

Cognitive Testing: Bedside Screening Tools and Formal Neuropsychological Assessment - A Comprehensive Academic Review

A Abyad

Correspondence:

A. Abyad, MD, MPH, MBA, DBA, AGSF

Consultant, Internal Medicine and Geriatrics, Dar Al Shifa Hospital, Kuwait

Chairman, Middle-East Academy for Medicine of Aging.

President, Middle East & North Africa Association on Aging & Alzheimer's

Coordinator, Middle-East Primary Care Research Network

Coordinator, Middle-East Network on Aging

Email: aabyad@cyberia.net.lb

Received: August 2026. Accepted: September 2026; Published: October 2026.

Citation: A. Abyad. Cognitive Testing: Bedside Screening Tools and Formal Neuropsychological Assessment

A Comprehensive Academic Review. World Family Medicine. October 2026; 24(7): 36 - 54 DOI: 10.5742/MEWFM.2026.241794

Abstract

Cognitive testing constitutes an indispensable component of contemporary clinical neuroscience, psychiatry, geriatrics, and primary care, providing the empirical foundation upon which diagnoses of mild and major neurocognitive disorder are established, staged, and monitored. This review provides a comprehensive, evidence-based examination of the two principal tiers of cognitive assessment employed in clinical practice: brief bedside cognitive screening instruments, exemplified by the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA), and formal neuropsychological testing conducted by licensed neuropsychologists. Drawing on diagnostic accuracy meta-analyses, clinical practice guidelines, and recent advances in digital and biomarker-based assessment, the review synthesizes evidence published predominantly between 2020 and 2026 alongside foundational diagnostic frameworks. Diagnostic criteria tables, comparative accuracy tables, and a stepwise clinical algorithm are presented to orient clinicians toward evidence-based test selection. The review further addresses performance validity assessment, differential diagnosis across dementia subtypes, cultural and educational bias in screening instruments, and emerging digital and blood-based biomarker technologies that

are reshaping the diagnostic pathway. The review concludes that bedside screening and formal neuropsychological testing occupy complementary, non-substitutable positions within a stepped, biologically informed model of cognitive evaluation, and that clinicians must select instruments deliberately according to clinical context, patient characteristics, and diagnostic yield.

Keywords: cognitive testing, screening tools, neuropsychological assessment

Introduction

Cognitive impairment represents one of the most consequential and rapidly growing challenges confronting global health systems. The Alzheimer's Association (2024) estimates that millions of individuals in the United States alone are living with Alzheimer's disease and related dementias, a figure projected to rise substantially as the population ages. Rajan et al. (2021) similarly project that the number of individuals with clinical Alzheimer's disease and mild cognitive impairment (MCI) in the United States will increase considerably between 2020 and 2060. On a global scale, Song et al. (2023) report that the pooled prevalence of MCI among older adults exceeds fifteen percent, with substantial regional variation attributable to demographic, educational, and methodological factors. These epidemiological trends underscore the urgency of accurate, accessible, and psychometrically sound cognitive assessment across the full continuum of clinical settings, from primary care offices to specialized memory clinics.

Cognitive functioning is not a unitary construct; rather, it comprises multiple interdependent domains, including episodic memory, working memory, executive function, language, visuospatial processing, attention, and processing speed. Because impairment can be circumscribed to a single domain or distributed across several, and because the pattern of impairment often carries substantial diagnostic significance, no single test can adequately characterize the entirety of an individual's cognitive profile. Clinical practice has therefore evolved a tiered, stepped approach to cognitive assessment. At the first tier, brief bedside screening instruments, administered in a matter of minutes by physicians, nurses, or other trained clinical personnel, serve a triage function: they flag the possible presence of cognitive impairment, provide a coarse index of severity, and inform decisions about further referral. At the second tier, formal neuropsychological testing, conducted by licensed neuropsychologists over several hours, provides a granular, domain-specific profile of cognitive strengths and weaknesses, integrates behavioural observation and collateral history, and supports differential diagnosis, prognosis, and treatment planning with a level of precision that screening instruments cannot achieve.

The clinical stakes of accurate cognitive assessment have grown substantially in recent years with the regulatory approval of disease-modifying therapies for early Alzheimer's disease, including lecanemab (van Dyck et al., 2023) and donanemab (Sims et al., 2023), both of which require biomarker confirmation of amyloid pathology and careful staging of cognitive-functional status prior to treatment initiation. This therapeutic context has elevated the diagnostic precision demanded of the cognitive assessment pathway, reinforcing the necessity of a well-coordinated stepped model that begins with screening and, where indicated, proceeds to formal neuropsychological evaluation and biomarker-based

diagnostic workup (Jack et al., 2024). This paper reviews the evidence base underlying bedside cognitive screening tools and formal neuropsychological testing, situates both within the broader diagnostic and nosological framework of neurocognitive disorders, and proposes an evidence-based clinical algorithm to guide assessment across the continuum of care.

Nosological and Diagnostic Framework of Cognitive Impairment

Before examining specific assessment instruments, it is necessary to situate cognitive testing within the broader diagnostic frameworks that determine how test results are interpreted and applied clinically. Two frameworks predominate in contemporary practice: the syndromic, function-based criteria articulated in the fifth edition, text revision, of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR; American Psychiatric Association, 2022), and the biologically anchored framework developed by the National Institute on Aging and Alzheimer's Association (NIA-AA), most recently revised by Jack et al. (2024).

1 DSM-5-TR Neurocognitive Disorders

The DSM-5-TR reorganizes the older construct of dementia under the broader heading of neurocognitive disorder (NCD), which is subdivided into mild NCD and major NCD (American Psychiatric Association, 2022). This nomenclature, first introduced in DSM-5 and retained in the text revision, reflects an effort to capture the full continuum of cognitive decline, including its earliest clinically meaningful stages, rather than restricting diagnosis to the more severe, function-impairing presentation historically labeled dementia. Diagnosis under this framework requires evidence of cognitive decline from a previous level of performance in one or more of six specified domains, complex attention, executive function, learning and memory, language, perceptual-motor function, and social cognition, established through standardized neuropsychological testing wherever possible, or, when formal testing is unavailable, through clinical assessment. Critically, the DSM-5-TR criteria distinguish mild from major NCD according to the degree of functional impact: major NCD requires that cognitive deficits interfere with independence in everyday activities, necessitating compensation or assistance, whereas mild NCD permits preserved independence, albeit with greater effort, compensatory strategies, or accommodation (American Psychiatric Association, 2022).

Etiological subtyping is appended to the core NCD diagnosis, with specified codes for NCD due to Alzheimer's disease, frontotemporal degeneration, Lewy body disease, vascular disease, traumatic brain injury, substance or medication use, HIV infection, prion disease, Parkinson's disease, Huntington's disease, another medical condition, multiple etiologies, or unknown etiology (American Psychiatric Association, 2022). This etiological specification underscores the centrality of differential

diagnosis to the neurocognitive assessment process and highlights why formal neuropsychological testing, capable of generating domain-specific profiles associated with particular etiologies, occupies an essential role beyond the detection function served by bedside screening.

2 The NIA-AA Biological Framework and 2024 Revised Criteria

In parallel with the DSM's syndromic approach, the NIA-AA has developed a biologically anchored diagnostic framework for Alzheimer's disease that has evolved considerably over the past decade and a half. The 2011 NIA-AA workgroups articulated separate criteria for the preclinical, MCI, and dementia stages of Alzheimer's disease (Albert et al., 2011; McKhann et al., 2011), establishing Alzheimer's disease as a clinical-biomarker construct. Building on the amyloid-tau-neurodegeneration (ATN) framework introduced in 2016, the 2018 NIA-AA research framework operationalized biomarkers for staging Alzheimer's disease across its full continuum, from asymptomatic biomarker positivity through overt dementia.

Most recently, Jack et al. (2024) published revised criteria for the diagnosis and staging of Alzheimer's disease on behalf of an Alzheimer's Association workgroup, incorporating blood-based biomarkers into the classification system for the first time and introducing a two-dimensional biological-clinical staging matrix. A central principle of the revised criteria is that Alzheimer's disease should be defined biologically rather than by clinical syndrome alone, consistent with practice in other areas of medicine such as oncology (Jack et al., 2024). Under this framework, the disease is conceptualized as a continuum, beginning with amyloid deposition in asymptomatic individuals (biological stage A, clinical stage 1), progressing through tau deposition and subtle cognitive impairment (biological stage B, clinical stage 2), and culminating in progressive neurodegeneration and dementia. Application of these criteria in drug development and clinical practice contexts has been discussed extensively, with commentators noting that although the criteria offer objective, biologically grounded measures for diagnosis and staging, their implementation depends heavily on the availability and accuracy of biomarkers, particularly increasingly accessible plasma-based assays (Jack et al., 2024/2025). The 2024 criteria have generated active discussion regarding their applicability to non-Alzheimer's dementia syndromes, including dementia with Lewy bodies, for which the relationship between amyloid co-pathology and clinical staging remains an area of ongoing investigation (Abdelnour, 2025).

The coexistence of the syndromic DSM-5-TR framework and the biological NIA-AA framework has direct implications for cognitive testing. Bedside screening and formal neuropsychological testing remain the primary instruments through which the syndromic, functional dimension of diagnosis is established, while biomarker studies (amyloid positron emission tomography, cerebrospinal fluid analysis, and increasingly, blood-based

biomarkers) establish the biological dimension. The two are complementary: a biologically positive amyloid scan in a cognitively normal individual does not, by itself, constitute a diagnosis of Alzheimer's dementia in the clinical sense, and a diagnosis of major NCD requires demonstration of functional impairment that biomarkers alone cannot establish (Shaughnessy & Weintraub, 2024).

Table 1. DSM-5-TR Diagnostic Criteria: Mild versus Major Neurocognitive Disorder

Criterion	Mild Neurocognitive Disorder	Major Neurocognitive Disorder
Cognitive decline	Modest decline from previous level in one or more cognitive domains, based on concern of the individual, informant, or clinician, and objective evidence of impairment (typically test performance in the 1–2 SD below appropriate norms range)	Significant decline from previous level in one or more cognitive domains, with objective impairment typically 2 or more SD below appropriate norms
Functional impact	Cognitive deficits do not interfere with capacity for independence in everyday activities (e.g., instrumental activities of daily living remain intact, though greater effort or compensatory strategies may be required)	Cognitive deficits interfere with independence in everyday activities, necessitating assistance
Delirium exclusion	Deficits do not occur exclusively in the context of delirium	Deficits do not occur exclusively in the context of delirium
Other mental disorder exclusion	Deficits not better explained by another mental disorder (e.g., major depressive disorder, schizophrenia)	Deficits not better explained by another mental disorder
Etiological specifier	Specified where possible (e.g., due to Alzheimer's disease, due to Lewy body disease, due to vascular disease, due to unknown etiology)	Specified where possible, with additional severity specifier (mild, moderate, severe) based on functional dependence
Typical assessment basis	Standardized neuropsychological testing preferred; clinical assessment when testing unavailable	Standardized neuropsychological testing preferred; clinical assessment when testing unavailable

Note. Adapted from diagnostic criteria described in American Psychiatric Association (2022).

Bedside Cognitive Screening Tools

Bedside cognitive screening instruments occupy the first tier of the stepped cognitive assessment pathway. They are characterized by brevity (typically five to fifteen minutes of administration time), minimal training requirements, and applicability across a wide range of clinical settings, including primary care offices, emergency departments, hospital wards, and community health screenings. The American Academy of Neurology practice guideline update specifies that when screening or assessing for MCI, clinicians should employ validated assessment tools, noting that a range of instruments demonstrate acceptable diagnostic accuracy without clear superiority of any single instrument (Petersen et al., 2018). Critically, the guideline emphasizes that because brief cognitive screening instruments are typically calibrated to maximize sensitivity at some cost to specificity, patients who screen positive should undergo further evaluation, potentially including formal neuropsychological testing with normative interpretation, before a diagnosis is confirmed (Petersen et al., 2018). This principle, that screening identifies candidates for further evaluation rather than establishing a definitive diagnosis, constitutes the conceptual foundation of the entire stepped assessment model and is a recurring theme throughout the evidence reviewed in this section.

1 The Mini-Mental State Examination (MMSE)

The MMSE, originally developed by Folstein, Folstein, and McHugh (1975), remains among the most extensively studied and internationally disseminated cognitive screening instruments in clinical use. The instrument comprises eleven items assessing orientation to time and place, registration of three words, attention and calculation (commonly operationalized through serial subtraction), delayed recall of the three registered words, language (naming, repetition, comprehension of a three-stage command, reading, writing), and visuoconstructional praxis (intersecting pentagon copy), yielding a total score ranging from zero to thirty points (Folstein et al., 1975).

The diagnostic accuracy of the MMSE has been the subject of extensive meta-analytic scrutiny. A Cochrane systematic review examining the MMSE for detection of dementia in community and primary care populations aged sixty-five and older found that, at a cutoff of twenty-four, the instrument achieved pooled sensitivity of approximately 0.85 and specificity of approximately 0.90 in community settings; at a cutoff of twenty-five, pooled sensitivity rose to approximately 0.87 with reduced specificity of approximately 0.82 (Creavin et al., 2016). Notably, when accuracy estimates were adjusted for level of education, sensitivity increased to approximately 0.97 while specificity fell to approximately 0.70, illustrating the substantial confounding influence of educational attainment on MMSE performance (Creavin et al., 2016). A separate Cochrane review examining the MMSE's utility in predicting conversion from MCI to dementia found highly variable accuracy, with sensitivities for conversion to Alzheimer's disease dementia ranging from twenty-seven to eighty-nine percent and specificities ranging from thirty-

two to ninety percent across included studies, leading the reviewers to conclude that there is insufficient evidence to support a substantial role for the MMSE as a stand-alone, single-administration predictor of conversion from MCI to dementia (Arevalo-Rodriguez et al., 2021).

Recent work has further probed the limitations of relying on the MMSE in isolation. Kim et al. (2025), employing a deep learning classifier trained on clinical data, demonstrated that a model relying solely on MMSE scores achieved an area under the receiver operating characteristic curve of 0.872 for predicting Clinical Dementia Rating category, whereas a composite model incorporating additional cognitive assessments achieved a significantly higher area under the curve of 0.934, indicating meaningfully improved diagnostic precision when the MMSE is supplemented with additional cognitive measures (Kim et al., 2025). The same investigation raised concerns regarding dementia overdiagnosis among younger, more highly educated individuals when the MMSE is used as the sole diagnostic instrument, reinforcing longstanding concerns regarding the interaction between educational attainment and MMSE performance (Kim et al., 2025).

Beyond its established limitations with respect to education and cultural background, the MMSE has been repeatedly criticized for relatively poor sensitivity to executive dysfunction and to the subtle deficits characteristic of MCI, for pronounced ceiling effects among highly educated respondents, and for the historical restriction of its use owing to copyright licensing requirements, a factor that contributed substantially to the increasing clinical adoption of alternative, freely available instruments such as the MoCA (Nasreddine et al., 2005).

2 The Montreal Cognitive Assessment (MoCA)

The MoCA was developed by Nasreddine et al. (2005) specifically to address the MMSE's reduced sensitivity to MCI. The instrument evaluates eight cognitive domains, visuospatial and executive function (via a trail-making task, a cube copy, and a clock-drawing task), naming, memory (five-word learning across two trials with delayed recall), attention (digit span, target detection, serial subtraction), language (sentence repetition and phonemic fluency), abstraction, delayed recall, and orientation, within a total administration time of approximately ten to fifteen minutes, yielding a score out of thirty points, with one additional point awarded to individuals with twelve years of education or fewer to partially correct for educational bias (Nasreddine et al., 2005).

The originally proposed optimal cutoff of twenty-six has been substantially revised by subsequent meta-analytic evidence. Islam et al. (2023), in a systematic review and meta-analysis pooling thirteen studies encompassing 2,158 participants, found that a cutoff of less than twenty-three maximized the combined sensitivity and specificity for detecting MCI, yielding pooled sensitivity of 73.5% and specificity of 91.3%; at the traditionally recommended cutoff of less than twenty-six, sensitivity rose to 83.8% but specificity fell to 70.8%, reflecting a substantial trade-off

between case detection and false-positive classification (Islam et al., 2023). These findings are broadly consistent with earlier systematic reviews and diverge notably from the original 2005 validation study, a discrepancy the authors attribute in part to differences in sample size and population characteristics between the original validation cohort and subsequent independent replications (Islam et al., 2023).

The MoCA's cross-cultural and cross-linguistic performance has received considerable recent attention. Systematic reviews addressing the instrument's validity in culturally diverse populations, including Spanish-speaking populations of Latin America and the Caribbean, have documented substantial heterogeneity in reported cutoffs, translation and adaptation practices, and psychometric reporting quality, with educational level consistently identified as a source of bias across populations studied (Cornea, Simu, & Rosca, 2022). The MoCA has also been extended well beyond its original purpose of detecting MCI in older adults to encompass cognitive screening in numerous other neurological and psychiatric populations, including subarachnoid hemorrhage, where systematic review evidence indicates good internal consistency and sensitivity to cognitive impairment in the subacute and chronic post-hemorrhage phases (Cornea, Simu, & Rosca, 2022).

Comparative studies incorporating remote and digital assessment modalities have further situated the MoCA within the evolving diagnostic landscape. Thompson et al. (2023) compared remote smartphone-based cognitive tasks, in-clinic digital tablet-based assessments, and the MoCA in their ability to distinguish cerebral amyloid status among cognitively healthy older adults, finding that certain digital episodic memory tasks achieved discriminative accuracy modestly exceeding that of the MoCA (area under the curve of 0.77 versus 0.74, respectively), suggesting that digital cognitive assessment may offer incremental sensitivity to very early, preclinical stages of Alzheimer's pathology beyond what conventional bedside screening captures (Thompson et al., 2023).

3 The Mini-Cog

The Mini-Cog, developed by Borson, Scanlan, Chen, and Ganguli (2003), combines a three-item delayed word recall task with a clock-drawing test, yielding administration times of approximately three to five minutes, making it among the briefest validated cognitive screening instruments in widespread use. In its original validation study, conducted within a population-based epidemiological sample, the Mini-Cog demonstrated sensitivity and specificity for dementia detection (76% and 89%, respectively) comparable to the MMSE at a cutpoint of twenty-five (79% and 88%, respectively) and to a conventional neuropsychological battery (75% and 90%, respectively) (Borson et al., 2003). However, a subsequent Cochrane systematic review evaluating the Mini-Cog for dementia detection in primary care settings identified substantial heterogeneity across studies and methodological limitations in all but one included investigation; the single low-risk-of-bias

study reported sensitivity of 76% and specificity of 73%, leading the Cochrane reviewers to conclude that there is presently insufficient evidence to routinely recommend the Mini-Cog as a standalone dementia screening test in primary care, notwithstanding its practical appeal (Seitz et al., 2021). The instrument's brevity nonetheless renders it a pragmatic option in time-constrained or resource-limited settings, provided that positive screens are followed by further evaluation.

4 The Saint Louis University Mental Status (SLUMS) Examination

The SLUMS examination, developed within the Veterans Affairs system, is an eleven-item, thirty-point instrument assessing orientation, memory, attention, and executive function through tasks including animal naming, digit span, figure recognition, clock drawing, and size differentiation, with administration requiring approximately seven minutes (Spencer, Noyes, Bair, & Ransom, 2022). The instrument was developed in part to address perceived limitations of the MMSE, including its heavy reliance on orientation items and its comparatively limited sensitivity among more highly educated examinees. A systematic review of the SLUMS's psychometric properties concluded that the instrument demonstrates adequate validity, sensitivity, and specificity for detecting cognitive impairment, while noting the continued need for more diverse normative samples and further test-retest reliability data before its use can be recommended without qualification (Spencer et al., 2022). Reported cutoffs, adjusted for educational attainment, generally classify scores of twenty-seven to thirty as within normal limits, twenty-one to twenty-six as suggestive of mild neurocognitive disorder, and scores at or below twenty as suggestive of major neurocognitive disorder, though sensitivity and specificity at these cutoffs vary appreciably across studies and populations (Spencer et al., 2022). A notable practical advantage of the SLUMS relative to the MMSE and MoCA is that it resides in the public domain and may be administered without licensing fees or mandatory training modules, a consideration of increasing relevance as healthcare systems seek to expand access to cognitive screening. Recent comparative work has begun to validate digital adaptations against the SLUMS in real-world clinical samples, reporting concurrent validity supportive of digital platforms as a scalable complement to traditional paper-based administration (Huynh, Patterson, & Huang, 2025).

5 Culture-Fair and Education-Independent Screening Instruments

A substantial and growing body of literature has documented that widely used cognitive screening instruments, including the MMSE and MoCA, are subject to systematic bias related to educational attainment, literacy, language proficiency, and cultural background. A comprehensive systematic review examining culture-fair cognitive screening tools identified eleven distinct instruments developed specifically to minimize such bias, including the Rowland Universal Dementia Assessment Scale (RUDAS), the Community Screening Instrument for Dementia (CSI-D), and the Visual Cognitive Assessment

Test (VCAT) (Chithiramohan et al., 2024). The review found that the RUDAS achieved diagnostic accuracy comparable to the MMSE (area under the curve ranging from 0.62 to 0.93) while demonstrating superior specificity, and that the RUDAS, CSI-D, and VCAT were minimally influenced by educational level, with the VCAT proving particularly useful for non-verbal patients or those with limited education or language proficiency (Chithiramohan et al., 2024). These findings carry direct implications for clinical practice in linguistically and culturally diverse populations, where reliance on standard MMSE or MoCA cutoffs without adaptation risks both under- and over-diagnosis of cognitive impairment.

6 Delirium Screening: The Confusion Assessment Method and Related Instruments

Although not a general cognitive screening tool in the conventional sense, delirium detection instruments occupy an essential adjunctive role within the bedside cognitive assessment armamentarium, as unrecognized delirium can profoundly confound the interpretation of cognitive screening results obtained in acute care settings. The Confusion Assessment Method (CAM), developed by Inouye and colleagues (1990), remains the most widely used and internationally validated delirium screening algorithm, evaluating four cardinal features: acute onset and fluctuating course, inattention, disorganized thinking, and altered level of consciousness. Diagnosis of delirium by CAM requires the presence of the first two features plus either of the remaining two (Inouye et al., 1990). Comparative studies of ultrabrief delirium screening instruments, including the 4 'A's Test (4AT) and modified CAM protocols adapted for emergency department use, have generally found favorable diagnostic accuracy and rapid administration times suitable for use by non-specialist personnel (Muser et al., 2022). Large-scale implementation research indicates that structured protocols can substantially increase delirium screening rates in hospital settings; one quality improvement study documented an increase in the proportion of hospital admissions screened for delirium from 74% to 98.7% over a four-year period following protocol implementation, although adherence to recommended daily screening frequency remained limited even after implementation, illustrating persistent challenges in translating validated instruments into consistent bedside practice (Penfold et al., 2024).

7 Digital and Remote Cognitive Screening

The proliferation of digital health technologies has substantially expanded the modalities through which bedside cognitive screening can be conducted, with particular acceleration following the COVID-19 pandemic's disruption of in-person clinical services. Owens et al. (2020) documented the rapid implementation of remote memory clinic services during and after the pandemic, demonstrating the feasibility of adapting cognitive assessment pathways to telehealth delivery. Subsequent validation research has examined the reliability of self-administered, remotely delivered digital cognitive assessments relative to clinician-administered or in-person

equivalents. Kochan et al. (2022) evaluated the reliability, validity, and user experience of remote unsupervised computerized neuropsychological assessments among community-living older adults, providing empirical support for the feasibility of unsupervised digital testing in this population. Morrissey, Gillings, and Hornberger (2024) similarly reported the feasibility and reliability of online relative to in-person cognitive testing among healthy older adults. Comparative work examining agreement between self-administered and research-coordinator-administered digital assessments has generally found moderate to good agreement, with intraclass correlation coefficients ranging from approximately 0.59 to 0.83 across cognitive domains assessed, and no statistically significant difference in performance between self- and coordinator-administered conditions after accounting for testing order and device type.

Digital assessment platforms have also demonstrated clinical utility beyond the detection of age-related cognitive decline. Shibata et al. (2024) applied a remote digital cognitive assessment battery to detect cognitive deficits associated with hippocampal atrophy in patients with autoimmune limbic encephalitis, finding significant impairment across memory, visuospatial, executive, and language domains relative to healthy controls, with global digital assessment scores correlating significantly with established neuropsychological measures. In oncology populations, Root et al. (2023) established adequate test-retest reliability and internal validity for a remote, computerized neurocognitive battery designed to assess breast cancer- and treatment-related cognitive dysfunction. In low-resource settings, Paddick et al. (2021) demonstrated the feasibility and effectiveness of app-based cognitive screening administered by lay health workers for dementia detection in rural Tanzania, illustrating the potential for digital tools to extend screening capacity into settings historically underserved by specialist neuropsychological services. Collectively, this body of evidence suggests that digital and remote cognitive screening modalities are approaching, and in some domains exceeding, the diagnostic performance of traditional paper-based bedside instruments, while simultaneously expanding access to underserved populations; however, questions regarding longitudinal practice effects, equity of access to requisite technology, and standardization of digital normative data remain areas of active investigation.

Table 2. Comparative Diagnostic Accuracy of Selected Bedside Cognitive Screening Instruments

Instrument	Target Condition	Reported Cutoff	Sensitivity	Specificity	Key Source
MMSE	Dementia (community)	<24	85%	90%	Creavin et al. (2016)
MMSE	Dementia (community)	<25	87%	82%	Creavin et al. (2016)
MoCA	MCI	<23	73.5%	91.3%	Islam et al. (2023)
MoCA	MCI	<26 (original)	83.8%	70.8%	Islam et al. (2023)
Mini-Cog	Dementia (population-based)	Algorithm-based	76%	89%	Borson et al. (2003)
Mini-Cog	Dementia (primary care, low-bias study)	Algorithm-based	76%	73%	Seitz et al. (2021)
SLUMS	Major NCD	≤20	93–98%	96–100%	Spencer et al. (2022)
SLUMS	Mild NCD	≤26	74–95%	65–76%	Spencer et al. (2022)
RUDAS	Dementia (culturally diverse samples)	Variable	52–94%	70–98%	Chithiramohan et al. (2024)

Note. Reported values reflect pooled or representative estimates from the cited sources; substantial heterogeneity exists across populations, settings, and study quality, and individual studies within each meta-analysis or review may report divergent estimates. NCD = neurocognitive disorder.

Formal Neuropsychological Testing

Formal neuropsychological testing represents the second, more intensive tier of the cognitive assessment pathway, distinguished from bedside screening by its administration by licensed neuropsychologists, its considerably greater time investment (typically two to eight hours, occasionally distributed across multiple sessions), its reliance on extensively normed and standardized instruments, and its capacity to generate a detailed, domain-specific cognitive profile rather than a single composite score. Shaughnessy and Weintraub (2024), summarizing the Alzheimer's Association's Diagnostic Evaluation, Testing, Counseling and Disclosure (DETeCD-ADRD) clinical practice guideline, emphasize that neuropsychological evaluation supports early and accurate diagnosis and staging, characterization of the individual clinical profile, longitudinal tracking of trajectory, and the formulation of recommendations tailored to the specific needs of the patient and care team. Formal testing is generally indicated when bedside screening results are ambiguous or discordant with clinical presentation, when differential diagnosis among competing etiologies is required, when a detailed baseline is needed prior to initiation of disease-modifying therapy, or when functional capacity must be characterized for legal, occupational, or caregiving purposes.

1 Domains Assessed and Representative Instruments

Comprehensive neuropsychological assessment systematically evaluates multiple cognitive domains using standardized, psychometrically validated instruments with published age-, education-, and in some cases sex- and race-adjusted normative data (Lezak, Howieson, Bigler, & Tranel, 2012). Attention and processing speed are commonly assessed using measures such as the Trail Making Test Part A and Digit Symbol/Symbol Digit Modalities Test. Executive functioning is assessed through instruments including the Trail Making Test Part B, the Wisconsin Card Sorting Test, and the Stroop Color-Word Test; the Wisconsin Card Sorting Test in particular has demonstrated utility in characterizing executive dysfunction following traumatic brain injury (Gómez-de-Regil, 2020). Memory is evaluated through list-learning paradigms such as the California Verbal Learning Test and the Rey Auditory Verbal Learning Test, as well as through visual memory measures such as the Rey-Osterrieth Complex Figure Test, which additionally captures visuoconstructional and organizational abilities. Language is assessed through confrontation naming tasks such as the Boston Naming Test and through verbal fluency paradigms. General intellectual functioning is typically characterized using a Wechsler intelligence scale, and mood and personality functioning are frequently assessed using instruments such as the Beck Depression Inventory or the Minnesota Multiphasic Personality Inventory, given the well-documented potential for affective disorders to confound the interpretation of cognitive test performance (Lezak et al., 2012).

A defining methodological feature of formal neuropsychological assessment, relative to bedside screening, is the integration of standardized psychometric data with detailed clinical history, behavioural observation during testing, and collateral informant report. This integrative approach allows the examining neuropsychologist to establish an individualized expected-performance profile based on the patient's demographic characteristics and premorbid functioning, and subsequently to evaluate whether observed performance falls below that expected level in a pattern consistent with a specific neurological or psychiatric etiology.

2 Diagnostic Frameworks: Conventional versus Neuropsychological (Actuarial) Criteria for MCI

A substantial body of recent research has examined the comparative diagnostic performance of conventional (Petersen/Winblad) MCI criteria, which rely primarily on subjective cognitive complaint corroborated by a single below-threshold cognitive test score, against more comprehensive neuropsychological (actuarial, or Jak/Bondi) criteria, which require impairment across multiple measures within a cognitive domain or across multiple domains before a diagnosis of MCI is assigned (Jak et al., 2009). Kim, Kim, Edmonds, and Bondi (2024) directly compared these two diagnostic approaches within a clinical sample, finding 87.82% concordance between conventional and neuropsychological diagnoses overall, but identifying a clinically meaningful subgroup of individuals classified as impaired under conventional criteria who were reclassified as cognitively normal under the more stringent, multi-measure neuropsychological criteria. This discordant subgroup exhibited higher depression scores and, counterintuitively, better objective memory and executive function performance than those classified as impaired under both systems, suggesting that reliance on conventional single-test criteria alone may lead to overdiagnosis of MCI, particularly among individuals whose single below-threshold score reflects normal test variability, transient affective symptoms, or measurement error rather than genuine incipient neurodegeneration (Kim et al., 2024). These findings have direct implications for the interpretation of both bedside screening results, which are inherently limited to a small number of measures per domain, and formal neuropsychological batteries, which by virtue of their breadth are better positioned to apply actuarial, multi-measure diagnostic criteria and thereby reduce misclassification.

3 Performance Validity Testing

A methodological feature unique to formal neuropsychological assessment, largely absent from bedside screening protocols, is the systematic evaluation of performance validity—that is, the degree to which obtained test results credibly reflect the examinee's genuine cognitive capacity rather than volitional underperformance, symptom exaggeration, or insufficient engagement with testing. Sherman, Slick, and Iverson (2020), in a twenty-year update of the influential Malingered Neurocognitive Dysfunction criteria, articulate a multidimensional framework for classifying the credibility

of neuropsychological data, incorporating evidence from dedicated performance validity tests, atypical patterns of performance on standard clinical measures, and behavioral observations made during the evaluation. Performance validity assessment has assumed particular prominence in forensic, disability, and personal injury evaluation contexts, where the potential for secondary gain necessitates careful evaluation of response validity before test results can be considered diagnostically meaningful (Bosso et al., 2022; Panico, Fonzo, Verde, Lancia, & Trojano, 2025).

Contemporary research in this domain has emphasized both the refinement of existing performance validity instruments and the development of novel measures designed to resist coaching and detection avoidance. Panico et al. (2025) describe the development of the Test of Malingering in Abstraction Skills, designed to balance sufficient apparent difficulty to avoid overt detection as an effort measure against sufficient actual ease to remain insensitive to genuine cognitive impairment, illustrating the ongoing psychometric sophistication required in this subfield. Finley (2024) further discusses the emerging role of digital technology, including computerized adaptive testing and artificial intelligence-assisted analysis, in enhancing the sensitivity and efficiency of performance validity assessment, an area of active methodological development. The integration of performance validity testing into standard neuropsychological practice represents a critical safeguard against both false-positive and false-negative diagnostic conclusions and constitutes one of the most significant methodological distinctions between formal neuropsychological assessment and bedside cognitive screening, which generally lacks any analogous validity-checking mechanism.

4 Differential Diagnosis Across Dementia Subtypes

A principal clinical contribution of formal neuropsychological testing lies in its capacity to generate domain-specific cognitive profiles that assist in differentiating among the major neurodegenerative dementia syndromes, each of which is associated with a relatively characteristic, though not invariant, pattern of cognitive impairment. Alzheimer's disease dementia typically presents with early and prominent impairment of episodic memory, particularly delayed recall and recognition, accompanied by relatively later involvement of executive, language, and visuospatial domains as the disease progresses, consistent with the biologically staged model articulated by Jack et al. (2024). Dementia with Lewy bodies, whose diagnostic consensus criteria were most recently updated in the fourth report of the DLB Consortium (McKeith et al., 2017), is characterized by prominent deficits in attention, executive function, and visuospatial processing, often with relatively preserved episodic memory in earlier stages, accompanied by core clinical features including fluctuating cognition, recurrent visual hallucinations, rapid eye movement sleep behavior disorder, and spontaneous parkinsonism. Behavioural variant frontotemporal dementia, for which revised diagnostic criteria were established by Rascovsky et al. (2011), is characterized primarily by early and prominent

behavioural and personality change alongside executive dysfunction, with relative sparing of episodic memory and visuospatial function in earlier stages, distinguishing it neuropsychologically from typical Alzheimer's disease presentations. The primary progressive aphasia syndromes, classified according to the framework established by Gorno-Tempini et al. (2011), are distinguished by predominant, isolated language impairment as the presenting and most prominent feature, with further subclassification into semantic, nonfluent/agrammatic, and logopenic variants based on the specific pattern of linguistic breakdown. Vascular cognitive impairment, for which consensus classification criteria were articulated by Skrobot et al. (2017), typically presents with prominent executive dysfunction and processing speed impairment, often with a stepwise rather than gradually progressive course and a temporal relationship to cerebrovascular events, distinguishing it from the more insidious onset characteristic of primary neurodegenerative dementias.

Table 3. Characteristic Neuropsychological Profiles across Major Dementia Subtypes

Dementia Subtype	Diagnostic Framework	Predominant Early Cognitive Deficits	Relatively Spared Domains (Early Stage)	Key Reference
Alzheimer’s disease	NIA-AA 2024 revised criteria	Episodic memory (encoding/delayed recall)	Executive function, language (until later stages)	Jack et al. (2024)
Dementia with Lewy bodies	DLB Consortium 4th consensus report	Attention, executive function, visuospatial processing	Episodic memory (relatively)	McKeith et al. (2017)
Behavioural variant frontotemporal dementia	Rascovsky revised criteria	Executive function, behaviour/personality change	Episodic memory, visuospatial function	Rascovsky et al. (2011)
Primary progressive aphasia (semantic, nonfluent, logopenic variants)	Gorno-Tempini classification	Language (variant-specific pattern)	Non-linguistic domains (until later stages)	Gorno-Tempini et al. (2011)
Vascular cognitive impairment	VICCCS consensus classification	Executive function, processing speed	Variable; often stepwise course	Skrobot et al. (2017)

Note. Profiles reflect characteristic, group-level patterns; substantial individual variability and mixed pathology are common in clinical practice. NIA-AA = National Institute on Aging–Alzheimer’s Association; DLB = dementia with Lewy bodies; VICCCS = Vascular Impairment of Cognition Classification Consensus Study.

5 Formal Neuropsychological Testing in Special Populations

Formal neuropsychological assessment plays a particularly critical role in populations for whom bedside screening instruments are known to be insensitive or otherwise inadequate. In traumatic brain injury (TBI), neuropsychological evaluation has been characterized as the best available predictor of functional outcome, given its capacity to characterize the specific pattern and severity of cognitive sequelae across attention, processing speed, executive function, and memory domains (Torregrossa et al., 2023). Torregrossa et al. (2023), in a comprehensive review with clinical recommendations, note that general cognitive screening instruments such as the MMSE may be appropriately administered during the acute and subacute phases following TBI, while more specific instruments, including the Wisconsin Card Sorting Test and Trail Making Test, are recommended for detailed characterization of executive function during postacute and chronic phases, allowing clinicians to tailor rehabilitation strategies to the individual's specific cognitive profile. Repeat neuropsychological assessment is common in TBI rehabilitation contexts to track recovery trajectory, though clinicians must account for practice effects, which can artificially inflate scores upon repeat administration and are typically mitigated through the use of alternate, psychometrically equivalent test forms. Cluster-analytic approaches applied to large samples of patients with acquired brain injury undergoing cognitive rehabilitation have identified distinct neuropsychological subgroups characterized by differential performance profiles, informing more individualized approaches to web-based cognitive rehabilitation programming (García-Rudolph et al., 2021).

Formal testing is likewise essential in populations characterized by low educational attainment, limited literacy, or cultural and linguistic backgrounds divergent from those in which standard cognitive instruments were normed and validated, given the demonstrated propensity of screening instruments to misclassify such individuals (Chithiramohan et al., 2024). In such populations, formal neuropsychological evaluation, particularly when informed by culturally and linguistically appropriate normative data, allows for more individualized interpretation than a fixed screening cutoff can provide. Similarly, in psychiatric populations, where cognitive complaints may reflect a complex interplay of primary cognitive impairment, medication effects, and mood or psychotic symptomatology, formal testing that incorporates concurrent assessment of affective and psychotic symptoms, along with performance validity indices, provides diagnostic clarity that brief screening cannot achieve. Recent work examining working memory precision and associative binding among individuals with MCI illustrates the continuing refinement of specialized neuropsychological paradigms capable of detecting subtle cognitive alterations not captured by conventional screening or even standard neuropsychological batteries (Jia, Woltering, Deutz, et al., 2023).

Clinical Algorithm for Cognitive Assessment

Synthesizing the evidence reviewed above, a stepwise clinical algorithm can be constructed to guide clinicians through the cognitive assessment pathway, from initial identification of concern through formal diagnosis and, where appropriate, biomarker confirmation. Such an algorithm is broadly consistent with the approach recommended in the American Academy of Neurology practice guideline (Petersen et al., 2018) and the Alzheimer's Association DETeCD-ADRD clinical practice guideline (Shaughnessy & Weintraub, 2024), both of which emphasize a graduated approach beginning with case identification and proceeding through progressively more specialized levels of evaluation as clinically indicated.

Table 4. Stepwise Clinical Algorithm for Cognitive Assessment

Step	Clinical Action	Instruments/Methods	Decision Point
1. Case identification	Elicit subjective cognitive concern from patient, informant, or clinical observation; review health risk assessment	Clinical interview; standardized health risk assessment tools	If concern present, proceed to Step 2; if absent, routine monitoring
2. Delirium/acute confounder exclusion (if applicable acute setting)	Rule out delirium or acute confounding medical condition before interpreting cognitive test results	Confusion Assessment Method (CAM); 4AT; medical workup	If delirium present, treat underlying cause before further cognitive testing; if absent, proceed to Step 3
3. Bedside cognitive screening	Administer brief, validated screening instrument appropriate to setting, education, and cultural/linguistic background	MMSE, MoCA, Mini-Cog, SLUMS, or culture-fair alternative (e.g., RUDAS)	Score within normal range: routine monitoring/reassessment interval; score suggestive of impairment: proceed to Step 4
4. Functional assessment	Evaluate impact of cognitive symptoms on instrumental and basic activities of daily living	Structured functional/IADL questionnaires; informant interview	Functional independence preserved: consistent with mild NCD/MCI; functional impairment present: consistent with major NCD
5. Medical and laboratory workup	Exclude reversible or contributory medical causes (e.g., thyroid dysfunction, vitamin B12 deficiency, depression, medication effects)	Laboratory studies; mood screening; medication review; structural neuroimaging	Reversible cause identified and treated: reassess cognition after treatment; no reversible cause: proceed to Step 6
6. Referral for formal neuropsychological testing	Refer when screening results are ambiguous, discordant with function, or detailed differential diagnosis is required	Comprehensive neuropsychological battery with performance validity testing	Profile consistent with specific etiology: proceed to targeted biomarker workup; diagnosis unclear: consider longitudinal reassessment
7. Biomarker and etiological workup	Confirm underlying biological etiology, particularly prior to consideration of disease-modifying therapy	Amyloid/tau PET, CSF biomarkers, blood-based biomarkers, structural/functional neuroimaging	Biomarker-confirmed Alzheimer's pathology with clinical impairment: eligible for staged diagnosis and treatment planning per NIA-AA criteria
8. Diagnosis, staging, and disclosure	Integrate clinical, cognitive, functional, and biomarker data into a formal diagnosis and staging	DSM-5-TR criteria; NIA-AA 2024 biological-clinical staging matrix	Diagnosis disclosed with counseling; care/treatment plan established

Note. Algorithm synthesized from recommendations in Petersen et al. (2018), Shaughnessy and Weintraub (2024), and Jack et al. (2024). IADL = instrumental activities of daily living; PET = positron emission tomography; CSF = cerebrospinal

Comparative Analysis: Bedside Screening versus Formal Neuropsychological Testing

Although bedside screening and formal neuropsychological testing are frequently discussed in tandem, they differ fundamentally across nearly every methodological and clinical dimension, and an appreciation of these differences is essential to appropriate instrument selection. Bedside screening instruments are optimized for speed, accessibility, and

broad applicability across non-specialist clinical settings, at the cost of diagnostic granularity and, frequently, specificity. Formal neuropsychological testing sacrifices brevity and accessibility in exchange for diagnostic precision, domain-specific characterization, integration of performance validity assessment, and the capacity to support differential diagnosis among etiologically distinct conditions that may present with superficially similar screening results. Table 5 summarizes these distinctions across key dimensions of clinical relevance.

Table 5. Comparative Summary: Bedside Cognitive Screening versus Formal Neuropsychological Testing

Dimension	Bedside Screening	Formal Neuropsychological Testing
Typical duration	5–15 minutes	2–8+ hours, potentially multiple sessions
Administering clinician	Physicians, nurses, medical assistants, trained lay health workers	Licensed neuropsychologists (doctoral-level training)
Primary purpose	Case detection, triage, longitudinal monitoring	Diagnosis, differential diagnosis, domain-specific profiling, capacity evaluation
Output	Single (or few) composite score(s)	Multidimensional profile across cognitive, behavioural, and emotional domains
Sensitivity to MCI	Variable; generally moderate, improved with lowered cutoffs	High, particularly with actuarial/multi-measure criteria
Specificity	Variable; often reduced when sensitivity maximized	Generally higher due to comprehensive, multi-measure assessment
Performance validity assessment	Not routinely incorporated	Integral component of standard practice
Cultural/educational adaptability	Limited unless culture-fair instrument specifically selected	Higher, given individualized normative comparison and clinical judgment
Accessibility/cost	Low cost; widely accessible, including via digital/remote platforms	Higher cost; limited by neuropsychologist availability, particularly in underserved regions
Role in biomarker-era diagnosis (e.g., anti-amyloid therapy eligibility)	Supports initial identification of candidates for further workup	Provides detailed functional-cognitive staging required alongside biomarker confirmation

Note. Synthesized from evidence reviewed in Sections 3 and 4, including Petersen et al. (2018), Lezak et al. (2012), and Shaughnessy and Weintraub (2024).

This comparative analysis reinforces the conclusion that screening and formal testing serve complementary rather than competing functions. Screening instruments efficiently triage large populations, directing limited specialist neuropsychological resources toward those individuals most likely to benefit from comprehensive evaluation, while formal testing provides the diagnostic precision necessary for individualized clinical decision-making, particularly in cases of diagnostic ambiguity, atypical presentation, or high-stakes treatment decisions such as eligibility for disease-modifying anti-amyloid therapy.

Emerging Directions in Cognitive Testing

The field of cognitive testing is undergoing rapid transformation, driven by three converging developments: the maturation of digital and remote assessment technologies, the emergence of blood-based biomarkers for Alzheimer's disease pathology, and the increasing application of artificial intelligence and machine learning methods to cognitive test interpretation.

The proliferation of digital cognitive assessment platforms, reviewed extensively in Section 3.7, is increasingly extending beyond simple digitization of traditional paper-based instruments toward the development of novel digital biomarkers with potentially superior sensitivity to preclinical and prodromal disease states. Thompson et al. (2023) demonstrated that certain digital cognitive tasks can achieve discriminative accuracy for amyloid status exceeding that of conventional bedside screening, while ecological momentary assessment and ambulatory, smartphone-based monitoring approaches allow for repeated, longitudinal characterization of cognitive fluctuation in naturalistic settings, a capability entirely absent from single-timepoint bedside or laboratory-based assessment. Continued methodological development is needed to establish robust normative datasets for digital instruments across diverse populations, to characterize practice effects arising from repeated remote administration, and to ensure equitable access to the technological infrastructure such assessments require, particularly among older adults and individuals in low-resource settings (Paddick et al., 2021).

The introduction of blood-based biomarkers into the revised NIA-AA diagnostic criteria (Jack et al., 2024) represents a further inflection point in the cognitive assessment pathway. Whereas biomarker confirmation of Alzheimer's pathology has historically required either positron emission tomography imaging or cerebrospinal fluid analysis via lumbar puncture, both of which are costly, invasive, or of limited availability outside specialized centers, blood-based biomarkers offer the prospect of substantially more accessible and scalable biological confirmation. Jack et al. (2024/2025), discussing the application of the revised criteria in drug development and clinical practice, note that improving the availability and accuracy of plasma biomarkers is expected to substantially enhance the

practical applicability of biologically anchored diagnostic criteria, particularly as disease-modifying therapies such as lecanemab (van Dyck et al., 2023) and donanemab (Sims et al., 2023) become more widely available and require biomarker-confirmed diagnosis prior to treatment initiation. The integration of blood-based biomarkers with cognitive testing data, whether obtained through bedside screening or formal neuropsychological assessment, will likely become an increasingly central feature of the diagnostic pathway over the coming years, potentially altering the sequencing and relative emphasis placed on cognitive versus biological data in future iterations of clinical practice guidelines.

Finally, artificial intelligence and machine learning methods are increasingly being applied both to enhance the diagnostic yield of existing cognitive instruments and to develop novel assessment paradigms. Kim et al. (2025) demonstrated that machine learning models incorporating MMSE data alongside additional cognitive measures achieve meaningfully improved diagnostic accuracy relative to MMSE scores considered in isolation, illustrating the potential for computational methods to extract additional diagnostic value from existing, widely available screening data without requiring entirely new instruments. Similarly, Finley (2024) discusses the emerging application of artificial intelligence to performance validity assessment within formal neuropsychological testing, suggesting that computational methods may eventually assist neuropsychologists in detecting subtle, non-credible response patterns not readily apparent through conventional clinical judgment alone. These developments collectively suggest a diagnostic future in which cognitive testing, whether at the bedside or within formal neuropsychological evaluation, is increasingly integrated with digital, biomarker, and computational data streams, potentially yielding diagnostic precision substantially exceeding that achievable through any single assessment modality in isolation.

Limitations and Methodological Considerations

Notwithstanding the substantial evidence base reviewed in this paper, several methodological limitations warrant explicit consideration by clinicians and researchers alike. First, the diagnostic accuracy literature for both bedside screening and formal neuropsychological instruments is characterized by considerable heterogeneity in study design, reference standards, population characteristics, and disease prevalence across settings, complicating direct comparison of sensitivity and specificity estimates across studies. The Cochrane reviews examining the MMSE (Creavin et al., 2016; Arevalo-Rodriguez et al., 2021) and the Mini-Cog (Seitz et al., 2021) each identified substantial methodological limitations, including high or unclear risk of bias in patient selection, across the majority of included primary studies, urging caution in the uncritical application of pooled accuracy estimates to individual clinical decision-making.

Second, the pervasive influence of educational attainment, literacy, language proficiency, and cultural background on cognitive test performance constitutes a persistent threat to the validity of cognitive assessment across both bedside screening and, to a lesser but still meaningful extent, formal neuropsychological testing. As documented extensively in Section 3.5, standard cutoffs derived from predominantly Western, highly educated validation samples frequently generalize poorly to populations with limited formal education, non-native language proficiency, or divergent cultural reference points, elevating risk of both false-positive and false-negative classification (Chithiramohan et al., 2024). While formal neuropsychological testing partially mitigates this concern through individualized normative comparison and clinical judgment, the underlying normative datasets themselves frequently suffer from similar demographic limitations, underscoring the ongoing need for more diverse and representative normative data development across all tiers of cognitive assessment.

Third, confounding by acute or subacute medical, affective, and pharmacological factors represents a persistent challenge to accurate cognitive assessment interpretation, irrespective of the tier of testing employed. Delirium, in particular, can profoundly distort cognitive test performance and, if unrecognized, may lead to inappropriate diagnosis of a chronic neurocognitive disorder in a patient experiencing a transient, reversible confusional state; the CAM and related delirium screening instruments (Inouye et al., 1990; Muser et al., 2022) exist precisely to guard against this form of misclassification, though implementation research indicates that consistent, protocol-adherent delirium screening remains incompletely achieved even in well-resourced hospital systems (Penfold et al., 2024). Similarly, depressive symptomatology can produce a pattern of cognitive underperformance, sometimes termed pseudodementia, that may be difficult to distinguish from genuine neurodegenerative impairment without careful integration of affective symptom assessment, as illustrated by the Kim et al. (2024) finding that individuals reclassified from impaired to normal under stringent neuropsychological criteria exhibited elevated depression scores.

Finally, the rapid pace of technological change in digital and biomarker-based assessment, while offering considerable promise, also introduces new sources of uncertainty, including limited long-term validation data, evolving normative standards, and unresolved questions regarding equity of access across socioeconomic and geographic strata (Paddick et al., 2021). Clinicians and health systems adopting these emerging technologies should do so with appropriate awareness of their still-maturing evidence base, alongside the well-established, extensively validated traditional instruments reviewed throughout this paper.

Discussion

The evidence reviewed in this paper collectively supports a stepped, context-sensitive approach to cognitive assessment in which bedside screening and formal neuropsychological testing occupy distinct, complementary, and non-interchangeable roles. Bedside instruments such as the MMSE, MoCA, Mini-Cog, and SLUMS provide efficient, broadly accessible means of identifying individuals who may warrant further evaluation, and their continued refinement, through revised cutoffs (Islam et al., 2023), culture-fair adaptation (Chithiramohan et al., 2024), and digital delivery (Kochan et al., 2022; Morrissey et al., 2024), has substantially expanded their clinical utility since their original development. At the same time, the accumulated evidence base makes clear that no bedside instrument, however well validated, can substitute for the diagnostic precision, domain-specific characterization, and performance validity safeguards afforded by formal neuropsychological testing, particularly in cases involving diagnostic ambiguity, atypical presentation, differential diagnosis among competing dementia etiologies, or high-stakes treatment eligibility determinations in the era of disease-modifying anti-amyloid therapy (van Dyck et al., 2023; Sims et al., 2023).

The integration of biologically anchored diagnostic frameworks, most recently articulated in the 2024 NIA-AA revised criteria (Jack et al., 2024), further underscores the continuing centrality of cognitive testing within a broader, increasingly multimodal diagnostic paradigm. Biomarker positivity alone does not constitute a clinical diagnosis of dementia or MCI in the functional sense embodied by both the DSM-5-TR (American Psychiatric Association, 2022) and prevailing clinical practice; rather, cognitive and functional assessment, whether through validated screening or comprehensive neuropsychological evaluation, remains indispensable to establishing the clinical-syndromic dimension of diagnosis that biomarkers alone cannot provide (Shaughnessy & Weintraub, 2024). This complementarity suggests that, rather than being rendered obsolete by advances in biomarker science, cognitive testing is likely to become increasingly integrated with biological data, with digital cognitive assessment in particular positioned to serve as a bridge between traditional bedside screening and biomarker-confirmed diagnosis, given its capacity for repeated, longitudinal, and potentially more sensitive characterization of subtle cognitive change (Thompson et al., 2023).

From a health systems perspective, the evidence also highlights persistent challenges in equitable implementation. The demonstrated vulnerability of standard cognitive screening instruments to educational, linguistic, and cultural bias (Chithiramohan et al., 2024), combined with the limited availability of formal neuropsychological services in many regions and healthcare systems, risks perpetuating disparities in the timely and accurate diagnosis of cognitive impairment

across demographic groups. The expansion of culture-fair screening instruments, lay-health-worker-administered digital tools (Paddick et al., 2021), and telehealth-delivered neuropsychological services (Owens et al., 2020) each represent partial responses to this challenge, though continued investment in the development, validation, and equitable deployment of such approaches remains necessary.

Conclusion

Cognitive testing spans a continuum from rapid, broadly accessible bedside screening to comprehensive, specialist-administered formal neuropsychological evaluation, with each tier fulfilling a distinct and indispensable function within the overall diagnostic pathway for neurocognitive disorders. The MMSE and MoCA remain foundational instruments for initial detection and longitudinal monitoring of cognitive impairment, notwithstanding well-documented limitations related to educational, cultural, and linguistic bias that have prompted the development of alternative and culture-fair instruments such as the SLUMS and RUDAS. Formal neuropsychological testing offers substantially greater diagnostic precision, domain-specific profiling, and performance validity assurance, positioning it as the essential second tier of assessment for cases requiring differential diagnosis, detailed functional characterization, or high-stakes treatment decision-making. The integration of biologically anchored diagnostic frameworks, exemplified by the 2024 NIA-AA revised criteria, alongside the rapid maturation of digital cognitive assessment technologies and blood-based biomarkers, is actively reshaping the cognitive assessment landscape, without diminishing the fundamental necessity of rigorous cognitive and functional evaluation. A stepped, context-sensitive approach, employing validated screening tools to identify individuals warranting further assessment, followed by formal neuropsychological evaluation and, where indicated, biomarker confirmation, continues to represent best practice in the evaluation of cognitive function across clinical settings.

References

- Abdelnour, C. (2025). The Alzheimer's Association revised criteria for diagnosis and staging of Alzheimer's disease—Where does Lewy body dementia stand? *Alzheimer's & Dementia*, 21(Suppl. 1). <https://doi.org/10.1002/alz.089685>
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroup. *Alzheimer's & Dementia*, 7(3), 270–279.
- Alzheimer's Association. (2024). 2024 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 20(5), 3708–4002.
- American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders (5th ed., text rev.). American Psychiatric Publishing.
- Arevalo-Rodriguez, I., Smailagic, N., Roqué-Figuls, M., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O. L., Bonfill Cosp, X., & Cullum, S. (2021). Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews*, 2021(7), CD010783. <https://doi.org/10.1002/14651858.CD010783.pub3>
- Borson, S., Scanlan, J. M., Chen, P., & Ganguli, M. (2003). The Mini-Cog as a screen for dementia: Validation in a population-based sample. *Journal of the American Geriatrics Society*, 51(10), 1451–1454.
- Bosso, T., Vischia, F., Keller, R., Vai, D., Imperiale, D., & Vercelli, A. (2022). A case report and literature review of cognitive malingering and psychopathology. *Frontiers in Psychiatry*, 13, Article 981475. <https://doi.org/10.3389/fpsyt.2022.981475>
- Chen, Y.-X., Liang, N., Li, X.-L., Yang, S.-H., Wang, Y.-P., & Shi, N.-N. (2021). Diagnosis and treatment for mild cognitive impairment: A systematic review of clinical practice guidelines and consensus statements. *Frontiers in Neurology*, 12, Article 719849. <https://doi.org/10.3389/fneur.2021.719849>
- Chithiramohan, T., Santhosh, S., Threlfall, G., Hull, L., Mukaetova-Ladinska, E., Subramaniam, H., & Beishon, L. (2024). Culture-fair cognitive screening tools for assessment of cognitive impairment: A systematic review. *Alzheimer's & Dementia Reports*, 8, Article ADR-230194.
- Cornea, A., Simu, M., & Rosca, E. C. (2022). Montreal Cognitive Assessment for evaluating cognitive impairment in subarachnoid hemorrhage: A systematic review. *Journal of Clinical Medicine*, 11(16), Article 4679. <https://doi.org/10.3390/jcm11164679>
- Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., Thom, V. M., Nash, K. J. E., Elhamoui, H., Milligan, R., Patel, A. S., Tsivos, D. V., Wing, T., Phillips, E., Kellman, S. M., Shackleton, H. L., Singleton, G. F., Neale, B. E., Watton, M. E., & Cullum, S. (2016). Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database of Systematic Reviews*, 2016(1), CD011145.
- Finley, J.-C. A. (2024). Performance validity testing: The need for digital technology and where to go from here. *Frontiers in Psychology*, 15, Article 1452462. <https://doi.org/10.3389/fpsyg.2024.1452462>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.
- García-Rudolph, A., García-Molina, A., Opisso, E., Tormos, J. M., Madai, V. I., Frey, D., & Bernabeu, M. (2021). Neuropsychological assessments of patients with acquired brain injury: A cluster analysis approach to address heterogeneity in web-based cognitive rehabilitation. *Frontiers in Neurology*, 12, Article 701946. <https://doi.org/10.3389/fneur.2021.701946>

- Gómez-de-Regil, L. (2020). Assessment of executive function in patients with traumatic brain injury with the Wisconsin Card-Sorting Test. *Brain Sciences*, 10(10), Article 699. <https://doi.org/10.3390/brainsci10100699>
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014.
- Huynh, D., Patterson, M., & Huang, B. (2025). Concurrent validity of BrainCheck compared to SLUMS in older adults using real-world clinical data. *Innovation in Aging*, 9(Suppl. 1). <https://doi.org/10.1093/geroni/igaf122.4090>
- Inouye, S. K., van Dyck, C. H., Alessi, C. A., Balkin, S., Siegel, A. P., & Horwitz, R. I. (1990). Clarifying confusion: The confusion assessment method. *Annals of Internal Medicine*, 113(12), 941–948.
- Islam, M., Iqbal, U., Walther, B. A., Nguyen, P.-A., Yang, H.-C., Poly, T. N., Huang, C.-W., Jian, W.-S., & Li, Y.-C. (2023). Accuracy of the Montreal Cognitive Assessment tool for detecting mild cognitive impairment: A systematic review and meta-analysis. *Alzheimer's & Dementia*, 19(8), 3436–3447. <https://doi.org/10.1002/alz.13040>
- Jack, C. R., Jr., Andrews, J. S., Beach, T. G., Buracchio, T., Dunn, B., Graf, A., Hansson, O., Ho, C., Jagust, W., McDade, E., Molinuevo, J. L., Okonkwo, O. C., Pani, L., Rafii, M. S., Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Teunissen, C. E., & Carrillo, M. C. (2024). Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's & Dementia*, 20(8), 5143–5169. <https://doi.org/10.1002/alz.13859>
- Jack, C. R., Jr., Andrews, J. S., Beach, T. G., Buracchio, T., Dunn, B., Graf, A., Hansson, O., Jagust, W., McDade, E., Molinuevo, J. L., Okonkwo, O. C., Pani, L., Rafii, M. S., Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Teunissen, C. E., Weber, C. J., & Carrillo, M. C. (2025). Application of the revised criteria for diagnosis and staging of Alzheimer's disease: Drug development and clinical practice. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 11(1), Article e70013. <https://doi.org/10.1002/trc2.70013>
- Jak, A. J., Bondi, M. W., Delano-Wood, L., Wierenga, C., Corey-Bloom, J., Salmon, D. P., & Delis, D. C. (2009). Quantification of five neuropsychological approaches to defining mild cognitive impairment. *American Journal of Geriatric Psychiatry*, 17(5), 368–375.
- Jia, Y., Woltering, S., Deutz, N., & colleagues. (2023). Working memory precision and associative binding in mild cognitive impairment. *Experimental Aging Research*, 49, 1–19.
- Kalbe, E., Folkerts, A.-K., Witt, K., Buhmann, C., & Liepelt-Scarfone, I. (2024). German Society of Neurology guidelines for the diagnosis and treatment of cognitive impairment and affective disorders in people with Parkinson's disease: New spotlights on diagnostic procedures and non-pharmacological interventions. *Journal of Neurology*, 271, 6160–6181. <https://doi.org/10.1007/s00415-024-12503-0>
- Kim, H., Kim, Y., Edmonds, E. C., & Bondi, M. W. (2024). Comparison of cognitive impairment diagnosis criteria in clinical settings: Conventional vs. neuropsychological. *Alpha Psychiatry*, 25(2), 212–219.
- Kim, H. G., et al. (2025). Dementia overdiagnosis in younger, higher educated individuals based on MMSE alone: Analysis using deep learning technology. *Journal of Korean Medical Science*, 40(9), Article e20. <https://doi.org/10.3346/jkms.2025.40.e20>
- Kochan, N. A., Heffernan, M., Valenzuela, M., Sachdev, P. S., Lam, B. C. P., Fiatarone Singh, M., & colleagues. (2022). Reliability, validity, and user-experience of remote unsupervised computerized neuropsychological assessments in community-living 55- to 75-year-olds. *Journal of Alzheimer's Disease*, 90(4), 1629–1645. <https://doi.org/10.3233/JAD-220665>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.
- Loewenstein, D. A., Curiel, R. E., Duara, R., & Buschke, H. (2018). Novel cognitive paradigms for the detection of memory impairment in preclinical Alzheimer's disease. *Assessment*, 25(3), 348–359.
- McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., ... Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, 89(1), 88–100.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroup. *Alzheimer's & Dementia*, 7(3), 263–269.
- Morrissey, S., Gillings, R., & Hornberger, M. (2024). Feasibility and reliability of online vs in-person cognitive testing in healthy older people. *PLOS ONE*, 19(8), Article e0309006. <https://doi.org/10.1371/journal.pone.0309006>
- Muser, O., Seiler, K., Bachnick, S., & colleagues. (2022). Delirium detection in hospitalized adults: The performance of the 4 'A's Test and the modified Confusion Assessment Method for the Emergency Department. A comparison study. *Bulletin of the National Research Centre*, 46, Article 179. <https://doi.org/10.1186/s42269-022-00863-4>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.

- Owens, A. P., Ballard, C., Beigi, M., Kalafatis, C., Brooker, H., Lavelle, G., Brown, R. G., & Zamboni, G. (2020). Implementing remote memory clinics to enhance clinical care during and after COVID-19. *Frontiers in Psychiatry*, 11, Article 579934.
- Paddick, S.-M., Yoseph, M., Gray, W. K., Ola, B., Dotchin, C., Chaya, E., Kisoli, A., Longdon, A., & Walker, R. (2021). Effectiveness of app-based cognitive screening for dementia by lay health workers in low resource settings: A validation and feasibility study in rural Tanzania. *Journal of Geriatric Psychiatry and Neurology*, 34(6), 613–621.
- Panico, F., Fonzo, E., Verde, A., Lancia, S., & Trojano, L. (2025). Proposal for a new tool assessing validity performance in forensic neuropsychological testing: The Test of Malingering in Abstraction Skills (TOMAS). *Neurological Sciences*. <https://doi.org/10.1007/s10072-025-08061-6>
- Penfold, C., & colleagues. (2024). Delirium detection tools show varying completion rates and positive score rates when used at scale in routine practice in general hospital settings: A systematic review. *Journal of the American Geriatrics Society*, 72(6), 1885–1901. <https://doi.org/10.1111/jgs.18751>
- Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., Gronseth, G. S., Marson, D., Pringsheim, T., Day, G. S., Sager, M., Stevens, J., & Rae-Grant, A. (2018). Practice guideline update summary: Mild cognitive impairment. Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology*, 90(3), 126–135.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, 56(3), 303–308.
- Rajan, K. B., Weuve, J., Barnes, L. L., McAninch, E. A., Wilson, R. S., & Evans, D. A. (2021). Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020–2060). *Alzheimer's & Dementia*, 17(12), 1966–1975.
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., van Swieten, J. C., Seelaar, H., Dopper, E. G. P., Onyike, C. U., Hillis, A. E., Josephs, K. A., Boeve, B. F., Kertesz, A., Seeley, W. W., Rankin, K. P., Johnson, J. K., Gorno-Tempini, M. L., Rosen, H., ... Miller, B. L. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456–2477.
- Root, J. C., Gaynor, A. M., Ahsan, A., Jung, D., Schofield, E., Ryan, E., Li, Y., & Ahles, T. A. (2023). Remote, computerised cognitive assessment for breast cancer and treatment-related cognitive dysfunction: Psychometric characteristics of the Cogsuite neurocognitive battery. *Archives of Clinical Neuropsychology*, 38(5), 699–713. <https://doi.org/10.1093/arclin/acac111>
- Seitz, D. P., Chan, C. C., Newton, H. T., Gill, S. S., Herrmann, N., Smailagic, N., Nikolaou, V., & Fage, B. A. (2021). Mini-Cog for the detection of dementia within a primary care setting. *Cochrane Database of Systematic Reviews*, 2021(7), CD011415. <https://doi.org/10.1002/14651858.CD011415.pub3>
- Shaughnessy, L. W., & Weintraub, S. (2024). The role of neuropsychological assessment in the evaluation of patients with cognitive-behavioral change due to suspected Alzheimer's disease and other causes of cognitive impairment and dementia. *Alzheimer's & Dementia*, 21(1), Article e14363. <https://doi.org/10.1002/alz.14363>
- Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, 35(6), 735–764.
- Shibata, K., Attaallah, B., Tai, X.-Y., Trender, W., Hellyer, P. J., Hampshire, A., Irani, S. R., Manohar, S. G., & Husain, M. (2024). Remote digital cognitive assessment reveals cognitive deficits related to hippocampal atrophy in autoimmune limbic encephalitis: A cross-sectional validation study. *eClinicalMedicine*, 69, Article 102437. <https://doi.org/10.1016/j.eclinm.2024.102437>
- Sims, J. R., Zimmer, J. A., Evans, C. D., Lu, M., Ardayfio, P., Sparks, J., & colleagues. (2023). Donanemab in early symptomatic Alzheimer disease: The TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA*, 330(6), 512–527.
- Skrobot, O. A., O'Brien, J., Black, S., Chen, C., DeCarli, C., Erkinjuntti, T., Ford, G. A., Kalaria, R. N., Pantoni, L., Pasquier, F., Roman, G. C., Wallin, A., Sachdev, P., Skoog, I., & Kehoe, P. G. (2017). The Vascular Impairment of Cognition Classification Consensus Study. *Alzheimer's & Dementia*, 13(6), 624–633.
- Song, W.-X., Wu, W.-W., Zhao, Y.-Y., Xu, H.-L., Chen, G.-C., Jin, S.-Y., & colleagues. (2023). Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. *Frontiers in Aging Neuroscience*, 15, Article 1227112.
- Spencer, R. J., Noyes, E. T., Bair, J. L., & Ransom, M. T. (2022). Systematic review of the psychometric properties of the Saint Louis University Mental Status (SLUMS) examination. *Clinical Gerontologist*, 45(3), 454–466. <https://doi.org/10.1080/07317115.2022.2032523>
- Thompson, L. I., & colleagues. (2023). Remote and in-clinic digital cognitive screening tools outperform the MoCA to distinguish cerebral amyloid status among cognitively healthy older adults. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 15(4), Article e12500. <https://doi.org/10.1002/dad2.12500>
- Torregrossa, W., Torrisi, M., De Luca, R., Casella, C., Rifici, C., Bonanno, M., & Calabrò, R. S. (2023). Neuropsychological assessment in patients with traumatic brain injury: A comprehensive review with clinical recommendations. *Biomedicine*, 11(7), Article 1991. <https://doi.org/10.3390/biomedicine11071991>
- van Dyck, C. H., Swanson, C. J., Aisen, P., Bateman, R. J., Chen, C., Gee, M., & colleagues. (2023). Lecanemab in early Alzheimer's disease. *New England Journal of Medicine*, 388(1), 9–21.