

OBSERVATIONAL STUDIES

AlloSure

Sayah et al (2020) conducted a pilot study investigating the ability of AlloSure dd-cfDNA testing to detect acute cellular rejection.⁷⁶ Biopsy-matched biorepository samples from 69 lung transplant recipients who had previously enrolled in the multicenter Lung Allograft Gene Expression Observational (LARGO) Study were evaluated. Diagnostic cohorts included patients with respiratory allograft infection (n=26), normal histopathology without infection or rejection (n=30), and acute cellular rejection without concurrent infection (n=13). Samples were obtained between > 14 days and < 1-year post-transplant, and samples associated with potential concurrent infection with rejection were excluded. Median dd-cfDNA levels were 0.485% (IQR, 0.220 to 0.790) in the normal cohort, 1.52% (IQR, 0.520 to 2.550) in the acute cellular rejection cohort, and 0.595% (IQR, 0.270 to 1.170) in the infection cohort. While dd-cfDNA levels were significantly higher in the acute cellular rejection cohort compared to the normal cohort (p =.026), samples associated with infection were not significantly different from the normal (p =.282) or acute cellular rejection (p =.100) cohorts. The AUC for detection of acute cellular rejection was 0.717 (95% CI, 0.547 to 0.887; p =.025). At a threshold of 0.87% dd-cfDNA and an estimated prevalence rate of 25%, sensitivity for acute cellular rejection was 73.1% (95% CI, 52.2% to 88.4%), specificity was 52.9% (95% CI, 27.8% to 77.0%), positive likelihood ratio was 1.55, negative likelihood ratio was 0.51, PPV was 34.1%, and NPV was 85.5%. The study is limited by the small sample size and use of archived samples and raises concerns regarding the ability of AlloSure dd-cfDNA testing to detect antibody-mediated rejection and to discriminate between infection and rejection.

Khush et al (2021) evaluated 107 biorepository plasma samples from 38 lung transplant recipients enrolled in the Genome Transplant Dynamics Study via AlloSure dd-cfDNA testing.⁷⁷ The study cohort included 14 patients (22 samples) with acute cellular rejection confirmed by histopathology, 6 patients (7 samples) treated for acute cellular rejection without a confirmed histopathological diagnosis, 6 patients (8 samples) with obstructive CLAD, 7 patients (9 samples) with antibody-mediated rejection (AMR), 22 patients (33 samples) with infection without rejection, and 18 patients (28 samples) with stable allografts. The median dd-cfDNA levels in the acute cellular rejection (0.91%; IQR, 0.39% to 2.07%) and CLAD (2.06%; IQR, 0.97% to 3.34%) cohorts were significantly higher compared to the stable cohort (p =.02, respectively). However, the antibody-mediated rejection cohort was not statistically different when compared with the stable cohort (p =.07). The median dd-cfDNA level in an aggregated rejection cohort, composed of acute cellular rejection, AMR, and CLAD samples, was approximately 3-fold higher when compared to the stable cohort (1.06% versus 0.38%). At a dd-cfDNA threshold of 0.85%, the sensitivity for this spectrum of rejection was 55.6%, specificity was 75.8%, PPV was 43.3%, and NPV was 83.6%. The study is limited by the small sample size and use of archived samples. The authors suggest that AlloSure dd-cfDNA testing may have clinical utility as a plasma marker of "tissue injury" and that the 0.85% dd-cfDNA threshold requires further prospective clinical validation.

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.

No RCTs assessing the clinical utility of dd-cfDNA (i.e., AlloSure) testing to diagnose lung allograft rejection were 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.

Because the clinical validity of dd-cfDNA (i.e., AlloSure) testing to assess for lung allograft rejection has not been established, a chain of evidence to support clinical utility cannot be constructed.

Section Summary: Donor-Derived Cell-Free DNA Testing for Lung Transplant

Two small diagnostic accuracy studies of the AlloSure dd-cfDNA test utilizing biorepository samples were identified.. At a threshold of 0.87% dd-cfDNA, the PPV and NPV for detecting acute cellular rejection in the first study were 34.1% and 85.5%, respectively. A second study reported a PPV of 43.3% and NPV of 83.6% at a dd-cfDNA cutoff of 0.85% for an aggregate rejection cohort composed of patients with acute cellular rejection, antibody-mediated rejection, and CLAD. These studies have raised concerns regarding the ability of dd-cfDNA testing to discriminate between rejection and infection or injury, and larger prospective studies validating the dd-cfDNA threshold for active rejection are needed to develop conclusions. At present, no studies evaluating the clinical utility for AlloSure dd-cfDNA testing were identified.

Summary of Evidence

For individuals who have chronic heart failure who receive the sST2 assay to determine prognosis and/or to guide management, the evidence includes correlational studies and 2 meta-analyses. Relevant outcomes are OS, quality of life, and hospitalization. Most of the evidence is from reanalysis of existing RCTs and not from studies specifically designed to evaluate the predictive accuracy of sST2, and prospective and retrospective cross-sectional studies made up a large part of 1 meta-analysis. Studies have mainly found that elevated sST2 levels are statistically associated with an elevated risk of mortality. A pooled analysis of study results found that sST2 significantly predicted overall mortality and cardiovascular mortality. Several studies, however, found that sST2 test results did not provide additional prognostic information compared with N-terminal pro B-type natriuretic peptide levels. Moreover, no comparative studies were identified on the use of the sST2 assay to guide the management of patients diagnosed with chronic heart

failure. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have heart transplantation who receive sST2 assay to determine prognosis and/or to predict acute cellular rejection, the evidence includes a small number of retrospective studies on the Presage ST2 Assay. Relevant outcomes are OS, morbid events, and hospitalization. No prospective studies were identified that provide high-quality evidence on the ability of sST2 to predict transplant outcomes. One retrospective study (n = 241) found that sST2 levels were associated with acute cellular rejection and mortality; another study (n = 26) found that sST2 levels were higher during an acute rejection episode than before rejection. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have heart transplantation who receive myTAI_{HEART} assay to determine acute cellular rejection, the evidence includes observational studies. A validation study using 158 matched endomyocardial biopsy-plasma pairs from 76 pediatric and adult heart transplant recipients (ages 2 months or older, and 8 days more post-transplant) found a donor-specific fraction cutoff (0.32%) that produced a 100% NPV for Grade 2 or higher acute cellular rejection. A prospective observational blinded study (n=174; pediatric=101, adult=73) using biopsy-paired samples found that myTAI_{HEART} level was associated with acute cellular and antibody-mediated rejection in both adult and pediatric heart transplant populations, and that an optimal donor fraction threshold (0.3%) ruled out the presence of either acute cellular rejection or antibody-mediated rejection. Both studies received industry funding. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have a heart transplant who receive a measurement of volatile organic compounds to assess cardiac allograft rejection, the evidence includes a diagnostic accuracy study. Relevant outcomes are OS, test validity, morbid events, and hospitalizations. The published study found that, for identifying grade 3 (now grade 2R) rejection, the NPV of the breath test the study evaluated (97.2%) was similar to endomyocardial biopsy (96.7%) and the sensitivity of the breath test (78.6%) was better than that for biopsy (42.4%). However, the breath test had a lower specificity (62.4%) and a lower PPV (5.6%) in assessing grade 3 rejection than a biopsy (specificity, 97%; PPV, 45.2%). The breath test was also not evaluated for grade 4 rejection. This single study is not sufficient to determine the clinical validity of the test measuring volatile organic compounds and no studies on clinical utility were identified. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have a heart transplant who receive GEP to assess cardiac allograft rejection, the evidence includes 2 diagnostic accuracy studies and several RCTs evaluating clinical utility. Relevant outcomes are OS, test validity, morbid events, and hospitalizations. The 2 studies, Cardiac Allograft Rejection Gene Expression Observation (CARGO, CARGO II) examining the diagnostic performance of GEP for detecting moderate-to-severe rejection lacked a consistent threshold for defining a positive GEP test (i.e., 20, 30, or 34) and reported a low number of positive cases. In the available studies, although the NPVs were relatively high (i.e., at least

88%), the performance characteristics were only calculated based on 10 or fewer cases of rejection; therefore, performance data may be imprecise. Moreover, the PPV in CARGO II was only 4.0% for patients who were at least 2 to 6 months post transplant and 4.3% for patients more than 6 months post transplant. The threshold indicating a positive test that seems to be currently accepted (a score of 34) was not prespecified; rather it evolved partway through the data collection period in the Invasive Monitoring Attenuation through Gene Expression (IMAGE) study. In addition, the IMAGE study had several methodologic limitations (e.g., lack of blinding); further, the IMAGE study failed to provide evidence that GEP offers an incremental benefit over biopsy performed on the basis of clinical exam or echocardiography. Patients at the highest risk of transplant rejection are patients within 1 year of the transplant, and, for that subset, there remains insufficient data on which to evaluate the clinical utility of GEP. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have a heart transplant who receive GEP with testing of dd-cfDNA to assess cardiac allograft rejection, the evidence includes 1 retrospective analysis of the HeartCare test and 1 diagnostic accuracy study of the AlloSure dd-cfDNA component of the HeartCare test. Relevant outcomes are OS, test validity, morbid events, and hospitalizations. The HeartCare analysis reported a 12.7% reduction in endomyocardial biopsy volume among patients undergoing routine surveillance. However, this observation is limited by lack of reporting on long-term health outcomes and incomplete assessment of diagnostic performance for combined testing, as patients with negative dd-cfDNA scores did not undergo biopsy regardless of GEP score per study protocol. The diagnostic accuracy of the AlloSure dd-cfDNA test was assessed separately in the Utility of Donor-Derived Cell Free DNA in Association With Gene Expression Profiling (D-OAR) study, revealing high NPVs (>96.6%). However, at a dd-cfDNA cutoff of 0.2%, PPVs were low overall (8.9%), in surveillance patients (8.1%), and in patients with clinical suspicion of rejection (11.6%). The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with a renal transplant who are undergoing surveillance or have clinical suspicion of allograft rejection who receive testing of dd-cfDNA to assess renal allograft rejection, the evidence includes small diagnostic accuracy studies. Relevant outcomes are OS, test validity, morbid events, and hospitalizations. One study examined the diagnostic performance of dd-cfDNA for detecting moderate-to-severe rejection; the NPV was moderately high (84%), and performance characteristics were calculated on 27 cases of active transplant rejection. The threshold indicating a positive test was not prespecified. A subsequent smaller single-center study that explored variation in clinical validity based on different rejection mechanisms found the strongest performance characteristics for AlloSure with antibody-mediated rejection. A retrospective single-center study of the Prospera dd-cfDNA test reported a PPV and NPV of 52% and 95%, respectively, for detection of active rejection among a combined cohort of patients undergoing surveillance or for-cause biopsies, using the 1% dd-cfDNA threshold previously proposed for the AlloSure test. Larger prospective studies validating the dd-cfDNA thresholds for active rejection are needed to develop conclusions for each test. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals with a lung transplant who receive testing of dd-cfDNA to assess lung allograft rejection, the evidence includes 2 small diagnostic accuracy studies utilizing biorepository samples. Relevant outcomes are OS, test validity, morbid events, and hospitalizations. One study examined the diagnostic performance of AlloSure dd-cfDNA testing at a threshold of 0.87% for detecting acute cellular rejection, yielding a PPV of 34.1% and a NPV of 85.5%. A second study reported a PPV of 43.3% and NPV of 83.6% for an aggregate rejection cohort composed of patients with acute cellular rejection, antibody-mediated rejection, and CLAD. These studies have raised concerns regarding the ability of dd-cfDNA testing to discriminate between rejection and infection or injury, and larger prospective clinical validation studies are required. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

SUPPLEMENTAL INFORMATION

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

Clinical Input From Physician Specialty Societies and Academic Medical Centers

While the various physician specialty societies and academic medical centers may collaborate with and make recommendations during this process, through the provision of appropriate reviewers, input received does not represent an endorsement or position statement by the physician specialty societies or academic medical centers, unless otherwise noted.

2012 Input

In response to requests, input was received from 7 academic medical centers and 1 specialty society while this policy was under review in 2012. Input was mixed on whether AlloMap should be investigational. Four reviewers agreed with the investigational status, 1 disagreed, and 3 indicated it was a split decision/other. Reviewers generally agreed that the sensitivity and specificity have not yet been adequately defined for AlloMap and that the negative predictive value was not sufficiently high to preclude the need for biopsy. There was mixed input about the need for surveillance cardiac biopsies to be performed in the absence of clinical signs and/or symptoms of rejection.

2008 Input

In response to requests, input was received from 2 academic medical centers and 2 physician specialty societies while this policy was under review in 2008. Three reviewers agreed that these approaches for monitoring heart transplant rejection are considered investigational. The American College of Cardiology disagreed with the policy, stating that the College considers the available laboratory tests to have good potential to diagnose heart transplant rejection and reduce the frequency of invasive biopsies performed on heart transplant patients, although questions remained as to their role in clinical practice.

Practice Guidelines and Position Statements

Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given

to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

American College of Cardiology et al

In 2017, the American College of Cardiology Foundation, American Heart Association, and Heart Failure Society published a focused update of their 2013 guideline on the management of heart failure.^{1,78} Part of the focus of the update was on biomarkers. The guidelines stated that soluble suppression of tumorigenicity-2 (sST2) is a biomarker for myocardial fibrosis that may predict hospitalization and death in patients with heart failure and provides additive prognostic information to natriuretic peptide levels. The guidelines were based on a class IIb recommendation (weak; benefit \geq risk) with level B-NR evidence (moderate-quality, nonrandomized) for the use of ST2 as an option to provide additive prognostic information to established clinical evaluation and biomarkers. The guidelines did not address other uses of ST2 or myTAI_{HEART}.

International Society of Heart and Lung Transplantation

In 2010, the International Society of Heart and Lung Transplantation issued guidelines for the care of heart transplant recipients.⁷⁹ The guidelines included the following recommendations (see Table 9).

Table 9. Guidelines for Postoperative Care of Heart Transplant Recipients

| Recommendation | COR | LOE |
|---|------------|------------|
| "The standard of care for adult HT recipients is to perform periodic EMB during the first 6 to 12 postoperative months for surveillance of HT rejection." | IIa | C |
| "After the first post-operative year, EMB surveillance for an extended period of time (e.g., every 4-6 months) is recommended in HT patients at higher risk for late acute rejection...." | IIa | C |
| "Gene Expression Profiling (AlloMap) can be used to rule out the presence of ACR of grade 2R or greater in appropriate low-risk patients, between 6 months and 5 years after HT." | IIa | B |

ACR: acute heart rejection; COR: class of recommendation; EMB: endomyocardial biopsy; HT: heart transplant; LOE: level of evidence.

Kidney Disease Improving Global Outcomes

The Kidney Disease Improving Global Outcomes (2009) issued guidelines for the care of kidney transplant recipients.⁸⁰ The guidelines included the following recommendations (see Table 10).

Table 10. Guidelines for Biopsy in Renal Transplant Recipients

| Recommendation | SOR | LOE |
|---|------------|------------|
| "We recommend kidney allograft biopsy when there is a persistent, unexplained increase in serum creatinine." | Level 1 | C |
| "We suggest kidney allograft biopsy when serum creatinine has not returned to baseline after treatment of acute rejection." | Level 2 | D |

| Recommendation | SOR | LOE |
|---|------------|------------|
| "We suggest kidney allograft biopsy every 7-10 days during delayed function." | Level 2 | C |
| "We suggest kidney allograft biopsy if expected kidney function is not achieved within the first 1-2 months after transplantation." | Level 2 | D |
| "We suggest kidney allograft biopsy when there is new onset of proteinuria." | Level 2 | C |
| "We suggest kidney allograft biopsy when there is unexplained proteinuria ≥ 3.0 g/g creatinine or ≥ 3.0 g per 24 hours." | Level 2 | C |

LOE: level of evidence; SOR: strength of recommendation.

U.S. Preventive Services Task Force Recommendations

Not applicable.

Ongoing and Unpublished Clinical Trials

Some currently ongoing and unpublished trials that might influence this review are listed in Table 11.

Table 11. Summary of Key Active Trials

| NCT No. | Trial Name | Planned Enrollment | Completion Date |
|--------------------------|--|---------------------------|--------------------------------------|
| AlloMap | | | |
| <i>Ongoing</i> | | | |
| NCT01833195 ^a | Outcomes AlloMap Registry: the Long-term Management and Outcomes of Heart Transplant Recipients With AlloMap Testing (OAR) | 2444 | Feb 2020 (active, not recruiting) |
| NCT02178943 ^a | Utility of Donor-Derived Cell-free DNA in Association With Gene-Expression Profiling (AlloMap®) in Heart Transplant Recipients (D-OAR) | 100 | Feb 2020 (unknown) |
| HeartCare | | | |
| <i>Ongoing</i> | | | |
| NCT03695601 ^a | Surveillance HeartCare Outcomes Registry (SHORE) | 3450 | Jun 2027 (recruiting) |
| AlloSure (Kidney) | | | |
| <i>Ongoing</i> | | | |
| NCT04566055 ^a | Assessing AlloSure dd-cfDNA Monitoring Insights of Renal Allografts With Longitudinal Surveillance (ADMIRAL) | 1000 | Oct 2020 (active, not recruiting) |
| NCT03326076 ^a | Evaluation of Patient Outcomes From the Kidney Allograft Outcomes AlloSure Registry (KOAR) | 4000 | Dec 2025 (recruiting) |

| NCT No. | Trial Name | Planned Enrollment | Completion Date |
|--------------------------|--|---------------------------|----------------------------------|
| NCT04601155 ^a | Transition of Renal Patients Using AlloSure Into Community Kidney Care (TRACK) | 3500 | Sep 2026 (recruiting) |
| AlloSure (Lung) | | | |
| <i>Ongoing</i> | | | |
| NCT04318587 ^a | Assessment of Donor Derived Cell Free DNA and Utility in Lung Transplantation | 50 | Sep 2023 (recruiting) |
| NCT05050955 ^a | AlloSure Lung Assessment and Metagenomics Outcomes Study (ALAMO) | 1500 | Dec 2026 (not yet recruiting) |
| Prospera (Kidney) | | | |
| <i>Ongoing</i> | | | |
| NCT04239703 ^a | INTERCOMEX DD-cfDNA-HLA-MMDx Study: Comparing the DD-cfDNA Test to MMDx Microarray Test, Central HLA Antibody Testx, and Histology | 300 | Dec 2022 (recruiting) |
| NCT04091984 ^a | The PROspera Kidney Transplant ACTIVE Rejection Assessment Registry (ProActive) | 3000 | Oct 2027 (recruiting) |
| NCT03984747 ^a | Study for the Prediction of Active Rejection in Organs Using Donor-derived Cell-free DNA Detection (SPARO) | 500 | Oct 2028 (recruiting) |
| Prospera (Heart) | | | |
| <i>Ongoing</i> | | | |
| NCT04707872 ^a | Trifecta-Heart cfDNA-MMDx Study: Comparing the DD-cfDNA test to MMDx Microarray Test and Central HLA Antibody Test | 300 | Dec 2022 (recruiting) |
| NCT05081739 ^a | Donor-Derived Cell-free DNA to Detect Rejection in Cardiac Transplantation (DETECT) | 600 | Jan 2025 (recruiting) |

NCT: national clinical trial.

^a Denotes industry-sponsored or cosponsored trial.

CODING

The following codes for treatment and procedures applicable to this policy are included below for informational purposes. This may not be a comprehensive list of procedure codes applicable to this policy.

Inclusion or exclusion of a procedure, diagnosis or device code(s) does not constitute or imply member coverage or provider reimbursement. 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.

The code(s) listed below are medically necessary ONLY if the procedure is performed according to the "Policy" section of this document.

| CPT/HCPCS | |
|------------------|--|
| 81595 | Cardiology (heart transplant), mRNA, gene expression profiling by real-time quantitative PCR of 20 genes (11 content and 9 housekeeping), utilizing subfraction of peripheral blood, algorithm reported as a rejection risk score. |
| 83006 | Growth stimulation expressed gene 2 (ST , Interleukin 1 receptor like-1 |
| 0018M | Transplantation medicine (allograft rejection, renal), measurement of donor and third-party-induced CD154+T-cytotoxic memory cells, utilizing whole peripheral blood, algorithm reported as a rejection risk score |
| 0055U | Cardiology (heart transplant), cell-free DNA, PCR assay of 96 DNA target sequences (94 single nucleotide polymorphism targets and two control targets), plasma [Test name-myTAIHEART, Manufacturer-Tai Diagnostics Inc.] |
| 0087U | Cardiology (heart transplant), mRNA gene expression profiling by microarray of 1283 genes, transplant biopsy tissue, allograft rejection and injury algorithm reported as a probability score [Test name- Molecular Microscope® MMDx—Heart; Manufacturer— Kashi Clinical Laboratories] |
| 0088U | Transplantation medicine (kidney allograft rejection), microarray gene expression profiling of 1494 genes, utilizing transplant biopsy tissue, algorithm reported as a probability score for rejection [Test name- Molecular Microscope® MMDx—Kidney; Manufacturer— Kashi Clinical Laboratories] |
| 0118U | Transplantation medicine, quantification of donor-derived cell-free DNA using whole genome next generation sequencing, plasma, reported as percentage of donor-derived cell-free DNA in the total cell-free DNA |
| 0221U | Red cell antigen (ABO blood group) genotyping (ABO), gene analysis, next generation sequencing, ABO (ABO, alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase) gene |

ICD-10 DIAGNOSES

Experimental / Investigational for all diagnoses related to this medical policy.

| REVISIONS | |
|------------------|--|
| 03-12-2013 | Policy added to the bcbsks.com web site. |
| 10-16-2013 | Description section reviewed |
| | In Policy section: <ul style="list-style-type: none"> ▪ Added to item B, "including but not limited to" to read, "B. The evaluation of genetic expression in the peripheral blood including, but not limited to, the detection of acute heart transplant rejection or graft dysfunction is considered experimental / investigational." |
| | Rationale section updated |
| | References updated |
| 01-01-2016 | Published 12-14-2015. Effective 01-01-2016. |
| | Description section updated |
| | In Policy section: <ul style="list-style-type: none"> ▪ In Item B removed "The evaluation of genetic expression in the peripheral blood" and added "The use of peripheral blood genetic profiling tests in the management of patients after heart transplantation" and "heart transplant" to read "The use of peripheral blood genetic profiling tests in the management of patients after heart transplantation including, but not limited to, the detection of acute heart transplant rejection or heart transplant graft dysfunction is considered experimental / investigational." These wording updates clarify the policy statement, but do not change the original intent. |
| | Rationale section updated |
| | In Coding section: <ul style="list-style-type: none"> ▪ Added CPT Code: 81595 (Effective January 1, 2016) ▪ Removed CPT Code: 86849 (Effective January 1, 2016) · Updated Coding notations |
| | References updated |
| 07-01-2018 | In Coding section: <ul style="list-style-type: none"> ▪ Added PLA Code: 0051U |
| 09-12-2018 | Description section updated |
| | In Policy section: <ul style="list-style-type: none"> ▪ In Item A removed "with the Heartsbreath test" and added "2R (formerly grade 3) to read "The measurement of volatile organic compounds to assist in the detection of grade 2R (formerly grade 3) heart transplant rejection is considered experimental / investigational." ▪ In Item B replaced "genetic" with "gene expression" · Updated Policy Guidelines |
| | Rationale section updated |
| | In Coding section: <ul style="list-style-type: none"> ▪ Removed CPT Code: 93505 (the code does not pertain to the policy) ▪ Updated Coding nomenclature: 0051U · Updated Coding notations. |

| REVISIONS | |
|------------------|--|
| | References updated |
| 12-20-2018 | Policy published 11-20-2018. Policy effective 12-20-2018. |
| | Title changed "Laboratory Tests for Heart and Kidney Transplant Rejection" from "Laboratory Tests for Heart Transplant Rejection" |
| | Description section updated |
| | In Policy section: <ul style="list-style-type: none"> ▪ Added "C. The use of peripheral blood measurement of donor-derived cell-free DNA in the management of patients after renal transplantation, including but not limited to the detection of acute renal transplant rejection or renal transplant graft dysfunction, is considered experimental / investigational." |
| | Rationale section updated |
| | In Coding section: <ul style="list-style-type: none"> ▪ Added CPT Code: 81599 ▪ Coding notations removed |
| | References updated |
| 07-01-2019 | In Coding section: <ul style="list-style-type: none"> ▪ Added PLA Codes: 0087U, 0088U |
| 07-01-2019 | In Coding section: <ul style="list-style-type: none"> ▪ Added PLA Code: 0118U |
| 05-18-2020 | Description section updated |
| | Rationale section updated |
| | References updated |
| 02-04-2022 | Changed Title To: "Laboratory Tests Post Transplant and for Heart Failure" |
| | Updated Description Section |
| | In Policy Section Added: <ul style="list-style-type: none"> ▪ Section A: The use of the Presage ST2 Assay to evaluate the prognosis of patients diagnosed with chronic heart failure is considered experimental / investigational. ▪ Section B: The use of the Presage ST2 Assay to guide management (e.g., pharmacologic, device-based, exercise) of patients diagnosed with chronic heart failure is considered experimental / investigational. ▪ Section C: The use of the Presage ST2 Assay in the post cardiac transplantation period, including but not limited to predicting prognosis and predicting acute cellular rejection, is considered experimental / investigational. ▪ Section D: The use of the myTAIHEART assay in the post cardiac transplantation period, including but not limited to predicting prognosis and predicting acute cellular rejection, is considered experimental / investigational. ▪ Section E: Added word "moderate" ▪ Section F: Added "or peripheral blood measurement of donor-derived cell-free DNA (dd-cfDNA), alone or in combination" ▪ Section H: The use of peripheral blood measurement of dd-cfDNA in the management of patients after lung transplantation, including but not limited to the |

| REVISIONS | |
|------------------|--|
| | detection of acute lung transplant rejection or lung transplant graft dysfunction, is considered experimental / investigational. |
| | Updated Rationale Section |
| | Updated Coding Section: <ul style="list-style-type: none"> ▪ Added: 83006, 0018M, 0221U ▪ Removed 81599, 0085T |
| | Updated References Section |

REFERENCES

1. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* Oct 15 2013; 62(16): e147-239. PMID 23747642
2. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics--2011 update: a report from the American Heart Association. *Circulation.* Feb 01 2011; 123(4): e18-e209. PMID 21160056
3. Virani SS, Alonso A, Benjamin EJ, et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. *Circulation.* Mar 03 2020; 141(9): e139-e596. PMID 31992061
4. Rohde LE, Beck-da-Silva L, Goldraich L, et al. Reliability and prognostic value of traditional signs and symptoms in outpatients with congestive heart failure. *Can J Cardiol.* May 15 2004; 20(7): 697-702. PMID 15197422
5. Marcus GM, Gerber IL, McKeown BH, et al. Association between phonocardiographic third and fourth heart sounds and objective measures of left ventricular function. *JAMA.* May 11 2005; 293(18): 2238-44. PMID 15886379
6. Stevenson LW, Perloff JK. The limited reliability of physical signs for estimating hemodynamics in chronic heart failure. *JAMA.* Feb 10 1989; 261(6): 884-8. PMID 2913385
7. Gaggin HK, Januzzi JL. Biomarkers and diagnostics in heart failure. *Biochim Biophys Acta.* Dec 2013; 1832(12): 2442-50. PMID 23313577
8. McMurray JJ, Adamopoulos S, Anker SD, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail.* Aug 2012; 14(8): 803-69. PMID 22828712
9. Lindenfeld J, Albert NM, Boehmer JP, et al. HFSA 2010 Comprehensive Heart Failure Practice Guideline. *J Card Fail.* Jun 2010; 16(6): e1-194. PMID 20610207
10. Savarese G, Trimarco B, Dellegrottaglie S, et al. Natriuretic peptide-guided therapy in chronic heart failure: a meta-analysis of 2,686 patients in 12 randomized trials. *PLoS One.* 2013; 8(3): e58287. PMID 23472172
11. Troughton RW, Frampton CM, Brunner-La Rocca HP, et al. Effect of B-type natriuretic peptide-guided treatment of chronic heart failure on total mortality and hospitalization: an

- individual patient meta-analysis. *Eur Heart J*. Jun 14 2014; 35(23): 1559-67. PMID 24603309
12. Bhardwaj A, Januzzi JL. ST2: a novel biomarker for heart failure. *Expert Rev Mol Diagn*. May 2010; 10(4): 459-64. PMID 20465500
 13. Chowdhury P, Kehl D, Choudhary R, et al. The use of biomarkers in the patient with heart failure. *Curr Cardiol Rep*. Jun 2013; 15(6): 372. PMID 23644993
 14. Ciccone MM, Cortese F, Gesualdo M, et al. A novel cardiac bio-marker: ST2: a review. *Molecules*. Dec 11 2013; 18(12): 15314-28. PMID 24335613
 15. Daniels LB, Bayes-Genis A. Using ST2 in cardiovascular patients: a review. *Future Cardiol*. Jul 2014; 10(4): 525-39. PMID 25301315
 16. Dieplinger B, Mueller T. Soluble ST2 in heart failure. *Clin Chim Acta*. Mar 30 2015; 443: 57-70. PMID 25269091
 17. Mueller T, Dieplinger B. The Presage((R)) ST2 Assay: analytical considerations and clinical applications for a high-sensitivity assay for measurement of soluble ST2. *Expert Rev Mol Diagn*. Jan 2013; 13(1): 13-30. PMID 23256700
 18. Shah RV, Januzzi JL. ST2: a novel remodeling biomarker in acute and chronic heart failure. *Curr Heart Fail Rep*. Mar 2010; 7(1): 9-14. PMID 20425491
 19. Xu D, Chan WL, Leung BP, et al. Selective expression of a stable cell surface molecule on type 2 but not type 1 helper T cells. *J Exp Med*. Mar 02 1998; 187(5): 787-94. PMID 9480988
 20. Weinberg EO, Shimpo M, De Keulenaer GW, et al. Expression and regulation of ST2, an interleukin-1 receptor family member, in cardiomyocytes and myocardial infarction. *Circulation*. Dec 03 2002; 106(23): 2961-6. PMID 12460879
 21. Stewart S, Winters GL, Fishbein MC, et al. Revision of the 1990 working formulation for the standardization of nomenclature in the diagnosis of heart rejection. *J Heart Lung Transplant*. Nov 2005; 24(11): 1710-20. PMID 16297770
 22. CareDx. Overview: AlloMap Testing: Answering Unmet Needs in Heart Transplant Surveillance. n.d.; <http://www.allomap.com/providers/overview/>. Accessed October 19, 2021
 23. Celec P, Vlkova B, Laukova L, et al. Cell-free DNA: the role in pathophysiology and as a biomarker in kidney diseases. *Expert Rev Mol Med*. Jan 18 2018; 20: e1. PMID 29343314
 24. North PE, Ziegler E, Mahnke DK, et al. Cell-free DNA donor fraction analysis in pediatric and adult heart transplant patients by multiplexed allele-specific quantitative PCR: Validation of a rapid and highly sensitive clinical test for stratification of rejection probability. *PLoS One*. 2020; 15(1): e0227385. PMID 31929557
 25. Duong Van Huyen JP, Tible M, Gay A, et al. MicroRNAs as non-invasive biomarkers of heart transplant rejection. *Eur Heart J*. Dec 01 2014; 35(45): 3194-202. PMID 25176944
 26. Patel PC, Hill DA, Ayers CR, et al. High-sensitivity cardiac troponin I assay to screen for acute rejection in patients with heart transplant. *Circ Heart Fail*. May 2014; 7(3): 463-9. PMID 24733367
 27. Organ Procurement and Transplantation Network. National Data. 2018; <https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/#>. Accessed October 20, 2021.
 28. Goldberg RJ, Weng FL, Kandula P. Acute and Chronic Allograft Dysfunction in Kidney Transplant Recipients. *Med Clin North Am*. May 2016; 100(3): 487-503. PMID 27095641

29. Ahmad I. Biopsy of the transplanted kidney. *Semin Intervent Radiol*. Dec 2004; 21(4): 275-81. PMID 21331139
30. Solez K, Colvin RB, Racusen LC, et al. Banff 07 classification of renal allograft pathology: updates and future directions. *Am J Transplant*. Apr 2008; 8(4): 753-60. PMID 18294345
31. Haas M. The Revised (2013) Banff Classification for Antibody-Mediated Rejection of Renal Allografts: Update, Difficulties, and Future Considerations. *Am J Transplant*. May 2016; 16(5): 1352-7. PMID 26696524
32. Natera. Prospera Transplant Assessment: Indications for Use. 2020; <https://www.natera.com/wp-content/uploads/2020/11/Organ-Health-Clinician-Indications-for-use.pdf>. Accessed October 20, 2021.
33. Natera. Prospera Transplant Assessment: Clinician's Guide to Results. 2019; <https://www.natera.com/wp-content/uploads/2020/11/Organ-Health-Clinician-Clinicians-guide-to-results.pdf>. Accessed October 19, 2021.
34. Parulekar AD, Kao CC. Detection, classification, and management of rejection after lung transplantation. *J Thorac Dis*. Sep 2019; 11(Suppl 14): S1732-S1739. PMID 31632750
35. Yusen RD, Edwards LB, Kucheryavaya AY, et al. The Registry of the International Society for Heart and Lung Transplantation: Thirty-second Official Adult Lung and Heart-Lung Transplantation Report--2015; Focus Theme: Early Graft Failure. *J Heart Lung Transplant*. Oct 2015; 34(10): 1264-77. PMID 26454740
36. Levine DJ, Glanville AR, Aboyou C, et al. Antibody-mediated rejection of the lung: A consensus report of the International Society for Heart and Lung Transplantation. *J Heart Lung Transplant*. Apr 2016; 35(4): 397-406. PMID 27044531
37. CareDx. Allosure Lung. 2021; [caredx.com/products-and-services/transplant-services/lung/](https://www.caredx.com/products-and-services/transplant-services/lung/). Accessed October 18, 2021.
38. Cohn JN, Tognoni G. A randomized trial of the angiotensin-receptor blocker valsartan in chronic heart failure. *N Engl J Med*. Dec 06 2001; 345(23): 1667-75. PMID 11759645
39. O'Connor CM, Whellan DJ, Lee KL, et al. Efficacy and safety of exercise training in patients with chronic heart failure: HF-ACTION randomized controlled trial. *JAMA*. Apr 08 2009; 301(14): 1439-50. PMID 19351941
40. Kjekshus J, Apetrei E, Barrios V, et al. Rosuvastatin in older patients with systolic heart failure. *N Engl J Med*. Nov 29 2007; 357(22): 2248-61. PMID 17984166
41. Januzzi JL, Rehman SU, Mohammed AA, et al. Use of amino-terminal pro-B-type natriuretic peptide to guide outpatient therapy of patients with chronic left ventricular systolic dysfunction. *J Am Coll Cardiol*. Oct 25 2011; 58(18): 1881-9. PMID 22018299
42. Ky B, French B, McCloskey K, et al. High-sensitivity ST2 for prediction of adverse outcomes in chronic heart failure. *Circ Heart Fail*. Mar 2011; 4(2): 180-7. PMID 21178018
43. Bayes-Genis A, de Antonio M, Galan A, et al. Combined use of high-sensitivity ST2 and NTproBNP to improve the prediction of death in heart failure. *Eur J Heart Fail*. Jan 2012; 14(1): 32-8. PMID 22179033
44. Broch K, Ueland T, Nymo SH, et al. Soluble ST2 is associated with adverse outcome in patients with heart failure of ischemic etiology. *Eur J Heart Fail*. Mar 2012; 14(3): 268-77. PMID 22302661
45. Felker GM, Fiuzat M, Thompson V, et al. Soluble ST2 in ambulatory patients with heart failure: Association with functional capacity and long-term outcomes. *Circ Heart Fail*. Nov 2013; 6(6): 1172-9. PMID 24103327

46. Gaggin HK, Motiwala S, Bhardwaj A, et al. Soluble concentrations of the interleukin receptor family member ST2 and -blocker therapy in chronic heart failure. *Circ Heart Fail.* Nov 2013; 6(6): 1206-13. PMID 24114865
47. Anand IS, Rector TS, Kuskowski M, et al. Prognostic value of soluble ST2 in the Valsartan Heart Failure Trial. *Circ Heart Fail.* May 2014; 7(3): 418-26. PMID 24622243
48. Zhang R, Zhang Y, An T, et al. Prognostic value of sST2 and galectin-3 for death relative to renal function in patients hospitalized for heart failure. *Biomark Med.* 2015; 9(5): 433-41. PMID 25985174
49. Dupuy AM, Curinier C, Kuster N, et al. Multi-Marker Strategy in Heart Failure: Combination of ST2 and CRP Predicts Poor Outcome. *PLoS One.* 2016; 11(6): e0157159. PMID 27311068
50. Aimo A, Vergaro G, Passino C, et al. Prognostic Value of Soluble Suppression of Tumorigenicity-2 in Chronic Heart Failure: A Meta-Analysis. *JACC Heart Fail.* Apr 2017; 5(4): 280-286. PMID 27816512
51. Januzzi JL, Horne BD, Moore SA, et al. Interleukin receptor family member ST2 concentrations in patients following heart transplantation. *Biomarkers.* May 2013; 18(3): 250-6. PMID 23557127
52. Pascual-Figal DA, Garrido IP, Blanco R, et al. Soluble ST2 is a marker for acute cardiac allograft rejection. *Ann Thorac Surg.* Dec 2011; 92(6): 2118-24. PMID 22035779
53. Richmond ME, Zangwill SD, Kindel SJ, et al. Donor fraction cell-free DNA and rejection in adult and pediatric heart transplantation. *J Heart Lung Transplant.* May 2020; 39(5): 454-463. PMID 31983667
54. Phillips M, Boehmer JP, Cataneo RN, et al. Heart allograft rejection: detection with breath alkanes in low levels (the HARDBALL study). *J Heart Lung Transplant.* Jun 2004; 23(6): 701-8. PMID 15366430
55. CareDx. HeartCare Interpretive Guide. April 2021; <https://caredx.com/wp-content/uploads/2021/09/LK-10482-Heartcare-Interpretive-Guide.pdf>. Accessed October 21, 2021.
56. Blue Cross Blue Shield Technology Evaluation Center (TEC). Gene expression profiling as a noninvasive method to monitor for cardiac allograft rejection. *TEC Assessment Program.* 2011;26(8).
57. Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am J Transplant.* Jan 2006; 6(1): 150-60. PMID 16433769
58. Crespo-Leiro MG, Stypmann J, Schulz U, et al. Clinical usefulness of gene-expression profile to rule out acute rejection after heart transplantation: CARGO II. *Eur Heart J.* Sep 01 2016; 37(33): 2591-601. PMID 26746629
59. Kanwar MK, Khush KK, Pinney S, et al. Impact of cytomegalovirus infection on gene expression profile in heart transplant recipients. *J Heart Lung Transplant.* Feb 2021; 40(2): 101-107. PMID 33341360
60. Gondi KT, Kao A, Linard J, et al. Single-center utilization of donor-derived cell-free DNA testing in the management of heart transplant patients. *Clin Transplant.* May 2021; 35(5): e14258. PMID 33606316

61. Khush KK, Patel J, Pinney S, et al. Noninvasive detection of graft injury after heart transplant using donor-derived cell-free DNA: A prospective multicenter study. *Am J Transplant*. Oct 2019; 19(10): 2889-2899. PMID 30835940
62. Kobashigawa J, Patel J, Azarbal B, et al. Randomized pilot trial of gene expression profiling versus heart biopsy in the first year after heart transplant: early invasive monitoring attenuation through gene expression trial. *Circ Heart Fail*. May 2015; 8(3): 557-64. PMID 25697852
63. Pham MX, Deng MC, Kfoury AG, et al. Molecular testing for long-term rejection surveillance in heart transplant recipients: design of the Invasive Monitoring Attenuation Through Gene Expression (IMAGE) trial. *J Heart Lung Transplant*. Aug 2007; 26(8): 808-14. PMID 17692784
64. Pham MX, Teuteberg JJ, Kfoury AG, et al. Gene-expression profiling for rejection surveillance after cardiac transplantation. *N Engl J Med*. May 20 2010; 362(20): 1890-900. PMID 20413602
65. CareDx. Allosure Kidney. 2021; <https://go.caredx.com/allosure/>. Accessed October 20, 2021.
66. Pai A, Swan JT, Wojciechowski D, et al. Clinical Rationale for a Routine Testing Schedule Using Donor-Derived Cell-Free DNA After Kidney Transplantation. *Ann Transplant*. Jul 02 2021; 26: e932249. PMID 34210952
67. Bloom RD, Bromberg JS, Poggio ED, et al. Cell-Free DNA and Active Rejection in Kidney Allografts. *J Am Soc Nephrol*. Jul 2017; 28(7): 2221-2232. PMID 28280140
68. Huang E, Sethi S, Peng A, et al. Early clinical experience using donor-derived cell-free DNA to detect rejection in kidney transplant recipients. *Am J Transplant*. Jun 2019; 19(6): 1663-1670. PMID 30725531
69. Stites E, Kumar D, Olaitan O, et al. High levels of dd-cfDNA identify patients with TCMR 1A and borderline allograft rejection at elevated risk of graft injury. *Am J Transplant*. Sep 2020; 20(9): 2491-2498. PMID 32056331
70. Sawinski DL, Mehta S, Alhamad T, et al. Association between dd-cfDNA levels, de novo donor specific antibodies, and eGFR decline: An analysis of the DART cohort. *Clin Transplant*. Sep 2021; 35(9): e14402. PMID 34184326
71. Jordan SC, Bunnapradist S, Bromberg JS, et al. Donor-derived Cell-free DNA Identifies Antibody-mediated Rejection in Donor Specific Antibody Positive Kidney Transplant Recipients. *Transplant Direct*. Sep 2018; 4(9): e379. PMID 30234148
72. Mayer KA, Doberer K, Tillgren A, et al. Diagnostic value of donor-derived cell-free DNA to predict antibody-mediated rejection in donor-specific antibody-positive renal allograft recipients. *Transpl Int*. Sep 2021; 34(9): 1689-1702. PMID 34448270
73. Puliyaanda DP, Swinford R, Pizzo H, et al. Donor-derived cell-free DNA (dd-cfDNA) for detection of allograft rejection in pediatric kidney transplants. *Pediatr Transplant*. Mar 2021; 25(2): e13850. PMID 33217125
74. Sigdel TK, Archila FA, Constantin T, et al. Optimizing Detection of Kidney Transplant Injury by Assessment of Donor-Derived Cell-Free DNA via Massively Multiplex PCR. *J Clin Med*. Dec 23 2018; 8(1). PMID 30583588
75. Bunnapradist S, Homkraisas P, Ahmed E, et al. Using both the Fraction and Quantity of Donor-Derived Cell-Free DNA to Detect Kidney Allograft Rejection. *J Am Soc Nephrol*. Oct 2021; 32(10): 2439-2441. PMID 34162734

76. Sayah D, Weigt SS, Ramsey A, et al. Plasma Donor-derived Cell-free DNA Levels Are Increased During Acute Cellular Rejection After Lung Transplant: Pilot Data. *Transplant Direct*. Oct 2020; 6(10): e608. PMID 33062841
77. Khush KK, De Vlaminc I, Luikart H, et al. Donor-derived, cell-free DNA levels by next-generation targeted sequencing are elevated in allograft rejection after lung transplantation. *ERJ Open Res*. Jan 2021; 7(1). PMID 33532456
78. Yancy CW, Jessup M, Bozkurt B, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Card Fail*. Aug 2017; 23(8): 628-651. PMID 28461259
79. Costanzo MR, Dipchand A, Starling R, et al. The International Society of Heart and Lung Transplantation Guidelines for the care of heart transplant recipients. *J Heart Lung Transplant*. Aug 2010; 29(8): 914-56. PMID 20643330
80. Kasiske BL, Zeier MG, Chapman JR, et al. KDIGO clinical practice guideline for the care of kidney transplant recipients: a summary. *Kidney Int*. Feb 2010; 77(4): 299-311. PMID 19847156
81. Centers for Medicare & Medicaid Services (CMS). National Coverage Determination for HEARTSBREATH Test for Heart Transplant Rejection (260.10). 2008; [https://www.cms.gov/medicare-coverage-database/details/ncd- details.aspx?NCDId=325](https://www.cms.gov/medicare-coverage-database/details/ncd-details.aspx?NCDId=325). Accessed October 20, 2021.