Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

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

When Services Are Eligible for Coverage
Coverage for eligible medical treatments or procedures, drugs, devices or biological products may be provided only if:

- Benefits are available in the member’s contract/certificate, and
- Medical necessity criteria and guidelines are met.

Based on review of available data, the Company may consider needle electromyography (NEMG) for the treatment of each of the following conditions to be eligible for coverage:

- Radiculopathies
- Plexopathies
- Neuropathies
- Nerve compression syndromes
- Neuromuscular junction disorders
- Myopathies

Based on review of available data, the Company may consider nerve conduction studies (NCSs) for the treatment of each of the following conditions to be eligible for coverage:

- Motor neuron diseases
- Myopathies
- Radiculopathies
- Plexopathies
- Neuropathies
- Nerve compression syndromes
- Neuromuscular junction disorders
- Carpal tunnel syndrome (CTS)
- Spinal cord injury
- Traumatic nerve lesion

When Services Are Considered Not Medically Necessary
Needle electromyography, which uses invasive needle electrodes, must be performed by a physician specifically trained in electrodiagnostic medicine, such as a doctor of medicine, doctor of osteopathy specializing in neurology or physical and rehabilitation medicine, or other provider specialties that have documented specific training in the use of NEMG. The Company considers NEMG to be not medically necessary if not directly performed and interpreted by the physician or another specifically trained provider. Nerve conduction studies should be either (a) performed directly by the physician or (b) performed by a trained individual under the direct supervision of the physician. The Company considers NCSs to be not medically necessary if not performed either by the physician or a trained individual under his direct supervision. Direct supervision is defined by the American Association of Electrodiagnostic Medicine.
Neurodiagnostics
Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

(AAEM) to mean that the physician trained in electrodiagnostic (EDX) medicine is in close physical proximity to the EDX laboratory while testing is underway, is immediately available to provide the trained individual with assistance and direction, and is responsible for selecting the appropriate studies to be performed.

Generally, the interpreting physician for both NEMG and NCS procedures should be a neurologist or physiatrist or physician with comparable supervised training within a residency or fellowship training program.

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

Based on review of available data, the Company considers surface electromyography (SEMG) to be investigational.*

Based on review of available data, the Company considers paraspinal SEMG to be investigational.*

Based on review of available data, the Company considers macro electromyography (EMG) to be investigational.*

Based on review of available data, the Company considers quantitative sensory testing (QST) to be investigational.*

Based on review of available data, the Company considers current perception threshold (CPT) to be investigational.*

Based on review of available data, the Company considers pressure specified sensory testing to be investigational.*

Based on review of available data, the Company considers automated nerve conduction tests or those with computer generated interpreted reports to be investigational.*

Based on review of available data, the Company considers needle electromyography (NEMG) or nerve conduction studies (NCSs) for the treatment of any condition other than those listed above to be investigational.*

Background/Overview
Automated Point-of-Care Nerve Conduction Tests
Portable devices have been developed to provide point-of-care nerve conductions studies. These devices have computational algorithms that are able to drive stimulus delivery, measure and analyze the response, and provide a report of study results. Automated nerve conduction could be used in various settings, including primary care, without the need for specialized training or equipment.
Nerve conductions studies and NEMG, when properly performed by a trained practitioner, are considered the gold standard of electrodiagnostic testing. However, the need for specialized equipment and personnel may limit the availability of electrodiagnostic testing for some patients. One proposed use of automated nerve conduction devices is to assist in the diagnosis of CTS. Carpal tunnel syndrome is a pressure-induced entrapment neuropathy of the median nerve as it passes through the carpal tunnel, resulting in sensorimotor disturbances. This syndrome is defined by its characteristic clinical symptoms, which may include pain, subjective feelings of swelling, and nocturnal paresthesia. A variety of simple diagnostic tools are available, and a positive response to conservative management (steroid injection, splints, and modification of activity) can confirm the clinical diagnosis. Electrodiagnostic studies may also be used to confirm the presence or absence of a median neuropathy at the wrist, assess the severity of the neuropathy, and assess alternate associated diagnoses. Nerve conduction is typically assessed prior to surgical release of the carpal tunnel, but the use of EMG in the diagnosis of CTS is controversial.

Point-of-care nerve conduction testing has also been proposed for the diagnosis of peripheral neuropathy and, in particular, for detecting neuropathy in patients with diabetes. Peripheral neuropathy is relatively common in patients with diabetes mellitus, and the diagnosis is often made clinically through the physical examination. Diabetic peripheral neuropathy can lead to important morbidity including pain, foot deformity, and foot ulceration. Clinical practice guidelines recommend using simple sensory tools such as the 10-g Semmes-Weinstein monofilament or the 128-Hz vibration tuning fork for diagnosis. These simple tests predict the presence of neuropathy defined by electrophysiological criteria with a high level of accuracy. Electrophysiological testing may be used in research studies and may be required in cases with an atypical presentation.

NC-stat® by NeuroMetrix is a portable nerve conduction test device designed to be used at the point-of-care. The system comprises a biosensor array, an electronic monitor, and a remote report generation system. The biosensor is a single use, preconfigured array consisting of a stimulation anode and cathode, skin surface digital thermometer, and response sensor. Biosensor arrays are available for assessment of sensory and motor nerves of the wrist (median and ulnar), and for the foot (peroneal, posterior tibial, and sural). A chip embedded in the biosensor panel measures skin surface temperature, the analysis algorithm adjusts for differences in temperature from 30°C, or if skin surface temperature is less than 23°C, the monitor will indicate that limb warming is necessary. Data are sent to a remote computer via a modem in the docking station, and the remote computer generates a report based on the average of 6 responses that is sent back by fax or email. In addition to the automated stimulus delivery and reporting, Nc-stat analysis adjusts the calculation for body temperature, height, and weight and uses the average of 6 responses. Sensitivity of the device for sensory nerve amplitude potentials is 2.1 µV; values lower than this are analyzed as zero, and responses with artifact are automatically eliminated from the analysis.

The Axon-II™ (PainDx) is an automated system that is being marketed for the detection of various sensory neurologic impairments caused by various pathologic conditions or toxic substance exposures, including signs of sympathetic dysfunction and detection of down-regulated A-delta function to locate injured nerve(s). The Axon-II software works with the Neural-Scan™ system (Neuro Diagnostics) and lists 7 automated studies (Cervical, Thoracic, Lumbar, Upper Extremities, Lower Extremities, Neuroma, Trigeminal), as well as a custom study. The Neural-Scan is a voltage-actuated sensory nerve conduction test device, which
measures the voltage amplitude necessary to cause a discernible nerve impulse. Results are adjusted and compared to population means; the most severe hypoesthesia is considered the primary lesion.

**Quantitative Sensory Testing**
Quantitative sensory testing systems are used for the noninvasive assessment and quantification of sensory nerve function in patients with symptoms of or the potential for neurologic damage or disease. Pain conditions evaluated may include diabetic neuropathy and uremic and toxic neuropathies, complex regional pain syndrome, CTS, and other nerve entrapment/compression disorders or damage.

Quantitative sensory testing has been investigated for a broad range of clinical applications, including evaluation of peripheral neuropathies, detection of CTS, spinal radiculopathy, evaluation of the effectiveness of peripheral nerve blocks, quantification of hypoesthetic and hyperesthetic conditions, and differentiation of psychogenic from neurologic disorders.

Quantitative sensory testing systems measure and quantify the amount of physical stimuli required for sensory perception to occur in the patient. As sensory deficits increase, the perception threshold of QST will increase, which may be informative in documenting progression of neurologic damage or disease. QST has not been established for use as a sole tool for diagnosis and management but has been used in conjunction with standard evaluation and management procedures (e.g., physical and neurologic examination, monofilament testing, pinprick, grip and pinch strength, Tinel sign and Phalen and Roos test) to enhance the diagnosis and treatment-planning process and confirm physical findings with quantifiable data. Stimuli used in QST includes touch, pressure, pain, thermal (warm and cold), or vibratory stimuli.

The criterion standard for evaluation of myelinated large fibers is the electromyographic nerve conduction study (EMG-NCS). However, the function of smaller myelinated and unmyelinated sensory nerves, which may show pathologic changes before the involvement of the motor nerves, cannot be detected by nerve conduction studies. Small fiber neuropathy has traditionally been a diagnosis of exclusion in patients who have symptoms of distal neuropathy and a negative NCS.

Depending on the type of stimuli used, QST can assess both small and large fiber dysfunction. Touch and vibration measure the function of large myelinated A-alpha and A-beta sensory fibers. Thermal stimulation devices are used to evaluate pathology of small myelinated and unmyelinated nerve fibers; they can be used to assess heat and cold sensation, as well as thermal pain thresholds. Pressure-specified sensory devices (PSSD) assess large myelinated sensory nerve function by quantifying the thresholds of pressure detected with light, static, and moving touch. Finally, current perception threshold testing involves the quantification of the sensory threshold to transcutaneous electrical stimulation. In current perception threshold testing, typically 3 different frequencies are tested: 5 Hz, designed to assess C fibers; 250 Hz, designed to assess A-delta fibers; and 2,000 Hz, designed to assess A-beta fibers. Results are compared with those of a reference population.

Because QST combines the objective physical sensory stimuli with the subject patient response, it is psychophysical in nature and requires patients who are alert, able to follow directions, and cooperative. In addition, to get reliable results, examinations need to be standardized with standardized instructions to the
patients, and stimuli must be applied in a consistent manner by trained staff. Psychophysical tests have greater inherent variability, making their results more difficult to standardize and reproduce.

**Paraspinal Surface Electromyography (SEMG) to Evaluate and Monitor Back Pain**

Surface electromyography, a noninvasive procedure that records the summation of muscle electrical activity, has been investigated as a technique to evaluate the physiologic functioning of the back. In addition, this procedure has been studied as a technique to evaluate abnormal patterns of electrical activity in the paraspinal muscles in patients with back pain symptoms, such as spasm, tenderness, limited range of motion (ROM), or postural disorders.

Back pain is an extremely common condition, affecting most individuals at some point in their lives. Identifying the pathogenesis of back pain is a challenging task, in part due to the complex anatomy of the back, which includes vertebrae, intervertebral discs, facet joints, spinal nerve roots, and numerous muscles. For example, back pain may be related to osteoarthritis, disc disease, subluxation, or muscular pathology, such as muscle strain or spasm. Moreover, due to referred pain patterns, the location of the pain may not be anatomically related to the pathogenesis of the pain. For example, buttock or leg pain may be related to pathology in the spine. In addition to the diagnostic challenges of back pain is the natural history of acute back pain. The majority of cases of acute low back pain will resolve with conservative therapy, such as physical therapy, and continuing normal activities within limits permitted by the pain. Thus, initial imaging or other diagnostic testing is generally not recommended unless “red flag” warning signs are present or the pain persists for longer than 4-6 weeks. Red flag findings include significant trauma, history of cancer, unrelenting night pain, fevers or chills, and progressive motor or sensory deficits.

Aside from physical examination, diagnostic tests include imaging technologies, such as magnetic resonance imaging (MRI), designed to identify pathology (e.g., bulging discs) or tests such as discography to localize the abnormality by reproducing the pain syndrome. However, due to their lack of specificity, all diagnostic tests must be carefully interpreted in the context of the clinical picture. For example, 5% of asymptomatic patients will have bulging discs as identified by MRI. Therefore, the presence of a bulging disc may only be clinically significant if well correlated with symptoms. Assessment of the musculature may focus on ROM or strength exercises.

In contrast to anatomic imaging, SEMG, which records the summation of muscle activity from groups of muscles, has been investigated as a technique to evaluate the physiologic functioning of the back. A noninvasive procedure, SEMG is contrasted with needle electromyography, an invasive procedure in which the electrical activity of individual muscles is recorded. Paraspinal SEMG, also referred to as paraspinal EMG scanning, has been explored as a technique to evaluate abnormal patterns of electrical activity in the paraspinal muscles in patients with back pain symptoms such as spasm, tenderness, limited ROM, or postural disorders. The technique is performed using 1 or an array of electrodes placed on the skin surface, with recordings made at rest, in various positions, or after a series of exercises. Recordings can also be made by using a handheld device, which is applied to the skin at different sites. Electrical activity can be assessed by computer analysis of the frequency spectrum (i.e., spectral analysis), amplitude, or root mean square of the electrical action potentials. In particular, spectral analysis that focuses on the median frequency has been used to assess paraspinal muscle fatigue during isometric endurance exercises.
Paraspinal SEMG has been researched as a technique to establish the etiology of back pain and also has been used to monitor the response to therapy and establish physical activity limits, such as assessing capacity to lift heavy objects or ability to return to work.

Paraspinal SEMG is an office-based procedure that may be most commonly used by physiatrists or chiropractors. The following clinical applications of the paraspinal SEMG have been proposed:

- Clarification of a diagnosis (i.e., muscle, joint, or disc disease)
- Selection of a course of medical therapy
- Selection of a type of physical therapy
- Preoperative evaluation
- Postoperative rehabilitation
- Follow-up of acute low back pain
- Evaluation of exacerbation of chronic low back pain
- Evaluation of pain management treatment techniques

**FDA or Other Governmental Regulatory Approval**

**U.S. Food and Drug Administration (FDA)**

**Nerve Conduction Tests**

Several devices are now being marketed for point-of-care neural conduction testing. NeuroMetrix received specific clearance to market NC-stat via the U.S. FDA 510(k) process in 1998, listing as predicate devices the TECA model-10 electromyograph and the Neurometer by Neurotron, which measures vibration threshold. The FDA-listed intended use was “to measure neuromuscular signals that are useful in diagnosing and evaluating systemic and entrapment neuropathies.” In addition, the approved application stated that “The NC-stat is intended to be used as an adjunct to and not a replacement for conventional electrodiagnostic measurements.” The 2004 510(k) added a sural biosensor for use in diagnosing neuropathies affecting the sural nerve (K041320). NeuroMetrix also received FDA clearance to market models with biosensors and engineering changes that enable the NC-stat to be used for motor and sensory nerves of the wrist (median, ulnar) and foot (peroneal, tibial, sural). The intended use as listed on the 510(k) approval from 2006 (K060584) is “to stimulate and measure neuromuscular signals that are useful in diagnosing and evaluating systemic and entrapment neuropathies.” The original NC-stat is no longer being produced. The current NeuroMetrix device, NC-stat DPN-Check™, assesses the sural nerve. The NeuroMetrix ADVANCE™ system received marketing clearance in 2008 (K070109). It is intended to perform NCSs using disposable surface electrodes (similar to NC-stat) with an additional module for invasive needle EMG. NeuroMetrix lists NC-stat as a predicate device for the ADVANCE system.

The Brevio® from Neurotron Medical received marketing clearance from the FDA in 2001. The Brevio is intended “for use for the measurement of nerve response latency and amplitude in the diagnosis and monitoring of peripheral neuropathies.” The XLTek Neuropath (Excel-Tech) received clearance for marketing through the FDA's 510(k) process in 2006; the indications are the same as those for NC-stat.
Quantitative Sensory Testing

A number of QST devices have been cleared for marketing by the FDA through the 510(k) process. They include devices for current perception threshold testing (e.g., Medi-Dx 7000®‡ Current Perception Threshold; Neuro Diagnostic Associates), pressure-specified sensory testing (e.g., NK Pressure-Specified Sensory Device ™‡; NK Biotechnical), vibration testing (e.g., Vibration Perception Threshold meter; Xilas Medical), and thermal testing (e.g., Thermal Sensory Analyzer, TSA and TSA-II; Medoc Corp., Israel).

SEMG

SEMG devices approved by the U.S. FDA include those that use a single electrode or a fixed array of multiple surface electrodes.

Several FDA-approved devices combine surface EMG along the spine with other types of monitors. For example, in 2007, the Insight Discovery (Fasstech; Burlington, MA) was cleared for marketing through the 510(k) process. The device contains 6 sensor types, 1 of which is surface EMG. The indications include measuring bilateral differences in surface EMG along the spine and measuring surface EMG along the spine during functional tasks. (Earlier Insight models had fewer sensor types.)

Centers for Medicare and Medicaid Services (CMS)

In February 2002, Medicare announced a national noncoverage policy on sensory nerve conduction threshold testing. Requests for reconsideration were submitted by both the makers of the Neurotron®‡ and the Medi-Dx 7000 devices. On reconsideration, Medicare affirmed its noncoverage policy, concluding that any use of sensory nerve conduction threshold testing to diagnose sensory neuropathies or radiculopathies is not reasonable and necessary. This decision was reaffirmed effective April 2004. Medicare has not addressed coverage for other types of quantitative sensory testing.

Rationale/Source

Automated Point-of-Care Nerve Conduction Tests

Assessment of a diagnostic technology typically focuses on 3 domains: 1) technical performance; 2) diagnostic accuracy (sensitivity, specificity, and positive and negative predictive value) in appropriate populations of patients; and 3) demonstration that the diagnostic information can be used to improve patient outcomes. This evaluation will focus on the technical performance of NC-stat, the first automated nerve conduction test device to be marketed, and its reported performance in diagnosing patients (validity) with suspected deficits of neuronal transmission (e.g., diabetic neuropathy, CTS, and lumbosacral radiculopathy).

Technical performance of a device is typically assessed with 2 types of studies, those that compare test measurements with a gold standard and those that compare results taken with the same device on different occasions (test-retest). The criterion standard for nerve conduction testing is the electrophysiologic NCS combined with needle EMG. Several studies have assessed the reliability and validity of NC-stat when used by personnel trained in electrophysiology. These studies, the majority of which are company-sponsored, are described below.
Technical Performance

Comparison to the Reference Standard

A 2006 study compared results for sensory nerve testing from NC-stat and the reference standard in median and ulnar nerves in 60 patients referred to an EMG laboratory for neck and shoulder pain who also volunteered to undergo testing with NC-stat. The reported correlations (Pearson correlation) between the NC-stat and the reference standard were high (0.91 for median nerve distal sensory latency [DSL], 0.70 for ulnar DSL, and 0.88 for the median ulnar difference of the DSL). However, this final correlation was calculated only with the responses obtained for 81 of 120 possible nerve pairs. The authors of this study report systematic differences between the two techniques and indicate that use of the NC-stat would require applicable reference ranges.

A study of motor nerve function compared NC-stat with standard nerve conduction tests of the wrist in a small study of 17 subjects with diabetes mellitus who had clinical evidence of peripheral neuropathy in either the upper or lower extremity. Again, Pearson correlation coefficients were relatively high and ranged from 0.70 for ulnar distal motor latency (DML) to 0.96 for median nerve DML.

Another NeuroMetrix-sponsored trial compared NC-stat and standard EMG results for peroneal and posterior tibial nerve conduction in 60 patients referred to an EMG laboratory. The report indicates that all patients referred to the laboratory were offered the opportunity to participate but does not provide the total number of referrals. F-wave latency (FLAT) was found to have the highest correlation (0.91, 0.90 Spearman correlation coefficient for peroneal and posterior tibial nerves, respectively), with moderate correlations for amplitude (0.86, 0.73) and distal motor latency (0.70, 0.45). The authors concluded that there was excellent criterion validity for the peroneal and posterior tibial FLAT and the peroneal amplitude; acceptable criterion validity for the peroneal DML and posterior tibial amplitude; but the validity of the posterior tibial DML could not be demonstrated. Although NC-stat results were significantly correlated with standard EMG tests in the study population as a whole, in a subgroup analysis of the most abnormal half of responses, the correlation coefficient for amplitude of the peroneal response was 0.82, and the correlation coefficient for distal motor latency was reduced to 0.32 for the posterior tibial nerve and 0.10 for the peroneal nerve. Thus, in this pathological subgroup analysis, criterion validity was lost for the peroneal distal motor latency and decreased from “excellent” to “acceptable” for the other parameters. The authors note that “this study did not address interpretations performed by physicians using NC-stat data, nor the validity of the reference ranges used or the way these were collected.”

In 2004, Rotman et al reported a Pearson correlation coefficient of 0.944 was reported for DML for 46 patients with CTS who had a nerve conduction study at a different time (average of 28 days’ difference). Another study compared results from NC-stat and standard NCS in a previously diagnosed patient population. This study compared DML of the median nerve in 72 patients (of 400 treated) with established CTS before and after surgical intervention, finding a correlation coefficient of 0.88 for the median nerve DML. However, a scatter plot indicates a poor correlation for longer latencies.

Test-Retest

NeuroMetrix reported intra-operator reliability in 15 healthy subjects who underwent measurements 7 days apart. The report states that “each upper and lower extremity nerve was tested twice by the same
Technician” and that 9 subjects participated in both upper and lower extremity studies. It is not clear from the report whether the upper and lower extremities were designed as separate studies, or if 12 of 42 (29%) measurements did not provide usable data. Of the data reported, the coefficient of variation ranged from 0.013 for F-wave latency to 0.298 for the compound muscle action potential amplitude of the peroneal nerve. A 2010 publication by NeuroMetrix reported test-retest reproducibility with the ADVANCE system in 30 subjects with symptoms suggestive of neuropathies; 29 subjects completed the study. Coefficients of variation ranged from 4.2% to 9.8% for tests measured 3-7 days apart. Between session intraclass correlation coefficients (ICCs) ranged from 0.98 for F-wave latency to 0.77 for sural sensory conduction velocity.

Diagnostic Accuracy
Diagnostic performance is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with a criterion standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true negative). Evaluation of diagnostic performance, therefore, requires independent assessment by the 2 methods in a population of patients who are suspected of disease but who do not all have the disease. An additional issue with NC-stat is that this device is designed to be used by minimally trained personnel (≈1 day for device-specific training), while the comparison standard is performed by specialists with extensive training in EMG and electrophysiology. Studies that do not meet these criteria (broad patient population and comparison of point-of-care [POC] use with the standard laboratory EMG) may be considered relevant to the technical performance of the device but are inadequate for evaluation of its diagnostic performance.

Carpal Tunnel Syndrome
In an early report of the NC-stat technology using DML to diagnose CTS, Leffler and colleagues reported that in 248 symptomatic hands (apparently a combination of an initial and validation group), compared with conventional diagnosis, testing using this device had a sensitivity of 86% and specificity of 90%. In the report by Rotman et al., the NC-stat DML was shown to have a sensitivity of 89% “at the predetermined specificity of 95%” for the diagnosis of CTS for “70 hands” that met the standardized CTS case definition. However, in a point-of-care study evaluating industrial workers for possible CTS using DML, many individuals who were identified with prolonged DML by NC-stat fell within the normal range (using 95% cutoff point) as defined by this study population. This study also comments on the importance of sensory nerve findings in the diagnosis of CTS, suggesting a need to better define “normal” values.

Diabetic Peripheral Neuropathy
Another study assessed the validity of NC-stat to diagnose diabetic peripheral neuropathy through sural nerve testing in patients from diabetes and diabetic neuropathy outpatient practices. Seventy-two consecutive patients (64 with type 2 diabetes) who completed a clinical evaluation, a conventional nerve conduction study, and a point-of-care NC-stat assessment were enrolled. The point-of-care assessment was independently conducted by non-technologist research staff following a single 1-hour lesson in the NC-stat protocol. The amplitude potential of the sural nerve was tested as an early indicator of diabetic neuropathy. Using a threshold of 6 µV, the authors report that the sensitivity and specificity of NC-stat for diagnosis of diabetic sensorimotor polyneuropathy, as defined by clinical and conventional
electrophysiological evaluation, was 92% and 82%, respectively. The Spearman correlation coefficient (compared with the reference standard) was 0.95. Further study is needed in a broad spectrum of patients, including those who present with atypical neuropathy in a clinical setting.

In 2015, Sharma et al assessed the technical accuracy of NC-stat DPN-Check in 162 patients with diabetes and 80 healthy controls. Based on the 10-point Neuropathy Disability Score, diabetic peripheral neuropathy was categorized as none, mild, moderate, or severe. Measurements with the POC device were conducted by blinded assessors. Receiver operating characteristic curves showed high overall accuracy in participants with either no neuropathy or severe neuropathy. However, for patients with mild neuropathy who would benefit most from early diagnosis, accuracy was substantially lower.

Further investigation is needed into specific approaches that include the POC NCS as a component of the clinical care of those with polyneuropathy.

**Lumbosacral Radiculopathy**

Fisher and colleagues explored the relationship between NC-stat and routine NCS/NEMG in 34 consecutive patients with a clinical history and/or examination consistent with lumbosacral radiculopathy. Inclusion in the study was based on chart review of symptoms from clinical history and/or examination (including low back pain or buttock pain, numbness, and/or paresthesias of one or both lower extremities) and having undergone testing with both NC-stat and routine electrodiagnostic studies. All testing was conducted by the principal investigator, and the reason for and timing of NC-stat testing was not specified. Of the 34 patients included in the study, 28 had MRI of the lumbosacral spine within 6 months of electrodiagnosis, 2 had a post-myelogram computed tomography (CT) scan, and 3 had lumbosacral spine radiographs. A neuroradiologist who was blinded to the clinical evaluation and electrodiagnostic results determined from MRI or CT that lumbosacral root injury was likely at the L4-5 and/or L5-S1 levels in 18 patients (60%). The study found some correlation between the electrodiagnostic testing and NC-stat. However, 6 of 10 patients who had unremarkable routine electrodiagnostic results had abnormal F-wave and compound muscle action potential (CMAP) amplitude abnormalities with NC-stat testing. The clinical implications of this finding are uncertain.

A 2011 report by Schmidt and colleagues assessed the accuracy of NC-stat diagnosis of lumbosacral radiculopathy in 50 patients and 25 controls with no prior history of lumbosacral radiculopathy. The patient cohort included patients referred to a tertiary referral EMG laboratory for testing of predominantly unilateral leg symptoms (pain, numbness, or weakness). Control subjects were recruited from clinic employees and from patients referred to the EMG laboratory for upper limb symptoms. All patients underwent focused history and physical examination and both standard and automated electrodiagnostic testing. Automated testing was performed by experienced technicians who were unaware of the electrodiagnostic test results. The data were transmitted to the manufacturer and compared with a large database of previously recorded data, which were adjusted for the age and height of the patient and subsequently determined to be normal or abnormal. In the patient cohort, the sensitivity of NC-stat was found to be 0% for L4 radiculopathy, 69% for L5 radiculopathy, and 64% for S1 radiculopathy compared with standard electrodiagnostic testing. By standard electrodiagnostic evaluation, 22 of the 50 symptomatic patients had findings consistent with L4, L5, or S1 radiculopathy, and 28 patients were found to be normal or to have a diagnosis other than
lumbosacral radiculopathy; NC-stat identified only 4 of these 28 cases (specificity of 14%). Standard electrodiagnostic testing also identified other important diagnoses in 9 patients (18%) that were not identified by the automated test, while NC-stat reported 6 other diagnoses in patients found to be normal by standard electrodiagnostic testing. All standard electrodiagnostic tests in the control group were normal, but the automated test found that 18 of these subjects were abnormal (specificity of 32%). The study found that the raw nerve conduction data were comparable for the 2 techniques; however, computer-generated interpretations by the automated device showed low specificity (numerous false-positives) in both symptomatic patients and normal control subjects. An accompanying editorial by England and Franklin states that the use of automated nerve conduction devices is controversial and that the use of NC-stat for lumbosacral radiculopathy would likely lead to a high misdiagnosis rate and potentially inappropriate treatment, including surgery. England and Franklin also conclude that an overly sensitive but not very specific test for CTS, or other mono- or polyneuropathies, cannot replace expert use and interpretation of conventional electrodiagnostic testing.

Mixed Population
A 2008 report assessed the diagnostic performance of NC-stat against the gold standard NCS in patients who had been referred for electrodiagnostic testing at one of several academic medical centers. Of 47 patients who were invited to participate in the study, 12 declined to participate, and records from 1 patient were missing, resulting in data analysis on 33 patients. The goal of the study was to compare the measurements of the two methods of nerve conduction testing as they would be used in standard practice, thus, patients were not excluded on the basis of the particular diagnosis for which they were referred. The diagnosis being tested was CTS in 25 patients (76%), with the remaining 8 patients having 8 other potential diagnoses, including ulnar neuropathy, upper extremity paresthesias, and C6 radiculopathy. NC-stat testing was independently performed by assistants (medical student, physical therapy assistant, or occupational therapy assistant) who were trained to operate the device following the manufacturer’s recommendations. NC-stat results could not be obtained for 2 patients for median motor studies and 3 patients for median sensory studies (15%). Based on the manufacturer’s suggested cutoff for abnormal nerve conduction, sensitivity was 100% for both the motor and sensory median-ulnar difference; specificity was 62–69% for the motor median-ulnar difference and 41% to 47% for the sensory median-ulnar difference. Pearson correlation coefficients ranged from 0.40 for the ulnar nerve to 0.91 for the median dorsal motor nerve. The ICCs had generally lower values than the Pearson coefficients, reflecting systematic bias due to methodologic differences in the two methods of NCS. The authors concluded that the recommended cutoff values for NC-stat may need to be adjusted, although the specific study results were limited by the small sample size. In addition, the authors noted that the study did not evaluate how well physicians can assign clinical relevance to the results and that while the device may be suited for research studies or screening of symptomatic patients, “in many clinical situations referral to a specialist for a more comprehensive evaluation would be prudent.”

Normative Values
In 2009, NeuroMetrix published a study of reference ranges for key nerve conduction parameters in healthy subjects. Data analyzed in the paper were pooled from 5 studies, including from 92 to 848 healthy subjects with data on the median, ulnar, peroneal, tibial, and sural nerves. Subject age and height were found to affect the parameters. In addition to providing reference ranges for clinicians to use (providing that NCS

©2017 Blue Cross and Blue Shield of Louisiana
An independent licensee of the Blue Cross and Blue Shield Association
No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without permission from Blue Cross and Blue Shield of Louisiana.
Page 11 of 25
techniques are consistent with those described in the paper), the authors stated that clinicians could use the same method to develop their own reference ranges. At this time, the proposed reference ranges have not been validated in a clinical patient population.

Clinical Outcomes

In 2011, Bourke et al. reported a nonrandomized comparison of clinic-based NC-stat versus referral to standard electrodiagnostic testing that evaluated efficiency of workup and costs. The study included 142 patients being considered for decompression surgery for CTS at a hand clinic. Seventy-one patients who accepted nerve conduction studies in a nurse-led clinic were compared with 71 historical controls who had been sent for nerve conduction studies at the regional neurophysiological unit. Patients with known or suspected complex neurological conditions were excluded from the study. Outcome measures were time from presentation to carpal tunnel decompression, the cost of each pathway, and the practicalities of using the device in the clinic. In the NC-stat group, 43 patients (61%) had a diagnosis of CTS confirmed by NC-stat and underwent decompression surgery, and 28 patients (39%) had normal or inconclusive tests. Of the 28, 12 were referred for electrodiagnostic testing, and 2 of the 12 were recommended for decompression surgery (3% false negative). In the referred group, 44 patients (62%) had confirmation of CTS and underwent decompression surgery. Use of NC-stat in the clinic reduced the time from presentation to surgery from 198 days to 102 days. Cost saving for NC-stat was reduced by the need to refer nearly 20% of patients for standard electrophysiological testing, but still favored the clinic-based approach. Health outcomes for the 2 approaches were not assessed.

The NeuroMetrix data registry was analyzed for all NC-stat studies performed over a period of 10 days that were coded for CTS and performed by a primary care provider. The initial data set consisted of studies on 1,190 patients performed by 613 different physician practices; studies that met CTS testing guidelines (82% met strict guidelines and 93% met less restrictive guidelines) were further analyzed. Thus, in nearly 1 of 5 patients (18.4%), the studies did not meet strict CTS testing guidelines. From the limited set, 31% were identified as normal, 53% exhibited CTS, 5% demonstrated an ulnar neuropathy, and 11% showed a nonspecific neuropathy. No comparison was made with standard nerve conduction testing nor was an assessment made of the impact of this testing on relevant clinical outcomes.

A 2007 study was identified that used NC-stat to assess the effect of a pharmaceutical agent on nerve conduction in patients with diabetic peripheral neuropathic pain.

One currently unpublished trial that might influence this policy is listed in Table 1.

Table 1. Summary of Key Trials

<table>
<thead>
<tr>
<th>NCT No.</th>
<th>Trial Name</th>
<th>Planned Enrollment</th>
<th>Completion Date</th>
</tr>
</thead>
</table>

NCT: national clinical trial.

a Denotes industry-sponsored or cosponsored trial.
Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

Summary
Studies have shown the correlation of portable automated nerve conduction test results with standard testing; however, questions remain about the diagnostic performance and clinical utility (i.e., impact on outcomes) of point-of-care automated testing. Particularly needed are data on the sensitivity and specificity of automated nerve conduction tests performed by non-specialists at the point-of-care in comparison with the “criterion standard” of laboratory NCS/EMG. One study from a tertiary care clinic found high sensitivity but low specificity for the diagnosis of lumbosacral radiculopathy. Another potential clinical use could be early identification of asymptomatic diabetic neuropathy to institute-appropriate clinical management before the onset of ulcerations, but no studies were identified that assessed the influence of point-of-care nerve conduction tests on clinical outcomes in this population. Overall, evidence addressing the utility of point-of-care nerve conduction tests in a clinical setting is limited. There is no peer-reviewed published medical literature on the use of voltage-actuated sensory nerve conduction tests and their impact on clinical outcomes. Overall, evidence remains insufficient to evaluate the effect of automated point-of-care nerve conduction tests on health outcomes. Therefore, automated point-of-care nerve conduction tests are considered investigational.

Quantitative Sensory Testing
Literature is reviewed on the various types of QST for which there are FDA-approved devices. This includes current perception threshold testing, PSSD testing, vibration perception threshold testing and thermal threshold testing.

For each of these, the following questions to the literature were asked:
• What is the technical performance of QST? (i.e., test-retest values, interoperator variability)
• What is the diagnostic performance of QST? (i.e., sensitivity and specificity of QST compared to conventional tests, using appropriate reference standards and conducted in appropriate populations)
• What is the clinical utility of QST? ie does QST change patient management or improve the net health outcome compared to standard testing?

In addition, systematic reviews and meta-analysis that evaluated studies using more than one type of QST were reviewed.

Systematic Reviews
A 2013 systematic review by Grosen and colleagues identified 14 studies that evaluated the association between QST findings and analgesic response. One study was conducted in healthy volunteers, 9 in surgical patients and 4 in patients with chronic pain. Study findings were not pooled, but were discussed for each of the patient populations. The authors reported that all of the studies in surgical patients were observational cohort studies, and analgesic response was not a primary outcome in any of the studies. Six of the 9 studies found a correlation between QST measurement (electrical, pressure and/or thermal stimulation) and consumption of analgesics. The article did not report whether or not the correlation was for all, or only some, of the outcomes related to analgesic consumption. The 4 studies on chronic pain patients were conducted as part of clinical drug trials, and QST was conducted at baseline prior to treatment. Two of the studies found a correlation between QST parameters and at least one analgesic response outcome.
Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

The authors concluded that the scientific evidence is not sufficiently robust that QST parameters are predictors of response to analgesic treatment.

**Current Perception Threshold Testing**

**Technical performance**

In 1999, the American Association of Electrodiagnostic Medicine (AAEM) published a technology review of the Neurometer® device. Much of the literature compared the results of Neurometer testing to nerve conduction studies in patients with known disease. In many instances, the results of the Neurometer testing demonstrated more numerous or pronounced abnormalities compared to nerve conduction studies, a finding that was consistent with the hypothesis that abnormalities of small nerve fibers precede those of large nerve fibers tested in nerve conduction studies. However, this observation could also be related to the fact that use of the Neurometer involves testing at multiple sites with 3 different frequencies and that any identified abnormality is considered significant. The AAEM assessment’s conclusions on the evidence of technical performance of current perception devices include the following:

- Reference values need to be established for well-characterized and representative populations.
- Reproducibility and interoperator variability of the Neurometer CPT normal values need to be established and expressed statistically in control subjects and patients with specific diseases.

In 2002, Yamashita and colleagues evaluated current perception thresholds using the Neurometer by comparing findings in 48 patients with lumbar radiculopathy and 11 healthy controls. The authors reported finding significantly higher current perception threshold values in the affected legs of patients with lumbar radiculopathy at 2,000, 250, and 5 Hz frequencies than in the unaffected legs. Current perception threshold values in the affected legs were also significantly higher in control subjects at 2,000 and 250 Hz frequencies but not significantly different at 5 Hz. The authors concluded that current perception threshold testing may be useful in quantifying sensory nerve dysfunction in patients with radiculopathy. However, this study did not establish standardized normal values or evaluate the reproducibility of QST measurements.

**Diagnostic performance**

Limited published evidence is available. Several studies have compared current perception threshold testing to other testing methods, but sensitivity and specificity have not been reported. For example, in 2012, Ziccardi and colleagues evaluated 40 patients presenting with trigeminal nerve injuries involving the lingual branch. Patients underwent current perception threshold testing, as well as standard clinical sensory testing. Statistically significant correlations were found between findings of electrical stimulation testing at 250 Hz and the reaction to pinprick testing (p=0.02), reaction to heat stimulation (p=0.01) and reaction to cold stimulation (p=0.004). In addition, significant correlations were found between electrical stimulation at 5 Hz and the reaction to heat stimulation (p=0.017), reaction to cold stimulation (p=0.004), but not the reaction to pinprick testing (p=0.096).

In addition, in 2001 Park and colleagues compared current perception threshold testing to standard references for thermal sensory testing and von Frey tactile hair stimulation in a randomized, double-blind, placebo-controlled trial with 19 healthy volunteers. The authors found that all current perception threshold measurements showed a higher degree of variability than thermal sensory testing and von Frey
measurements but concluded that there is some evidence that similar fiber tracts may be measured, especially C-fiber tract activity at 5 Hz, with current perception threshold, thermal sensory, and von Frey testing methods. This study was limited in that only healthy volunteers were included.

**Clinical Utility**

No comparative studies evaluating the impact of current perception testing on patient management decisions or health outcomes were identified.

A related study, published in 2009, utilized the Neurometer device in individuals with hand-arm vibration exposure. The primary purpose of the study was to evaluate the utility of a staging scale (the Stockholm sensorineural scale), not to determine the accuracy of quantitative sensory testing, so it does not provide additional evidence on the clinical utility of current perception testing as part of the initial evaluation of individuals with possible hand-arm vibration syndrome.

**Pressure-Specified Sensory Testing**

**Technical performance**

No published studies were identified.

**Diagnostic performance**

Standard evaluation and management of patients with potential nerve compression, disease, or damage consists of physical examination techniques and may include Semmes-Weinstein monofilament testing and, in some more complex cases, nerve conduction velocity (NCV) testing. Several studies have compared performance of these tests and PSSD. For example, a 2000 study by Weber and colleagues evaluated the sensitivity and specificity of PSSD and NCV testing in a total of 79 patients including 26 healthy controls. The NCV test had a sensitivity of 80% and a specificity of 77%. The PSSD had a sensitivity of 91% and a specificity of 82%; the difference between the two tests was not significantly different.

A 2010 study by Nath and colleagues evaluated 30 patients with winged scapula and upper trunk injury and 10 healthy controls. They used the U.S. FDA-cleared PSSD by Sensory Management Services to measure the minimum perceived threshold in both arms for detecting 1-point static (1PS) and 2-point static (2PS) stimuli. The authors used a published standard reference threshold value for the dorsal hand first web (DHFW) skin and calculated threshold values for both the DHFW and the deltoid using the upper limit of the 99% normal confidence interval. No published threshold values were available for the deltoid location. PSSD testing was done on both arms of all participants, and EMG testing only on the affected arms of symptomatic patients. Using calculated threshold values, patients with normal EMG results had positive PSSD results on 50% (8/16) of 1PS deltoid, 71% (10/14) of 2PS deltoid, 65% (11/17) of 1PS DHFW, and 87% (13/15) of 2PS DHFW tests. The authors stated that the study findings suggest that PSSD is more sensitive than needle EMG in detecting brachial plexus upper trunk injury. These findings should be confirmed in additional studies. In addition, the thresholds used to categorize a PSSD finding as positive for the deltoid should be validated in future reports.

A 2013 systematic review by Hubscher and colleagues evaluated studies on the relationship between QST and self-reported pain and disability in patients with spinal pain. Twenty-eight of 40 studies identified used
Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

PSSD. In their overall analysis, the investigators found low or no correlation between pain thresholds, as assessed by QST and self-reported pain intensity or disability. For example, the pooled estimate of the correlation between pain threshold and pain was -0.15 (95% confidence interval [CI]: -0.18 to -0.11) and between pain threshold and disability was -0.16 (95% CI: -0.22 to -0.10). The findings suggest low accuracy of QST as a tool for diagnosing patients' level of spinal pain and disability.

Clinical Utility
No clinical trials were identified that demonstrated that use of the PSSD resulted in changes in patient management or improved patient outcomes. In 2012, Suokas and colleagues published a systematic review of studies evaluating QST in painful osteoarthritis; the majority of studies used pressure testing. The authors did not report finding any studies that evaluated the impact of QST on health outcomes.

Vibration Testing
Technical Performance
A multicenter study funded by a pharmaceutical company compared vibration threshold testing (CASE IV, biothesiometer, C64 graduated tuning fork) with standard nerve conduction studies in 195 (86% follow-up) subjects with diabetes mellitus. The tests were performed independently by trained technicians; all standard nerve conduction evaluations were sent to a central reading center. Intraclass correlation coefficients for the tests ranged from 0.81 to 0.95, indicating excellent to highly reproducible results. Correlation coefficients for the various vibration QST instruments were moderate at -0.55 (CASE IV vs. tuning fork) to 0.61 (CASE IV vs. biothesiometer). In contrast, the correlation coefficient between CASE IV and a composite score for nerve conduction was low (r: 0.24). These results indicate that vibration threshold testing could not replace standard nerve conduction testing but might provide a complementary outcome measure.

Diagnostic Performance
In 2010, a study from India evaluated 100 patients with type 2 diabetes using a vibration perception threshold device, the Sensitometer (Dhansai Lab), which is produced in Mumbai and is not FDA-approved. The authors reported sensitivities and specificities compared to standard nerve conduction studies. For vibration testing, a positive finding (i.e., presence of neuropathy) was defined as patient reporting of no vibration sensation at a voltage of more than 15V. According to findings of nerve conduction studies, 70 of 100 patients had evidence of neuropathy. Vibration perception thresholds had a sensitivity of 86% and a specificity of 76%. Semmes-Weinstein monofilament testing, which was also done, had a higher sensitivity than vibration testing, 98.5%, and a lower specificity, 55%. Finally, a diabetic neuropathy symptom score determined by responses to a patient questionnaire, had a sensitivity of 83% and a specificity of 79%. The authors commented that the simple neurologic examination score appeared to be as accurate as vibration testing. The Sensitometer is not available in the United States, and it is not known how similar this device is to FDA-cleared vibration threshold testing devices.

Clinical Utility
No clinical trials were identified that demonstrated that use of vibration testing resulted in changes in patient management or improved patient outcomes.
Thermal Testing
Technical Performance
A 2012 systematic review by Moloney et al examined the literature on the reliability of thermal QST. A total of 21 studies met the review’s inclusion criteria (using an experimental design, assessing reliability, comparing thermal QST with other methods of assessment, testing at least twice). The investigators used a quality appraisal checklist to evaluate the reliability of the studies identified. Only 5 of the 21 studies were considered to be high quality. The review authors found considerable variation in the reliability of thermal QST; this included the 5 studies considered to be of high quality. The authors also noted several methodologic issues that could be improved in future studies, including better description of raters and their training, blinding and randomization, and standardization of test protocols. A 2015 study by Vuilleumier et al evaluated reliability of QST in a low back pain population; it included thermal QST using an FDA-approved device by Medoc. A total of 89 patients participated in 2 QST sessions conducted at least 7 days apart. The median of 3 thermal perception trials on the first day was compared with the median on the second day (between-session reliability). Several measures of reliability were reported (ie, coefficient of variability [CV]), ICC, coefficient of reliability). The reliability of heat pain detection and tolerance at the arm and leg were considered to be acceptable, with between-session CVs ranging from 1.8% to 6.1%. However, cold pain detection at the arm or leg did not have acceptable reliability, with between-session CVs ranging from 44% to 87%.

Diagnostic Performance
In 2008, Devigili et al. published a retrospective review of 486 patients referred for suspected sensory neuropathy. The study used an FDA-approved Medoc, Ltd. thermal perception-testing device. A total of 150 patients met the entry criteria for the study, which included symptoms suggesting sensory neuropathy and availability of 1) clinical examination, including spontaneous and stimulus-evoked pain, 2) a sensory and motor nerve conduction study, 3) warm and cooling thresholds assessed by quantitative sensory testing, and 4) skin biopsy with distal intraepidermal nerve fiber (IENF) density. Based on the combined assessments, neuropathy was ruled out in 26 patients; 124 patients were diagnosed with sensory neuropathy, and of these, 67 patients were diagnosed with small nerve fiber neuropathy. Using a cutoff of 7.63 IENF/mm at the distal leg (based on the 5th percentile of controls), 59 patients (88%) were considered to have abnormal IENF (small nerve fiber) density. Only 7.5% of patients had abnormal results for all 3 examinations (clinical, QST, skin biopsy), 43% of patients had both abnormal skin biopsy and clinical findings, and 37% of patients had both abnormal skin biopsy and QST results. The combination of abnormal clinical and QST results was observed in only 12% of patients. These results indicate that most of the patients evaluated showed an IENF density of less than 7.63 together with either abnormal spontaneous or evoked pain (clinical examination) or abnormal thermal thresholds (QST). The authors of this study recommended a new diagnostic “gold standard” based on the presence of at least 2 of 3 abnormal results (clinical, QST, and IENF density).

Clinical Utility
No clinical trials were identified that demonstrated that use of thermal testing resulted in changes in patient management or improved patient outcomes. A 2014 study by Ahmed et al addressed how QST might be used in practice, although it did not directly study clinical utility. The study was prospective and included 124 opioid-naïve patients scheduled for abdominal myomectomy or hysterectomy. Patients underwent
preoperative thermal QST, conducted by the same investigator. Tests included warm and cold sensation in which patients activated a button as soon as they felt a temperature change and warm and cold pain modalities in which the button was pressed when pain thresholds were reached.

The primary outcome was morphine use in the 24 hours after surgery. Intravenous morphine was administered postsurgery such that an individual’s pain level rated less than 4 on a 0 to 10 scale; pain was assessed on arrival and every 6 hours thereafter. In addition, a patient-controlled analgesia system was used to deliver morphine in the first 24 hours. Preoperative heat and cold pain thresholds were significantly correlated with 24-hour morphine consumption. Patients with initial heat pain thresholds above the median and cold pain thresholds above the median used more morphine (median, \( \geq 19 \) mg, \( p=0.004 \)). The authors stated that findings could be used to stratify patients preoperatively based on their baseline thermal QST scores and manage patients more or less aggressively, depending on their QST test findings. However, because the study did not prospectively manage patients and opioid administration was individualized to each patient; it is not clear how management would differ if QST scores had been incorporated in the management strategy.

### Clinical Input Received through Physician Specialty Societies and Academic Medical Centers

In response to the request for input from physician specialty societies and academic medical centers, input was received from 1 specialty society and 1 academic medical center regarding use of quantitative sensory testing while the policy was under review in 2008. Input from both sources agreed with the policy statement that QST is considered investigational, as adopted in the policy in April 2008.

### Summary

There is insufficient evidence that the use of quantitative sensory testing for the noninvasive assessment and quantification of sensory nerve function is as accurate as conventional tests. Questions remain about reference values in normal populations and the reproducibility of test results. In addition, there is a lack of evidence that use of quantitative sensory testing impacts patient management or improves the net health outcome. Therefore, this technology is considered investigational.

### Paraspinal Surface Electromyography to Evaluate and Monitor Back Pain

Surface electromyography has been used as a research tool to evaluate the performance of paraspinal muscles in patients with back pain and to further understand the etiology of low back pain. Preliminary research has also been performed on which SEMG parameters best differentiate between individuals with and without back pain. However, validation of its use as a clinical diagnostic technique involves a sequential 3-step procedure as follows:

1. Technical performance of a device is typically assessed by studies that compare test measurements with a gold standard and those that compare results taken with the same device on different occasions (test-retest).

2. Diagnostic performance is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with the gold standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true-positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true-negative).
Evaluation of diagnostic performance, therefore, requires independent assessment by the 2 methods in a population of patients who are suspected of disease but who do not all have the disease.

3. Evidence related to improvement of clinical outcomes with use of this testing assesses the data linking use of a test to changes in health outcomes (clinical utility). While in some cases, tests can be evaluated adequately using technical and diagnostic performance, when a test identifies a new or different group of patients with a disease; randomized trials are needed to demonstrate impact of the test on the net health outcome.

The following discussion focuses on these three steps as they apply to SEMG.

**Technical performance**

A 2014 systematic review by Mohseni Bandpei et al identified 12 studies on the test-retest reliability of paraspinal SEMG. Seven of the studies included only healthy individuals. The remaining 5 studies evaluated SEMG in patients with low back pain; 3 of these included a healthy control group. Overall, the studies reported that interrater reliability, as measured by intraclass coefficient, varied widely (from 0.26 to 0.91), with most of the values in the moderate to high range. Studies were heterogeneous in terms of methodology and SEMG parameters used. This evidence demonstrates that reliability of SEMG is at least moderate in the assessment of back muscle fatigue, but does not address the accuracy or validity of the test.

**Diagnostic performance**

No articles that compare the results of SEMG (which tests groups of muscles) with needle electromyography (which tests individual muscles) for diagnosing any specific muscle pathology were identified in literature searches. However, the pathology of individual muscles (i.e., radiculopathy, neuropathy, etc.) may represent a different process than the pathology of muscle groups (i.e., muscle strain, spasm, etc.), and thus SEMG may be considered by its advocates as a unique test for which there is currently no gold standard. Nevertheless, even if one accepts this premise, there are inadequate data to evaluate the diagnostic performance of SEMG. For example, no articles were identified in the published peer-reviewed literature that established definitions of normal or abnormal SEMG. In some instances, asymmetrical electrical activity may have been used to define abnormality; results may be compared to a “normative data base.” However, there was no published literature defining what degree of asymmetry would constitute abnormality or how a normative database was established.

In the absence of a criterion standard diagnostic test, correlation with the clinical symptoms and physical exam is critical. De Luca has published a series of studies investigating a type of SEMG called the Back Analysis System (BAS), consisting of surface electrodes and other components to measure the electrical activity of muscles during isometric exercises designed to produce muscle fatigue. Using physical exam and clinical history as a gold standard, the author found that BAS was able to accurately identify control and back pain patients 84% and 91% of the time, respectively, with the values increasing to 100% in some populations of patients. (Accuracy is the sum of true-positive and true-negative results.) However, these
studies were not designed as a clinical diagnostic tool per se but were intended to investigate the etiology of back pain and to investigate muscular fatigue patterns in patients with and without back pain.

Hu et al in Hong Kong published 2 articles on dynamic topography, an approach to analyzing SEMG findings. The studies had similar protocols. Both included low back pain patients and healthy controls; all participants underwent SEMG at study enrollment and then back pain patients participated in a rehabilitation program. The first study found different dynamic topography at baseline between healthy people and people with back pain, eg, a more symmetric pattern in healthy controls. After physical therapy, the dynamic topography images of back pain patients were more similar to the healthy controls on some of the parameters that were assessed. In the second study, following rehabilitation, back pain patients were classified as responders or nonresponders based on changes in back pain severity. Some associations were found between baseline SEMG parameters and response to rehabilitation. SEMG was not repeated following the rehabilitation program, and thus it is not clear whether there are any significant associations between continued symptoms and SEMG abnormalities. Moreover, it is not clear how SEMG analysis would affect treatment decisions for low back pain patients.

Improvement of Clinical Outcomes

Several articles describe the use of SEMG as an aid in classifying low back pain. Much of this research has focused on the use of spectral analysis to assess muscle fatigability rather than how information from SEMG could enhance patient management. While SEMG may be used to objectively document muscle spasm or other muscular abnormalities, it is unclear how such objective documentation would supplant or enhance clinical evaluation, or how this information would be used to alter the treatment plan. In part, the difficulty in clinical interpretation understands the extent to which the SEMG abnormalities are primary or secondary. In addition, as noted in the Background section, no specific workup is recommended for acute low back pain without warning signs.

There are no data regarding the impact of SEMG on the final health outcome. For example, SEMG has been proposed as a technique to differentiate muscle spasm from muscle contracture, with muscle spasm treated with relaxation therapy and contracture treated with stretching exercises. However, there are no data to validate that such treatment suggested by SEMG resulted in improved outcomes.

A review of spinal muscle evaluation in low back pain patients, published in 2007, indicates that the validity of SEMG remains controversial. The authors note that although many studies show increased fatigability of the paraspinal muscles in patients with low back pain, it is not known whether these changes are causes or consequences of the low back pain. Also, "the considerable inter-individual variability and the absence of normative data complicate the description of normal or abnormal profiles, thereby limiting the diagnostic usefulness of SEMG."

Summary

The evidence on paraspinal surface electromyography for evaluation and monitoring back pain includes a systematic review of interrater reliability and several studies on using findings to classify back pain. Relevant outcomes are test accuracy, test validity, symptoms, functional outcomes, quality of life, and
resource utilization. There are no data on the impact of SEMG on patient management or health outcomes. The evidence is insufficient to determine the effects of the technology on health outcomes.

**References**

Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017


Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

01/19/2011 Medical Policy Implementation Committee approval. Coverage eligibility unchanged.
03/01/2012 Medical Policy Committee review
03/21/2012 Medical Policy Implementation Committee approval. Coverage eligibility unchanged.
01/23/2013 Coding updated
03/07/2013 Medical Policy Committee review
03/20/2013 Medical Policy Implementation Committee approval. Added “Based on review of available data, the Company considers automated nerve conduction tests to be investigational.”
03/06/2014 Medical Policy Committee review
03/19/2014 Medical Policy Implementation Committee approval. No change to coverage.
04/02/2015 Medical Policy Committee review
04/20/2015 Medical Policy Implementation Committee approval. No change to coverage.
08/03/2015 Coding update: ICD10 Diagnosis code section added; ICD9 Procedure code section removed.
04/07/2016 Medical Policy Committee review
04/20/2016 Medical Policy Implementation Committee approval. Investigational statement for NEMG or NCS for treatment of any diagnosis other than those listed in coverage statement.
01/01/2017 Coding update: Removing ICD-9 Diagnosis Codes
04/06/2017 Medical Policy Committee review
04/19/2017 Medical Policy Implementation Committee approval. No change to coverage.
Next Scheduled Review Date: 04/2018

Coding

The five character codes included in the Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines are obtained from Current Procedural Terminology (CPT®), copyright 2016 by the American Medical Association (AMA). CPT is developed by the AMA as a listing of descriptive terms and five character identifying codes and modifiers for reporting medical services and procedures performed by physician. The responsibility for the content of Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines is with Blue Cross and Blue Shield of Louisiana and no endorsement by the AMA is intended or should be implied. The AMA disclaims responsibility for any consequences or liability attributable or related to any use, nonuse or interpretation of information contained in Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines. Fee schedules, relative value units, conversion factors and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein. Any use of CPT outside of Blue Cross Blue Shield of Louisiana Medical Policy Coverage Guidelines should refer to the most current Current Procedural Terminology which contains the complete and most current listing of CPT codes and descriptive terms. Applicable FARS/DFARS apply.

CPT is a registered trademark of the American Medical Association.

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

<table>
<thead>
<tr>
<th>Code Type</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0106T, 0107T, 0108T, 0109T, 0110T, 95860, 95861, 95863, 95864, 95865, 95866, 95867, 95868, 95869, 95870, 95872, 95875, 95885, 95886, 95887, 95905, 95907, 95908, 95909, 95910, 95911, 95912, 95913, 95925, 95926, 95927, 95999</td>
</tr>
<tr>
<td>HCPCS</td>
<td>G0255, S3900</td>
</tr>
</tbody>
</table>
Neurodiagnostics

Policy # 00186
Original Effective Date: 05/15/2006
Current Effective Date: 04/19/2017

<table>
<thead>
<tr>
<th>ICD-10 Diagnosis</th>
<th>All related diagnoses</th>
</tr>
</thead>
</table>

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

A. Whether the medical treatment, procedure, drug, device, or biological product can be lawfully marketed without approval of the U.S. FDA and whether such approval has been granted at the time the medical treatment, procedure, drug, device, or biological product is sought to be furnished; or

B. Whether the medical treatment, procedure, drug, device, or biological product requires further studies or clinical trials to determine its maximum tolerated dose, toxicity, safety, effectiveness, or effectiveness as compared with the standard means of treatment or diagnosis, must improve health outcomes, according to the consensus of opinion among experts as shown by reliable evidence, including:

1. Consultation with the Blue Cross and Blue Shield Association technology assessment program (TEC) or other nonaffiliated technology evaluation center(s);
2. Credible scientific evidence published in peer-reviewed medical literature generally recognized by the relevant medical community; or
3. Reference to federal regulations.

**Medically Necessary (or “Medical Necessity”) - Health care services, treatment, procedures, equipment, drugs, devices, items or supplies that a Provider, exercising prudent clinical judgment, would provide to a patient for the purpose of preventing, evaluating, diagnosing or treating an illness, injury, disease or its symptoms, and that are:

A. In accordance with nationally accepted standards of medical practice;

B. Clinically appropriate, in terms of type, frequency, extent, level of care, site and duration, and considered effective for the patient's illness, injury or disease; and

C. Not primarily for the personal comfort or convenience of the patient, physician or other health care provider, and not more costly than an alternative service or sequence of services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of that patient's illness, injury or disease.

For these purposes, "nationally accepted standards of medical practice" means standards that are based on credible scientific evidence published in peer-reviewed medical literature generally recognized by the relevant medical community, Physician Specialty Society recommendations and the views of Physicians practicing in relevant clinical areas and any other relevant factors.

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

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