Neuropsychological Testing Under the Medical Benefit

Policy Number: CS083.R
Effective Date: September 1, 2024

➲ Instructions for Use

Table of Contents **Page**

Application 1
Coverage Rationale 1
Definitions 3
Applicable Codes 3
Description of Services 3
Clinical Evidence 5
U.S. Food and Drug Administration 26
References 26
Policy History/Revision Information 33
Instructions for Use 33

Related Community Plan Policy
- Maximum Dosage and Frequency

Commercial Policy
- Neuropsychological Testing Under the Medical Benefit

Optum Clinical Guideline
- Psychological and Neuropsychological Testing

Application

This Medical Policy does not apply to the states listed below; refer to the state-specific policy/guideline, if noted:

State	Policy/Guideline
Idaho	Neuropsychological Testing Under the Medical Benefit (for Idaho Only)
Indiana	None
Kansas	Neuropsychological Testing Under the Medical Benefit (for Kansas Only)
Kentucky	Neuropsychological Testing Under the Medical Benefit (for Kentucky Only)
Louisiana	Neuropsychological Testing Under the Medical Benefit (for Louisiana Only)
Nebraska	Neuropsychological Testing Under the Medical Benefit (for Nebraska Only)
New Jersey	Neuropsychological Testing Under the Medical Benefit (for New Jersey Only)
New Mexico	Neuropsychological Testing Under the Medical Benefit (for New Mexico Only)
Ohio	Neuropsychological Testing Under the Medical Benefit (for Ohio Only)
Pennsylvania	Neuropsychological Testing Under the Medical Benefit (for Pennsylvania Only)
Tennessee	Neuropsychological Testing Under the Medical Benefit (for Tennessee Only)

Coverage Rationale

Neuropsychological testing is proven and medically necessary for evaluating individuals with the following conditions when the results of testing will be used to support a diagnosis, prognosis, or treatment plan:

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

  Note: 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
- Demyelinating disorders including multiple sclerosis
- Intellectual disability or intellectual developmental disorder when all of the following are present:
    - The intellectual disability or intellectual developmental disorder is 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)
        - Onset of intellectual and adaptive deficits during the developmental period

  Note: 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
- Neurocognitive disorders including mild cognitive impairment (MCI), dementia or symptoms of dementia such as memory impairment or memory loss (including Alzheimer's and 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
- Neurotoxin exposure with at least one of the following:
    - Demonstrated serum levels of neurotoxins
    - Individual with one or more of the following:
        - Documented prenatal alcohol, drug, or toxin exposure
        - History of radiation therapy or chemotherapy
- Seizure disorder including individuals with epilepsy
- Stroke
- Traumatic Brain Injury (TBI)
- The individual is being considered for a medical or surgical procedure that may affect brain function (e.g., epilepsy surgery, resection of brain tumors or arteriovenous malformations, deep brain stimulation, stem cell, or organ transplants)

The following are unproven and not medically necessary due to insufficient evidence of efficacy:
- Baseline neuropsychological testing in asymptomatic individuals at risk for sport-related concussions
- Computerized cognitive testing such as to Cognivue®, Mindstreams® Cognitive Health Assessment, BrainCare™, and QbTest
- Computerized neuropsychological testing when used as a stand-alone test for evaluating concussions
- Neuropsychological testing for the following diagnoses alone without other proven conditions as noted above:
    - Headaches including migraine headache
    - History of myocardial infarction
    - Intermittent explosive disorder
- Neuropsychological testing that is comprised exclusively of self-administered or self-scored inventories, or as screening tests of cognitive function or neurological disease whether paper-and-pencil or computerized (e.g., AIMS, Folstein Mini-Mental Status Examination)
- Neuropsychological testing that is used as a routine screening tool
- Neuropsychological testing that is administered for educational or vocational purposes that do not alter or direct medical or health management
- Repeat neuropsychological testing that is not required for medical decision-making

- The individual is neurologically, cognitively, or psychologically unable to participate in a meaningful way in the neuropsychological testing process
- The individual has been diagnosed previously with brain dysfunction, such as Alzheimer's disease, and there is no expectation that neuropsychological testing would impact the individual's medical, functional, or behavioral management

Definitions

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, 2021)

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 Guidelines may apply.

CPT Code	Description
96116	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]), by physician or other qualified health care professional, both face-to-face time with the patient and time interpreting test results and preparing the report; first hour
96121	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]), by physician or other qualified health care professional, both face-to-face time with the patient and time interpreting test results and preparing the report; each additional hour (List separately in addition to code for primary procedure)
96132	Neuropsychological testing evaluation services by physician or other qualified health care professional, including integration of patient data, interpretation of standardized test results and clinical data, clinical decision making, treatment planning and report, and interactive feedback to the patient, family member(s) or caregiver(s), when performed; first hour
96133	Neuropsychological testing evaluation services by physician or other qualified health care professional, including integration of patient data, interpretation of standardized test results and clinical data, clinical decision making, treatment planning and report, and interactive feedback to the patient, family member(s) or caregiver(s), when performed; each additional hour (List separately in addition to code for primary procedure)
96136	Psychological or neuropsychological test administration and scoring by physician or other qualified health care professional, two or more tests, any method; first 30 minutes
96137	Psychological or neuropsychological test administration and scoring by physician or other qualified health care professional, two or more tests, any method; each additional 30 minutes (List separately in addition to code for primary procedure)
96138	Psychological or neuropsychological test administration and scoring by technician, two or more tests, any method; first 30 minutes
96139	Psychological or neuropsychological test administration and scoring by technician, two or more tests, any method; each additional 30 minutes (List separately in addition to code for primary procedure)
96146	Psychological or neuropsychological test administration, with single automated, standardized instrument via electronic platform, with automated result only

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, and validity of particular procedures for the specific function being measured.

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, Repeatable Battery for Assessment of Neuropsychological Status (RBANS), 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 individuals various stimuli or puzzles to solve using a mouse or a keypad. The Mindstreams system automatically generates a report that details the individual'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. Individuals 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. The QbTest is an online computerized test that measures activity, attention, and impulsivity for assessment of attention-deficit hyperactivity disorder (ADHD).

Computerized neuropsychological tests have been proposed to be used as part of the overall medical 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 doctoral-level psychologist who is licensed to practice independently and demonstrates sufficient training and experience
- 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(s) adhere to the prevailing national professional and ethical standards regarding test administration, scoring, and interpretation
- A board-certified neurologist

Refer to the following Optum Supplemental Clinical Criteria for more information:

- Psychological and Neuropsychological Testing (to access this guideline, go to: Optum Provider Express > Clinical Resources > Clinical Criteria and Guidelines > Optum National Behavioral Health Clinical Criteria. Accessed May 15, 2024

Clinical Evidence

Attention Deficit Hyperactivity Disorder (ADHD)

Peterson et al. (2024) conducted a systematic review for The Agency for Healthcare Research and Quality (AHRQ) to assess the evidence on the diagnosis, treatment, and monitoring of ADHD in children and adolescents to inform a planned update of the American Academy of Pediatrics guidelines. The review included 550 studies reported in 1, 097 publications (231 studies addressed diagnosis, 312 studies addressed treatment, and 10 studies addressed monitoring). The studies that addressed diagnosis included studies that evaluated the diagnostic performance of parental ratings, teacher rating scales, teen/child self-reports, clinician tools, neuropsychological tests, EEG approaches, imaging, and biomarkers. Studies that evaluated neuropsychological tests of executive functioning used study-specific combinations of individual cognitive measures which made it more difficult to compare performance across studies. The authors stated that evaluation of specific tools was difficult as many of the studies assessed a large number of different tools within broader categories, used different components of the same tools, or combined components in a variety of ways. Of the 231 studies that addressed diagnosis, there were 74 that included neuropsychological tests, assessing executive function, and/or encompassing a variety of cognitive assessments, including continuous performance tests, to diagnose ADHD with high variability in study designs, inclusion/exclusion criteria, and reported sensitivity, specificity, and accuracy in their outcomes. The authors reported that about a third of the studies included race and ethnicity demographics with seven studies having populations with more than 50% of the participants being White, and with the proportion of girls in the study populations ranging from none to 74%. The authors reported that multiple approaches showed promising diagnostic performance; however, the strength of evidence (SoE) was generally low with few studies reporting estimates for children under the age of seven. The authors concluded that there are many diagnostic tools available to aid in the diagnosis of ADHD and that further research is needed to assess performance of diagnostic tools for ADHD in children under the age of seven years.

Bausela-Herreras et al. (2024) conduced a systematic review and meta-analysis to determine the usefulness of the Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P) and to analyze the possible ceiling and floor effect of its scores in the assessment of executive function (EF) in preschoolers (age range between two and six years) with signs compatible with a possible diagnosis of ADHD. The study included seven articles with 2311 participants under 6 years of age with ADHD-compatible symptoms (n = 1459) compared to typically developing (TD) participants (n = 791). One study also included children with oppositional defiant disorder (ODD; n = 51) and ODD + ADHD (n = 10). The floor effect was described as when most participants score very low on an evaluation, indicating that the measure used may not be sensitive enough to capture variations in performance and the ceiling effect occurs when most participants score very high, suggesting that the measure may not be able to distinguish adequately between higher levels of ability or performance. The authors reported that executive deficits in early-age individuals with symptoms compatible with ADHD are more extensive than just deficits in working memory and that a floor effect was found in tests associated with the Flexibility clinical scale and the Emotional Control index, both hot executive functions, and a ceiling effect was associated with two of the BRIEF-P indices: Global Executive Functioning and the Emergent Metacognition index, configured by the clinical scales Working Memory and Planning/Organization, which are indices associated with the cold dimensions of executive functions. The authors concluded that the BRIEF-P is an instrument that facilitated obtaining a sensitive and discriminative executive profile, although it should be used in combination with other neuropsychological performance tests. Limitations of the study included including the small sample sizes of six of the seven studies, the average age of the participants (before ADHD diagnosis) precluded considering the influence of factors related to developmental stage and maturity and the inclusion of participants with ODD.

Becke et al. (2023) administered a comprehensive neuropsychological test battery in an analogue study that included 57 adults with ADHD, and 211 university students who were divided into two groups with 60 students in the control group, and 151 students in the simulator group to evaluate individual test's utility in detecting noncredible performance. Participants in the simulator group were then divided to receive one of three sets of instructions: naive simulators received general instructions to feign ADHD and no additional information, symptom-coached simulators were given the DSM diagnostic criteria of ADHD, and fully coached simulators received information on both the neuropsychological assessment of ADHD and its diagnostic criteria. Analysis by the authors of the test results demonstrated that the Simulation Group showed a higher median number of test results falling into the suspect range based on the newly derived cut-off scores than the ADHD Group and Control Group. The authors reported that all of the tests ensured at least 90% specificity in the ADHD Group but that sensitivity differed significantly between tests, ranging from 0% to 64.9%. The

authors concluded that tests focusing on selective attention, vigilance, and inhibition were most useful in detecting the instructed simulation of adult ADHD, while tests focusing on figural fluency and task switching lacked sensitivity. Limitations of the study include the single-center and simulation study design, the lack of heterogeneity in the pool used to select the control and simulation groups and the risks associated with the embedded validity indicators in some of the tests and from the risk of overfitting.

Pagán et al. (2023) performed a systematic review to assess the diagnostic utility of the Conner's Continuous Performance Test (CCPT) for diagnosing ADHD in adults. Their review and analysis included 35 published studies with sample sizes ranging between 24 and 413 participants that assessed symptoms for both childhood and adults. The authors stated that there was moderate reliability, subpar discriminant and ecological validity, and mixed sensitivity and specificity for the CCPT. They concluded that their review gave support to previous critiques of the CCPT's diagnostic and utility as a treatment measure and stated that clinicians should assess information from multiple sources when diagnosing ADHD in an adult patient. Limitations of the study include the exclusion of adults with comorbidities from most of the included studies, the heterogeneity in study designs of the studies reviewed including how ADHD was assessed between studies, the instruments being used, the diagnostic measures and outcomes, and the lack of control groups in many of the included studies.

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/Impairment, and Mild Cognitive Impairment (MCI)

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 (AD) 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 AD 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 AD. Dementia due to AD can be distinguished from dementia due to vascular disease by differences in pattern of memory impairment and the progressive nature of AD (Costa et al., 2017). Careful interpretation of test results, taken in conjunction with medical findings, allows differentiation of AD from normal memory loss due to aging, and from vascular dementia.

Quaranta et al. (2023) conducted a single-center, longitudinal study with 253 patients (129 (51%) women; mean age 72.95 years) with Mild Cognitive Impairment (MCI) to assess the possible role of neuropsychological profiles at the baseline in predicting the rate of progression from MCI to overt dementia. Participants underwent evaluation every six months for six years with a neuropsychological battery that included the Mini-Mental State Examination (MMSE), Rey Auditory Verbal Learning Test (RAVLT), copy and delayed recall of the Rey–Osterrieth Complex Figure (ROCF), phonological verbal fluency, categorial verbal fluency, copy of figures with and without landmarks, Raven's Colored Progressive Matrices, Stroop's test, digit span forward and backward, objects naming, and the Multiple Features Targets Cancellation. Each participant also underwent brain MRI and PET-FDG examination to confirm neurodegenerative pathology and to exclude other pathologies. Participants were assessed by a neurologist blinded to the results of the baseline neuropsychological examination and the diagnosis of dementia was formulated if the clinical criteria for dementia due to AD were met. The diagnosis of AD was confirmed in a following visit after six months. The authors reported that 186 converted to dementia and 67 remained stable at the sixth year of follow-up. There were 27 patients who progressed in the first two years (fast converters), 107 converted in the third and fourth year (intermediate converters) and 51 who converted after the fourth year of follow-up (slow converters). The authors reported that participants with stable MCI

performed better than fast decliners in MMSE), several long-term memory scores, and category verbal fluency test (CFT) while stable and intermediate converters differed only in the MMSE and CFT tests, and stable and slow converters differed only in MMSE and phonological/semantic discrepancy score. The authors concluded that early impairment of semantic memory could predict the evolution to AD before the onset of episodic memory disorders, and that the discrepancy between phonological and semantic verbal fluency could be able to detect impairment in advance in respect of simple CFT tests. Limitations of the study include the single-center design and the small sample size.

In a test validation study on the cross-cultural dementia (CCD) screening test for diagnosing AD and Parkinson's disease, Delgado-Álvarez et al. (2023) recruited 150 participants from a single outpatient center and divided them into three groups with 30 participants with AD with mild dementia (AD-D), 30 participants with AD in mild cognitive impairment (AD-MCI), 30 participants with mild cognitive impairment associated with Parkinson's disease (PD-MCI) and 60 participants in the healthy control (HC) group (50% for comparisons with AD, 50% for comparisons with PD-MCI) with no significant differences in age, education, and sex. A comprehensive neuropsychological test battery and the CCD screening test were completed for each participant. The authors reported that intergroup differences were found according to the cognitive profile of each clinical condition and that the CCD test described differences in executive functions and speed scores comparing AD-MCI and PD-MCI. They also noted correlations between standardized neuropsychological tests and CCD measures which they stated support the convergent validity of the CCD test. The authors concluded that the CCD test showed good discrimination properties and cut-off scores for dementia and that the CCD test would be useful as a novel cognitive tool in the assessment of patients with cognitive impairment in different neurological conditions. Limitations of the study included the single-center design, the lack of a group of participants with Parkinson's disease that were cognitively preserved, the generally low level of education of the participants, and the lack of evaluation of cognitive reserve.

The Agency for Healthcare Research and Quality (AHRQ) wanted to identify which individual cognitive tests or combinations of cognitive tests are most accurate for clinically diagnosing clinical Alzheimer's-type dementia (CATD). AHRQ noted that there were no evidence-based guidelines about the merits of either brief cognitive testing or comprehensive neuropsychological testing in this patient population and that access to comprehensive neuropsychological testing is limited in many clinical settings. Fink et al. (2020) completed a Comparative Effectiveness Review for AHRQ in which they analyzed 56 studies on the accuracy of brief cognitive tests for CATD and found that multiple brief cognitive tests were highly sensitive and specific for distinguishing CATD from normal cognition, but less so for distinguishing mild CATD from normal cognition or CATD from mild cognitive impairment (MCI).

In a systematic review and meta-analysis, Belleville et al. (2017) determined the extent to which cognitive measures can predict progression from mild cognitive impairment (MCI) to Alzheimer's type dementia (AD), assessed the predictive accuracy of different cognitive domain categories, and determined whether accuracy varies as a function of age and length of follow-up. The authors systematically reviewed and meta-analyzed data from longitudinal studies reporting sensitivity and specificity values for neuropsychological tests to identify individuals with MCI who will develop AD. Twenty-eight studies met the eligibility criteria (2,365 participants) and reported predictive values from 61 neuropsychological tests with a 31-month mean follow-up. Values were pooled to provide combined accuracy for 14 cognitive domains. Many domains showed very good predictive accuracy with high sensitivity and specificity values. Verbal memory measures and many language tests yielded very high predictive accuracy. Other domains (e.g., executive functions, visual memory) showed better specificity than sensitivity. Predictive accuracy was highest when combining memory measures with a small set of other domains or when relying on broad cognitive batteries. The authors concluded that neuropsychological assessment can strongly contribute to predicting dementia while individuals are still in the MCI phase. According to the authors, cognitive tests are excellent at predicting MCI individuals who will progress to dementia and should be a critical component of any toolkit intended to identify AD at the pre-dementia stage.

Pedersen et al. (2017) 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.

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 multi-centered, longitudinal study that evaluated the sensitivity of the NIH Toolbox Cognition Battery (NIHTB-CB) to detect developmental changes in persons with intellectual disability (ID), Shields et al. (2023) administered the NIHTB-CB and the Stanford-Binet Intelligence Scales, Fifth Edition (SB5), a reference standard cross-validation measure to 256 individuals (ages six to 25 years old at visit 1 with a minimum mental age equivalent of three years) with fragile X syndrome (FXS), Down syndrome (DS), and other ID (OID). After two years (median time between visits was 25.92 months), the authors retested 197 of the participants at visit 2 and assessed group developmental changes in each cognitive domain of the NIHTB-CB and SB5 at three age points (10, 15 and 22 years). The authors reported that the effect sizes of growth measured by the NIHTB-CB tests were comparable with or exceeded those of SB5 and that the NIHTB-CB showed significant gains in almost all domains in OID at younger ages (10 years) with continue gains at 15 years and stability in early adulthood (22 years) and that some significant improvements were seen in mid-adolescence and early adulthood, such as attention/inhibitory control in all groups, episodic memory in DS, working memory, receptive vocabulary, and oral reading in OID and DS, and processing speed in FXS and DS. Limitations of the study included the use of only two assessment periods, the difficulty in recruiting participants in the lower and upper ranges, and the inherent challenges in assessing the youngest participants with significant ID. The authors concluded that the NIHTB-CB was sensitive to developmental changes in individuals with ID etiologies and recommended future research to evaluate the battery more to establish sensitivity to change within the context of treatment studies and delineation of clinically meaningful changes in NIHTB-CB scores.

Tricket et al. (2022) conducted a parallel mediation analysis with cross-sectional data from 152 extremely premature (EP; < 27 weeks of gestation) children and 120 term-born controls who were assessed at age 11 to identify specific cognitive mechanisms that are associated with poor academic attainment in children born preterm for the development of interventional strategies. Mathematics and reading attainment were evaluated to assess the following mediators: verbal working memory, visuospatial working memory, verbal processing speed, attention, and visuospatial processing. The authors reported that children born EP had significantly lower mean composite mathematics and reading scores than controls equating to a deficit of −1.1 SD in reading and −1.4 SD in mathematics, after adjusting for sex and socio-economic status. When children with severe neurodevelopmental disability were excluded, the difference in means, adjusted for sex and socio-economic status, remained significant for reading (−0.4 SD) and mathematics (−1.0 SD). Lower scores were also reported by the authors for verbal working memory (-0.5 SD), visuospatial working memory (-0.7 SD), attention (-0.6 SD), visuospatial processing (-1.1 SD) and verbal processing speed (-0.6 SD) although the magnitude of difference in all five neuropsychological skills were decreased but still significant when children with severe disability were excluded. The authors concluded that children born before 27 weeks of gestation had substantially poorer attainment in reading and mathematics compared to children born at term and that their study identified that a combination of neuropsychological skills including verbal working memory, visuospatial working memory and visuospatial processing may be especially important to target in interventions to improve mathematics and reading outcomes for EP children with average to moderately low IQ. Limitations of the study include the inability of some EP children with severe neurodevelopmental disabilities to complete the entire battery of tests, the recruitment of controls from mainstream schools as this may not be reflective of the general population, the limited time available to assess each child within school, and the use of cross-sectional data for mediation models.

Johnson et al. (2016) explored comorbidity in intellectual disability (ID) and learning disabilities (LD) in children born extremely preterm (EP; < 26 + 0 weeks' gestation). A UK national cohort of 161 EP children and 153 term-born controls without neurosensory impairments was assessed at 11 years of age (the EPICure Study). Neuropsychological abilities commonly affected by EP birth were assessed using the NEPSY Developmental Neuropsychological Test. Overall, 75 (47%) EP children and 7 (4.6%) controls had ID or LD. Comorbidity in ID/LD was more common among EP children than controls (24% vs. 0%). EP children with comorbid ID/LD had significantly poorer neuropsychological abilities and curriculum-based attainment than EP children with an isolated disability or no disabilities. LD was associated with a three times increased risk for special educational needs (SEN). However, EP children with ID alone had poorer neuropsychological abilities and curriculum-based attainment than children with no disabilities, yet there was no increase in SEN provision among this group. The authors concluded that EP children are at high risk for comorbid intellectual and learning disabilities. According to the authors, education professionals should be aware of the complex nature of EP children's difficulties and the need for multi-domain assessments to guide intervention.

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 (Ekdahl et al., 2023; Filipčíková et al., 2022; 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 (Pertichetti et al., 2023;Söderström et al., 2022; Pranckeviciene et al., 2017; Meskal et al., 2016; Walsh et al., 2016; Cochereau et al., 2016)
- Demyelinating disease including multiple sclerosis (Ezegbe et al., 2023; Delgado-Álvarez et al., 2022; Fuchs et al., 2022; Tekin et al., 2022; Böttrich et al., 2020; Wojcik et al., 2019; von Bismarck et al., 2018; Ruet and Brochet, 2018; Vollmer et al., 2016)
- Encephalopathy (Sigurdardottir et al., 2022; Rayes et al., 2019; Moore et al., 2017; Burton et al., 2017)
- Epilepsy and seizure disorders (Fox et al. 2024; Allebone et al. 2023; Silberg et al., 2020; Parra-Díaz and García-Casares, 2017; Grau-López et al., 2017; Wilson et al., 2015; Filippini et al., 2016; Patrikelis et al., 2016)
- Neurotoxin exposure (Nascimento et al., 2016)
- Stroke (Kusec et al. 2023; Zuo et al., 2022; Lo Buono et al., 2018; Tan et al., 2017; Zweifel-Zehnder et al., 2015; Chen et al., 2015)

Computerized Neuropsychological Testing for Concussion

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.

Wilmoth et al. (2023) conducted a systematic review of the psychometric properties of standard neuropsychological tests and computerized tools. The review included 103 articles that focused on adolescent or young adult participants with the Immediate Post-Concussion Assessment and Cognitive Test being the most widely evaluated in 65 of the 103 studies (63%). There were 13 studies that evaluated a hybrid battery, with both computerized cognitive and traditional neuropsychological tests. Risk of bias for most studies (n = 76) indicated a moderate risk of bias, while two studies had a high risk of bias. The authors reported that test-retest reliability estimates ranged from 0.14 to 0.93 for computerized tools and 0.02 to 0.95 for standard neuropsychological tests with the strongest correlations on processing speed tasks for both modalities and that reliability was improved with a 2-factor model (processing speed and memory). The authors also reported that sensitivity to decreased cognitive performance within 72 hours of injury ranged from 45%-93% for computerized tools and 18%-80% for standard neuropsychological tests and that the method for classifying cognitive decline affected sensitivity estimates. The authors concluded that combining computerized tools and standard neuropsychological tests with the strongest psychometric performance provided the greatest value in clinical assessment. Limitations included common methodological processes (such as use of convenience samples, limited controls for confounding factors, and unreported effect sizes), protocol deviations, non-standardized instructions, homogeneity of the athletic participants, continued development of cognitive functions during adolescents that may impact long test-retest intervals, and the heterogeneity of the standard neuropsychological test batteries, and methods used to classify impairment. Studies by Broglio (2018), Hang (2015), MacDonald (2015), McCrory (2017), Nelson (2016), and Weissberger (2017) are included in this systematic review.

Ivins et al. (2022) completed an initial psychometric analysis of the Brain Gauge (BG) personal computer-based test battery to evaluate its potential use for evaluating patients with acute mild traumatic brain injury (mTBI). The study

participants were 73 military service members (SM) who were assessed within 7 days of their injury at military medical treatment facilities [emergency department (30.1%), primary care (41.1%) or a TBI specialty clinic (28.8%)] and 100 healthy service members as a control group. Prior to completing the BG, participants were administered a demographic and military questionnaire, the Neurobehavioral Symptom Inventory (NSI), a PTSD Checklist (PCL-5), and a Patient Health Questionnaire (PHQ-9). The authors reported that SMs with mTBI had statistically significant worse performance on both BG Reaction Time (RT) tests and the Sequential Amplitude Discrimination test as well as having a significantly lower whole-battery composite. The authors stated that, while particular subtests of BG are sensitive to the effects of acute mTBI, there was questionable clinical utility of these scores and that the mTBI group performed worse on some tests than the control group. The authors noted that the base rate analysis revealed that a minority of those with mTBI had multiple scores at or near potentially clinically meaningful performance thresholds, contradicting the very high diagnostic accuracy statistics published by BG's developers which raised concerns about the use of an aggregate score from the BG test battery. Limitations of the study include the small sample size, the delay of up to 7 days post injury which may have influenced results due to possible cognitive recovery that had occurred and the inclusion of only SMs in the study. The authors concluded that their study did find that SMs with acute mTBI on average performed worse than healthy control SMs on the BG Cortical Metric Symptom Score, the BG RT tests and the Sequential Amplitude Discrimination test but that the results also demonstrated that overall, BG does not distinguish mTBI cases from controls at a clinically meaningful rate, and not nearly at the rates previously reported in the literature.

In a prospective longitudinal observational cohort study, Takagi et al. (2019) examined whether cognitive functioning (measured by CogSport) has prognostic value for predicting rapid versus slow recovery. Data were collected at 1-4, 14, and 90-days post-injury. Eligible children were aged ≥ 5 and < 18 years presenting to the Emergency Department having sustained a concussion within 48 hours. Concussion was defined according to the Zurich/Berlin Consensus Statement on Concussion in Sport. Dependent variables were reaction times and error rates on the CogSport Brief Battery. In total, 220 cases were analyzed: 98 in a rapid recovery group [asymptomatic at 14 days post-injury, mean age 11.5 (3.2), 73.5% male] and 122 in a slow recovery group [symptomatic at 14 days post-injury, mean age 12.0 (3.1), 69.7% male]. Longitudinal GEE analyses modeled the trajectories of both mean log10-transformed reaction time and error rates between groups over time (1-4, 14 and 90 days). Both group main and interaction (time by group) terms for all models were non-significant ($p > .05$). The authors concluded that cognitive functioning, measured by CogSport and assessed within 1-4 days of concussion, does not predict prolonged recovery in a pediatric sample. Further, there were no significant group differences at any time point. The authors stated that considering the widespread use and promotion of Computerized neuropsychological tests (CNTs), it is important that clinicians understand the significant limitations of the CogSport battery.

Ivins et al. (2019) assessed the agreement between the following four brief computerized neurocognitive assessment tools by comparing rates of low scores: Automated Neuropsychological Assessment Metrics (ANAM); CogState, also known as CogSport or Axon Sports; Central Nervous System Vital Signs (CNSVS); and Immediate Post-concussion Assessment and Cognitive Test (ImPACT). Four hundred and six U.S. Army service members (SMs) with and without acute mTBI completed two randomly assigned CNTs with order of administration also randomly assigned. A base rate analysis was performed for each CNT to determine the proportions of SMs in the control and mTBI groups who had various numbers of scores that were 1.0+, 1.5+, and 2.0+ standard deviations below the normative mean. These results were used to identify a hierarchy of low score levels ranging from poorest to least poor performance. The agreement was compared between every low score level from each CNT pair administered to the SMs. More SMs in the mTBI group had low scores on all CNTs than SMs in the control group. As performance worsened, the association with mTBI became stronger for all CNTs. Most if not all SMs who performed at the worst level on any given CNT also had low scores on the other CNTs they completed but not necessarily at an equally low level. The authors indicated that the psychometric comparability and clinical utility of these CNTs are not well understood and until such studies are done it will not be possible to make any judgments about which CNT, if any, is superior to the others. The authors state that until more evidence emerges, these CNTs should be used cautiously and only as one source of information from among many other types of clinical assessments. None of them should be used as a definitive or standalone diagnostic tool. An important limitation of this study is that there were relatively small numbers of SMs in each CNT pair who performed at the poorest levels so the results of this study should be treated as preliminary. Another limitation is that the data is from military service members and these findings may not be generalizable to other populations such as high school and college athletes.

Cole et al. (2018) 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 272 healthy control active duty service members and to 231 service members within 7 days of an mTBI. 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 with effect sizes indicating small to no meaningful effect. The authors concluded that although 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.

The Centers for Disease Control and Prevention (CDC) National Center for Injury Prevention and Control Board of Scientific Counselors, a federal advisory committee, established the Pediatric Mild Traumatic Brain Injury Guideline Workgroup and developed a guideline based on a previous systematic review of the literature (Lumba-Brown et al., 2018) to obtain and assess evidence toward developing clinical recommendations for health care professionals related to the diagnosis, prognosis, and management/treatment of pediatric mTBI. The CDC guideline included the recommendations on the diagnosis, prognosis, and management/treatment of pediatric mTBI that were assigned a level of obligation (i.e., must, should, or may) based on confidence in the evidence. Regarding computerized cognitive testing, the CDC stated that health care professionals may use validated, age-appropriate computerized cognitive testing in the acute period of injury as a component of the diagnosis of mTBI (moderate; level C).

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 (McCrory et al., 2017).

Farnsworth et al. (2017) analyzed reliability data for computerized neurocognitive tests (CNTs) using meta-analysis and examined moderating factors that may influence reliability. Studies were included in the meta-analysis if they met all of the following criteria: used a test-retest design, involved at least 1 CNT, provided sufficient statistical data to allow for effect-size calculation, and were published in English. Two independent reviewers investigated each article to assess inclusion criteria. Eighteen studies involving 2,674 participants were retained. Intraclass correlation coefficients were extracted to calculate effect sizes and determine overall reliability. Moderator analyses were conducted to evaluate the effects of the length of the test-retest interval, intraclass correlation coefficient model selection, participant demographics, and study design on reliability. Heterogeneity was evaluated using the Cochran Q statistic. The proportion of acceptable outcomes was greatest for the Axon Sports CogState Test (75%) and lowest for the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) (25%). Moderator analyses indicated that the type of intraclass correlation coefficient model used significantly influenced effect-size estimates, accounting for 17% of the variation in reliability. The authors concluded that the Axon Sports CogState Test, which has a higher proportion of acceptable outcomes and shorter test duration relative to other CNTs, may be a reliable option; however, future studies are needed to compare the diagnostic accuracy of these instruments [The Nakayama et al. (2014) study which was previously cited in this policy, is included in the Farnsworth et al. (2017) meta-analysis].

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 ('naïve'). 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.

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 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 hrs. 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.

In a systematic review, Alsalaheen et al. (2016) assessed the literature on the reliability of the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT). Ten studies that met the inclusion criteria were included in the review. With the exception of processing speed, all composite scores consistently exhibited poor to moderate reliability. When considering two time points, participants who were misclassified as experiencing a “reliable change” in any score ranged between 5% and 26% for verbal memory, 2.2% and 19.6% for visual memory, 4% and 24% for processing speed, and 4% and 23.2% for reaction time. The authors concluded that the majority of ImPACT composite scores did not consistently demonstrate good reliability. According to the authors, clinicians should be cautious when ImPACT is used as a criterion for medical clearance to return to play after concussion. Because of its widespread use in concussion-related clinical research, researchers must exercise due diligence when utilizing ImPACT to evaluate outcomes after concussion or to validate other outcome measures. (Cited in Farnsworth et al., 2017).

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 3,960 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. A brief cognitive evaluation tools cannot substitute for formal neuropsychological assessment. According to the authors, there is insufficient evidence to recommend the widespread routine use of baseline neuropsychological testing.

Baseline Neuropsychological Testing for Concussion

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.

D’Alessio et al. (2024) conducted a case-control study with 6495 middle and high school–aged student-athletes (average age 14.9 ±1.3 years, 59.4% male) in a single county in Ohio who completed a baseline Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) test between ages 10 and 18 years old and who could be linked to the Child-Household Integrated Longitudinal Data (CHILD) system via a probabilistic matching process to examine the relationship between individual-level experiences of adversity and baseline test validity of ImPACT. Validity of the first baseline test (as determined by ImPACT’s built-in validity scoring, age, and home address at the first baseline test) along with any concussion history and diagnosis of learning or developmental disorders (ADHD, dyslexia, autism) were extracted from the ImPACT testing records where 435 (6.7%) participants were found to have an invalid test. The authors reported that social mobility may play an important role in baseline validity and that those with upward social mobility (i.e., poverty or neighborhood deprivation in early childhood only) were not different from youth without such experiences;

however, youth with persistent adversity across childhood or downward social mobility had 50% to 72% lower odds of achieving a valid baseline test. The authors also reported that maltreatment had no significant effect on test validity. The authors concluded that certain patterns of adversity may predispose youth to invalid baseline testing scores, potentially increasing their risk of inappropriate injury management and poor outcomes and that without adequate retesting, youth with particular patterns of adversity may face compounded disadvantages, both from the adversities themselves and from inaccurate neurocognitive diagnostic tests. Limitations of the study include the single county design, the use of data gathered from the child welfare system as children with minoritized racial identities are more likely to be referred or investigated for maltreatment, the use of program enrollment for the Supplemental Nutrition Assistance Program and the Temporary Assistance for Needy Families Program may have limited the authors' understanding of poverty in the study population, and the use of the mother's race as a proxy for the child's race since race is not recorded on birth certificates in Ohio.

Cosgrave et al. (2023) completed a prospective cohort study with 135 school-aged (15 to 19 years) rugby players from 5 schools to explore whether the Sports Concussion Assessment Tool (SCAT), CogState Brief Battery (CBB) and the King-Devick test (K-D test) can be used to monitor concussion status through to full recovery. In this study, all participants completed baseline tests in the preseason where it was found that 61 (45.2%) reported a prior lifetime history of sport related concussion (SRC) and 64 (48%) participants reported symptoms on their baseline SCAT (mean 3.3; range 1-16) with the most common symptoms being fatigue/low energy (31%), neck pain (16%) and irritability (14%). The season consisted of 25 training weeks and 18 games on average across the five teams. During the season, 16 participants experienced 18 SRCs with 9 (56.3%) of the participants having a prior history of SRC. These participants underwent weekly post-concussion testing starting within 1-7 days of injury (mean 3.9 days) with the full battery of tests and an individualized rehabilitation program until recovered. One participant remained symptomatic at 87 days and was referred to a neurologist. Of the remaining 17 concussions, mean severity was 20 days (range 4-42 days). Participants with SRC underwent 52 post-concussion CBB assessments with results consistent with clinically assessed recovery status on 27 (51.9%) occasions. The CBB had a false positive rate of 33% and test specificity was 67%. On 7 (13.5%) occasions participants failed the CBB when clinically they were deemed to be recovered from their concussion. The CBB had a false negative rate of 58% and test sensitivity was 42%. On 18 (34.6%) occasions participants passed the CBB when clinically they were deemed not recovered from their concussion. There were 50 post-concussion K-D tests performed that had results consistent with clinically assessed recovery status on 32 (64%) occasions. The K-D test had a false positive rate of 11% and test specificity was 90%. On 2 (4%) occasions participants failed the K-D when clinically they were deemed to be recovered from their concussion. The K-D test had a false negative rate of 52% and test sensitivity was 48%. On 16 (32%) occasions participants passed the K-D when clinically they had been deemed not recovered from their concussion. The authors reported that the CBB and K-D tests were poorly associated with clinical assessment and produced high false negative rates of 0.58 and 0.52, respectively. The authors concluded that analysis of clinical recovery with CBB and K-D test revealed a relatively poor ability to accurately monitor concussion status compared to clinical assessment. The authors stated that their findings suggest that these tools not be used in isolation for monitoring SRC recovery in adolescents. Limitations of the study include the lack of objective measures of concussion recovery, dependence of concussion detection on school medical staff, the small sample size and concussion incidence and the lack of a follow-up comparison of the tests on the participants who did not sustain an SRC.

Tsushima et al. (2019) identified valid, invalid (identified by five embedded Invalidity Indicators), and sandbagging (identified by three "red flags") results in the ImPACT baseline test scores of 6,346 high school athletes. In addition, the ImPACT post-concussion scores of 266 athletes who sustained a concussion during the school year were evaluated to compare the baseline-to-post concussion changes of valid versus a combined group of invalid and sandbagging scorers. There were 3,299 (51.99%) athletes who had valid baseline scores, 269 (4.24%) had invalid scores, and 3,009 (47.42%) had sandbagging scores. (There were 231 who obtained both invalidity and sandbagging scores.) The overall difference in baseline-to-post concussion changes between the valid scorers and the combined group of invalid and sandbagging scorers was statistically significant. The authors stated that the high rate of athletes who had invalid and sandbagging scores raised concern that the underperformance of baseline testing occurs more commonly than is probably realized by those who utilize computerized neuropsychological testing with high school athletes. Accordingly, efforts are needed to improve test administration procedures so that maximal attention and effort can be maintained among the test takers. According to the authors, increased caution is called for in employing the baseline-to-post concussion paradigm when return-to-play decisions are made.

Abeare et al. (2018) assessed the prevalence of invalid performance on baseline neurocognitive testing using embedded measures within computerized tests and individually administered neuropsychological measures and examined the influence of incentive status and performance validity on neuropsychological test scores. A total of 83 collegiate football athletes completing their preseason baseline assessment within the university's concussion management program and a control group of 140 non-athlete students were included in the study. The cross-sectional design of the study was based on differential incentive status: motivated to do poorly to return to play more quickly after sustaining a concussion

(athletes) versus motivated to do well due to incentivizing performance (students). The main measures of the study included Immediate Post-Concussion and Cognitive Testing (ImPACT), performance validity tests, and measures of cognitive ability. Half of the athletes failed at least one embedded validity indicator within ImPACT (51.8%), and the traditional neuropsychological tests (49.4%), with large effects for performance validity on cognitive test scores, incentive status, and the combination of both factors on measures of attention and processing speed. The authors concluded that invalid performance on baseline assessment is common (50%), consistent across instruments (ImPACT or neuropsychological tests) and settings (one-on-one or group administration), increases as a function of incentive status (risk ratios: 1.3-4.0) and results in gross underestimates of the athletes' true ability level, complicating the clinical interpretation of the postinjury evaluation and potentially leading to premature return to play.

In a retrospective, cross-sectional study, Abeare et al. (2018) assessed the prevalence of invalid performance on baseline testing and assessed whether the prevalence varies as a function of age and validity indicator. Participants included 7,897 consecutively tested equivalently proportioned male and female athletes aged 10 to 21 years, who completed baseline neurocognitive testing for the purpose of concussion management. Baseline assessment was conducted with the Immediate Post concussion Assessment and Cognitive Testing (ImPACT). Base rates of failure on published ImPACT validity indicators were compared within and across age groups. Hypotheses were developed after data collection but prior to analyses. Of the 7,897 study participants, 4,086 (51.7%) were male, mean (SD) age was 14.71 years, 7,820 (99.0%) were primarily English speaking, and the mean (SD) educational level was 8.79 years. The base rate of failure ranged from 6.4% to 47.6% across individual indicators. Most of the sample (55.7%) failed at least 1 of 4 validity indicators. The base rate of failure varied considerably across age groups [117 of 140 (83.6%) for those aged 10 years to 14 of 48 (29.2%) for those aged 21 years], representing a risk ratio of 2.86. The authors indicated that the results for base rate of failure were surprisingly high overall and varied widely depending on the specific validity indicator and the age of the examinee. The strong age association, with 3 of 4 participants aged 10 to 12 years failing validity indicators, suggests that the clinical interpretation and utility of baseline testing in this age group is questionable. According to the authors, these findings underscore the need for close scrutiny of performance validity indicators on baseline testing across age groups.

Computerized Cognitive Testing Such as Cognivue, Mindstreams, BrainCare, and QbTest

Available clinical trials fail 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.

McMurray et al (2024) conducted a study to assess the validity, reliability, and applicability of a short digitally administered test, the BrainEx SCREEN, as a screening tool for mild cognitive impairment (MCI) in an aging population in primary care settings. The study included 147 asymptomatic adults aged 55 or older from five primary care Family Health Teams (FHTs). Participants included health care practitioners, patients and FHT administrative executives with no history of MCI or diagnosis of dementia. Each participant underwent the BrainEx SCREEN, the Quick Mild Cognitive Impairment (QMCI) screen, and the Geriatric Anxiety Scale–10 in a minimum of two screening sessions spaced three months apart with a maximum of four screening sessions over a year. It was noted that the emergence of the COVID-19 pandemic resulted in the some of the assessments being done virtually while others were done face-to-face. The authors reported that their study found that the SCREEN's longer administration time and complex scoring algorithm had challenges such as the internal consistency being below the acceptable threshold and that the test-retest reliability had limitations with moderate intraclass correlation coefficient and inadequate k values. The authors also reported that sensitivity and specificity were consistent between cross-tabulation and discrepant analysis. Limitations noted by the authors included the demographic homogeneity (65.3% female, well-educated participants), the absence of a comprehensive gold standard for MCI diagnosis, and financial constraints that limited the inclusion of confirmatory neuropsychological testing. Other limitations included the small study population, and the absence of verbal fluency testing and memory tests. The authors concluded that the SCREEN did not meet the necessary criteria for an optimal MCI screening tool in a primary care setting due to the longer administration time and lower reliability.

Hayes (2023) published an Evidence Analysis Research Brief on the Cognitrax (CNS Vital Signs LLC) neurocognitive assessment battery indicated for diagnosing and evaluating cognitive, behavioral, and socioemotional functioning to detect neurocognitive disorders. The battery consists of four subtests: finger tapping, verbal memory, symbol digit coding, and the continuous performance test. These tests are utilized to evaluate mental processes, including simple motor performance, attention, memory, and processing speed. A patient's neurocognitive function is computed by tracking the number of correct and incorrect responses as well as reaction time with each of the four subsets. In their brief, Hayes found that there currently is not enough published peer-reviewed literature to evaluate the evidence related to Cognitrax for assessment of behavioral and cognitive changes in a full assessment as they were not able to identify any studies that evaluated this assessment battery.

Bellato et al. (2023) conducted a systematic review and meta-analysis of 15 studies that evaluated accuracy and clinical utility of the QbTest, a computerized cognitive test that is intended to support the clinical assessment of ADHD. There was a total of 2,058 participants (48.6% with ADHD, 53% adults, 20% less than 10 years of age) included in the studies. The authors noted that nine of the 15 studies used QbTest+ (for adolescents > 12 years old, and adults), one used QbTest 6–12 (children ≤ 12 years old), three used both QbTest+ and QbTest 6–12 (since they had mixed samples of children, adolescents and adults), one used QbCheck (conducted at participants' home, based on both QbTest+ and QbTest 6–12 due to wide age range of the sample), and one used QbMini (for preschoolers). The risk analysis showed potential risk of bias in patient selection in seven of the studies. The authors reported that the QbTest total scores showed acceptable, rather than good, sensitivity and specificity in the nine studies that reported both sensitivity and specificity, while subscales showed low-to-moderate sensitivity and moderate-to-good specificity. The authors also reported that the pooled Area under the Receiver Operating Characteristics Curve (AUC) showed moderate-to-acceptable discriminative ability. Subgroup analyses were not able to be conducted by the authors due to the scarcity of the studies. Limitations of the study include the heterogeneity of the versions of the QbTest used, age groups included in the studies, patient selection criteria and study design, the reporting parameters and comparison groups included in the studies, the homogeneity of the high-income countries included in the study, and the small number of studies available for inclusion. The authors concluded that the QbTest scores are not sufficiently accurate in discriminating between ADHD and non-ADHD clinical cases when they are used on their own. The authors recommended that the QbTest should not be used as a stand-alone screening or diagnostic tool or as a triage system for accepting individuals on the waiting list for clinical services; however, when the QbTest is used as an adjunct to support a full clinical assessment, it can produce efficiencies in the assessment pathway and reduce the time to diagnosis. This systematic review included the Hollis 2018 and Brunkhorst-Kanaan 2020 studies previously included in this policy.

In a systematic review and meta-analysis of diagnostic tests for the screening of mild cognitive impairment (MCI) and dementia, Chan et al. (2022) included 90 studies with 22,567 participants to evaluate diagnostic performance among different types of digital drawing tests and paper-and-pencil drawing tests. Seventy-six of the included studies included participants with MCI or dementia in an outpatient clinic or from the community while the rest of the studies recruited participants in a hospital or long-term care setting. The digital drawing tests included in their review and analysis included the digital clock drawing test (CDT), digital pentagon drawing test, digital Rey-Osterrieth complex figure (ROCF), digital tree drawing test, digital house drawing test, and digital spiral test while the paper-and-pencil drawing tests included the CDT, pentagon drawing test, cube drawing test, and ROCF. Six of the studies used digital CDT and 80 of the studies used paper-and-pencil CDT. The primary outcome of the study was the diagnostic performance of the CDT for the screening of MCI and dementia and the secondary outcome was the diagnostic performance of the other types of drawing tests. The authors reported that the performances of the digital and paper-and-pencil pentagon drawing tests were comparable in the screening of dementia, but that the digital CDT demonstrated better sensitivity and specificity diagnostic performance than paper-and-pencil CDT for MCI. Other types of digital drawing tests showed comparable performance with paper-and-pencil formats. Limitations of this study include the lack of head-to-head comparisons, and that the number of studies to compare diagnostic performance of drawing tests are limited. The authors stated that the benefits of digital drawing tests may be stronger if there were more studies available for this meta-analysis.

Romero-Garcia et al. (2022) completed a single-center, prospective cohort study to assess if cognitive impairments would be apparent in a young and high functioning cohort and that app-based cognitive screening would complement traditional neuropsychological assessments. Their study included 17 patients with diffuse gliomas who completed a neuropsychological battery of tests that took 2 -3 hours to complete and the OCS-BRIDGE assessment, an app-based touchscreen assessment that could be completed in 30 minutes. The traditional neuropsychological assessment was administered pre-operatively while the OCS-BRIDGE was administered pre- and post-operatively at the 3- and 12-month follow-ups. The authors reported that the traditional assessment showed that 79% of participants had an impairment in at least one domain, and an average of 2.88 cognitive impairments per participant before surgery, and that, after surgery, all but one participant had at least one impairment with a mean of 4.5 impairments per participants. The OCS-BRIDGE touchscreen assessment showed that 59% of participants had an impairment in at least one domain with a mean of 0.94 impairments per participant before surgery while longitudinal post-operative changes showed that 44% had a reduced number of impairments by their last assessment, 25% had the same, and 31% showed an increased number of impairments. Overall, the traditional neuropsychological tests detected 44 preoperative impairments among the 17 participants in the four combined domains of attention, memory, verbal skills, and non-verbal skills. OCS-BRIDGE detected 13 impairments and 28 possible impairments pre-operatively The authors recognized that the traditional assessment using multiple items across the difficulty range proved more sensitive than the brief touchscreen assessment; however, they also noted that the capacity of the screening app to capture reaction times enhanced its sensitivity in the area of non-verbal function. The authors concluded that a combined approach, using traditional assessment in those areas where brief screening, may be less sensitive, and OCS-BRIDGE style measures for reaction time and perceptual tasks may be most effective and recommended robust, objective, and accessible assessment across multiple centers. Limitations of the study include the small sample size and single-center design, logistical and technical limitations to the

assessments, heterogeneity of tumor location, size and treatment and the potential for practice effects due to reuse of the cognitive assessment tools.

A statistical analysis by Ye et al. (2022) was performed to evaluate BrainCheck, a computerized cognitive testing battery, for its diagnostic accuracy and ability to distinguish the severity of cognitive impairment. A total of 99 participants diagnosed with dementia, mild cognitive impairment (MCI), or normal cognition (NC) completed the BrainCheck battery. Statistical analyses compared participant performances on BrainCheck based on their diagnostic group. BrainCheck battery performance showed differences between the NC, MCI, and dementia groups, achieving 88% or higher sensitivity and specificity (i.e., true positive and true negative rates) for separating dementia from NC, and 77% or higher sensitivity and specificity in separating the MCI group from the NC and dementia groups. Three-group classification found true positive rates of 80% or higher for the NC and dementia groups and true positive rates of 64% or higher for the MCI group. The authors concluded that BrainCheck was able to distinguish between diagnoses of dementia, MCI, and NC, providing a potentially reliable tool for early detection of cognitive impairment. A small sample size makes it difficult to decide whether these conclusions can be generalized to a larger population. Further research with randomized controlled trials is needed to validate these findings.

Chan et al. (2021) performed a systematic review to evaluate the diagnostic performance of digital cognitive tests for mild cognitive impairment (MCI) and dementia in older adults. Literature searches were systematically performed in the OVID databases. Validation studies that reported the diagnostic performance of a digital cognitive test for MCI or dementia were included. The main outcome was the diagnostic performance of the digital test for the detection of MCI or dementia. A total of 56 studies with 46 digital cognitive tests were included in this study. Most of the digital cognitive tests were shown to have comparable diagnostic performances with the paper-and-pencil tests. Twenty-two digital cognitive tests showed a good diagnostic performance for dementia, with a sensitivity and a specificity over 0.80, such as the Computerized Visuo-Spatial Memory test and Self-Administered Tasks Uncovering Risk of Neurodegeneration. Eleven digital cognitive tests showed a good diagnostic performance for MCI such as the Brain Health Assessment. However, all the digital tests only had a few validation studies to verify their performance. The authors concluded that digital cognitive tests showed good performances for MCI and dementia, and that the digital test can collect digital data that is far beyond the traditional ways of cognitive tests. Further research with randomized controlled trials is needed to validate these findings.

An observational study by Rubin et al. (2021) was performed to determine the feasibility of implementing an iPad-based cognitive impairment screening tool, the psychometric properties of the tool, and predictors of cognitive impairment among adults seeking HIV care. A convenience sample of participants completed Brain Baseline Assessment of Cognition and Everyday Functioning (BRACE), which included (1) Trail Making Test Part A, measuring psychomotor speed; (2) Trail Making Test Part B, measuring set-shifting; (3) Stroop Color, measuring processing speed; and (4) the Visual–Spatial Learning Test. Global neuropsychological function was estimated as mean T score performance on the 4 outcomes. Impairment on each test or for the global mean was defined as a T score ≤ 40. Subgroups of participants repeated the tests 4 weeks or > 6 months after completing the first test to evaluate intrapersonal test-retest reliability and practice effects (improvements in performance due to repeated test exposure). An additional subgroup completed a lengthier cognitive battery concurrently to assess validity. Relevant factors were abstracted from electronic medical records to examine predictors of global neuropsychological function. The study population consisted of 404 people with HIV (age: mean 53.6 years; race: 332/404, 82% Black; 34/404, 8% White, 10/404, 2% American Indian/Alaskan Native; 28/404, 7% other and 230/404, 58% male; 174/404, 42% female) of whom 99% (402/404) were on antiretroviral therapy. Participants completed BRACE in a mean of 12 minutes (SD 3.2), and impairment was demonstrated by 34% (136/404) on Trail Making Test A, 44% (177/404) on Trail Making Test B, 40% (161/404) on Stroop Color, and 17% (67/404) on Visual-Spatial Learning Test. Global impairment was demonstrated by 103 out of 404 (25%). Test-retest reliability for the subset of participants (n = 26) repeating the measure at 4 weeks was 0.81 and for the subset of participants (n = 67) repeating the measure almost 1 year later (days: median 294, IQR 50) was 0.63. There were no significant practice effects at either time point (p = .20 and p = .68, respectively). With respect for validity, the correlation between global impairment on the lengthier cognitive battery and BRACE was 0.63 (n = 61; p < .001), with 84% sensitivity and 94% specificity to impairment on the lengthier cognitive battery. The authors concluded that they were able to successfully implement BRACE and estimate cognitive impairment burden in the context of routine clinic care. BRACE was also shown to have good psychometric properties. This easy-to-use tool in clinical settings may facilitate the care needs of people with HIV as cognitive impairment continues to remain a concern in people with HIV. Further research with randomized controlled trials is needed to validate these findings.

Wilson et al. (2021) performed a systematic review of literature to evaluate the benefits, limitations, and validity of computerized neuropsychological assessment devices (CNADs) in the evaluation of HIV-associated neurocognitive disorder (HAND). Following a comprehensive search, the abstracts of relevant articles were compiled and then reviewed for the use of digital neuropsychological testing in the setting of HIV. The articles that met these criteria were read, and their reference lists further examined to compile a more inclusive review. The review was limited to peer-reviewed

English-language journals published within the past 20 years, with no other restrictions, such as sample size or analysis type. Eight CNADs that have undergone validity testing in the setting of HIV were identified and included in the review. The studies included CNADs modeled after traditional testing batteries as well as non-traditional cognitive batteries with advanced technology features including simulated or virtual realities and quick, daily mobile phone assessments, which were reviewed. This review suggests that these computerized neuropsychological assessment devices remain in the early stages of development. The authors concluded that these digital batteries do not have the ability to supplant gold standard neuropsychological tests in screening for HAND. However, many have the potential to become effective clinical screening tools. This review reveals most of these validity studies do not employ large enough sample sizes (fewer than 100) to conclusively determine their ability to detect HAND, creating a degree of uncertainty in external validity. A small sample size makes it difficult to decide whether these conclusions can be generalized to a larger population. The findings of this study need to be validated by well-designed studies. Further investigation is needed before clinical usefulness of this procedure is proven.

Cahn-Hidalgo et al. (2020) determined the cut-off scores for classification of cognitive impairment and assessed Cognivue safety and efficacy in a validation study. Adults (age 55-95 years) at risk for age-related cognitive decline or dementia were invited via posters and email to participate in two cohort studies conducted at various outpatient clinics and assisted- and independent-living facilities. In the cut-off score determination study (n = 92), optimization analyses by positive percent agreement (PPA) and negative percent agreement (NPA), and by accuracy and error bias were conducted. In the clinical validation study (n = 401), regression, rank linear regression, and factor analyses were conducted. Participants in the clinical validation study also completed other neuropsychological tests. For the cut-off score determination study, 92 participants completed St. Louis University Mental Status (SLUMS, reference standard) and Cognivue tests. Analyses showed that SLUMS cut-off scores of < 21 (impairment) and > 26 (no impairment) corresponded to Cognivue scores of 54.5 (NPA = 0.92; PPA = 0.64) and 78.5 (NPA = 0.5; PPA = 0.79), respectively. Therefore, conservatively, Cognivue scores of 55-64 corresponded to impairment, and 74-79 to no impairment. For the clinical validation study, 401 participants completed ≥ 1 testing session, and 358 completed 2 sessions 1-2 wk. apart. Cognivue classification scores were validated, demonstrating good agreement with SLUMS scores (weighted κ 0.57; 95% CI: 0.50-0.63). Reliability analyses showed similar scores across repeated testing for Cognivue (R 2 = 0.81; r = 0.90) and SLUMS (R 2 = 0.67; r = 0.82). Psychometric validity of Cognivue was demonstrated vs. traditional neuropsychological tests. Scores were most closely correlated with measures of verbal processing, manual dexterity/speed, visual contrast sensitivity, visuospatial/executive function, and speed/sequencing. The investigators concluded that Cognivue scores ≤ 50 avoid misclassification of impairment, and scores ≥ 75 avoid misclassification of un-impairment. According to the investigators, this validation study demonstrates good agreement between Cognivue and SLUMS; superior reliability; and good psychometric validity. A limitation of these studies is the use of a single reference standard, SLUMS. Longitudinal follow-up studies are needed to evaluate the ability of Cognivue to monitor cognitive deterioration over time.

Groppell et al., (2019) determined the accuracy and validity of BrainCheck Memory as a diagnostic aid for age-related cognitive impairment, as compared against physician diagnosis and other commonly used neurocognitive screening tests, including the Saint Louis University Mental Status (SLUMS) exam, the Mini-Mental State Examination (MMSE), and the Montreal Cognitive Assessment (MoCA). A total of 583 volunteers over the age of 49 were tested from various community centers and living facilities. The volunteers were divided into five cohorts: a normative population and four comparison groups for the SLUMS exam, the MMSE, the MoCA, and physician diagnosis. Each comparison group completed their respective assessment and BrainCheck Memory. A total of 398 subjects were included in the normative population. A total of 84 participants were in the SLUMS exam cohort, 51 in the MMSE cohort, 35 in the MoCA cohort, and 18 in the physician cohort. BrainCheck Memory assessments were significantly correlated to the SLUMS exam, with coefficients ranging from .5 to .7. Correlation coefficients for the MMSE and BrainCheck and the MoCA and BrainCheck were also significant. Of the 18 subjects evaluated by a physician, 9 (50%) were healthy, 6 (33%) were moderately impaired, and 3 (17%) were severely impaired. A significant difference was found between the severely and moderately impaired subjects and the healthy subjects (p = .02). The investigators found that the BrainCheck Memory composite score showed stronger correlations with the standard assessments as compared to the individual BrainCheck assessments. Receiver operating characteristic (ROC) curve analysis of this composite score found a sensitivity of 81% and a specificity of 94%. The investigators concluded that BrainCheck Memory provides a sensitive and specific metric for age-related cognitive impairment in older adults, with the advantages of a mobile, digital, and easy-to-use test. According to the authors, some participants were unable to complete BrainCheck's entire battery of assessments. While this was accounted for during the analysis, the missing data may have limited statistical power. The investigators also indicated that due to the study's small sample size, more research is needed to compare and validate BrainCheck against physician diagnosis.

Aslam et al. (2018) conducted a systematic review to determine whether automated computerized tests accurately identify patients with progressive cognitive impairment and, if so, to investigate their role in monitoring disease progression and/or response to treatment. Six electronic databases were searched from January 2005 to August 2015, to identify papers for inclusion. Studies assessing the diagnostic accuracy of automated computerized tests for mild cognitive impairment (MCI)

and early dementia against a reference standard were included. Where possible, sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratios were calculated. The Quality Assessment of Diagnostic Accuracy Studies tool was used to assess risk of bias. Sixteen studies assessing 11 diagnostic tools for MCI and early dementia were included. No studies were eligible for inclusion in the review of tools for monitoring progressive disease and response to treatment. The overall quality of the studies was good. However, the wide range of tests assessed and the non-standardized reporting of diagnostic accuracy outcomes meant that statistical analysis was not possible. The authors concluded that some tests have shown promising results for identifying MCI and early dementia. However, concerns over small sample sizes, lack of replicability of studies, and lack of evidence available make it difficult to make recommendations on the clinical use of the computerized tests for diagnosing, monitoring progression, and treatment response for MCI and early dementia.

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 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 AD. 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.

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%).

Intermittent Explosive Disorder

There is insufficient clinical evidence to demonstrate that the use of neuropsychological testing for individuals with intermittent explosive disorder without associated cognitive disorders can be used effectively for clinical decision making to improve patient management of this condition.

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-5), 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

There is insufficient clinical evidence to demonstrate that the use of neuropsychological testing for individuals with migraine or other headaches without associated cognitive disorders can be used effectively for clinical decision making to improve patient management of this condition.

In their omnibus meta-analysis that evaluated performance on clinical measures of neuropsychological function in people with migraine (PwM) and healthy controls (HC), Pizer et al. (2024) included 58 studies with 5,452 PwM (mean age 66 years, 60% male) and 16,647 HCs (mean age 69 years, 55% male). Types of migraines reported were unspecified (n = 2,381), chronic (n = 234), episodic (n = 19), with aura (n = 846) and without aura (2,046). Neuropsychological test scores were obtained from 144 different neuropsychological outcome measures across the 10 content domains. Study quality was considered good for 14 (24%) of the 58 included studies, fair for 38 (66%), and poor for 6 (10%). The authors reported that there was a lower overall cognitive performance in PwM than in HCs, although there was high between-study heterogeneity. The authors found significant domain-specific negative effects in global cognition, executive function, processing speed, visuospatial/construction, simple/complex attention, learning memory, and language while orientation, motor, and intelligence were not significant. Moderator analyses by the authors showed that age (particularly younger HCs), samples drawn from health care facilities versus community-based populations, and higher attack duration were associated with larger (negative) effects and accounted for a significant proportion of between-study heterogeneity in effects. The authors stated that PwM without aura yielded stronger (negative) effects when compared to those PwM with aura, although the aura status did not account for heterogeneity between the studies. The authors concluded that PwM demonstrated worse neurocognition than HCs based on neuropsychological tests, especially on cognitive screening tests and tests within executive functioning and processing speed domains, while the effects were generally small to moderate in magnitude and evident only in clinic samples, and that an aura was not meaningfully associated with neurocognitive impairment. Limitations of the study include the heterogeneity of the study populations, types of migraines and neuropsychological tests used, the lack of reporting of medication utilization for moderator analysis, the inability to confirm that potential cognitive deficits were attributable to the neural substrates of migraine, the potential for mis categorization of cognitive outcomes, and the lack of potential for experimenter bias due to lack of preregistration of the protocol for this study.

Lozano-Soto et al. (2023) conducted a case-control study to examine the presence of neuropsychological deficits in chronic migraine (CM) patients during the interictal phase. The study included 39 CM patients recruited from a single outpatient center and 20 age-, sex-, and education-matched healthy controls (HCs). All study participants underwent clinical, neuropsychiatric, and neuropsychological evaluation by a clinical neurologist to evaluate cognitive domains, including sustained attention (SA), information processing speed (IPS), visuospatial episodic memory, working memory (WM), and verbal fluency (VF), as well as depressive and anxiety symptoms. CM patients exhibited higher scores than HCs for all clinical and neuropsychiatric measures, but no differences were found in personality characteristics. The authors reported that more than half of the CM patients (54%) showed mild-to-severe neuropsychological impairment (NI) with 35.9% classified as having mild NI, 12.8% with moderate NI and 5.1% with severe NI. Additional exploratory analysis showed that more than half (54%) of CM patients with mild, moderate, or severe NI took two or more than two medications and that the severity of NI was associated with the number of treatments received. The authors reported that CM patients exhibited variable NI during periods between acute migraine attacks and that the patients demonstrated cognitive impairment in SA, verbal episodic memory, and Stroop-like interference. Limitations of the study included the small sample size, the single center design, and the large variety of the treatments that the patients received. The authors concluded that CM can be accompanied by a variety of cognitive symptoms during the interictal phase and that these cognitive impairments were most likely related to the mechanisms underlying migraine-induced disability.

In another study that investigated the cognitive impairment of migraineurs, Qin et al. (2022) enrolled 117 adult patients with primary headaches, including 87 with migraine, 30 with tension-type headache (TTH) and 30 healthy controls. No significant differences were found in age, sex, or years of education among the three groups. The authors reported that the Montreal Cognitive Assessment (MoCA) total score and the scores of visuospatial and executive functions, language, and delayed recall in the migraine and TTH groups were significantly lower than those in the healthy control group (all $p < 0.05$) while no significant differences were observed in naming, attention, abstraction, and orientation between the

patients and healthy controls. Limitations of the study included the questionnaire and scales to assess the study subjects, the risk of recall bias in the evaluation of the subjects' anamnestic description of migraine history, the small sample size, the short-term follow-up period the lack of supplementary examinations and the lack of specific information regarding antimigraine medication use and type. The authors concluded the study confirmed cognitive impairment in patients with migraine and TTH and that the duration of attack had an effect on cognitive function in migraineurs.

A cross-sectional study by Chen et al. (2021) was performed to assess whether patients with migraine without aura (MwoA) during the interictal period have attention impairment and to identify the migraine characteristics related to attention deficits. Forty-four subjects with MwoA (4 males, 40 females) and 20 controls matched for age, gender, and literacy education were included in the study. The attention network test (ANT) and a battery of neuropsychological tests, including the trail-making test (TMT), the digit span test (DST), and the Stroop test, were administered to the participants during the headache-free period. Patients in MwoA were more anxious ($p = 0.007$) and depressed ($p = 0.001$) than healthy subjects. Significant differences between the two groups were detected in the executive network ($p = 0.006$) but not in the alerting and orienting networks of ANT. Mean reaction time of ANT in the MwoA group was significantly longer than that in the control group ($p = 0.028$). Patients showed worse performance on DST-forward ($p < 0.001$), DST-backward ($p < 0.001$), DS Total ($p < 0.001$), TMT-A ($p < 0.001$), TMT-B ($p < 0.001$) and TMT-d ($p = 0.002$). Differences found in executive functions between the two groups were unrelated to gender, age, literacy, anxiety, and depression. Multiple regression analysis revealed no relation between clinical characteristics of headache and scores on the executive function with MwoA. The authors concluded that the study suggested that patients in MwoA present worse performances on the executive control of attention networks during the headache-free period, which appear not be associated with measures of migraine severity. The authors also stated although more studies are needed in this area, the results could be useful to find a specific neuropsychological biomarker for migraine pathophysiology. A small sample size makes it difficult to decide whether these conclusions can be generalized to a larger population. Further research with randomized controlled trials is needed to validate these findings.

Foti et al. (2017) identified 16 studies evaluating the association between migraine and cognitive impairment. The authors found that these studies 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.

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.

History of Myocardial Infarction

There is insufficient clinical evidence to demonstrate that the use of neuropsychological testing for individuals with a history of myocardial infarction without associated cognitive disorders can be used effectively for clinical decision making to improve patient management of this condition.

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 (Gallagher et al., 2023; Sauvé 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, eight 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.

Self-Administered or Self-Scored Neuropsychological Testing

The evidence is insufficient to establish the validity and reliability of self-administered or self-scored neuropsychological testing, whether paper-and-pencil or computerized, as a screening tool of cognitive function or for diagnosing neurological disease. Prospective controlled trials are needed to demonstrate the clinical utility of these tests to detect impairment in neuropsychological processing.

Sloane et al. (2023) conducted a single-center, prospective cohort study to evaluate the feasibility and effectiveness of the Miro tablet-based, self-administered neurocognitive mobile application with cognitive games that are intended to detect and classify cognitive deficits as effectively as traditional in-person neuropsychological testing. This study included 79 patients (mean age 62.9 years old, 61% male, 70% had at least college education) who were randomized to either undergo pencil-and-paper or tablet testing. Participants were evaluated and diagnosed with cognitive impairment based on clinical assessment, neuroimaging evaluation and neuropsychological testing prior to study enrollment. The cohort included cognitively impaired adults with right and left hemispheric strokes, mild cognitive impairment, various neurodegenerative disorders such as Alzheimer's dementia, Parkinsonism related cognitive impairment, or Primary Progressive Aphasia (PPA). The study also included 29 age-matched healthy controls with the same inclusion and exclusion criteria as patients but without a neurological diagnosis. All participants completed both pencil-and-paper and tablet assessments. The authors reported that there were statistically significant Pearson correlations between the neuropsychological tests and the tablet equivalents in all domains with moderate or strong correlations in 16 of 17 tests. The authors also reported that all tablet-based subtests differentiated healthy controls from neurologically impaired patients except for the Spatial Span Forward and Finger Tapping modules. Limitations of the study include the single center design, the small sample size, the heterogeneity of diagnostic groups included, and the cross-sectional view of participant performance. The authors concluded that the Miro mobile application was found to be widely acceptable to participants and that it supported the validity of the assessments in the differentiation of healthy controls from patients with neurocognitive deficits in a variety of cognitive domains and across multiple neurological disease etiologies.

Oliva et al. (2022) performed a validation study of the NAIHA Neuro Cognitive Test (NNCT), a computerized, self-administered neuropsychological screening test designed for elderly people with and without cognitive impairment via digitized cognitive assessments. The study included 147 adults over 65 years of age. The authors reported that the validity of the NNCT was demonstrated by correlating outcomes from the Mini Mental State Exam (MMSE), the Clock Drawing Test (CDT) and the Cambridge Cognitive Examination-Revised (CAMCOG-R) test as all subscales of the NNCT test correlated significantly and positively with some of these tests. The authors reported that the NNCT also discriminated correctly to assign the participants into the three groups, Healthy Older Adults (HOA; n = 70), mild cognitive impairment (MCI; n = 44) and Alzheimer's dementia (AD; n = 33) and that the test can be used for screening and for diagnostic support. Limitations of the study include the small sample size, the single-center design, and the inability of some participants to complete all phases of the study due to poor vision and limited ability to use tablet devices.

A cross-sectional study by Paterson et al. (2022) was performed to validate the online Brain Health Assessment (BHA) for detection of amnestic mild cognitive impairment (aMCI) compared to gold-standard neuropsychological assessment, the Montreal Cognitive Assessment (MoCA). Using a cross-sectional design, community-dwelling older adults completed a neuropsychological assessment, were diagnosed as normal cognition (NC) or aMCI and completed the BHA and MoCA. Both logistic regression (LR) and penalized logistic regression (PLR) analyses determined BHA and demographic variables predicting aMCI; MoCA variables were similarly modeled. Diagnostic accuracy was compared using area under the receiver operating characteristic curve (ROC AUC) analyses. Ninety-one participants met inclusion criteria (51 aMCI, 40 NC). PLR modeling for the BHA indicated Face-Name Association, Spatial Working Memory, and age-predicted aMCI [ROC AUC = 0.76; 95% confidence interval (CI): 0.66-0.86]. Optimal cut-points resulted in 21% classified as aMCI (positive), 23% negative, and 56% inconclusive. For the MoCA, digits, abstraction, delayed recall, orientation, and age predicted aMCI (ROC AUC = 0.71; 95% CI: 0.61-0.82). Optimal cut-points resulted in 22% classified positive, 8% negative, and 70% inconclusive (LR results presented within). The BHA model classified fewer participants into the inconclusive category and more as negative for aMCI, compared to the MoCA model (Stuart-Maxwell p = .004). The authors concluded that self-administered BHA provides similar detection of aMCI as a clinician-administered screener (MoCA), with fewer participants classified inconclusively. The BHA has the potential to save practitioners time and decrease unnecessary referrals for a comprehensive assessment to determine the presence of aMCI. Further research with randomized controlled trials (RCTs) is needed to validate these findings.

A cross-sectional study by Vyshedskiy et al. (2022) was performed to assess test scores and the correlation between Boston Cognitive Assessment (BOCA) and Montreal Cognitive Assessment (MoCA) test scores. BOCA is a self-administered 10-minute at-home test intended for longitudinal cognitive monitoring, and MoCA, a gold standard pen-and-paper test of global cognition. BOCA uses randomly selected non-repeating tasks to minimize practice effects. BOCA evaluates eight cognitive domains: 1) Memory/Immediate Recall, 2) Combinatorial Language Comprehension/Prefrontal Synthesis, 3) Visuospatial Reasoning/Mental Rotation, 4) Executive Function/Clock Test, 5) Attention, 6) Mental Math, 7)

Orientation, and 8) Memory/Delayed Recall. A total of 100 patients were included in the study. BOCA was administered to patients with cognitive impairment (n = 50) and age- and education-matched controls (n = 50). Test scores were significantly different between patients and controls ($p < 0.001$) suggesting good discriminative ability. The Cronbach's alpha was 0.87 implying good internal consistency. BOCA demonstrated strong correlation with MoCA ($r = 0.90$, $p < .0.001$). The study revealed strong ($r = 0.94$, $p < 0.001$) test-retest reliability of the total BOCA score one week after participants' initial administration. The practice effect tested by daily BOCA administration over 10 days was insignificant ($\beta = 0.03$, $p = 0.68$). The effect of the screen size tested by BOCA administration on a large computer screen and re-administration of the BOCA to the same participant on a smartphone was insignificant ($\beta = 0.82$, $p = 0.17$; positive β indicates greater score on a smartphone). The authors concluded that BOCA has the potential to reduce the cost and improve the quality of longitudinal cognitive tracking essential for testing novel interventions designed to reduce or reverse cognitive aging. The authors also state that additionally, the test can be used to assess the effect of anesthesia, long-term effect of cancer drugs, COVID fog, and other conditions know to affect cognition. Further research with randomized controlled trials is needed to validate these findings.

A randomized clinical trial was completed by Mahncke et al. (2021) to evaluate the efficacy of self-administered computerized cognitive training. A multisite randomized double-blind clinical trial of a behavioral intervention with an active control was conducted from September 2013 to February 2017, including assessments at baseline, post-training, and after a 3-month follow-up period. The goal of this study was to evaluate the efficacy of a self-administered computerized plasticity-based cognitive training programs in primarily military/veteran participants with a history of mTBI and cognitive impairment. Participants self-administered cognitive training (experimental and active control) programs at home, remotely supervised by a healthcare coach, with an intended training schedule of 5 days per week, 1 hour per day, for 13 weeks. Participants (149 contacted, 83 intent-to-treat) were confirmed to have a history of mTBI (mean of 7.2 years post-injury) through medical history/clinician interview and persistent cognitive impairment through neuropsychological testing and/or quantitative participant reported measure. The experimental intervention was a brain plasticity-based computerized cognitive training program targeting speed/accuracy of information processing, and the active control was composed of computer games. The primary cognitive function measure was a composite of nine standardized neuropsychological assessments, and the primary directly observed functional measure a timed instrumental activities of daily living assessment. Secondary outcome measures included participant-reported assessments of cognitive and mental health. The treatment group showed an improvement in the composite cognitive measure larger than that of the active control group at both the post-training [+6.9 points, confidence interval (CI) +1.0 to +12.7, $p = 0.025$, $d = 0.555$] and the follow-up visit (+7.4 points, CI +0.6 to +14.3, $p = 0.039$, $d = 0.591$). Both large and small cognitive function improvements were seen twice as frequently in the treatment group than in the active control group. No between-group effects were seen on other measures, including the directly observed functional and symptom measures. Statistically equivalent improvements in both groups were seen in depressive and cognitive symptoms. Further investigation is needed before clinical usefulness of this procedure is proven.

A cohort study by Scharre et al. (2021) was performed to compare longitudinal Self-Administered Gerocognitive Examination (SAGE) test scores to non-self-administered Mini-Mental State Examination (MMSE) scores in 5 different diagnostic subgroups. A cohort study evaluating annual rates of change was performed on 665 consecutive patients from Ohio State University Memory Disorders Clinic. Patients with at least two visits 6 months apart evaluated with SAGE and MMSE and classified according to standard clinical criteria as subjective cognitive decline (SCD), mild cognitive impairment (MCI), or AD dementia were included. The pattern of change in SAGE scores was compared to MMSE. One way and repeated measures ANOVA and linear regression models were used. Four hundred twenty-four individuals [40 SCD, 94 MCI non-converters to dementia, 70 MCI converters to dementia (49 to AD dementia and 21 to non-AD dementia), 220 AD dementia] met inclusion criteria. SAGE and MMSE scores declined respectively at annual rates of 1.91 points/year ($p < 0.0001$) and 1.68 points/year ($p < 0.0001$) for MCI converters to AD dementia, and 1.82 points/year ($p < 0.0001$) and 2.38 points/year ($p < 0.0001$) for AD dementia subjects. SAGE and MMSE scores remained stable for SCD and MCI non-converters. Statistical decline from baseline scores in SAGE occurred at least 6 months earlier than MMSE for MCI converters to AD dementia (14.4 vs. 20.4 months), MCI converters to non-AD dementia (14.4 vs. 32.9 months), and AD dementia individuals (8.3 vs. 14.4 months). The authors concluded that SAGE detects MCI conversion to dementia at least 6 months sooner than MMSE. Being self-administered, SAGE also addresses a critical need of removing some barriers in performing cognitive assessments. Limitations of this study include potential referral and sampling biases. Repetitively administering SAGE and identifying stability or decline may provide clinicians with an objective cognitive biomarker impacting evaluation and management choices. Further research with randomized controlled trials is needed to validate these findings.

Clinical Practice Guidelines

American Academy of Child and Adolescent Psychiatry (AACAP)

Practice parameters from the AACAP (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).

In their 2007 Practice Parameter on the assessment and treatment of children and adolescents with ADHD, Pliszka et al. (2007) reported that patients with ADHD have impairments in the executive functioning domains of response inhibition, vigilance, working memory, and some measures of planning, although not all patients with ADHD show executive function deficits. Their recommendation states that psychological and neuropsychological tests are not mandatory for the diagnosis of ADHD but should be performed if the patient's history suggests low general cognitive ability or low achievement in language or mathematics relative to the patient's intellectual ability.

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 Academy of Neurology (AAN)

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, reaffirmed 2022).

A practice guideline update for disorders of consciousness (DoC) developed by the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research indicates that clinicians should use standardized neurobehavioral assessment measures that have been shown to be valid and reliable to improve diagnostic accuracy for the purpose intended in patients with DoC (Giacino et al. 2018, reaffirmed 2021).

In a practice guideline update summary for mild cognitive impairment (MCI), the AAN recommends that when performing a Medicare Annual Wellness Visit, clinicians should not rely on historical report of subjective memory concerns alone when assessing for cognitive impairment (Level B). Various assessment instruments have acceptable diagnostic accuracy for detecting MCI, with no instrument being superior to another. The guideline states that because brief cognitive assessment instruments are usually calibrated to maximize sensitivity rather than specificity, patients who test positive for MCI should then have further assessment (e.g., more in-depth cognitive testing, such as neuropsychological testing with interpretation based on appropriate normative data) to formally assess for this diagnosis (Petersen et al., 2018; reaffirmed 2021).

American Academy of Pediatrics (AAP)

A joint statement for learning disabilities, dyslexia, and vision from the AAP, 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; Reaffirmed 2014).

In 2018, the AAP updated the clinical report guidance for sport-related concussion (SRC) in children and adolescents. The authors of the report indicate that there are numerous studies evaluating the reliability of various computerized neurocognitive tests (CNTs) platforms; however, studies conducted independently of the developers of the tests have questioned the overall reliability of testing from year to year. The reliability of pencil-and-paper testing has also been questioned. The authors indicate that ideally, neurocognitive testing is performed and interpreted by a neuropsychologist. However, given the large number of athletes participating in sports and the relative scarcity of and limited access to neuropsychologists, a widespread CNT program would not be practical or possible. If a non-neuropsychologist is using CNTs, collaboration with a neuropsychologist to aid in test administration and interpretation may be beneficial. CNTs or baseline testing is not specifically addressed in the conclusion or recommendation sections of the report (Halstead et al., 2018).

American Heart Association (AHA) and American Stroke Association

In a guideline for healthcare professionals from the AHA and the American Stroke Association, Winstein 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 Medical Society for Sports Medicine (AMSSM)

The AMSSM position statement on concussion in sport states that baseline testing may be useful in some cases but is not necessary, required or an accepted standard of care for the appropriate management of sport-related concussion (Harmon et al., 2019).

American Psychological Association (APA)

The APA published a Psychological and Neuropsychological Testing Billing and Coding Guide (American Psychological Association, 2019). The guide states that neuropsychological testing is considered medically necessary where initial assessment or assessment over time is needed to:

- Measure cognitive or behavioral deficits related to known or suspected CNS impairment, trauma, or neuropsychiatric disorders, including when the information will be useful in determining a diagnosis, prognosis, or informing treatment planning.
- Evaluate primary symptoms of impaired attention and concentration that can occur in many neurological and psychiatric conditions.
- Determine the potential impact of substances that may cause cognitive impairment (e.g., radiation, chemotherapy, prescribed or illicit drugs, toxins) or result in measurable improvement in cognitive function, including when this information is used to determine treatment planning.
- Conduct pre-surgical or treatment-related measurement of cognitive function to determine whether it is safe to proceed with a medical or surgical procedure that may impact brain function (e.g., deep brain stimulation, resection of brain tumors or arteriovenous malformations, epilepsy surgery, stem cell or organ transplant) or significantly alter a patient's functional status.
- Design, administer, and/or monitor outcomes of cognitive rehabilitation procedures, such as compensatory memory training for brain-injured patients.
- Measure cognitive or functional deficits in children and adolescents based on an inability to develop expected knowledge, skills or abilities as required to adapt to cognitive, social, emotional, or physical demands.
- Evaluate primary symptoms of impaired attention and concentration that can occur in many neurological and psychiatric conditions.

The American Psychological Association Guide (2019) also indicates that neuropsychological testing is not considered reasonable and necessary when:

- Administered for educational or vocational purposes that do not inform medical or health management.

- Comprised exclusively of self-administered or self-scored inventories, or as screening tests of cognitive function or neurological disease (whether paper-and-pencil or computerized; e.g., AIMS, Folstein Mini-Mental Status Examination).
- The patient is neurologically, cognitively, or psychologically unable to participate in a meaningful way in the testing process.
- Used as a routine screening tool given to the individual or to general populations.
- Repeat testing is not required for medical decision-making.
- Administered when the patient is currently under the influence or impaired by alcohol, drugs (prescription or illicit), or other substances.
- The patient has been diagnosed previously with brain dysfunction, such as AD, and there is no expectation that the testing would impact the patient's medical, functional, or behavioral management.

The APA 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; updated 2021):

- 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 aware that standardized psychological and neuropsychological tests are important tools in the assessment of dementia and age-related cognitive change.
- 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 healthcare access but may also disadvantage older persons with limited experience and expertise interacting with technology.

National Academy of Neuropsychology (NAN)

The NAN 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).

National Institute for Health and Care Excellence (NICE)

In their guideline for the management of Multiple Sclerosis, NICE (2022) advises that the practitioner should be aware that the symptoms of MS can include cognitive problems, including memory problems. The guideline recommends that cognition should be assessed as part of the person's comprehensive review and that the assessment should be tailored to

the person's needs, which may include a clinic interview or brief formal assessment, or consideration of a referral for a full neuropsychological assessment if needed.

Dementia assessment, management and support guidelines published by NICE (2018) state that following initial assessment and diagnosis of suspected dementia, patients are to be referred to a specialist once reversible causes of cognitive decline have been ruled out. Following standard, validated criteria use and assessment, neuropsychological testing may be considered if it is unclear:

- Whether the patient has cognitive impairment, or
- Whether their cognitive impairment is caused by dementia, or
- What the correct subtype diagnosis is

U.S. Preventive Services Task Force (USPSTF)

A recommendation statement published by the USPSTF (2020) on screening for cognitive impairment in older adults, including neuropsychological testing, concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening for cognitive impairment.

U.S. Food and Drug Administration (FDA)

This section is to be used for informational purposes only. FDA approval alone is not a basis for coverage.

In March 2021, the FDA cleared the ANAM Test system for Computerized Cognitive Assessment Aid for concussion. The ANAM system is an assessment aid in the management of concussion. The device consists of a software program that administers a battery of neurocognitive tests to an individual to assess their cognitive status. The device may be used with an off-the-shelf computer or a novel device. Refer to the following websites for more information:

- https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpcd/classification.cfm?id=3918
- https://www.accessdata.fda.gov/cdrh_docs/pdf20/K201376.pdf

(Accessed May 15, 2024)

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. Refer to the following website for more information: https://www.accessdata.fda.gov/cdrh_docs/pdf13/DEN130033.pdf. (Accessed May 15, 2024)

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." Refer to the following websites for more information:

- http://www.accessdata.fda.gov/cdrh_docs/pdf15/DEN150037.pdf
- http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm517526.htm

(Accessed May 15, 2024)

References

Abeare C, Messa I, Whitfield C, et al. Performance validity in collegiate football athletes at baseline neurocognitive testing. J Head Trauma Rehabil. 2018 Nov 28.

Abeare CA, Messa I, Zuccato BG, et al. Prevalence of invalid performance on baseline testing for sport-related concussion by age and validity indicator. JAMA Neurol. 2018 Jun 1;75(6):697-703.

Ahto M, Isoaho R, Puolijoki H, et al. Cognitive impairment among elderly coronary heart disease patients. Gerontology. 1999;45(2):87-95.

Allebone J, Kanaan RA, Rayner G, et al. Neuropsychological function in psychosis of epilepsy. Epilepsy Res. 2023 Oct; 196:107222.

Almeida OP, Garrido GJ, Beer C, et al. Coronary heart disease is associated with regional grey matter volume loss: implications for cognitive function and behavior. Intern Med J. 2008 Jul;38(7):599-606.

Alsalaheen B, Stockdale K, Pechumer D, et al. Measurement error in the Immediate Post concussion Assessment and Cognitive Testing (ImPACT): systematic review. J Head Trauma Rehabil. 2016 Jul-Aug;31(4):242-51.

American Academy of Pediatrics, Section on Ophthalmology, Council on Children with Disabilities; American Academy of Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; American Association of Certified Orthoptists. Joint statement - Learning disabilities, dyslexia, and vision. Pediatrics. 2009 Aug;124(2):837-44. Reaffirmed 2014. Available at: http://pediatrics.aappublications.org/content/124/2/837.long. Accessed May 15, 2024.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596. Accessed May 15, 2024

American Psychological Association. APA guidelines for the evaluation of dementia and age-related cognitive change. Am Psychol. 2021, Feb. Available at: https://www.apa.org/practice/guidelines/guidelines-dementia-age-related-cognitive-change.pdf. Accessed May 15, 2024.

American Psychological Association. Psychological and neuropsychological testing billing and coding guide. 2019. Available at: https://www.apaservices.org/practice/reimbursement/health-codes/testing/billing-coding. Accessed May 15, 2024.

Aslam RW, Bates V, Dundar Y, et al. A systematic review of the diagnostic accuracy of automated tests for cognitive impairment. Int J Geriatr Psychiatry. 2018 Apr;33(4):561-575.

Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a theory of mind? Cognition. 1985;21:37-46.

Bauer RM, Iverson GL, Cernich AN, et al. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. Clin Neuropsychol. 2012;26(2):177-96.

Bausela-Herreras E, Alonso-Esteban Y, Alcantud-Marín F. Behavior rating inventory of executive function in preschool (BRIEF-P) and attention-deficit and hyperactivity disorders (ADHD): A systematic review and meta-analysis of floor and ceiling effects. Children (Basel). 2024 Jan; 11(1): 58.

Bechtel N, Kobel M, Penner IK, et al. Attention-deficit/hyperactivity disorder in childhood epilepsy: a neuropsychological and functional imaging study. Epilepsia. 2012 Feb;53(2):325-33.

Becke M, Tucha L, Butzbach M, et al. Feigning adult ADHD on a comprehensive neuropsychological test battery: An analogue study. Int J Environ Res Public Health. 2023 Feb 24;20(5):4070.

Bellato A, Hall CL, Groom MJ, et al. Practitioner review: Clinical utility of the QbTest for the assessment and diagnosis of attention-deficit/hyperactivity disorder - a systematic review and meta-analysis. J Child Psychol Psychiatry. 2023 Oct 6.

Belleville S, Fouquet C, Hudon C, et al.; Consortium for the early identification of Alzheimer's disease-Quebec. Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: a systematic review and meta-analysis. Neuropsychol Rev. 2017 Dec;27(4):328-353.

Böttrich N, Mückschel M, Dillenseger A, et al. On the reliability of examining dual-tasking abilities using a novel E-health device-A proof of concept study in multiple sclerosis. J Clin Med. 2020 Oct 25;9(11):3423.

Broglio SP, Katz BP, Zhao S, et al.; CARE Consortium Investigators. Test-retest reliability and interpretation of common concussion assessment tools: Findings from the NCAA-DoD CARE Consortium. Sports Med. 2018 May;48(5):1255-1268.

Brunkhorst-Kanaan N, Verdenhalven M, Kittel-Schneider S, et al. The quantified behavioral test-a confirmatory test in the diagnostic process of adult ADHD? Front Psychiatry. 2020 Mar 20; 11:216.

Burton KL, Williams TA, Catchpoole SE, et al. Long-term neuropsychological outcomes of childhood onset acute disseminated encephalomyelitis (ADEM): a Meta-Analysis. Neuropsychol Rev. 2017 Mar 31.

Cahn-Hidalgo D, Estes PW, Benabou R. Validity, reliability, and psychometric properties of a computerized, cognitive assessment test (Cognivue®). World J Psychiatry. 2020 Jan 19;10(1):1-11.

Cameron J, Kure CE, Pressler SJ, et al. Diagnostic accuracy of cognitive screening instruments in heart failure: a systematic review. J Cardiovasc Nurs. 2016 Sep-Oct;31(5):412-24.

Carlozzi NE, Kirsch NL, Kisala PA, et al. An examination of the Wechsler Adult Intelligence Scales, Fourth Edition (WAIS-IV) in individuals with complicated mild, moderate and severe traumatic brain injury (TBI). Clin Neuropsychol. 2015;29(1):21-37.

Centers for Disease Control and Prevention. Traumatic brain injury & concussion. May 13, 2021. Available at: http://www.cdc.gov/TraumaticBrainInjury/index.html. Accessed May 15, 2024.

Chan JYC, Bat BKK, Wong A, et al. Evaluation of digital drawing tests and paper-and-pencil drawing tests for the screening of mild cognitive impairment and dementia: A systematic review and meta-analysis of diagnostic studies. Neuropsychol Rev. 2022 Sep;32(3):566-576.

Chan JYC, Yau STY, Kwok TCY, et al. Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: a systematic review. Ageing Res Rev. 2021 Dec; 72:101506.

Chen C, Dong X, Gu P, et al. Attention impairment during the interictal state in migraineurs without aura: a cross-sectional study. J Pain Res. 2021 Oct 6; 14:3073-3083.

Chen X, Wong A, Ye R, et al. Validation of NINDS-CSN neuropsychological battery for vascular cognitive impairment in Chinese stroke patients. BMC Neurol. 2015 Mar 3; 15:20.

Cochereau J, Herbet G, Duffau H. Patients with incidental WHO grade II glioma frequently suffer from neuropsychological disturbances. Acta Neurochir (Wien). 2016 Feb;158(2):305-12.

Cole WR, Arrieux JP, Ivins BJ, et al. A comparison of four computerized neurocognitive assessment tools to a traditional neuropsychological test battery in service members with and without mild traumatic brain injury. Arch Clin Neuropsychol. 2018 Feb 1;33(1):102-119.

Cosgrave C, Fuller C, Kung S, et al. A comparison of clinical assessment with common diagnostic tools for monitoring concussion recovery in adolescent rugby union players. Phys Ther Sport. 2023 Apr 6; 61:165-171.

Costa A, Bak T, Caffarra P, et al. The need for harmonization and innovation of neuropsychological assessment in neurodegenerative dementias in Europe: consensus document of the Joint Program for Neurodegenerative Diseases Working Group. Alzheimer's Res Ther. 2017 Apr 17;9(1):27.

D'Alessio AS, Salas Atwell M, et al. Experiences of adversity and validity of baseline concussion testing. J Athl Train. 2024 Apr 1;59(4):373-380.

Delgado-Álvarez A, Delgado-Alonso C, Goudsmit M, et al. Validation of a brief cross-cultural cognitive screening test in Multiple Sclerosis. Mult Scler Relat Disord. 2022 Nov; 67:104091.

Dresler T, Lürding R, Paelecke-Habermann Y, et al. Cluster headache and neuropsychological functioning. Cephalalgia. 2012 Aug;32(11):813-21.

Echemendia RJ, Iverson GL, McCrea M, et al. Advances in neuropsychological assessment of sport-related concussion. Br J Sports Med. 2013 Apr;47(5):294-8.

Ekdahl N, Möller MC, Deboussard CN, et al. Investigating cognitive reserve, symptom resolution and brain connectivity in mild traumatic brain injury. BMC Neurol. 2023 Dec 20;23(1):450.

Ezegbe C, Zarghami A, van der Mei I, et al. Instruments measuring change in cognitive function in multiple sclerosis: A systematic review. Brain Behav. 2023 Jun;13(6): e3009.

Farnsworth JL 2nd, Dargo L, Ragan BG, et al. Reliability of computerized neurocognitive tests for concussion assessment: a meta-analysis. J Athl Train. 2017 Sep;52(9):826-833.

Filipčíková M, Quang H, Cassel A, et al. Exploring neuropsychological underpinnings of poor communication after traumatic brain injury: The role of apathy, disinhibition and social cognition. Int J Lang Commun Disord. 2022 Dec 21.

Filippini M, Ardu E, Stefanelli S et al., Neuropsychological profile in new-onset benign epilepsy with centrotemporal spikes (BECTS): Focusing on executive functions. Epilepsy Behav. 2016 Jan; 54:71-9.

Fink HA, Hemmy LS, Linskens EJ, et al. Diagnosis and treatment of clinical Alzheimer's-type dementia: A systematic review. Comparative Effectiveness Review No. 223. (Prepared by the Minnesota Evidence-based Practice Center under Contract No. 290- 2015-00008-I.) AHRQ Publication No. 20-EHC003. Rockville, MD: Agency for Healthcare Research and Quality; April 2020.

Foti M, Lo Buono V, Corallo F, et al. Neuropsychological assessment in migraine patients: a descriptive review on cognitive implications. Neurol Sci. 2017 Apr;38(4):553-562.

Fox J, Bollig MK, Mishra M, Jacobs M. Neuropsychological characteristics of drug resistant epilepsy patients with and without comorbid functional seizures. Epilepsy Res. 2024 Mar; 201:107340.

Fuchs TA, Gillies J, Jaworski MG 3rd, et al. Repeated forms, testing intervals, and SDMT performance in a large multiple sclerosis dataset. Mult Scler Relat Disord. 2022 Dec; 68:104375.

Gallagher R, Ouyang ML, Tofler G, et al. Sensitivity and specificity of 5 min cognitive screening tests in patients with acute coronary syndrome. Eur J Cardiovasc Nurs. 2023 Mar 1;22(2):166-174.

Gaudet CE, Weyandt LL. Immediate Post-Concussion and Cognitive Testing (ImPACT): a systematic review of the prevalence and assessment of invalid performance. Clin Neuropsychol. 2017 Jan;31(1):43-58.

Giacino JT, Katz DI, Schiff ND, et al. Practice guideline update recommendations summary: Disorders of consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. Neurology. 2018 Sep 4;91(10):450-460. Reaffirmed September 18, 2021.

Giza CC, Kutcher JS, Ashwal S, et al. Summary of evidence-based guideline update: Evaluation and management of concussion in sports: Report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology. 2013 Mar 18. Reaffirmed April 30, 2022.

Grau-López L, Jiménez M, Ciurans J, et al. Importance of neuropsychological and clinical features to predict seizure control in medically treated patients with mesial temporal epilepsy and hippocampal sclerosis. Epilepsy Behav. 2017 Feb 24; 69:121-125.

Grubb NR, Simpson C, Fox KA. Memory function in patients with stable, moderate to severe cardiac failure. Am Heart J. 2000 Jul;140(1): E1-5.

Groppell S, Soto-Ruiz KM, Flores B, et al. A rapid, mobile neurocognitive screening test to aid in identifying cognitive impairment and dementia (BrainCheck): Cohort Study. JMIR Aging. 2019 Mar 21;2(1): e12615.

Halstead ME, Walter KD, Moffatt K; Council on Sports Medicine and Fitness. Sport-related concussion in children and adolescents. Pediatrics. 2018 Dec;142(6).

Hang B, Babcock L, Hornung R, et al. Can computerized neuropsychological testing in the emergency department predict recovery for young athletes with concussions? Pediatr Emerg Care. 2015 Oct;31(10):688-93.

Hanks RA, Jackson AM, Crisanti LK. Predictive validity of a brief outpatient neuropsychological battery in individuals 1-25 years post traumatic brain injury. Clin Neuropsychol. 2016 Oct;30(7):1074-86.

Happe F, Frith U. The weak coherence account: detail-focused style in autism spectrum disorders. J Autism Dev Disord. 2006; 36:5-25.

Harmon KG, Clugston JR, Dec K, et al. American Medical Society for Sports Medicine position statement on concussion in sport. Clin J Sport Med. 2019 Mar;29(2):87-100.

Hartman E, Houwen S, Scherder E, et al. On the relationship between motor performance and executive functioning in children with intellectual disabilities. J Intellect Disabil Res. 2010 May;54(5):468-77.

Hayes, Inc. Evidence Analysis Research Brief. Cognitrax (CNS Vital Signs LLC) for assessment of behavioral and cognitive changes. Lansdale, Pa: Hayes, Inc.; August 2023.

Hollis C, Hall CL, Guo B, et al.; the AQUA Trial Group. The impact of a computerized test of attention and activity (QbTest) on diagnostic decision-making in children and young people with suspected attention deficit hyperactivity disorder: single-blind randomized controlled trial. J Child Psychol Psychiatry. 2018 Dec;59(12):1298-1308.

Ivins BJ, Arrieux JP, Cole WR. An initial psychometric analysis of the Brain Gauge tactile-based test battery and its potential for clinical use assessing patients with acute mild traumatic brain injury. Arch Clin Neuropsychol. 2022 Oct 19;37(7):1564-1578.

Ivins BJ, Arrieux JP, Schwab KA, et al. Using rates of low scores to assess agreement between brief computerized neuropsychological assessment batteries: A clinically based approach for psychometric comparisons. Arch Clin Neuropsychol. 2019 Feb 23. pii: acz004.

Johnson S, Strauss V, Gilmore C, et al. Learning disabilities among extremely preterm children without neurosensory impairment: Comorbidity, neuropsychological profiles and scholastic outcomes. Early Hum Dev. 2016 Dec; 103:69-75.

Kontos AP, Braithwaite R, Dakan S, et al. Computerized neurocognitive testing within 1 week of sport-related concussion: meta-analytic review and analysis of moderating factors. J Int Neuropsychol Soc. 2014 Mar;20(3):324-32.

Kusec A, Milosevich E, Williams OA, et al. Long-term psychological outcomes following stroke: the OX-CHRONIC study. BMC Neurol. 2023 Nov 30;23(1):426.

Lo Buono V, Bonanno L, Palmeri R, et al. Relation among psychopathological symptoms, neuropsychological domains, and functional disability in subacute poststroke rehabilitation. J Stroke Cerebrovasc Dis. 2018 Feb 5. pii: S1052-3057(17)30707-3.

Lozano-Soto E, Cruz-Gómez ÁJ, Rashid-López R, et al. Neuropsychological and neuropsychiatric features of chronic migraine patients during the interictal phase. J Clin Med. 2023 Jan 9;12(2):523.

Lumba-Brown A, Yeates KO, Sarmiento K, et al. Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury among children. JAMA Pediatr. 2018a Nov 1;172(11): e182853.

MacDonald J, Duerson D. Reliability of a computerized neurocognitive test in baseline concussion testing of high school athletes. Clin J Sport Med. 2015 Jul;25(4):367-72.

Mahncke HW, DeGutis J, Levin H, et al. A randomized clinical trial of plasticity-based cognitive training in mild traumatic brain injury. Brain. 2021 Aug 17;144(7):1994-2008.

McCrory P, Meeuwisse W, Dvorak J, et al. Consensus statement on concussion in sport - the 5th international conference on concussion in sport held in Berlin, October 2016. Br J Sports Med. 2017 Apr 26.

McMurray J, Levy A, Pang W, Holyoke P. Psychometric evaluation of a tablet-based tool to detect mild cognitive impairment in older adults: Mixed methods study. J Med Internet Res. 2024 Apr 19;26: e56883.

Meskal I, Gehring K, Rutten GJ, et al. Cognitive functioning in meningioma patients: a systematic review. J Neurooncol. 2016 Jun;128(2):195-205.

Moore RC, Paolillo EW, Heaton A, et al. Clinical utility of the UCSD Performance-Based Skills Assessment-Brief (UPSA-B) in adults living with HIV: Associations with neuropsychological impairment and patient-reported everyday functioning difficulties. PLoS One. 2017 Aug 24;12(8): e0183614.

Moser RS, Iverson GL, Echemendia RJ, et al; NAN Policy and Planning Committee. Neuropsychological evaluation in the diagnosis and management of sports-related concussion. Arch Clin Neuropsychol. 2007 Nov;22(8):909-16.

Nakayama Y, Covassin T, Schatz P, et al. Examination of the test-retest reliability of a computerized neurocognitive test battery. Am J Sports Med. 2014 Jun 6;42(8):2000-2005.

Nascimento S, Baierle M, Göethel G, et al. Associations among environmental exposure to manganese, neuropsychological performance, oxidative damage and kidney biomarkers in children. Environ Res. 2016 May; 147:32-43.

National Institute for Health and Care Excellence (NICE). Dementia: assessment, management and support for people living with dementia and their careers. Available at: http://www.nice.org.uk/guidance/ng97. Accessed May 15, 2024.

National Institute for Health and Care Excellence (NICE). Multiple sclerosis in adults: management. Available at: https://www.nice.org.uk/guidance/ng220/resources/multiple-sclerosis-in-adults-management-pdf-66143828948677. Accessed May 15, 2024.

Nelson LD, Furger RE, Gikas P, et al. Prospective, head-to-head study of three computerized neurocognitive assessment tools part 2: utility for assessment of mild traumatic brain injury in emergency department patients. J Int Neuropsychol Soc. 2017 Apr;23(4):293-303.

Nelson LD, LaRoche AA, Pfaller AY, et al. Prospective, head-to-head study of three computerized neurocognitive assessment tools (CNTs): reliability and validity for the assessment of sport-related concussion. J Int Neuropsychol Soc. 2016 Jan;22(1):24-37.

Oliva I, Losa J. Validation of the Computerized Cognitive Assessment Test: NNCT. Int J Environ Res Public Health. 2022 Aug 23;19(17):10495.

Ozonoff S, Pennington BF, Rogers SJ. Executive function deficits in high functioning autistic individuals: relationship to theory of mind. J Child Psychol Psychiatry. 1991; 32:1081-1105.

Pagán AF, Huizar YP, Schmidt AT. Conner's Continuous Performance Test and adult ADHD: A systematic literature review. J Atten Disord. 2023 Feb;27(3):231-249.

Parra-Díaz P, García-Casares N. Memory assessment in patients with temporal lobe epilepsy to predict memory impairment after surgery: A systematic review. Neurologia. 2017 Apr 19. pii: S0213-4853(17)30150-0.

Paterson TSE, Sivajohan B, Gardner S, et al. Accuracy of a self-administered online cognitive assessment in detecting amnestic mild cognitive impairment. J Gerontol B Psychol Sci Soc Sci. 2022 Feb 3;77(2):341-350.

Patrikelis P, Gatzonis S, Siatouni A, et al. Preoperative neuropsychological presentation of patients with refractory frontal lobe epilepsy. Acta Neurochir (Wien). 2016 Jun;158(6):1139-50.

Pedersen KF, Larsen JP, Tysnes OB, et al., Natural course of mild cognitive impairment in Parkinson disease: A 5-year population-based study. Neurology. 2017 Feb 21;88(8):767-774.

Pertichetti M, Corbo D, Belotti F, et al. Neuropsychological evaluation and functional magnetic resonance imaging tasks in the preoperative assessment of patients with brain tumors: A systematic review. Brain Sci. 2023 Sep 28;13(10):1380.

Peterson BS, Trampush J, Maglione M, et al. ADHD diagnosis and treatment in children and adolescents [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2024 Mar. Report No.: 24-EHC003Report No.: 2023-SR-03. PMID: 38657097. Available at: https://www.ncbi.nlm.nih.gov/books/NBK603001/. Accessed May 26, 2024.

Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology. 2018 Jan 16;90(3):126-135. Reaffirmed January 30, 2021. Available at: https://www.neurology.org/doi/10.1212/WNL.0000000000004826. Accessed May 15, 2024.

Pizer JH, Aita SL, Myers MA, et al. Neuropsychological function in migraine headaches: An expanded comprehensive multidomain meta-analysis. Neurology. 2024 Feb 27;102(4): e208109.

Pliszka S; AACAP Work Group on Quality Issues. Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2007 Jul;46(7):894-921. Available at: https://www.jaacap.org/article/S0890-8567(09)62182-1/fulltext. Accessed May 28, 2024.

Pranckeviciene A, Deltuva VP, Tamasauskas A, et al. Association between psychological distress, subjective cognitive complaints and objective neuropsychological functioning in brain tumor patients. Clin Neurol Neurosurg. 2017 Dec; 163:18-23.

Qin T, Chen C. Cognitive dysfunction in migraineurs. Medicina (Kaunas). 2022 Jun 29;58(7):870.

Quaranta D, Caraglia N, L'Abbate F, et al. Neuropsychological predictors of conversion from mild cognitive impairment to dementia at different timepoints. Brain Behav. 2023 Sep;13(9): e3098.

Racine AM, Clark LR, Berman SE, et al. Associations between performance on an abbreviated cogstate battery, other measures of cognitive function, and biomarkers in people at risk for Alzheimer's disease. J Alzheimers Dis. 2016 Oct 18;54(4):1395-1408.

Rayes HA, Tani C, Kwan A, et al. What is the prevalence of cognitive impairment in lupus and which instruments are used to measure it? A systematic review and meta-analysis. Semin Arthritis Rheum. 2018 Oct;48(2):240-255.

Roebuck-Spencer TM, Glen T, Puente AE, et al. Cognitive screening tests versus comprehensive neuropsychological test batteries: A National Academy of Neuropsychology education paper. Arch Clin Neuropsychol. 2017 Mar 10:1-8.

Romero-Garcia R, Owen M, McDonald A, et al. Assessment of neuropsychological function in brain tumor treatment: a comparison of traditional neuropsychological assessment with app-based cognitive screening. Acta Neurochir (Wien). 2022 Aug;164(8):2021-2034.

Rubin LH, Severson J, Marcotte TD, et al. Tablet-based cognitive impairment screening for adults with HIV seeking clinical care: observational study. JMIR Ment Health. 2021 Sep 9;8(9): e25660.

Ruet A, Brochet B. Cognitive assessment in patients with multiple sclerosis: From neuropsychological batteries to ecological tools. Ann Phys Rehabil Med. 2018 Feb 17. pii: S1877-0657(18)30012-5.

Sauvé MJ, Lewis WR, Blankenbiller M, et al. Cognitive impairments in chronic heart failure: a case controlled study. J Card Fail. 2009 Feb;15(1):1-10.

Scharre DW, Chang SI, Nagaraja HN, et al. Self-administered gerocognitive examination: longitudinal cohort testing for the early detection of dementia conversion. Alzheimers Res Ther. 2021 Dec 6;13(1):192.

Shields RH, Kaat A, Sansone SM, et al. Sensitivity of the NIH Toolbox to detect cognitive change in individuals with intellectual and developmental disability. Neurology. 2023 Feb 21;100(8): e778-e789.

Shopin L, Shenhar-Tsarfaty S, Ben Assayag E, et al. Cognitive assessment in proximity to acute ischemic stroke/transient ischemic attack: comparison of the Montreal cognitive assessment test and mindstreams computerized cognitive assessment battery. Dement Geriatr Cogn Disord. 2013;36(1-2):36-42.

Sigurdardottir AH, Knudtzen FC, Nymark A, Bang M. Fatigue and cognitive impairment in neuroborreliosis patients posttreatment - A neuropsychological retrospective cohort study. Brain Behav. 2022 Sep;12(9):e2719.

Silberg T, Ahoniska-Assa J, Bord A, et al. In the eye of the beholder: Using a multiple-informant approach to examine the mediating effect of cognitive functioning on emotional and behavioral problems in children with an active epilepsy. Seizure. 2020 Nov; 82:31-38.

Silver CH, Blackburn LB, Arffa S, et al. The importance of neuropsychological assessment for the evaluation of childhood learning disorders NAN Policy and Planning Committee. Arch Clin Neuropsychol. 2006 Oct;21(7):741-4.

Sloane KL, Fabian R, Wright A, et al. Supervised, self-administered tablet-based cognitive assessment in neurodegenerative disorders and stroke. Dement Geriatr Cogn Disord. 2023;52(2):74-82.

Söderström H, Brocki K, Kleberg JL, et al. Neurocognitive functions before and after radiotherapy in pediatric brain tumor survivors. Pediatr Neurol. 2022 Aug; 133:21-29.

Takagi M, Hearps SJC, Babl FE, et al. Does a computerized neuropsychological test predict prolonged recovery in concussed children presenting to the ED? Child Neuropsychol. 2020 Jan;26(1):54-68.

Talwalker S, Overall JE, Srirama MK, et al. Cardinal features of cognitive dysfunction in Alzheimer's disease: a factor-analytic study of the Alzheimer's Disease Assessment Scale. J Geriatr Psychiatr Neurol. 1996; 9:39-46.

Tan HH, Xu J, Teoh HL, et al. Decline in changing Montreal Cognitive Assessment (MoCA) scores is associated with post-stroke cognitive decline determined by a formal neuropsychological evaluation. PLoS One. 2017 Mar 27;12(3): e0173291.

Tekin S, Bir LS, Oncel C, et al. Evaluation of cognitive dysfunction by the clock drawing test in multiple sclerosis and clinically isolated syndrome patients: Correlation with other neuropsychological tests. Neurosciences (Riyadh). 2022 Oct;27(4):251-256.

Trickett J, Bernardi M, Fahy A, et al. Neuropsychological abilities underpinning academic attainment in children born extremely preterm. Child Neuropsychol. 2022 Aug;28(6):746-767.

Tsushima WT, Yamamoto MH, Ahn HJ, et al. Invalid baseline testing with ImPACT: Does sandbagging occur with high school athletes? Appl Neuropsychol Child. 2019 Aug 13:1-10.

U.S. Preventive Services Task Force, Owens DK, Davidson KW, et al. Screening for cognitive impairment in older adults: U.S. Preventive Services Task Force recommendation statement. JAMA. 2020 Feb 25;323(8):757-763. Available at: https://www.ncbi.nlm.nih.gov/books/NBK554654/. Accessed May 28, 2024.

Volkmar F, Siegel M, Woodbury-Smith M, et al. American Academy of Child and Adolescent Psychiatry (AACAP) Committee on Quality Issues (CQI). Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. J Am Acad Child Adolesc Psychiatry. 2014 Feb;53(2):237-57. Available at: https://www.jaacap.org/article/S0890-8567(13)00819-8/pdf. Accessed May 28, 2024.

Vollmer T, Huynh L, Kelley C, et al. Relationship between brain volume loss and cognitive outcomes among patients with multiple sclerosis: a systematic literature review. Neurol Sci. 2016 Feb;37(2):165-79.

von Bismarck O, Dankowski T, Ambrosius B, et al. Treatment choices and neuropsychological symptoms of a large cohort of early MS. Neurol Neuroimmunol Neuroinflamm. 2018 Mar 1;5(3):e446.

Vyshedskiy A, Netson R, Fridberg E, et al. Boston cognitive assessment (BOCA) - a comprehensive self-administered smartphone - and computer-based at-home test for longitudinal tracking of cognitive performance. BMC Neurol. 2022 Mar 15;22(1):92.

Walsh KS, Noll RB, Annett RD, et al. Standard of care for neuropsychological monitoring in pediatric neuro-oncology: Lessons from the Children's Oncology Group (COG). Pediatr Blood Cancer. 2016 Feb;63(2):191-5.

Weissberger GH, Strong JV, Stefanidis KB, et al. Diagnostic accuracy of memory measures in Alzheimer's dementia and mild cognitive impairment: A systematic review and meta-analysis. Neuropsychol Rev. 2017 Dec;27(4):354-388.

Wilmoth K, Brett BL, Emmert NA, et al. Psychometric properties of computerized cognitive tools and standard neuropsychological tests used to assess sport concussion: A systematic review. Neuropsychol Rev. 2023 Dec;33(4):675-692.

Wilson S, Milanini B, Javandel S, et al. Validity of digital assessments in screening for HIV-related cognitive impairment: a review. Curr HIV/AIDS Rep. 2021 Dec;18(6):581-592.

Wilson SJ, Baxendale S, Barr W, et al. Indications and expectations for neuropsychological assessment in routine epilepsy care: Report of the ILAE Neuropsychology Task Force, Diagnostic Methods Commission, 2013-2017. Epilepsia. 2015 May;56(5):674-81.

Winstein CJ, Stein J, Arena R, et al.; American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, and Council on Quality of Care and Outcomes Research. Guidelines for Adult Stroke Rehabilitation and Recovery: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2016 Jun;47(6): e98-e169.

Wojcik CM, Beier M, Costello K, et al; National MS Society Cognition Work Team. Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review. Mult Scler. 2019 Dec;25(14):1848-1869.

Ye S, Sun K, Huynh D, et al. A computerized cognitive test battery for detection of dementia and mild cognitive impairment: instrument validation study. JMIR Aging. 2022 Apr 15;5(2): e36825.

Zuo L, Dong Y, Liao X, et al. Risk factors for decline in Montreal Cognitive Assessment (MoCA) scores in patients with acute transient ischemic attack and minor stroke. J Clin Hypertens (Greenwich). 2022 Jul;24(7):851-857.

Zweifel-Zehnder AE, Stienen MN, Chicherio C, et al.; Swiss SOS study group. Call for uniform neuropsychological assessment after aneurysmal subarachnoid hemorrhage: Swiss recommendations. Acta Neurochir (Wien). 2015 Sep;157(9):1449-58.

Policy History/Revision Information

Date	Summary of Changes
06/01/2025	**Application** ***Idaho*** and ***Kansas*** • Added language to indicate this Medical Policy does not apply to the states of Idaho and Kansas; refer to the state-specific policy versions
09/01/2024	**Supporting Information** • Updated *Clinical Evidence* and *References* sections to reflect the most current information • Archived previous policy version CS083.Q

Instructions for Use

This Medical Policy provides assistance in interpreting UnitedHealthcare standard benefit plans. When deciding coverage, the federal, state or contractual requirements for benefit plan coverage must be referenced as the terms of the federal, state or contractual requirements for benefit plan coverage may differ from the standard benefit plan. In the event of a conflict, the federal, state or contractual requirements for benefit plan coverage govern. Before using this policy, please check the federal, state or contractual requirements for benefit plan coverage. UnitedHealthcare reserves the right 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 InterQual® criteria, to assist us in administering health benefits. The UnitedHealthcare Medical Policies 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.