NEUROPSYCHOLOGICAL TESTING UNDER THE MEDICAL BENEFIT

Policy Number: CS083.E  Effective Date: October 1, 2017

INSTRUCTIONS FOR USE

This Medical Policy provides assistance in interpreting UnitedHealthcare benefit plans. When deciding coverage, the federal, state or contractual requirements for benefit plan coverage must be referenced. The terms of the federal, state or contractual requirements for benefit plan coverage may differ greatly from the standard benefit plan upon which this Medical Policy is based. In the event of a conflict, the federal, state or contractual requirements for benefit plan coverage supersede this Medical Policy. All reviewers must first identify member eligibility, any federal or state regulatory requirements, and the contractual requirements for benefit plan coverage prior to use of this Medical Policy. Other Policies and Coverage Determination Guidelines may apply. UnitedHealthcare reserves the right, in its sole discretion, to modify its Policies and Guidelines as necessary. This Medical Policy is provided for informational purposes. It does not constitute medical advice.

UnitedHealthcare may also use tools developed by third parties, such as the MCG™ Care Guidelines, to assist us in administering health benefits. The MCG™ Care Guidelines are intended to be used in connection with the independent professional medical judgment of a qualified health care provider and do not constitute the practice of medicine or medical advice.

BENEFIT CONSIDERATIONS

Before using this policy, please check the federal, state or contractual requirements for benefit coverage.

Some benefit documents within UnitedHealthcare exclude neuropsychological testing for some or all indications. The member’s state specific documents must be consulted in order to determine benefit coverage for neuropsychological testing.

Neuropsychological testing for attention-deficit/hyperactivity disorder (ADHD) is a medical benefit service when medically referred and related or secondary to a known/suspected organic-medical condition resulting from brain injury or disease process (e.g., concussion, intractable seizure disorder, cancer treatment effects). Neuropsychological testing for ADHD is a mental health benefit service when representing a developmental condition not due to specific brain injury or disease process, where there are suspected organic functional impairments.

The scope of the criteria for attention-deficit/hyperactivity disorders and developmental disorders or significant developmental delays is applicable only to neuropsychological testing that is covered by the medical benefit.

COVERAGE RATIONALE

Neuropsychological testing is proven and medically necessary for evaluating patients with the following conditions when the result of testing will influence clinical decision making:

- Attention-deficit/hyperactivity disorder (ADHD) when all of the following are present:

     1. Attention-deficit/hyperactivity disorder (ADHD)
     2. Medical referral for neuropsychological testing
     3. Clinical decision making influenced by test results
Specific neurocognitive behavioral deficits related to ADHD need to be evaluated and testing has been recommended by a physician and is related or secondary to a known or suspected organic-medical condition resulting from brain injury or disease process (e.g., concussion, intractable seizure disorder, cancer treatment effects, genetic disorders, inborn errors of metabolism).

*The scope of these criteria is applicable only to neuropsychological testing that is covered by the medical benefit. These criteria do not apply to evaluate or determine educational interventions.*

- Confirmed space-occupying brain lesion including but not limited to the following:
  - Brain abscess
  - Brain tumors
  - Arteriovenous malformations within the brain
- Dementia or symptoms of dementia such as memory impairment or memory loss (including extrapyramidal disorders such as Parkinson’s disease) that is associated with a new onset or progressive memory loss and a decline in at least one of the following cognitive domains (DSM-5):
  - Complex attention
  - Executive function
  - Learning and memory
  - Language
  - Perceptual-motor
  - Social cognition
- Demyelinating disorders including multiple sclerosis
- Intellectual disability or intellectual developmental disorder when all of the following are present:
  - The intellectual disability or intellectual developmental disorders associated with a known or suspected medical cause (e.g., traumatic brain injury, in utero toxin exposure, early seizure disorder, sickle cell disease, genetic disorders) and
  - The intellectual disability or intellectual developmental disorder meets all of the following criteria (DSM-5):
    - Deficits in intellectual function, such as reasoning, problem solving, planning, abstract thinking, judgment, academic learning, and learning from experience, confirmed by both clinical assessment and individualized, standardized intelligence testing
    - Deficits in adaptive functioning that result in failure to meet developmental and sociocultural standards for personal independence and social responsibility. Without ongoing support, the adaptive deficits limit functioning in one or more activities of daily life, such as communication, social participation, and independent living across multiple environments, such as home, school, work and community, and
    - Onset of intellectual and adaptive deficits during the developmental period

*The scope of these criteria is applicable only to neuropsychological testing that is covered by the medical benefit. These criteria do not apply to evaluate or determine educational interventions.*

- Encephalopathy including acquired immunodeficiency syndrome (AIDS) encephalopathy, human immunodeficiency virus (HIV) encephalopathy, hepatic encephalopathy, Lyme disease encephalopathy including neuroborreliosis, Wernicke’s encephalopathy and systemic lupus erythematosus (SLE) encephalopathy.
- Neurotoxin exposure with at least one of the following:
  - Demonstrated serum levels of neurotoxins
  - Individual with documented significant prenatal alcohol, drug, or toxin exposure
- Seizure disorder including patients with epilepsy and patients being considered for epilepsy surgery
- Stroke
- Traumatic brain injury (TBI): TBI is defined as a bump, blow, or jolt to the head or a penetrating head injury that disrupts the normal function of the brain (Centers for Disease Control and Prevention). See the following website for more information: [http://www.cdc.gov/TraumaticBrainInjury/index.html](http://www.cdc.gov/TraumaticBrainInjury/index.html). (Accessed May 30, 2017)

**Baseline neuropsychological testing is unproven and not medically necessary in asymptomatic persons at risk for sport-related concussions or brain injuries.**

There is insufficient evidence to indicate that the use of baseline neuropsychological testing in athletes or other individuals alters risk from concussion. There is insufficient evidence that baseline tests influence physician decision-making or outcomes of treatment of concussion.

**Computerized neuropsychological testing is unproven and not medically necessary for evaluating concussions or brain injuries.**

The evidence is insufficient to establish the validity and reliability of computerized tests to evaluate concussions. Prospective controlled trials are needed to demonstrate the clinical utility of these tests to detect impairment following concussion.

**Neuropsychological testing is unproven and not medically necessary for the following diagnoses alone without other proven conditions as noted above:**

- Headaches including migraine headache
- History of myocardial infarction
• Intermittent explosive disorder

There is insufficient clinical evidence to demonstrate that the use of neuropsychological testing for patients with myocardial infarction, migraine or other headaches or intermittent explosive disorder without associated cognitive disorders can be used effectively for clinical decision making to improve patient management of those conditions.

**Computerized cognitive testing including but not limited to Cognivue®, Mindstreams® Cognitive Health Assessment and BrainCare™ is unproven and not medically necessary for diagnosing dementia or mild cognitive impairment.**

Available clinical trials have failed to document a beneficial effect of computerized cognitive testing on long-term clinical outcomes. The evidence is insufficient to establish the validity of computerized cognitive testing compared with traditional tests for the assessment of dementia and cognitive impairment.

### APPLICABLE CODES

The following list(s) of procedure and/or diagnosis codes is provided for reference purposes only and may not be all inclusive. Listing of a code in this policy does not imply that the service described by the code is a covered or non-covered health service. Benefit coverage for health services is determined by federal, state or contractual requirements and applicable laws that may require coverage for a specific service. The inclusion of a code does not imply any right to reimbursement or guarantee claim payment. Other Policies and Coverage Determination Guidelines may apply.

<table>
<thead>
<tr>
<th>CPT Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>96116</td>
<td>Neurobehavioral status exam (clinical assessment of thinking, reasoning and judgment, e.g., acquired knowledge, attention, language, memory, planning and problem solving, and visual spatial abilities), per hour of the psychologist's or physician's time, both face-to-face time with the patient and time interpreting test results and preparing the report</td>
</tr>
<tr>
<td>96118</td>
<td>Neuropsychological testing (e.g., Halstead-Reitan Neuropsychological Battery, Wechsler Memory Scales and Wisconsin Card Sorting Test), per hour of the psychologist's or physician's time, both face-to-face time administering tests to the patient and time interpreting these test results and preparing the report</td>
</tr>
<tr>
<td>96119</td>
<td>Neuropsychological testing (e.g., Halstead-Reitan Neuropsychological Battery, Wechsler Memory Scales and Wisconsin Card Sorting Test), with qualified health care professional interpretation and report, administered by technician, per hour of technician time, face-to-face</td>
</tr>
<tr>
<td>96120</td>
<td>Neuropsychological testing (e.g., Wisconsin Card Sorting Test), administered by a computer, with qualified health care professional interpretation and report</td>
</tr>
</tbody>
</table>

_CPT® is a registered trademark of the American Medical Association_

### DESCRIPTION OF SERVICES

Neuropsychological testing is a set of formal procedures utilizing diagnostic tests specifically focused on identifying the presence of brain damage, injury or dysfunction and any associated functional deficits. Measurement of deficits cannot be based on single test results and should always be assessed in the context of the medical and neurological examination. Neuropsychological testing is customarily associated with neurological diagnoses rather than behavioral health diagnoses.

Neuropsychological tests are administered in a variety of contexts including paper-and-pencil, computers, and visual aids. Following an initial clinical interview with a neuropsychologist, tests are strategically selected to identify specific deficits and preserved abilities. Standardized tests are then administered by a trained technician or neuropsychologist. Some tests offer multiple forms making them useful for repeated administration to the same patient, thereby minimizing practice effects. In light of the numerous procedures available for assessment of different neurocognitive functions, test selection is based on familiarity of the examiner with certain tests, availability of appropriate normative data, ability of the patient to participate in testing (e.g., quadriplegic or hemiplegic patients may not be able to participate in psychomotor testing), and validity of particular procedures for the specific function being measured. For developmental disorders, neuropsychological tests are useful as part of a complete clinical decision making process and do not unilaterally make the diagnosis of autism spectrum disorder (Zwaigenbaum, 2009).

Neuropsychological tests include but are not limited to the following: Boston Diagnostic Aphasia Examination (BDAE), Conners' Continuous Performance Test (CCPT), Controlled Oral Word Association Test (COWAT), Delis-Kaplan Test Battery, Freedom from Distractibility Index (FFDI) from the Wechsler Intelligence Scales, Gordon Diagnostic System (GDS), Halstead-Reitan Neuropsychological Battery, Rey Auditory Verbal Learning Test (RAVLT), Rey-Osterreith...
Complex Figure Test, Stroop Color and Word Test, Test of Variables of Attention (TOVA), Trail Making Tests, Wechsler Adult Intelligence Scale-Revised (WAIS-III/IV), Wide Range Assessment of Memory and Learning (WRAML), and Wisconsin Card Sorting Test (WCST). At times, neurocognitive measures are supplemented by emotional functioning and personality testing and include but are not limited to the following: Minnesota Multiphasic Personality Inventory-2 (MMPI-2)/Minnesota Multiphasic Personality Inventory-A (MMPI-A), Personality Assessment Inventory (PAI), Geriatric Rating Scale, Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), and Rorschach Inkblot Method.

Computerized testing for dementia and cognitive impairment including the Mindstreams® Cognitive Health Assessment (NeuroTrax® Corp.) uses computer-based assessments in an attempt to identify cognitive impairment in the elderly. The software programs give patients various stimuli or puzzles to solve using a mouse or a keypad. The Mindstreams system automatically generates a report that details the patient's performance in the standard cognitive domains, or areas, e.g., memory, attention, executive function, visual spatial perception, verbal skills, motor planning, and information processing. According to NeuroTrax, BrainCare™ is the current version of the original MindStreams product. Cognivue (Cerebral Assessment Systems, Inc.) is another computerized cognitive test that is intended for early detection of dementia signs. Patients take the 10-minute test using the Cognivue mobile computer workstation to assess visuomotor coordination, perceptual processing, and memory. Cognivue is intended to help identify patients who may be in the early stages of dementia and should undergo further evaluation.

Computerized neuropsychological tests have been proposed to be used as part of the overall management of concussion to monitor recovery. Most computer-based cognitive assessment tools are designed to detect the speed and accuracy of attention, memory, and thinking ability. Currently available computerized tests include ImPACT (Immediate Post-Concussion Assessment and Cognitive Testing, ImPACT Applications, Inc.), ANAM (Automated Neuropsychological Assessment Metrics, the United States Army Medical Department), CogState Sport (Axon Sports, Ltd.), and HeadMinder (Headminder, Inc.). These tests are being investigated for baseline testing of asymptomatic persons and managing concussions once they occur.

Neuropsychological testing is within the scope of the provider’s professional training and licensure when the provider is any of the following:

- A doctorial-level psychologist who is licensed to practice independently, and demonstrates sufficient training and experience.
- A psychometrist or psychometrician who administers and scores psychological tests under the supervision of a licensed doctoral-level psychologist, and whose services are billed by the supervising psychologist. The supervising psychologist must have face-to-face contact with the member at intake and during the feedback session and is responsible for final test interpretation, report writing and final signature of approval.
- A credentialed psychiatrist who meets the following requirements:
  - Recognized certification in neurology through the American Board of Psychiatry and Neurology; Accreditation in behavioral neurology and neuropsychiatry through the American Neuropsychiatric Association;
  - State medical licensure specifically allowing for the provision of neuropsychological testing service(s);
  - Evidence of professional training and expertise in the specific tests and/or assessment measures for which authorization is requested;
  - Physician and supervised psychometrician adhere to the prevailing national professional and ethical standards regarding test administration, scoring and interpretation.
  - A board-certified neurologist

See the following Optum Coverage Determination Guideline for more information:

- Psychological and Neuropsychological Testing (to access this guideline, go to: Optum Provider Express > Clinical Resources > Guidelines/Policies/Manuals > Coverage Determination Guidelines)

**CLINICAL EVIDENCE**

**Attention Deficit Hyperactivity Disorder (ADHD)**

In a systematic review, Hall et al. (2016) describe the current evidence base for the use of neuropsychological continuous performance tests (CPTs) and objectively measured activity to support the diagnostic procedure and medication management for children with attention-deficit hyperactivity disorder (ADHD). Four databases (PsycINFO, Medline, Allied and Complementary Medicine (AMED), and PsycARTICLES) were systematically searched to understand the current evidence base for (1) the use of CPTs to aid clinical assessment of ADHD; (2) the use of CPTs to aid medication management; and (3) the clinical utility of objective measures of activity in ADHD. Sixty relevant articles were identified. The search revealed six commercially available CPTs that had been reported on for their clinical use. There were mixed findings with regard to the use of CPTs to assess and manage medication, with contrasting evidence on their ability to support clinical decision-making. There was a strong evidence base for the use of objective measures of activity to aid ADHD/non-ADHD group differentiation, which appears sensitive to medication effects and would also benefit from further research on their clinical utility. The findings suggest that combining CPTs and an objective measure of activity may be particularly useful as a clinical tool and worthy of further pursuit.
Bechtel et al. (2012) evaluated whether boys with epilepsy-related ADHD and developmental ADHD share a common behavioral, pharmaco-responsive, and neurofunctional pathophysiology. Seventeen boys with diagnosed combined epilepsy/ADHD, 15 boys with developmental ADHD, and 15 healthy controls (aged 8-14 years) performed on working memory tasks (N-back) while brain activation was recorded using functional magnetic resonance imaging. On a behavioral level, boys with epilepsy-related ADHD as well as those with developmental ADHD performed similarly poorly on tasks with high cognitive load when compared to healthy controls. On the functional level, both patient groups showed similar reductions of activation in all relevant parts of the functional network of working memory when compared to controls. The study data showed strong similarities between epilepsy-related and developmental ADHD on the behavioral, pharmaco-responsive, and neural level, favoring the view that ADHD with and without epilepsy shares a common underlying neurobehavioral pathophysiology.

Dementia, Possible Dementia, Memory Loss, and Memory Impairment
For memory impairment or dementia screening, a single test of global measures of function or a measure of cognitive function is usually administered along with a test of behavioral or emotional symptoms. In addition to brief screening tests, for some patients, comprehensive neuropsychological testing may be indicated to confirm a diagnosis, evaluate effects of treatment, and assist in designing rehabilitative or intervention strategies for the patient. Standardized test batteries are too long for most patients with dementia; specialized dementia batteries or an individualized test battery is usually more appropriate.

A definitive diagnosis of Alzheimer’s disease is based on the presence of memory deficits along with deficits in at least one other aspect of cognition, and in some cases is made on neuropsychological test results alone (Talwalker, 1996). Impairment in primary (short-term) memory alone is not a useful diagnostic marker for Alzheimer’s disease in the early stages. Tests of delayed recall (long-term memory) and retrieval of facts of common knowledge have been shown to be the most useful measures to distinguish normal aging and early Alzheimer’s disease. Dementia due to Alzheimer’s disease can be distinguished from dementia due to vascular disease by differences in pattern of memory impairment and the progressive nature of Alzheimer’s disease (Costa et al., 2017). Careful interpretation of test results, taken in conjunction with medical findings, allows differentiation of Alzheimer’s disease from normal memory loss due to aging, and from vascular dementia.

Pedersen et al. (2016) examined the incidence, progression, and reversion of mild cognitive impairment in patients with Parkinson disease (PD-MCI) over 5 years. A population-based cohort of patients with incident PD underwent repeated neuropsychological testing of attention, executive function, memory, and visuospatial abilities at baseline (n=178), 1 year (n=175), 3 years (n=163), and 5 years (n=150). Patients were classified as PD-MCI and diagnosed with dementia according to published criteria. Thirty-six patients (20.2%) fulfilled criteria for PD-MCI at baseline. Among those with normal cognition at baseline (n=142), the cumulative incidence of PD-MCI was 9.9% after 1 year, 23.2% after 3 years, and 28.9% after 5 years of follow-up. Overall, 39.1% of patients with baseline or incident PD-MCI progressed to dementia during the 5-year study period. The conversion rate to dementia was 59.1% in patients with persistent PD-MCI at 1 year vs 7.2% in those with normal cognition during the first year. A total of 27.8% of patients with baseline PD-MCI and 24.2% of those with incident PD-MCI had reverted to normal cognition at study end, but the reversion rate decreased to 9.4% in those with persistent PD-MCI at 2 consecutive visits. Compared with cognitively normal patients, PD-MCI reverters within the first 3 years of follow-up were at increased risk of subsequently developing dementia. The authors concluded that early PD-MCI, regardless of persistence or reversion to normal cognition, has prognostic value for predicting dementia in patients with PD.

Yoon et al. (2015) evaluated whether olfactory and neuropsychological tests can aid in the differentiation of dementia with Lewy bodies (DLB) from Alzheimer’s disease (AD) at the mild cognitive impairment (MCI) stage since the early differentiation of DLB from AD may be important to delay disease progression. The study included 122 MCI patients who were monitored until they developed dementia or until their condition stabilized; the follow-up period averaged 4.9 years (range: 3.9-6.2 years). Baseline olfactory function as measured with the Cross-Cultural Smell Identification (CCSI) test and neuropsychological data were compared. During the follow-up period, 32 subjects developed probable AD (MCI-AD), 18 had probable DLB (MCI-DLB), 45 did not convert to dementia (MCI-stable), and eight developed a non-AD/DLB dementia. The mean CCSI score was significantly lower than that of MCI-AD patients and MCI-stable patients. The area under the curve of the receiver operating characteristic to discriminate MCI-DLB from MCI-AD using CCSI scores was 0.84. Frontal-executive function and visuospatial ability was worse in patients with MCI-DLB, while verbal recognition memory impairment was greater in those with MCI-AD. The authors concluded that olfactory and neuropsychological tests can help predict conversion to DLB or AD in patients with MCI.

Madureira et al. (2010) determined the extent to which the performance in neuropsychological tests would be able to predict the clinical diagnosis of dementia. The LADIS (Leukoaraiosis and Disability) is a prospective study that evaluates the impact of white matter changes (WMC) on the transition of independent elderly to disability. The subjects were evaluated at baseline and yearly during 3 years with a comprehensive clinical, functional and neuropsychological protocol. At each visit, dementia was classified according to clinical criteria. The performance in
the neuropsychological batteries was compared according to the clinical diagnosis of dementia. From the initially enrolled 639 subjects, 480 were evaluated at year 3. Dementia was diagnosed in 90 participants. The demented subjects had worse performance in almost all the baseline cognitive tests. Using receiver operating characteristic curves, the investigators found that the Vascular Dementia Assessment Scale (VADAS) battery had higher sensitivity and specificity rates to identify dementia compared with the Mini-Mental State Examination (MMSE) and Alzheimer's Disease Assessment Scale. Worse performances on baseline MMSE were predictors of dementia. The investigators concluded that performance on the MMSE and the VADAS battery were important predictors of dementia at a 3-year period.

Pseudodementia, a dementia of "nonorganic" etiology, is due to profound depression and can be difficult to differentiate from true dementia. The Geriatric Depression Scale is commonly used for evaluating depression in elderly people. Prospective studies have shown increased accuracy in differentiating pseudodementia from true dementia with repeated testing 12-18 months later (Yousef, 1998). This is a vital distinction to make, as organic dementia is often progressive and is usually not reversible, while dementia associated with depression may reverse or resolve with treatment.

**Developmental Disorders**

In general, empirical data, rather than evidence from prospective studies with long-term follow-up, support the use of neuropsychological testing for developmental disorders in infants and children.

In a national cohort of extremely low birth weight (ELBW) children, neuropsychological test profiles were assessed in 4 groups defined according to a neurological examination at 5 years of age: normal neuromotor status (N=56), motor coordination problems (N=32), multiple subtle neuromotor signs including both motor coordination problems and deviant reflexes (N=20), and spastic diplegia (N=12). The neurocognitive assessment included a test of intelligence, the Wechsler Primary and Preschool Scale of Intelligence–Revised (WPPSI-R) and 14 subtests of attention and executive functions, verbal functions, manual motor functions, visuoconstructive functions and verbal learning. The children with normal neuromotor status performed within the average range; children with motor coordination problems had widespread impairment; and children with spastic diplegia and children with multiple minor neuromotor signs had uneven test profiles with stronger verbal results but weaknesses in attention and executive functions, and in manual motor and visuoconstructual tasks. According to the investigators, very early gestation children with neuromotor signs, including motor coordination problems, are at risk for neurocognitive impairment, in spite of average intelligence. More impaired children have more irregular test profiles. Follow-up and neuropsychological assessments of very preterm children with minor neuromotor signs are therefore indicated (Korkman et al., 2008).

Hartman et al. (2010) examined the motor skills and executive functions in school-age children with borderline and mild intellectual disabilities (ID). Sixty-one children aged between 7 and 12 years diagnosed with borderline ID (33 boys and 28 girls; 71 < IQ < 79) and 36 age peers with mild ID (24 boys and 12 girls; 54 < IQ < 70) were assessed. Their abilities were compared with those of 97 age- and gender-matched typically developing children. Qualitative motor skills, i.e. locomotor ability and object control, were evaluated with the Test of Gross Motor Development (TGMD-2). Executive functioning (EF), in terms of planning ability, strategic decision-making and problem solving, was gauged with the Tower of London (TOL) task. Compared with the reference group, the full ID cohort scored significantly lower on all assessments. According to the investigators, the study results support the notion that besides being impaired in qualitative motor skills, intellectually challenged children are also impaired in higher-order executive functions. The authors conclude that deficits in the two domains are interrelated, so early interventions boosting their motor and cognitive development are recommended.

**Traumatic Brain Injury**

Longitudinal and case controlled studies along with numerous case reports support the use of neuropsychological tests to assess the severity of injury and the prognosis for patients with closed head trauma, to monitor progression, and to provide measures of outcome for determining degree of recovery (Hanks et al., 2016; Carlozzi et al., 2015).

**Other Disorders**

Neuropsychological testing may have a role in the clinical management of the following medical disorders:

- Brain lesions including abscesses, tumors, and arteriovenous malformations in the brain (Meskal et al., 2016; Walsh et al., 2016; Cochereau et al., 2016; Iuvone et al., 2011; Visani et al., 2006)
- Demyelinating disease including multiple sclerosis (Vollmer et al., 2016; Glanz et al., 2012)
- Encephalopathy (Burton et al., 2017; Poh and Chang, 2012; Stewart et al., 2010)
- Epilepsy and seizure disorders (Parra-Díaz and García-Casares, 2017; Grau-López et al., 2017; Wilson et al., 2015; Filippini, 2016; Patrikelis et al.,)
- Neurotoxin exposure (Nascimento et al., 2016)
- Stroke (Tan et al., 2017; Zweifel-Zehnder et al., 2015; Chen et al., 2015)
Computerized Neuropsychological Testing for Concussion

Gaudet and Weyandt (2017) conducted a systematic review of existing research investigating Immediate Post-Concussion and Cognitive Testing (ImPACT) and the prevalence of invalid baseline results including the effectiveness of ImPACT’s embedded invalidity indicators in detecting suspect effort. Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines were followed in order to systematically structure a search across four databases and analysis of studies that presented data related to the prevalence of invalid performance and/or the effectiveness of ImPACT’s embedded invalidity indicators. A total of 17 studies included prevalence rates of invalid performances or examined the effectiveness of ImPACT’s invalidity indicators. Of the 17 studies, 12 included prevalence rates of invalid baseline results; and across this group of studies (after removing an outlier), the weighted prevalence rate of invalid baseline results was 6%. Four of the 17 studies examined the effectiveness of ImPACT’s embedded invalidity indicators. ImPACT’s embedded invalidity indicators correctly identified suboptimal effort in approximately 80% of individuals instructed to perform poorly and avoid detection (‘coached’) or instructed to perform poorly (‘naive’). According to the authors, these findings raise a number of issues pertaining to the use of ImPACT. Invalid performance incidence may increase with large group versus individual administration, use in nonclinical settings, and among those with Attention Deficit-Hyperactivity Disorder or learning disability. Additionally, the older desktop version of ImPACT appears to be associated with a higher rate of invalid performances than the online version. Although ImPACT’s embedded invalidity indicators detect invalid performance at a rate of 6% on average, known group validity studies suggest that these measures miss invalid performance approximately 20% of the time when individuals purposefully underperform.

Kontos et al. (2014) performed a meta-analysis assessing the effects of sport-related concussion as measured by computerized neurocognitive tests (NCT) 1-week post injury. Thirty-seven studies involving 3960 participants between 2000 and 2011 were included. Code substitution, visual memory, processing speed, and memory tasks demonstrated negative effects for concussion. Younger adolescents had lower NCT performance than older adolescents and college aged athletes. ImPACT studies demonstrated a negative effect for concussion as did those involving contact sports. The authors found that computerized neurocognitive testing results suggest athletes suffer impairments within one week of a concussion. Several factors such as age, type of neurocognitive test, and test administrator may lead to more pronounced impairments. The authors indicated that no single tool can or should be used to measure the effect of concussion. Instead, clinicians and researchers should adopt a comprehensive approach to assessing this injury.

Echemendia et al. (2013) critically reviewed the literature from the past 12 years regarding key issues in sports-related neuropsychological assessment of concussion. Based on the review of the literature, the authors concluded that traditional and computerized neuropsychological tests are useful in the evaluation and management of concussion. Brief cognitive evaluation tools are not substitutes for formal neuropsychological assessment. According to the authors, there is insufficient evidence to recommend the widespread routine use of baseline neuropsychological testing.

Bruce et al. (2014) examined the 1 year test-retest reliability of ImPACT in a multilingual sample of professional hockey players. A total of 305 professional hockey players were tested 1 year apart using ImPACT. Reliable change confidence intervals were calculated and test-retest reliability was measured using Pearson and Intraclass correlation coefficients. Results indicated that the 1-year test-retest reliabilities for the Visual Motor and Reaction Time Composites ranged from low to high (.52 to .81). In contrast, 1-year test-retest reliabilities for the Verbal and Visual Memory Composites were low (.22 to .58). According to the authors, the 1-year test-retest results provided mixed support for the use of Visual Motor and Reaction Time Composites in select samples; in contrast, the Verbal and Visual Memory Composites may not be sensitive to clinical change.

Nakayama et al. (2014) examined the test-retest reliability of the ImPACT between baseline, 45 days, and 50 days. Eighty-five physically active college students (51 male, 34 female) volunteered for this study. Participants completed the ImPACT as well as a 15-item memory test at baseline, 45 days, and 50 days. Intraclass correlation coefficients (ICCs) were calculated for ImPACT composite scores, and change scores were calculated using reliable change indices (RCIs) and regression-based methods (RBMs) at 80% and 95% confidence intervals (CIs). The respective ICCs for baseline to day 45, day 45 to day 50, baseline to day 50, and overall were as follows: verbal memory (0.76, 0.69, 0.65, and 0.78), visual memory (0.72, 0.66, 0.60, and 0.74), visuomotor (processing) speed (0.87, 0.88, 0.85, and 0.91), and reaction time (0.67, 0.81, 0.71, and 0.80). All ICCs exceeded the threshold value of 0.60 for acceptable test-retest reliability. All cases fell well within the 80% CI for both the RCI and RBM, while 1% to 5% of cases fell outside the 95% CI for the RCI and 1% for the RBM. According to the authors, the study results suggest that the ImPACT is a reliable neurocognitive test battery at 45 and 50 days after the baseline assessment. The current findings agree with those of other reliability studies that have reported acceptable ICCs across 30-day to 1-year testing intervals, and they support the utility of the ImPACT for the multidisciplinary approach to concussion management. The authors state that when managing concussed athletes, the ImPACT should not be used as a stand-alone measure.

Hang et al. (2015) determined if computerized neurocognitive testing (Immediate Post-Concussion Assessment and Cognitive Testing [ImpACT]) in the emergency department (ED) can be used as a prognostic tool to detect young athletes at risk of having protracted concussive symptoms. This was a prospective cohort study of athletes aged 11 to

Neuropsychological Testing Under the Medical Benefit
UnitedHealthcare Community Plan Medical Policy
Proprietary Information of UnitedHealthcare. Copyright 2017 United HealthCare Services, Inc.
18 years who presented to an ED less than 24 hours after sustaining a sports-related concussion. ImPACT was administered in the ED, and performance was categorized as "poor" if the athlete had 3 (of 4) or greater low domain scores. Participants completed the Post-Concussion Symptom Scale (PCSS) in the ED and by phone at 1 and 2 weeks after injury. Athletes were symptomatic if their PCSS score was more than 6 in males and more than 8 in females. One hundred nine patients were enrolled; 60% and 36% remained symptomatic at 1 and 2 weeks after injury, respectively. "Poor" ImPACT performance was not particularly useful in predicting athletes with protracted symptoms. In bivariate analysis, a higher ED PCSS score was associated with protracted symptoms. The authors concluded that computerized neurocognitive testing in the ED has limited usefulness in predicting protracted symptoms. Total acute symptom burden may be a useful prognostic tool in the ED evaluation of concussed young athletes, yet further research is necessary.

Nelson et al. (2017) evaluated the reliability and validity of three computerized neurocognitive assessment tools (Automated Neuropsychological Assessment Metrics [ANAM], Defense Automated Neurobehavioral Assessment [DANA], and Immediate Post-Concussion Assessment and Cognitive Testing [ImPACT]) for assessing mild traumatic brain injury (mTBI). The study included mTBI (n=94) and matched trauma control (n=80) subjects recruited from a level I trauma center emergency department (ED) completed symptom and neurocognitive assessments within 72 hours of injury and at 15 and 45 days post-injury. Concussion symptoms were also assessed via phone at 8 days post-injury. Computerized neurocognitive assessment tools (CNTs) did not differentiate between groups at any time point. Roughly a quarter of stability coefficients were over .70 across measures and test-retest intervals in controls. The authors concluded that the CNTs evaluated, developed and widely used to assess sport-related concussion, did not yield significant differences between patients with mTBI versus other injuries. Symptom scores better differentiated groups than CNTs, with effect sizes weaker than those reported in sport-related concussion studies. According to the authors, nonspecific injury factors, and other characteristics common in ED settings, likely affect CNT performance across trauma patients as a whole and thereby diminish the validity of CNTs for assessing mTBI in this patient population. The authors indicated that this investigation had several limitations. First, subjects were evaluated in a laboratory setting within 72 hr of injury; thus, it is possible that stronger group differences in clinical assessment measures would have been found had subjects been assessed more acutely (such as within the ED). Second, the study design (i.e., assignment of two of three CNTs to each subject) and presence of loss to follow-up (16% at 45 days post-injury) contributed to smaller sample sizes (<50) for some CNT measures and at some time points.

Cole et al. (2017) investigated the validity of four computerized neurocognitive assessment tools (NCATs): the ANAM4, CNS-VS, CogState, and ImPACT. Two NCATs were randomly assigned and a battery of traditional neuropsychological (NP) tests administered to healthy control active duty service members (n=272) and to service members within 7 days of an mTBI (n=231). Analyses included correlations between NCAT and the NP test scores to investigate convergent and discriminant validity, and regression analyses to identify the unique variance in NCAT and NP scores attributed to group status. Effect sizes (Cohen's f2) were calculated to guide interpretation of data. Only 37 (0.6%) of the 5,655 correlations calculated between NCATs and NP tests are large. The majority of correlations are small, with no clear patterns suggestive of convergent or discriminant validity between the NCATs and NP tests. Though there are statistically significant group differences across most NCAT and NP test scores, the unique variance accounted for by group status is minimal (i.e. semipartial R2 ≤ 0.033, 0.024, 0.062, and 0.011 for ANAM4, CNS-VS, CogState, and ImPACT, respectively), with effect sizes indicating small to no meaningful effect. The authors concluded that though the results are not overly promising for the validity of the four NCATs investigated, traditional methods of investigating psychometric properties may not be appropriate for computerized tests.

In a cross-sectional study, Resch et al. (2013) evaluated test-retest reliability for the computerized neuropsychological test battery (ImPACT) using 2 different clinically relevant time intervals. Study participants included Group 1 (n=46) which consisted of 25 men and 21 women (age = 22.4 ± 1.89 years) and Group 2 (n=45) which consisted of 17 men and 28 women (age = 20.9 ± 1.72 years). Both groups completed ImPACT forms 1, 2, and 3, which were delivered sequentially either at 1-week intervals (group 1) or at baseline, day 45, and day 50 (group 2). Group 2 also completed the Green Word Memory Test (WMT) as a measure of effort. Intraclass correlation coefficients (ICCs) were calculated for the composite scores of ImPACT between time points. Repeated-measures analysis of variance was used to evaluate changes in ImPACT and WMT results over time. The authors found variable test-retest reliability for ImPACT metrics. Visual motor speed and reaction time demonstrated greater reliability than verbal and visual memory. According to the authors, current data support a multifaceted approach to concussion assessment using clinical examinations, symptom reports, cognitive testing, and balance assessment.

Elbin et al. (2011) investigated the 1-year test-retest reliability of the ImPACT online version in a sample of high school athletes. A total of 369 varsity high school athletes completed 2 mandatory preseason baseline cognitive assessments approximately 1 year apart as required by their respective athletics program. No diagnosed concussion occurred between assessments. Intraclass correlation coefficients (ICCs) for ImPACT online indicated that motor processing speed (.85) was the most stable composite score, followed by reaction time (.76), visual memory (.70), and verbal memory (.62). Unbiased estimates of reliability were consistent with ICCs: motor processing speed (.85), reaction time (.76), visual memory (.71), and verbal memory (.62). The authors concluded that the online ImPACT
baseline is a stable measure of neurocognitive performance across a 1-year time period for high school athletes. This was an uncontrolled case series and this limits the validity of the study.

Maerlender et al. (2010) compared scores on the ImPACT battery to a comprehensive battery of traditional neuropsychological measures and several experimental measures used in the assessment of sports-related concussion in 54 healthy male athletes. Convergent validity was demonstrated for four of the five ImPACT domain scores. Two cognitive domains often compromised as a result of mild TBI were not directly identified by the ImPACT battery: sustained attention and auditory working memory. Affective symptoms correlated with performance on measures of attention and working memory. In this healthy sample the correlations between the domains covered by ImPACT and the neuropsychological battery supports ImPACT as a useful screening tool for assessing many of the cognitive factors related to mild traumatic brain injury. However, the narrow construct structure of ImPACT would limit interpretation, particularly with regard to the important functions of working memory and response accuracy. This may make ImPACT testing difficult to interpret for the untrained professional. According to the investigators, the study suggests that other sources of data such as a traditional neuropsychological testing including verbal memory, visual memory, and working memory need to be considered when identifying and managing concussions.

In a consensus statement, the 5th International Conference on Concussion in Sport states that the use of neuropsychological testing (NP) contributes significant information in concussion assessment. Brief computerized cognitive evaluation tools are a commonly utilized component of these assessments worldwide given the logistical limitation in accessing trained neuropsychologists. However, it should be noted that these are not substitutes for complete NP assessment. For children, it is recommended that age-specific validated symptom-rating scales be used in sport-related concussion (SRC) assessment, and further research is required to establish the role and utility of computerized NP testing in this age group. The consensus statement suggests that baseline testing may be useful, but is not necessary for interpreting post-injury scores (McCroryet al., 2017).

**Baseline Neuropsychological Testing for Concussion**

Nelson et al. (2015) examined the rates and predictors of invalid baseline performance for 3 computerized neurocognitive tests (CNTs): Automated Neuropsychological Assessment Metrics (ANAM), Axon Sports, and Immediate Post-Concussion and Cognitive Testing (ImPACT). High school and collegiate athletes (N=2063) completed 2 of 3 CNTs each during preseason evaluations. All possible pairings were present across the sample, and the order of administration was randomized. Examiners provided 1-on-1, scripted pretest instructions, emphasizing the importance of good effort. Profile validity was determined by the manufacturers' standard criteria. The overall percentage of tests flagged as of questionable validity was lowest for ImPACT (2.7%) and higher for ANAM and Axon (10.7% and 11.3%, respectively). The majority of invalid baseline profiles were flagged as such because of failure on only 1 validity criterion. Several athlete and testing factors (eg, attention deficit hyperactivity disorder [ADHD], estimated general intellectual ability, administration order) predicted validity status for 1 or more CNTs. Considering only first CNT administrations and participants without ADHD and/or a learning disability (n=1835) brought the rates of invalid baseline performances to 2.1%, 8.8%, and 7.0% for ImPACT, ANAM, and Axon, respectively. Invalid profiles on the Medical Symptom Validity Test (MSVT) were rare (1.8% of participants) and demonstrated poor correspondence to CNT validity outcomes. The investigators concluded that the validity criteria for these CNTs may not identify the same causes of invalidity or be equally sensitive to effort. According to the investigators, the validity indicators may not be equally appropriate for some athletes (e.g., those with neurodevelopmental disorders).

Haran et al. (2016) examined differences between the baseline-referenced and norm-referenced approaches for determining decrements in Automated Neuropsychological Assessment Metrics Version 4 TBI–MIL (ANAM) performance following mild traumatic brain injury (mTBI). ANAM data were reviewed for 616 US Service members, with 528 of this sample having experienced an mTBI and 88 were controls. Post-injury change scores were calculated for each sub-test: (1) normative change score = in-theater score - normative mean and (2) baseline change score = in-theater score - pre-deployment baseline. Reliable change cut-scores were applied to the change and the resulting frequency distributions were compared using McNemar tests. Receiver operator curves (ROC) using both samples (i.e., mTBI and control) were calculated for the change scores for each approach to determine the discriminate ability of the ANAM. There were no statistical differences between the approaches. When the area under the curve for the ROCs were averaged across sub-tests, there were no significant differences between either the norm-referenced (0.65) or baseline-referenced (0.66) approaches. The investigators concluded that overall, the findings suggest there is no clear advantage of using the baseline-referenced approach over norm-referenced approach.

MacDonald and Duerson (2015) examined the test-retest reliability of a computerized neurocognitive test used for baseline assessments in high school athletes over 1 year. Study participants included high school athletes (N=117) participating in American football or soccer. All study participants completed 2 baseline computerized neurocognitive tests taken 1 year apart at their respective schools. The test measures performance on 4 cognitive tasks: identification speed (Attention), detection speed (Processing Speed), one card learning accuracy (Learning), and one back speed (Working Memory). Reliability was assessed by measuring the intraclass correlation coefficient (ICC) between the repeated measures of the 4 cognitive tasks. Pearson and Spearman correlation coefficients were
calculated as a secondary outcome measure. The measure for identification speed performed best and the measure for one card learning accuracy performed worst. All tests had marginal or low reliability. The authors concluded that in a population of high school athletes, computerized neurocognitive testing performed in a community setting demonstrated low to marginal test-retest reliability on baseline assessments 1 year apart. The authors stated that further investigation should focus on (1) improving the reliability of individual tasks tested, (2) controlling for external factors that might affect test performance, and (3) identifying the ideal time interval to repeat baseline testing in high school athletes. According to the authors, this study adds to the evidence that suggests in this population baseline testing may lack sufficient reliability to support clinical decision making.

In a study conducted by Schmidt et al. (2012), 1,060 collegiate student-athletes completed baseline testing as part of an ongoing clinical program. Gender-specific normative means were obtained from a subset of 673 athletes with no history of self-reported concussion, learning disabilities, or attention deficit disorders. Concussions were later diagnosed in 258 athletes who had completed baseline testing. Athletes completed a computerized neurocognitive test (Automated Neuropsychological Assessment Metrics), postural control assessment, and a 15-item graded symptom checklist at baseline and again following injury. Two post-concussion difference scores were computed for each outcome measure: (1) Baseline comparison = post-concussion score - individualized baseline score; and (2) Normative comparison = post-concussion score - normative mean. Athletes were considered impaired if post-concussion difference exceeded the reliable change parameters. The baseline comparison method identified 2.6 times more impairments than the normative comparison method for Simple Reaction Time-Test 1. The normative comparison method identified 7.6 times more impairments than the baseline comparison method for Mathemati Processing. No other disagreements were observed for postural control or symptom severity. The authors concluded that when using these concussion assessment tools, clinicians may consider using normative data in lieu of individualized baseline measures. This may be especially useful to clinicians with limited resources and an inability to capture baselines on all athletes.

Randolph (2011) reviewed the risks associated with sport-related concussion, and the clinical validity and reliability data for the most commonly used baseline test, the ImPACT program. The authors found no published prospective controlled study of the current version of ImPACT that would allow a determination to be made as to whether ImPACT is capable of detecting impairment in a significant percentage of athletes once they are symptom free. According to the authors, the bulk of the evidence suggests that ImPACT is not particularly sensitive to the effects of concussion, particularly once subjective symptoms have resolved. The poor sensitivity and low reliability of this test is associated with a high false negative rate (i.e., classifying a player's neurocognitive status is normal, when in fact, it is not). The use of baseline neuropsychological testing, therefore, is not likely to diminish risk. The clinical utility of baseline testing in clinical decision-making was not addressed.

Computerized Cognitive Testing such as Mindstreams and BrainCare

Shopin et al. (2013) compared a computerized battery of neuropsychological tests for memory, attention and executive functions (MindStreams) with the Montreal Cognitive Assessment (MoCA) to detect mild-to-moderate cognitive impairments in poststroke patients. A total of 454 patients with transient ischemic attack (TIA) or stroke enrolled to the TABASCO (Tel Aviv Brain Acute Stroke Cohort) study, a prospective study which includes consecutive first-ever mild-to-moderate stroke patients, were included. All participants underwent neurological and cognitive evaluations. The patients' mean MoCA and MindStreams scores were lower than normal; however, the TIA group presented significantly better scores using either method. The correlation between the MoCA and the computerized global score was 0.6. A significant correlation was found between the subcategory scores (executive function, memory and attention). However, the MoCA identified many more subjects with low scores (<26) compared to the MindStreams (70.6 vs. 15.7%).

Punchik et al. (2015) retrospectively analyzed patients referred to a comprehensive geriatric assessment unit with screening instruments to determine if further comprehensive cognitive assessment was necessary. Cognitive screening and assessment included visual-spatial components: the Mini Mental State Examination, the Clock Drawing Test, the Montreal Cognitive Assessment Test, and the Neurotrax (Mindstreams) computerized cognitive assessment battery. The average age of the 190 eligible patients was 81.09±5.42 years. Comparing the individual tests with that of the visual-spatial index of Neurotrax, the investigators found the Trail Making B test to be most sensitive (72.4%) and the Cube Test to have the highest specificity (72.8%). A combination of tests resulted in higher sensitivity and lower specificity. The authors concluded that the use of a combination of visual-spatial tests for screening in neurocognitive disorders should be evaluated in further prospective studies.

Dwolatzky et al. (2010) examined the validity of the Mindstreams battery designed specifically for evaluating those with moderate cognitive impairment. One hundred and seventy participants over the age of 60 years performed the computerized battery in addition to standard clinical evaluation. Staging was according to the Clinical Dementia Rating Scale (CDR) on the basis of clinical data but independent of computerized cognitive testing results, thus serving as the gold standard for evaluating the discriminant validity of the computerized measures. Seven participants received a global Clinical Dementia Rating (CDR) score of 0 (not impaired), 76 were staged as CDR 0.5 (very mildly impaired),
58 as CDR 1 (mildly impaired), 26 as CDR 2 (moderately impaired), and 3 as CDR 3 (severely impaired). Mindstreams Global Score performance was significantly different across CDR groups, reflecting poorer overall battery performance for those with greater impairment. This was also true for the domain summary scores, with Executive Function and Memory distinguishing best between CDR 0.5 and 1, and Orientation best differentiating among CDR 1 and 2. The investigators concluded that the Mindstreams battery for moderate impairment differentiates among varying degrees of cognitive impairment in older adults, providing detailed and distinct cognitive profiles. Limitations of this study include lack of a control group and small sample size.

Achiron et al. (2007) compared the Mindstreams test battery with the Neuropsychological Screening Battery for Multiple Sclerosis (NSBMS), which is considered the reference standard for cognitive screening in MS, in patients with MS (n=58) and in a control group of healthy volunteers (n=71) who were matched for age, education, gender, handedness, and computer use. The 71 controls were randomly selected from 410 individuals who were used to establish normative values for the Mindstreams system. Five of the 7 index scores (memory, executive function, attention, information processing, and motor skills) significantly discriminated MS patients from controls, while visuospatial and verbal-function indexes did not. However, the NSBMS system was not assessed in a similar manner; only correlation coefficients of the Mindstreams index scores and NSBMS system outcomes were presented. As with the study by Ritsner et al. (2006), all of the correlations were statistically significant, but the magnitude of the correlation coefficients indicates only moderate correlation at best. This study, therefore, demonstrates the capability of the Mindstreams system to differentiate MS patients from healthy volunteers across 5 of 7 cognitive domains, but the data are insufficient to establish the equivalence of the Mindstreams system to the standard of care or to demonstrate a benefit of Mindstreams assessment on clinical outcomes.

Racine et al. (2016) conducted a study that included 469 late middle-aged participants from the Wisconsin Registry for Alzheimer’s Prevention (mean age 63.8±7 years at testing; 67% female; 39% APOE4+) to evaluate whether computerized cognitive assessments, like the CogState battery, are sensitive to preclinical cognitive changes or pathology in people at risk for Alzheimer’s disease (AD). The study examined relationships between a CogState abbreviated battery (CAB) of seven tests and demographic characteristics, traditional paper-based neuropsychological tests as well as a composite cognitive impairment index, cognitive impairment status (determined by consensus review), and biomarkers for amyloid and tau (CSF phosphorylated-tau/Aβ42 and global PET-PiB burden) and neural injury (CSF neurofilament light protein). CSF and PET-PiB were collected in n = 71 and n = 91 participants, respectively, approximately four years prior to CAB testing. For comparison, three traditional tests of delayed memory in parallel were examined. Similar to studies in older samples, the CAB was less influenced by demographic factors than traditional tests. CAB tests were generally correlated with most paper-based cognitive tests examined and mapped onto the same cognitive domains. Greater composite cognitive impairment index was associated with worse performance on all CAB tests. Cognitively impaired participants performed significantly worse compared to normal controls on all but one CAB test. Poorer One Card Learning test performance was associated with higher levels of CSF phosphorylated-tau/Aβ42. The authors concluded that these results support the use of the CogState battery as measures of early cognitive impairment in studies of people at risk for Alzheimer’s disease. However, according to the authors, the study also suggests that CogState at a single time point may not substantially improve preclinical AD detection over traditional neuropsychological tests.

Overall, the available evidence is insufficient to establish the validity of computerized cognitive testing such as Mindstreams and BrainCare compared with traditional tests for the assessment of cognitive impairment.

**Intermittent Explosive Disorder**

There are no clear underlying medical issues associated with intermittent explosive disorder, nor are there published clinical trials that support the use of neuropsychological testing for this disorder. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), published by the American Psychiatric Association, the following criteria must be met in order for a patient to be diagnosed with intermittent explosive disorder:

- Recurrent behavioral outbursts that represent a failure to control aggressive impulses as manifested by one of the following:
  - Verbal aggression (e.g., temper tantrums, tirades, verbal arguments or fights) or physical aggression towards property, animals, or other individuals, occurring, on average, twice weekly for a period of three months. The physical aggression does not result in damage or destruction of property and does not result in physical injury to animals or other individuals.
  - Three behavioral outbursts involving damage or destruction of property and/or physical assault with physical injury against animals or other individuals occurring within a 12-month period.
- The magnitude of aggressiveness expressed during the recurrent outbursts is grossly out of proportion to the provocation or any precipitating psychosocial stressors.
- The recurrent aggressive outbursts are not premeditated (i.e., are impulsive) and are not committed to achieve some tangible objective (e.g., money, power, intimidation).
- The recurrent aggressive outbursts cause either marked distress in the individual or impairment in occupational or interpersonal functioning, or are associated with financial or legal consequences.


- Chronological age is at least 6 years (or equivalent developmental level).
- The recurrent aggressive outbursts are not better explained by another mental disorder and are not attributable to another medical condition or to physiological effects of a substance.

**Headaches Including Migraine**

Literature addressing the neuropsychological consequences of migraine headaches is not conclusive. Studies on the relationship between migraine and cognitive functioning have demonstrated conflicting results. Some studies show a detrimental effect of migraine on cognitive skills and other studies have shown no difference in cognitive skills for patients with migraine headaches (Foti et al., 2017).

Dresler et al. (2012) evaluated three neuropsychological tests (Trail Making Test (TMT), Go/Nogo Task and Stroop Task) that were completed by four headache patient samples (chronic CH, episodic CH in the active or inactive period, and migraine patients) and compared to healthy controls. Analyses revealed that patients with chronic and active episodic CH appeared particularly impaired in tests relying more on intact executive functioning (EF) than on basal cognitive processes. Within the CH groups performance decreased linearly with increasing severity. The authors stated that impaired EF could also result from medication and sleep disturbances due to active CH. The authors went on to say that because decreased performance was also present outside the attacks it may hint at generally altered brain function, but does not necessarily reflect clinically relevant behavior.

There is insufficient clinical evidence to conclude that the use of neuropsychological testing for patients with migraine headaches without associated cognitive disorders can be used effectively for clinical decision making to improve management of this condition. No published clinical trials were found that support the use of neuropsychological testing for clinical decision making to improve management for patients with other types of headaches who did not have associated cognitive disorders.

**History of Myocardial Infarction**

Literature addressing the neuropsychological consequences of myocardial infarction is not conclusive. Studies on the relationship between myocardial infarction and cognitive functioning have demonstrated conflicting results. Some studies show a detrimental effect of myocardial infarction on cognitive skills (Sauve et al., 2009; Almeida et al., 2008). Other studies have shown no difference in cognitive skills for patients with myocardial infarctions (Ahto et al., 1999, Grubb et al., 2000).

In a systematic review, Cameron et al. (2016) evaluated the diagnostic accuracy of cognitive screening instruments in screening for mild cognitive impairment (MCI) in heart failure (HF) patients. Inclusion criteria for the review were as follows: primary studies examining cognitive impairment in HF, administration of a cognitive screening instrument and neuropsychological test battery, and cognitive impairment indicated by performance on neuropsychological tests 1.5 SDs less than that of normative data. The precision, accuracy, and receiver operating characteristic curves of the Mini Mental State Examination were computed. From 593 citations identified, 8 publications met inclusion criteria. Risk of bias included selective HF patient samples, and no study examined the diagnostic test accuracy of the cognitive screening instruments. The Mini Mental State Examination had low sensitivity (26%) and high specificity (95%) with a score of 28 or less as the optimal threshold for MCI screening. The authors concluded that screening for cognitive impairment in HF is recommended; however, future studies need to establish the diagnostic accuracy of screening instruments of MCI in this population.

Neuropsychological data were gathered from 46 healthy controls, 42 cardiac patients referred for percutaneous coronary intervention (PCI), and 43 cardiac patients referred for coronary artery bypass grafting (CABG). Fourteen cognitive function tests were utilized at baseline and at three time points after surgery (3 weeks, 4 months, 1 year). No clear pattern of group differences or change at follow-up emerged. A greater percentage of CABG patients than controls worsened in seven tests (three at 1 year), but a greater percentage of PCI patients than controls also worsened in seven tests (three at 1 year). Generalized estimating equations showed only two tests (Wechsler Adult Intelligence Scale, Third Edition, Digit Symbol, and Hopkins Verbal Learning Test, Revised, Total Recall) to be significantly different between groups from baseline to 1 year. Compared with healthy controls, more PCI patients than CABG patients worsened in the former of those two tests, whereas more PCI and CABG patients improved on the latter. The investigators concluded that current CABG procedure does not appear to create cognitive decline (Sweet, 2008).

There is insufficient clinical evidence to conclude that the use of neuropsychological testing for patients with myocardial infarction without associated cognitive disorders can be used effectively for clinical decision making to improve management of this condition.
In an evidence-based guideline update for the evaluation and management of concussion in sports, the AAN states that it is likely that neuropsychological testing of memory performance, reaction time, and speed of cognitive processing, regardless of whether administered by paper-and-pencil or computerized method, is useful in identifying the presence of concussion (sensitivity 71%-88% of athletes with concussion). This is based on evidence from 1 Class II study and multiple Class III studies. The AAN also states that both types of testing (paper-and-pencil or computerized) generally require a neuropsychologist for accurate interpretation, although the tests may be administered by a non-neuropsychologist. According to AAN, there is insufficient evidence to support conclusions about the use of neuropsychological testing in identifying concussion in preadolescent age groups. The AAN goes on to say that inexperienced licensed health care providers (LHCPs) should be instructed in the proper administration of standardized validated sideline assessment tools. This instruction should emphasize that these tools are only an adjunct to the evaluation of the athlete with suspected concussion and cannot be used alone to diagnose concussion (Level B - probably effective). The AAN further states that LHCPs caring for athletes might utilize individual baseline scores on concussion assessment tools, especially in younger athletes, those with prior concussions, or those with preexisting learning disabilities/attention deficit/hyperactivity disorder, as doing so fosters better interpretation of postinjury scores (Level C - Possibly effective) (Giza et al., 2013).

In a position statement on sports concussion, the AAN made the following recommendations (AAN, 2010, updated 2013):

- Any athlete who is suspected to have suffered a concussion, regardless of severity, is to be removed immediately from participation in a game or practice.
- A licensed health care professional, such as a neurologist, whose scope of practice includes proper training in the evaluation and management of concussion, must clear the youth athlete before he or she can return to play. This includes sports recognized by high school athletic associations as well as youth and recreational leagues run by other entities.

The AAN published a report regarding neuropsychological testing of adults. This report indicates that neuropsychological testing is most useful for management planning in patients with suspected dementia, multiple sclerosis, Parkinson's disease, traumatic brain injury, stroke, and HIV encephalopathy. It is also useful for detecting deficits in patients with particularly high premorbid intelligence levels in which bedside-type clinical testing may be insensitive to mild alterations. Neuropsychological testing also has an important role in evaluating patients undergoing epilepsy surgery (Assessment: neuropsychological testing of adults. Considerations for neurologists. Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. Neurology, 1996).

The Quality Standards Subcommittee of the AAN published an evidence-based review for the early detection of dementia. ( Petersen et al., 2001) The recommendations state that neuropsychologic batteries are useful to identify patients with dementia, particularly when administered to an increased-risk population (i.e., those with cognitive impairment).

A practice parameter for the screening and diagnosis of autism developed by the American Academy of Neurology and the Child Neurology Society indicates that neuropsychological, behavioral, and academic assessments should be performed as needed, in addition to the cognitive assessment, to include social skills and relationships, educational functioning, problematic behaviors, learning style, motivation and reinforcement, sensory functioning, and self-regulation for the diagnosis of autism (Filipek et al., 2000).

In a practice parameter update on the evaluation and management of driving risk in dementia, the AAN states that there is insufficient evidence to recommend neuropsychological testing to predict driving capability among patients with dementia (Iverson et al., 2010).

In a practice parameter update on the care of the patient with amyotrophic lateral sclerosis (ALS), the AAN states that the domain of cognitive and behavioral impairment in ALS is a rapidly evolving field and there is little consensus regarding diagnostic criteria and assessment methods. Screening tests of executive function may be considered to detect cognitive impairment in patients with ALS prior to confirmation with formal neuropsychological evaluation (Level C) (Miller et al., 2009).

In a guideline for Healthcare Professionals from the American Heart Association and the American Stroke Association, Winstead et al. (2016) provided a synopsis of best clinical practices in the rehabilitative care of adults recovering from stroke. According to the guideline, a formal neuropsychological examination (including assessment of language, neglect, praxis, memory, emotional responses, and specific cognitive syndromes) may be helpful after the detection of cognitive impairment with a screening instrument. Neuropsychological protocols must be sensitive to a wide range of abilities, especially the assessment of executive and attentional functions. These guidelines state that screening for
cognitive deficits is recommended for all stroke patients before discharge home (class I, level B evidence). The guidelines also indicate that when screening reveals cognitive deficits, a more detailed neuropsychological evaluation to identify areas of cognitive strength and weakness may be beneficial (class IIa, level C evidence).

**American Psychological Association (APA)**
The American Psychological Association published updated guidelines for the evaluation of dementia and age-related cognitive change. The guidelines include the following information regarding neuropsychological testing for this condition (American Psychological Association, 2012):

- Neuropsychological evaluation and cognitive testing remain among the most effective differential diagnostic methods in discriminating pathophysiological dementia from age-related cognitive decline, cognitive difficulties that are depression-related, and other related disorders. Even after reliable biological markers have been discovered, neuropsychological evaluation and cognitive testing will still be necessary to determine the onset of dementia, the functional expression of the disease process, the rate of decline, the functional capacities of the individual, and hopefully, response to therapies.
- Comprehensive neuropsychological evaluations for dementia and cognitive change include tests of multiple cognitive domains, typically including memory, attention, perceptual and motor skills, language, visuospatial abilities, reasoning, and executive functions. Measures of mood and personality may be relevant in many cases. Psychologists are encouraged to refer to current compendia resources and the clinical research literature in selecting assessment instruments. Psychologists are encouraged to use standardized, reliable, and valid tests.
- Technology assisted assessments (e.g., computer administered cognitive batteries, tele-health visits) are rapidly advancing but appropriate psychometric properties and normative data are nascent. These technologies may have significant advantages for older persons with limited mobility or health-care access, but may also disadvantage older persons with limited experience and expertise interacting with technology.

**American Psychiatric Association**
In its guidelines on the treatment of Alzheimer’s disease and other dementias, the American Psychiatric Association states that neuropsychological testing may be helpful in deciding whether a patient with subtle or atypical symptoms actually has dementia. Neuropsychological testing is particularly useful in the evaluation of individuals who present with mild cognitive impairment, which requires evidence of memory and/or other cognitive difficulties in the presence of intact functioning, and in the evaluation of individuals with the onset of dementia early in life. Testing may help to characterize the extent of cognitive impairment, to distinguish among the types of dementias, and to establish baseline cognitive function. The American Psychiatric Association also states that a variety of research definitions for mild cognitive impairment are in place, but there is no consensus on the optimal definition. The most widely accepted definition requires the following (American Psychiatric Association, 2007):

- Subjective cognitive complaints,
- Evidence of objective deficits in cognitive function based on age- and education-adjusted norms on standardized neuropsychological tests,
- Intact daily functioning,
- Evidence of cognitive decline from a prior level, and
- Evidence of not meeting the criteria for dementia.

**American Academy of Pediatrics (AAP)**
A joint statement for learning disabilities, dyslexia, and vision from the American Academy of Pediatrics, Section on Ophthalmology, Council on Children with Disabilities; American Academy of Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; and the American Association of Certified Orthoptists states that children who exhibit signs of learning disabilities should be referred for educational, psychological, neuropsychological, and/or medical diagnostic assessments (AAP, 2009).

In a policy statement on sport-related concussion in children and adolescents, the AAP states that neuropsychological testing can be helpful to provide objective data to athletes and their families after a concussion. Neuropsychological testing is one tool in the complete management of a sport-related concussion and alone does not make a diagnosis or determine when return to play is appropriate. According to the AAP, testing is performed by using one of several computerized neuropsychological tests including ANAM (Automated Neuropsychological Assessment Metrics), CogState, HeadMinder, and ImPACT or through pencil-and-paper testing administered by a neuropsychologist. Each of the computerized tests has published data on test-retest reliability, and all have demonstrated deficits in concussed athletes compared with their baseline assessments. One critique of the computerized tests is that the vast majority of studies have been conducted by the developers of the tests, which raises some concern for bias, because some independent study results have suggested slightly less reliable results. More rigorous pencil-and-paper testing conducted formally by a neuropsychologist is also an option, although test-retest reliability has been questioned. If an athlete is suffering from postconcussive symptoms over several months or has had multiple concussions, formal assessment by a neuropsychologist may be beneficial, specifically to identify areas for which the athlete may need academic accommodations (Halstead and Walter, 2010; Reaffirmed August 2014).
American Academy of Child and Adolescent Psychiatry (AACAP)
Practice parameters from the American Academy of Child and Adolescent Psychiatry (Volkmar et al. 2014) state that neuropsychological correlates of autism spectrum disorder include impairments in executive functioning (e.g., simultaneously engaging in multiple tasks) (Ozonoff et al. 1991), weak central coherence (integrating information into meaningful wholes) (Happe and Frith 2006), and deficits in theory-of-mind tasks (taking the perspective of another person) (Baron-Cohen et al. 1985).

International Society for Hepatic Encephalopathy and Nitrogen Metabolism (ISHEN)
A guideline prepared by the Commission on Neuropsychological Assessment of Hepatic Encephalopathy appointed by the ISHEN states that neuropsychological testing is an established methodology for quantifying cognitive impairment due to various forms of encephalopathy, including low-grade or minimal hepatic encephalopathy (Randolph, 2009).

National Academy of Neuropsychology (NAN)
The National Academy of Neuropsychology developed an education paper to provide information to clinicians, healthcare administrators, and policy developers about the purpose, strengths, and limitations of computerized cognitive screening tests versus comprehensive neuropsychological evaluations. Screening tests are generally brief and narrow in scope, they can be administered during a routine clinical visit, and they can be helpful for identifying individuals in need of more comprehensive assessment. Some screening tests can also be helpful for monitoring treatment outcomes. Comprehensive neuropsychological assessments are multidimensional in nature and used for purposes such as identifying primary and secondary diagnoses, determining the nature and severity of a person’s cognitive difficulties, determining functional limitations, and planning treatment and rehabilitation. Cognitive screening tests are expected to play an increasingly important role in identifying individuals with cognitive impairment and in determining which individuals should be referred for further neuropsychological assessment. However, limitations of existing cognitive screening tests are present and cognitive screening tests should not be used as a replacement for comprehensive neuropsychological testing (Roebuck-Spencer et al, 2017).

In a policy for the evaluation of childhood learning disorders, the NAN states that when comprehensive information about a child’s brain-related strengths and weaknesses is necessary to understand potential sources of the problem and implications for functioning, a neuropsychological evaluation is most often the best choice (Silver et al., 2006).

In a position paper on the diagnosis and management of sports-related concussion, the NAN states that neuropsychological evaluation is recommended for the diagnosis, treatment, and management of sports-related concussion at all levels of play (Moser et al., 2007).

American Academy of Clinical Neuropsychology (AACN) and National Academy of Neuropsychology (NAN)
A joint position paper of the AACN and NAN sets forth their position on appropriate standards and conventions for computerized neuropsychological assessment devices (CNADs). The authors state that CNADs are subject to, and should meet, the same standards for the development and use of educational, psychological, and neuropsychological tests (American Psychological Association, 1999) as are applied to examiner-administered tests. The authors also state that those employing CNADs have the education, training, and experience necessary to interpret their results in a manner that will best meet the needs of the patients they serve (Bauer et al., 2012).

American Medical Society for Sports Medicine
In a position statement for concussion in sport the American Medical Society for Sports Medicine provided an evidence-based, best practises summary to assist physicians with the evaluation and management of sports concussion (Harmon et al. 2013). The following statements were made regarding neuropsychological (NP) testing:

- NP tests are an objective measure of brain-behavior relationships and are more sensitive for subtle cognitive impairment than clinical exam. Most concussions can be managed appropriately without the use of NP testing.
- Computerized neuropsychological (CNP) testing should be interpreted by healthcare professionals trained and familiar with the type of test and the individual test limitations, including a knowledgeable assessment of the reliable change index, baseline variability and false-positive and false-negative rates.
- Paper and pencil NP tests can be more comprehensive, test different domains and assess for other conditions which may masquerade as or complicate assessment of concussion.
- NP testing should be used only as part of a comprehensive concussion management strategy and should not be used in isolation.
- The ideal timing, frequency and type of NP testing have not been determined. In some cases, properly administered and interpreted NP testing provides an added value to assess cognitive function and recovery in the management of sports concussions.
- It is unknown if use of NP testing in the management of sports concussion helps prevent recurrent concussion, catastrophic injury or long-term complications.
- Comprehensive NP evaluation is helpful in the post-concussion management of athletes with persistent symptoms or complicated courses.
NeuroTrax Corp., the manufacturer of Mindstreams, submitted a 510k premarket notification to the FDA on March 9, 2007, however, the FDA determined that the Mindstreams device was not substantially equivalent to other devices and declined the application in a letter dated December 21, 2007, which explained that the FDA considered the device to be Class III device. NeuroTrax Corp. appealed this decision on April 10, 2008, but the FDA maintained its earlier decision and notified the company of such on December 4, 2008. On March 15, 2012, the FDA sent a warning letter to NeuroTrax Corp. stating that the Mindstreams device was being marketed in the United States without FDA marketing clearance or approval. NeuroTrax Corp. was required to cease marketing in the United States within 15 days of receipt of the warning letter. See the following website for more information: http://www.fda.gov/ICECI/EnforcementActions/WarningLetters/2012/ucm296778.htm. (Accessed April 23, 2017)

In June 2015, the FDA cleared Cognivue through the de novo classification pathway. The de novo pathway is used for low- to moderate-risk medical devices that are not equivalent to an already legally marketed device. FDA identifies Cognivue as a “Computerized Cognitive Assessment Aid.” According to the FDA, this test is indicated as an adjunctive tool for evaluating perceptual and memory function in individuals aged 55 to 95 years old. See the following website for more information: https://www.accessdata.fda.gov/cdrh_docs/pdf13/DEN130033.pdf. (Accessed May 24, 2107)

On August 22, 2016, the FDA began to allow the marketing of two computerized neurocognitive tests for assessing individuals immediately following a suspected brain injury or concussion: ImPACT and ImPACT Pediatric (ImPACT Applications). Both tests were reviewed via the agency’s de novo classification process, a pathway to market for certain “first-of-a-kind” and low- to moderate-risk medical devices. ImPACT and ImPACT Pediatric are computerized cognitive assessment aids intended for use in conjunction with standard medical evaluation for signs and symptoms of a head injury. ImPACT is designed to assess people 12 to 59 years of age, while ImPACT Pediatric is designed for children aged 5 to 11 years. The FDA states that these tests should not be used to “rule out a concussion or determine whether an injured player should return to a game.” See the following websites for more information:
- http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm517526.htm
(Accessed April 23, 2017)

CENTERS FOR MEDICARE AND MEDICAID SERVICES (CMS)

Medicare does not have a National Coverage Determination (NCD) for neuropsychological testing. Local Coverage Determinations (LCDs) exist; see the LCDs for Outpatient Psychiatry and Psychology Services, Psychiatry and Psychology Services, Psychological and Neuropsychological Testing and Therapy and Rehabilitation Services. (Accessed May 25, 2017)

REFERENCES


**POLICY HISTORY/REVISION INFORMATION**

<table>
<thead>
<tr>
<th>Date</th>
<th>Action/Description</th>
</tr>
</thead>
</table>
| 10/01/2017   | • Revised coverage rationale:  
  o Replaced language indicating:  
    ▪ “Baseline neuropsychological testing is unproven and not medically necessary in asymptomatic persons at risk for sport-related concussions” with “baseline neuropsychological testing is unproven and not medically necessary in asymptomatic persons at risk for sport-related concussions or brain injuries”  
    ▪ “Computerized neuropsychological testing is unproven and not medically necessary when used alone for evaluating concussions” with “computerized neuropsychological testing is unproven and not medically necessary when used alone for evaluating concussions or brain injuries”  
<p>|</p>
<table>
<thead>
<tr>
<th>Date</th>
<th>Action/Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>necessary for evaluating concussions or brain injuries”</td>
</tr>
<tr>
<td></td>
<td>- “Computerized cognitive testing such as Mindstreams® Cognitive Health Assessment and BrainCare™ is investigational, unproven, and not medically necessary for diagnosing dementia or mild cognitive impairment” with “computerized cognitive testing including but not limited to Cognivue®, Mindstreams® Cognitive Health Assessment, and BrainCare™ is unproven and not medically necessary for diagnosing dementia or mild cognitive impairment”</td>
</tr>
<tr>
<td></td>
<td>o Modified language pertaining to clinical evidence/study findings for computerized neuropsychological testing to indicate:</td>
</tr>
<tr>
<td></td>
<td>- The evidence is insufficient to establish the validity and reliability of computerized tests to evaluate concussions</td>
</tr>
<tr>
<td></td>
<td>- Prospective controlled trials are needed to demonstrate the clinical utility of these tests to detect impairment following concussion</td>
</tr>
<tr>
<td></td>
<td>o Modified language pertaining to clinical evidence/study findings for computerized cognitive testing indicating; removed language indicating no U.S. Food and Drug Administration (FDA) clearance was located in the FDA database for Mindstreams® Cognitive Health Assessment or for BrainCare™</td>
</tr>
<tr>
<td></td>
<td>- Updated supporting information to reflect the most current description of services, clinical evidence, FDA information, and references</td>
</tr>
<tr>
<td></td>
<td>- Archived previous policy version CS083.D</td>
</tr>
</tbody>
</table>