2023-08-09

The Clinical Utility of a Computerized Cognitive Assessment to Predict Incident Amnestic Mild Cognitive Impairment and Alzheimer’s Disease

Docteur, Natalia Gabriela

UNIVERSITY OF CALGARY

The Clinical Utility of a Computerized Cognitive Assessment to Predict Incident Amnestic Mild Cognitive Impairment and Alzheimer's Disease

by

Natalia Gabriela Docteur

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN CLINICAL PSYCHOLOGY

CALGARY, ALBERTA

AUGUST, 2023

Abstract

Detecting the initial signs of neurodegeneration is integral for early diagnosis and intervention. Computerized cognitive assessments are accessible, efficient, and precise tools for identifying cognitive impairment and risk of neurodegeneration. While computerized instruments can be feasibly administered repeatedly for longitudinal cognitive monitoring, their clinical utility compared to conventional paper-and-pencil tools is yet unknown. The present study examined the utility of a computerized task, the One Card Learning (OCL) test, to detect conversion to dementia and associate with amyloid (Aβ) imaging markers using single and repeated test administration compared to the Montreal Cognitive Assessment (MoCA) and Rey Auditory Verbal Learning Test (RAVLT). The primary and secondary outcomes were conversion from cognitively normal (CN) to amnestic mild cognitive impairment (aMCI) or Alzheimer's disease (AD) over a four-year study period and positron emission tomography estimates of Aβ, respectively. Data were collected from the Alzheimer's Disease Neuroimaging Initiative 3 longitudinal cohort study. Participants were older adults aged 56 to 98 years who were CN at baseline. Results showed that the OCL did not better predict conversion to aMCI or AD from cognitive health compared to the MoCA or RAVLT when assessed at baseline or over repeated administrations. Unadjusted baseline OCL performance associated with Aβ status comparably to the MoCA and RAVLT. While repeated MoCA scores provided the strongest estimate of Aβ accumulation, OCL score trajectories uniquely detected diminished practice effects associated with pathological Aβ accumulation. The OCL may offer distinct clinical utility to detect preclinical AD biomarker accumulation. Future research is needed to examine the application of computerized assessments before they are fully integrated into clinical practice.

Preface

This thesis is original, unpublished, independent work by the author, N. Docteur. The data for the current study were collected by the Alzheimer's Disease Neuroimaging Initiative.

Acknowledgements

I would like to thank my supervisor, Brandy Callahan, for her support throughout the last two years and for always being available to advise me. Thank you to my committee members and external examiner, Zahinoor Ismail, Giuseppe Iaria, and Cameron Clark, for their guidance and valuable feedback on my thesis. Thank you to my labmates for helping me prepare for my defense and for sharing their snacks. Thank you to my cohort for being allies, both in our work and outside of the program. Thank you to my mom and sister for your patience and encouragement while I make my own path. Thank you to my partner for listening and for being a consistent source of support and joy.

Table of Contents

List of Tables

List of Figures

Introduction

Contemporary changes to society from technological advances have created new opportunities for the field of neuropsychology. Among the proposed innovations to current neuropsychological methods is the incorporation of computerized technology into assessment procedures (Bilder, 2011; Minnesota 2022 Update Conference Planning Commission, 2022). Computerized assessments are novel instruments that capitalize on the efficiency and widespread availability of digital tools while addressing limitations of conventional paper-and-pencil tests (Binng, Splonskowski, & Jacova, 2020; Koo & Vizer, 2019; Sternin, Burns, & Owen, 2019). While the application of computerized cognitive assessments (CCAs) into clinical practice has the potential to improve access to, and quality of, assessment services for millions of Canadians, their comparative utility remains to be substantiated empirically.

It is primarily important to define CCAs and describe how they differ from paper-and-pencil tests. CCAs measure thinking and memory abilities based on observable performance, including domains of executive function, attention, visuo-spatial navigation, processing speed, language, and episodic and working memory (De Roeck, De Deyn, Dierckx, & Engelborghs, 2019; Staffaroni, Tsoy, Taylor, Boxer, & Possin, 2020). Like conventional tests, digital assessments use narrow-band tasks with standardized administration procedures to interpret performance using normative data. However, CCAs are unique in that task instructions and stimuli are presented via computerized means, and participant responses, latency measurements, and norming of scores are recorded automatically (Sternin et al., 2019). In contrast to conventional tools, some CCAs may be administered without supervision from a neuropsychologist or trained professional, meaning that they can be conducted remotely outside of hospital or clinic settings.

Advantages of CCAs

CCAs demonstrate improved accessibility, efficiency, and precision over conventional tests. The flexibility to administer CCAs outside of healthcare settings provides enhanced accessibility to populations who live rurally, have mobility limitations, or are at increased health risk of infectious diseases (Binng et al., 2020; O'Connell, Vellani, Robertson, O'Rourke, & McGilton, 2021; Staffaroni et al., 2020; Tsoy, Zygouris, & Possin, 2021). Broader accessibility particularly benefits equity-seeking communities that face added barriers to assessment services in Canada (Reitmanova & Gustafson, 2009; Thomson, Chaze, George, & Guruge, 2015; Williams, 2001). CCAs have the capacity to improve testing efficiency, minimize costs, and reduce patient burden. Unsupervised testing and digitally mediated stimuli reduce operational costs associated with conventional testing (Ohman, Hassenstab, Berron, Scholl, & Papp, 2021). Testing time and fatigue can also be cut down using computerized adaptive testing, a process that selects items based on the examinee's proficiency level (Huebner, 2010). Finally, standardized administration, scoring, and recording procedures via automated processes result in enhanced precision and reduced risk of human error (Koo & Vizer, 2019; Sternin et al., 2019). CCAs can feasibly capture more fine-grained information on response styles, item-by-item timing, and validity data via person-fit statistics, maximizing the value of test data (Bellone & Van Patten, 2020; Huebner, 2010).

Broad Psychometric Support for CCAs

Given these advantages, research has found support for the use of CCAs to screen for cognitive status in neurocognitive disorder populations. Validation studies comparing the concurrent and prognostic validity of conventional and computerized assessments show analogous performance to conventional tests within multiple sclerosis, Alzheimer's disease (AD), mild cognitive

impairment (MCI), stroke, frontotemporal dementia, and mixed clinical samples (Binng et al., 2020; Lapshin, O'Connor, Lanctot, & Feinstein, 2012; Ohman et al., 2021; Staffaroni et al., 2020; Woodhouse et al., 2013). Among the most widely studied CCAs, there is strong criterion validity to distinguish normal controls from early MCI and dementia, and associate with brain biomarkers of dementia neuropathology (Binng et al., 2020; De Roeck et al., 2019; Denboer, Nicholls, Corte, & Chestnut, 2014; Staffaroni et al., 2020).

Reliability and feasibility statistics for identifying cognitive impairment in aging suggest that CCAs provide acceptable psychometrics in clinical and research settings. There is empirical support for adequate to excellent internal consistency (Cronbach's α = .72-.96) and test-retest reliability (r = .70-.97) in the most widely studied CCAs (Binng et al., 2020; De Roeck et al., 2019; Staffaroni et al., 2020; Tsoy et al., 2021). Feasibility research, which is particularly important with unsupervised administration, indicates that completion rates for self-administered tests range from 60-97% over repeated testing sessions and are similarly acceptable across sociodemographic stratifications (Collerton et al., 2007; Koo & Vizer, 2019; Mielke et al., 2015; Tsoy et al., 2020). However, individual factors like low computer familiarity, negative attitudes toward technology, and decreased digital experience may affect feasibility estimates in older adults (Iverson, Brooks, Ashton, Johnson, & Gualtieri, 2009; Koo & Vizer, 2019; Ohman et al., 2021; Staffaroni et al., 2020; Tierney et al., 2014; Valdes, Sadeq, Harrison Bush, Morgan, & Andel, 2016).

Mild Cognitive Impairment in the Neurodegenerative Continuum

MCI is an early-stage neurocognitive syndrome characterized by objective deficits in cognition without loss of functional independence (Petersen, 2004). While the typical healthy aging trajectory shows small but steady decreases in some thinking and memory abilities, MCI

encompasses cognitive decline beyond expected changes (Jack et al., 2018). MCI therefore represents an early stage of measurable cognitive dysfunction prior to acute deficits suggestive of a dementia diagnosis. It is estimated that 6.7-25.2% of adults aged ≥60 years meet criteria for an MCI diagnosis, with rates increasing with age and lower education (Cheng, Chen, & Chiu, 2017; Jongsiriyanyong & Limpawattana, 2018; Langa & Levine, 2014; Petersen et al., 2018; Suzuki et al., 2013). MCI prevalence is expected to increase as Canada's population continues to age (Alzheimer Society of Canada, 2016). MCI diagnosis is associated with increased medical comorbidity, conversion to dementia, and mortality (Bae et al., 2018; Haaksma et al., 2017; McGrattan et al., 2022; Petersen et al., 2018; Stephan et al., 2011; Vassilaki et al., 2015). Thus, early detection of impaired cognition may not only identify risk of future neurodegeneration, but also flag those who may require clinical evaluation for medical concerns and supportive care.

Examination of MCI subtypes and progression to dementia has improved our knowledge of the neurodegenerative continuum and conferred earlier diagnoses. Accelerated diagnosis is important for addressing cognitive and behavioural symptoms of MCI, initiating interventional approaches that may delay cognitive decline, and aiding patients to engage in long-term planning for their care (Kasper et al., 2020; Lissek & Suchan, 2021; Petersen et al., 2018). Amnestic MCI (aMCI) is cognitive impairment that includes episodic memory dysfunction and is linked to 1.5 to 10 times increased risk of developing Alzheimer's disease (AD) compared to non-amnestic MCI (naMCI) (Ferman et al., 2013; Jungwirth, Zehetmayer, Hinterberger, Tragl, & Fischer, 2012; Michaud, Su, Siahpush, & Murman, 2017). Relative risk estimates for progressing from aMCI and naMCI to AD vary depending on the setting sampled (clinic versus community), follow-up period assessed, and diagnostic criteria utilized. aMCI is generally considered a prodromal stage of AD whereby patients transition from preclinical "normal" cognition to significant cognitive and

functional deficits with accompanying neuropathology. Annual conversion rates from aMCI to AD in the community are estimated at 6-17% (Ferman et al., 2013; Landau et al., 2010; Michaud et al., 2017; Mitchell & Shiri-Feshki, 2009). In contrast, naMCI, which is cognitive dysfunction in alternative domains such as executive function, attention, etc., but not memory, is associated with an increased likelihood of developing other neurodegenerative disorders such as fronto-temporal and Lewy-body dementias (Molano et al., 2010). Differentiation of MCI subtypes is thus an important marker to inform diagnostic decision-making and targeted clinical interventions.

Biological Definitions of AD

In addition to neuropsychological profiles, biomarkers have emerged as important signals of AD pathology and future decline. Briefly, the National Institute on Aging and Alzheimer's Association organizes current AD biomarkers into three broad categories: amyloid beta (Aβ) accumulation, or plaque deposits, that lead to synaptic dysfunction and neural death; tau (T) aggregation, or neurofibrillary tangles, that disrupt axonal conduction and nutrient delivery to neurons; and neurodegeneration (N), such as cerebral hypometabolism and structural volume loss that are detected via neuroimaging (Jack et al., 2018; Murphy & LeVine, 2010; Schraen-Maschke et al., 2008). Together, these three biomarkers are often referred to as ATN, where individuals can be classified as 'positive' (+) if the biomarker is found to be present, or 'negative' (-) if it is not. Genetic markers such as apolipoprotein E4 (ApoE ε4) are also implicated in determining AD pathology (Liu, Kanekiyo, Xu, & Bu, 2013). ATN biomarkers have shown strong predictive power for the onset of cognitive and neuropsychiatric AD symptoms, where biomarker positivity now defines the preclinical AD stage in the absence of cognitive or behavioural symptoms (Bucci, Chiotis, Nordberg, & Alzheimer's Disease Neuroimaging, 2021; Ebenau et al., 2020; Jack et al.,

2018; Miao et al., 2022). Emphasis on biomarkers to define AD pathology is advantageous for predicting decline prior to the emergence of detectable symptoms and for driving therapeutic research that targets mechanisms of neurodegeneration in vivo (Silverberg, Elliott, Ryan, Masliah, & Hodes, 2018). While there is some evidence that biomaker data may not significantly improve prediction of dementia over demographic and cognitive markers (Callahan et al., 2015; Glymour et al., 2018), which are more accessible and cost-efficient data points, biomarkers remain important for assessing AD pathology in the current literature (Silverberg et al., 2018).

Current Neuropsychological Assessment Procedures for Cognitive Decline

Current neuropsychological methods rely on conventional instruments administered at a single time point to identify MCI. Conventional cross-sectional methods characterize cognitive functioning, where scores of 1 to 1.5 standard deviations (SD) below expected group-level normative performance typically signals the presence of deficits, although not all MCI diagnostic criteria specify absolute cutoffs for clinically significant impairment (Albert et al., 2011; American Psychiatric Association, 2013; Jak et al., 2009; Petersen et al., 2014). Reliance on cross-sectional methods using mean- or regression-based normative comparisons to infer progressive decline can result in limited specificity and significant bias in detecting deficits. Cross-sectional cutoffs to determine dysfunction may over-pathologize normal cognitive performance as 30.8-39.0% of healthy older adults perform ≤1.5 SD below adjusted norms on at least one memory test (Brooks, Iverson, Holdnack, & Feldman, 2008; Brooks, Iverson, & White, 2007). Adding to the risk of false positives is the non-specificity of below average cognitive performance. Numerous non-neurodegenerative disorders including depression, bipolar disorder, attention-deficit/hyperactivity disorder, and schizophrenia are associated with objective cognitive deficits that may mimic

neurodegeneration if evaluated only at the cross-sectional level (Aprahamian, Nunes, & Forlenza, 2013; Callahan, Shammi, Taylor, Ramakrishnan, & Black, 2021; Lanza, Sejunaite, Steindel, Scholz, & Riepe, 2020; Morimoto & Alexopoulos, 2013; Mukku et al., 2021; Murante & Cohen, 2017; Rajji & Mulsant, 2008; Schouws, Comijs, Dols, Beekman, & Stek, 2016). Moreover, individual-level characteristics – such as ethno-linguistic diversity and intellectually high- and low baseline performance – are poorly represented in normative samples but significantly affect cross-sectional cognitive performance and may contribute to misidentification of decline (Briceno et al., 2020; Brooks, Holdnack, & Iverson, 2011; Brooks, Iverson, & White, 2009; Castora-Binkley, Peronto, Edwards, & Small, 2015; Gross et al., 2015; Iverson & Karr, 2021).

Longitudinal Neuropsychological Assessment

Clinical and research-based guidelines for identifying MCI are shifting focus towards longitudinal measurement of cognitive trajectories due to the above-mentioned limitations of cross-sectional methods (Albert et al., 2011; American Psychological Association, 2021; Jack et al., 2018). Longitudinal assessment is theorized to be more sensitive to detect AD pathology by measuring rates of decline that better account for baseline cognition and individual-level characteristics that impact test performance (Sternin et al., 2019). Emerging research on longitudinal cognitive testing indicates that accounting for change over time is indeed an effective strategy for predicting conversion to MCI/AD (Mortamais et al., 2017). A study by Nation et al. (2019) found that over 12-months, repeated neuropsychological testing significantly predicted incident AD over and above baseline cognition, demographics, and ApoE ε4 status in cognitively normal (CN) older adults with AD biomarkers (odds ratio (OR)= 2.84). In a meta-analysis examining relationships between Aβ and cognition, the effect of biomarker positivity on episodic memory was stronger for

longitudinal decline (*d*= 0.24) than cross-sectional impairment (*d*= 0.15) in CN older adults (Baker et al., 2017). In terms of the temporal detection of longitudinal decline, episodic memory decline can be observed four years prior to MCI due to AD diagnosis (Mistridis, Krumm, Monsch, Berres, & Taylor, 2015), and seven years preceding AD diagnosis (Grober et al., 2008).

The effect of learning on performance across repeat test administrations, historically considered an assessment confound, has gained recent attention for its potential utility as a marker for AD. Practice effects describe improvement in test scores over serial assessments that are due to familiarity with items, acquired strategies, or enhanced test-taking comfort rather than changes in the target construct being measured (Calamia, Markon, & Tranel, 2012). Within CN adults, practice effects are seen in episodic memory tasks, demonstrating interference with scores over six years at one-year testing intervals, with maximum effect sizes ranging from Cohen's *d*= .20-.72 (Calamia et al., 2012; Goldberg, Harvey, Wesnes, Snyder, & Schneider, 2015). When comparing practice effects between stable CN to MCI/AD converters, converters exhibited equivalent practice effects to controls on a memory test across the first two administrations over 15 months, but showed weakened practice effects over the next four years (Machulda et al., 2013). A recent literature review on practice effects in the AD continuum found that diminished practice effects in episodic memory tests predicted progressive cognitive decline and incident dementia (Jutten et al., 2020). Less robust practice effects were associated with ApoE ε4 status and Aβ status, although there is some evidence that the absence of learning is predicted by hippocampal neurodegeneration rather than amyloidosis (Jutten et al., 2020; Machulda et al., 2017). In addition, reduced practice effects can detect concurrent abnormal Aβ accumulation over three years (Ihara et al., 2018). Failure to account for practice effects can impede detection of MCI as previous test exposures mask the magnitude of true decline over time (Elman et al., 2018; Sanderson-Cimino et al., 2020).

To summarize, repeated cognitive assessment methods may be more sensitive to detect the earliest signs of incipient neurodegeneration compared to a single administration. Repeated monitoring allows for identification of early pathological decline before patients demonstrate objective deficits severe enough to warrant an MCI diagnosis at predetermined normative cutoffs. By accounting for individual-level characteristics that impact performance at any single administration, thereby minimizing within-subject error variance, repeated teasing can better detect subtle memory changes that indicate non-normative decline. Specifically, repeated testing can uniquely identify cognitive performance trajectories that show lack of improvement via attenuated practice effects, a novel cognitive marker of AD neuropathology.

CogState Brief Battery

The CogState Brief Battery (CBB) is among the most widely studied CCAs for measuring cognitive functioning in healthy and impaired adults (Maruff et al., 2013; Tsoy et al., 2021). The CBB measures psychomotor reaction time, attention, executive function and visual learning and memory across four subtests (Hammers et al., 2012; Lim et al., 2012). It can be delivered via computer or tablet, requiring no supervision to administer (Fredrickson et al., 2010). The CBB demonstrates strong psychometric properties in healthy adults and those with mild traumatic brain injury, multiple sclerosis, schizophrenia, HIV-associated dementia, MCI, and AD (De Meijer et al., 2018; De Roeck et al., 2019; Lim et al., 2012; Maruff et al., 2013; Maruff et al., 2009; Mielke et al., 2015; Tsoy et al., 2021; Wojcik et al., 2019; Zarshenas & Cullen, 2018). It has been used in several prospective cohort studies including the Alzheimer's Disease Neuroimaging Initiative, the Australian Imaging, Biomarkers, and Lifestyle Study, and Mayo Clinic Study of Aging.

Baseline Assessment with the CBB

The majority of research has utilized single CBB administrations to measure cognition and classify diagnostic groups. The CBB memory subtest which assesses visual learning and recognition, and the executive function subtest which measures working memory and visual attention, have both demonstrated concurrent and convergent validity with demographic variables and conventional tests in neurocognitive disorders (Lim et al., 2012; Mackin et al., 2018; Maruff et al., 2013; Maruff et al., 2009; Racine et al., 2016). In terms of diagnostic validity, the CBB can successfully differentiate CN from MCI (area under the curve (AUC)= 0.75-0.91, sensitivity= 70-80.4%, specificity= 70-84.7%), and AD (sensitivity= 100%, specificity= 84.7%), although cross-sectional thresholds for impairment fall within the normal range of performance (-0.21 to –1 SD), calling into question the practical utility of implementing such cutoffs clinically (Alden et al., 2021; Lim et al., 2012; Mackin et al., 2018; Maruff et al., 2013; Racine et al., 2016).

To date, only two studies have examined the prognostic validity of a single CBB administration to predict incident dementia. In a sample of patients aged ≥50 years who were CN at baseline, CBB memory scores predicted incident MCI or dementia (AUC= 0.67, sensitivity= 60%, specificity= 70%) comparably to the Rey Auditory Verbal Learning Test (RAVLT), a gold-standard conventional verbal memory test (AUC= 0.70, sensitivity= 70%, specificity= 61%) over approximately 3.2 years (Stricker, Lundt, Albertson, et al., 2020). Similarly, Pudumjee et al. (2021) found that a single CBB memory score, when assessed at baseline, predicted incident MCI/AD over approximately 4.8 years (AUC= 0.64, sensitivity= 70%, specificity= 56%) in older adults aged ≥65 years who were CN at baseline. They also compared baseline performance to a second memory score assessed 30 months later, finding that prognostic validity improved while specificity fell (AUC= 0.66, sensitivity= 86%, specificity= 41%). Although there was a small

increase in prediction between the baseline and 30-month testing time points, the loss of 2.5 years to intervene may offset the potential benefit of delayed monitoring.

Previous research has also examined associations between a single CBB administration and biomarkers of AD pathology. Lim et al. (2016) demonstrated that a CBB memory and executive function composite score successfully distinguished between CN and MCI Aβ-, as well as MCI Aβ- and Aβ+ groups. The same composite, when optimized by maximizing true positives and minimizing false positives, also effectively differentiated Aβ-T- from Aβ+T+ participants with MCI (AUC= 0.86) (Alden et al., 2021; Unal, 2017). However, in CN older adults aged ≥65 years, baseline CBB memory performance differentiated Aβ+T+ from Aβ-T- participants less effectively (AUC= 0.64) (Pudumjee et al., 2021). Limited studies have compared the CBB and conventional test performance to associate with AD biomarkers, with mixed results. In one study, neither computerized (CBB memory) nor conventional (RAVLT/Logical Memory (LM) test) baseline performance associated with positron emission tomography (PET) estimates of Aβ status, however all three cognitive test scores did predict a small but similar amount of variance in PET hypometabolism (Mielke et al., 2014). Another study found that among CN adults aged ≥50 years, the RAVLT better differentiated between CN Aβ-T- and Aβ+T+ groups (AUC= 0.66) compared to the CBB composite (AUC= 0.59), although the CBB was more sensitive (55% vs. 45%) (Stricker, Lundt, Albertson, et al., 2020). Divergent findings may be due to the inclusion of both aMCI and naMCI participants in some studies or variations in cognitive domains assessed (Bondi et al., 2008; Mickes et al., 2007). Given these inconsistencies, further investigation is warranted to better understand the utility of the CBB applied at a single time point to detect Aβ early in the AD continuum.

Repeated Assessment with the CBB

Less is known about the efficacy of repeated testing with the CBB to track cognitive decline and predict conversion. Memory decline as measured by the CBB subtest associates with older age, lower education, family history of memory impairment, cognitive complaints, MCI, and ApoE ε4 status (Banh et al., 2022; Darby et al., 2011; Lim, Ellis, et al., 2013). Previous research indicates repeated CBB memory performance improves model prediction of cognitive status over baseline scores alone (Banh et al., 2022). Only one study has investigated serial CBB assessments to predict incident MCI/AD. This study also evaluated the comparative predictive power of repeated and single CBB. Pudumjee et al. (2021) found that over 30 months, CBB memory change better identified MCI/AD converters (OR= 2.02, AUC= 0.68) compared to baseline performance (OR= 1.70, AUC= 0.64), although the latter had higher sensitivity (70% vs. 64%). Further examination on the prognostic validity of the CBB to identify incident aMCI or AD from cognitive health compared to current conventional tools will provide valuable information on the clinical utility of a novel CCA and inform future applications of computerized tools in clinical practice.

In addition to detecting clinical markers of dementia, there is also evidence that repeated CBB performance associates with AD biomarkers. In a study comparing single and repeated CBB assessments to predict Aβ and T status, Pudumjee et al. (2021) found that serial CBB memory testing (AUC= 0.69) better differentiated Aβ+T+ from Aβ-T- participants compared to baseline scores alone (AUC= 0.64), although the difference was not statistically significant. CBB memory decline has also demonstrated efficacy to predict cortical Aβ positivity over 24 months among CN adults aged ≥50 years (OR= 6.34) (Darby et al., 2011). The rate of CBB memory decline differentiated Aβ+ aMCI from Aβ- CN and aMCI participants, as well as Aβ+ CN and MCI from Aβ- CN controls (Lim, Ellis, et al., 2013; Lim et al., 2014). One study demonstrated that in Aβ+

individuals, both CN and MCI, CBB memory scores declined at an equivalent and consistent rate over 36 months compared to Aβ- participants, both CN and MCI (Lim et al., 2014). In that study, the CBB detected significant differences in practice effects between Aβ- and Aβ+ groups. Further research showed that over 72 months, the CBB memory and executive function composite tracked improved performance over time in Aβ- CN and MCI participants, while differentiating from MCI Aβ+ individuals that showed significant cognitive decline (*d*= 0.8) and greater hippocampal volume loss (HVL) (*d*= 1.7) (Lim et al., 2016). In a unique study examining longitudinal accumulation of AD biomarkers, Lim et al. (2015) found that repeated CBB memory scores associated with increased HVL, which then predicted higher rate of Aβ accumulation over time. This complex body of research suggests that a cost-efficient and feasible CCA can successfully identify AD neuropathology and track biomarker accumulation as it emerges.

In sum, previous research demonstrates that the CBB can accurately classify cognitive status and distinguish biomarker groups when administered at a single time point. It has also been established that serial CBB assessment can identify individuals who will convert to dementia and is sensitive to detect subtle cognitive changes related to AD biomarker status and accumulation. A gap in this literature is the need to identify whether the CBB's clinical utility, when administered once or repeatedly, equals, or surpasses that of conventional paper-and-pencil cognitive screening tools currently used in clinical practice. If so, this would support its use in primary care settings as a feasible and cost-effective measure of cognition to identify dementia risk.

Objectives and Hypotheses

The present study **(A)** investigated the utility of a CCA to predict incident aMCI or AD over a four-year follow-up period compared to conventional neuropsychological tests, when assessed at

baseline and over repeated administrations. This study also **(B)** examined associations between CCA performance and PET Aβ compared to conventional neuropsychological tests, using single and repeated assessment. It was hypothesized that changes in serial CCA performance would better predict incident aMCI or AD compared to conventional tests, but that baseline performance would show equivalent utility to identify incident dementia over the follow-up period. For the secondary objective, it was hypothesized that repeated CCA performance would more strongly associate with PET Aβ accumulation than repeated conventional tests, but that single assessment associations with PET Aβ status would be statistically similar between a CCA and conventional tests.

Methods

Study Design

This study was a within-subjects longitudinal observational design. Data were extracted from the Alzheimer's Disease Neuroimaging Initiative 3 (ADNI3), a prospective cohort study that measured emergent cognitive, clinical, biological, and neuroimaging markers in the AD continuum (https://adni.loni.usc.edu). Running from 2017-2023, ADNI3 followed CN, aMCI and AD participants over four years. Participants underwent review of demographics, diagnostic assessment, conventional neuropsychological battery, in-clinic CBB, and Aβ PET imaging, among other tasks, at enrollment (baseline). The schedule of assessments varied by diagnostic status. For CN participants, in-clinic visits occurred biennially with repeat diagnostic review, conventional neuropsychological and CBB testing, and Aβ PET imaging. For participants with aMCI or AD, in-clinic visits occurred annually, with repeat diagnostic review, and conventional neuropsychological and CBB testing occurring each time, whereas Aβ PET imaging continued to

occur biennially. CN and aMCI participants were prompted via email to complete remote CBB administrations four times per year between in-clinic visits for the duration of the study.

Recruitment and Inclusion Criteria

Participants were recruited from nearly 60 sites across the United States. Individuals enrolled in previous ADNI trials (“rollover" participants) were eligible to continue into ADNI3. All ADNI3 participants were 55-90 years old at enrollment. Inclusion criteria for the present study mirrors that of ADNI3. For CN participants, this included absence of subjective memory complaints, normal memory functioning based on the delayed Logical Memory (LM) Test, Mini Mental State Examination (MMSE)= 24-30, Clinical Dementia Rating (CDR)= 0, preserved daily functioning, and four-week pharmacological stability. For aMCI, inclusion criteria were subjective or informant-reported memory complaints, impaired delayed LM, MMSE= 24-30, CDR= 0.5, absence of impaired daily functioning indicative of AD, and 12-weeks of stable permitted medications. For AD participants, this included subjective or informant memory complaints, significantly impaired delayed LM, MMSE= 20-24, CDR= 0.5-1.0, impaired daily functioning based on NINCDS/ADRDA criteria, and 12-weeks of stable medications (Albert et al., 2011).

Demographics

In line with previous research on demographic correlates of CBB performance, age, sex, education, race/ethnicity, and ApoE ε4 status were collected as covariates. Age was calculated based on completion date of the relevant cognitive test for each analysis. Ethnoracial categories were coded as Asian, Black, Indigenous, mixed, or White. ApoE ε4+ was defined as ≥1 ApoE ε4 allele.

CBB One Card Learning Test

The CBB is a brief computerized cognitive screener. The One Card Learning (OCL) subtest of the CBB was the primary computerized measure of interest. The OCL is characterized as a visual learning and recognition test (Hammers et al., 2012; Lim et al., 2012). In this task, participants were shown four randomly selected target playing cards face-up from a normal deck, one at a time (White et al., 2021). After the targets were presented, 80 cards were presented serially, including both targets and distractors (Maruff et al., 2013). Participants were prompted to key "yes" or "no" for each card to indicate whether it was one of the four target cards presented earlier (Hammers et al., 2012; Lim et al., 2012). The proportion of correct responses formed an accuracy score, which was normalized using an arcsine square-root transformation for a more normal distribution. Average performance for cognitively healthy older adults is approximately set at 1.0 with a standard deviation (*SD*) = 0.10 (White et al., 2021). In addition, the CBB includes in-built integrity measures. Responses were flagged as invalid when OCL item reaction time $<3.24_{\log}10$ (1.75 seconds) or if <75% of trials were completed or for total accuracy scores <50% (Fredrickson et al., 2010; Stricker et al., 2019). Participants retook the OCL a second time if their performance failed integrity checks. All OCLs were administered on a PC. OCL administration time is approximately 5 minutes. In ADNI3, the OCL was conducted at each in-clinic visit and participants were additionally prompted to perform the OCL remotely, four times annually. If a participant was diagnosed with AD, they were no longer prompted to complete the CBB at home.

The CBB demonstrates strong psychometric properties in CN and MCI adult populations. Test-retest values range from *r*= .68-.87 at 3-months intervals, .90-.96 at 4-month intervals, and .65-.91 at 12-month intervals (Lim, Jaeger, et al., 2013; Tsoy et al., 2021). In terms of construct and convergent validity, the OCL correlates with age, sex, education, premorbid IQ, global

cognition, and subjective cognitive impairment, as well as conventional verbal and visual memory tests like the RAVLT (*r*= .32), Logical Memory test (*r*= .20), Spatial Span (*r*= .69), Brief Visuospatial Memory Test (*r*= .25-.83), and Rey Complex Figure Test (*r*= .79) (Lim et al., 2012; Mackin et al., 2018; Maruff et al., 2013; Maruff et al., 2009; Racine et al., 2016).

Rey Auditory Verbal Learning Test

The RAVLT is a conventional verbal episodic memory test. In this test, examinees are read a list of 15 words over five trials, after which they are prompted to repeat the list back. A second, distractor list is then read to the patient, and they are prompted to repeat the distractor items followed by a final recall trial for the original list of words. Thirty minutes after the final immediate recall trial, examinees are asked to recall as many words as they can from the first list only. The total number of words recalled after the delay represents the delayed recall score, the RAVLT outcome of interest for this study. The RAVLT is a gold-standard conventional tool for measuring verbal learning and recall. Test-rest reliability ranges from .61-.86 at one-month intervals and .38-.70 at one-year (Lezak & Lezak, 2004). Criterion validity measures indicate that delayed recall performance between the RAVLT and a similar test, the California Verbal Learning Test (CVLT), covaries in healthy adult (*r*= .37) and cognitively impaired (*r*= .81) groups (Crossen & Wiens, 1994; Stallings, Boake, & Sherer, 1995). Previous factor analytic research shows that the RAVLT loads with other verbal and visual memory tests (Strauss, Sherman, Spreen, & Spreen, 2006).

The decision to compare the OCL to the RAVLT was two-fold. First, existing research often compares OCL performance to the RAVLT or a similar verbal learning task (i.e., LM, CVLT) (Lim, Ellis, et al., 2013; Mielke et al., 2014; Stricker, Lundt, Albertson, et al., 2020). Selecting the RAVLT allowed for continuity to compare results with the current literature. Second,

RAVLT performance is indeed an efficacious and sensitive tool for measuring episodic memory dysfunction in neurocognitive disorder populations, making it a worthy criterion measure (Balthazar, Yasuda, Cendes, & Damasceno, 2010; Estevez-Gonzalez, Kulisevsky, Boltes, Otermin, & Garcia-Sanchez, 2003). However, the RAVLT is not validated for use as a screening tool, and its relatively long administration time and specialized training required to conduct, score and norm it limits it's feasibility for primary care screening. Thus, the OCL was also compared to a conventional tool widely used in clinical settings, the Montreal Cognitive Assessment (MoCA).

Montreal Cognitive Assessment

The MoCA is specifically designed to detect early signs of MCI. The screener assesses visuospatial reasoning, language, memory, executive functioning, attention, and orientation (Hobson, 2015). Together, scores form a composite representing global cognitive function, with a maximum score of 30. Generally, a cutoff < 26 is indicative of impaired cognitive performance, although research suggests that optimal cut points vary by demographic stratifications (Milani, Marsiske, Cottler, Chen, & Striley, 2018). The test must be conducted by a trained professional and administration time is approximately 10 minutes (Nasreddine et al., 2005). In ADNI3, the MoCA was conducted in-clinic according to the same schedule as the RAVLT. Psychometric research indicates that it is a reliable and valid tool for detecting early cognitive dysfunction. Test-retest reliability values range from .88 over two weeks, .75-.92 over one to two months, and .33-.48 at one to four years (Cooley et al., 2015; Koski, 2013; Lee, Lin, & Chiu, 2021; Nasreddine et al., 2005). The MoCA covaries with scores from validated cognitive assessments such as the MMSE in MCI (r= .60) and AD groups (r= .70) (Freitas, Simoes, Alves, & Santana, 2013). The MoCA shows strong diagnostic

accuracy to differentiate CN from MCI (AUC= .86, sensitivity= 81%, specificity= 77%) and AD (AUC= .98, sensitivity= 88%, specificity= 98%) (Nasreddine et al., 2005).

Selection of the MoCA allowed for improved examination of the comparative clinical utility of the OCL as a screening tool. The MoCA is a gold-standard brief instrument for use in primary care that was specifically designed to have high sensitivity to detect MCI. However, the MoCA is subject to notable limitations of conventional tools discussed previously (i.e., non-specificity to progressive degenerating disorders, limited norms for ethno-linguistically diverse groups, does not account for practice effects over serial assessments).

Aβ Imaging

Central Aβ levels were quantified using amyloid PET neuroimaging. Briefly, Aβ PET scanning identifies the presence of plaque deposits in vivo and is currently utilized in clinical practice to determine risk for, or confirm, AD diagnosis (Suppiah, Didier, & Vinjamuri, 2019). Participants newly enrolled in ADNI3 received Florbetaben (FBB) amyloid PET scanning (300 MBq +/- 10%) to quantify Aβ. Rollover participants underwent Florbetapir (FBP) amyloid PET scanning (370 MBq +/- 10%) for consistency of comparison over serial neuroimaging scans. Comparison between scan types was enabled by converting regional standardized uptake values into [^{11}C]Pittsburgh Compound referenced retention values (Weiner et al., 2017). For the single assessment analyses, cross-sectional Aβ status was determined using the summary standardized uptake value ratio (SUVR) based on whole cerebellum referenced region with a binary positivity cutoff ≥ 1.08 for FBB scans and ≥ 1.11 for FBP scans (≥2 SD above control based mean SUVR) (Royse et al., 2021). The repeated measures analyses determined Aβ accumulation with a cortical composite SUVR based on eroded subcortical white matter, brainstem, and whole cerebellum

referenced regions with a binary positivity cutoff ≤ 0.74 for FBB scans and ≤ 0.78 for FBP scans, a linear regression-based longitudinal threshold (Royse et al., 2021).

Statistical Analyses

All statistical analyses were completed using SPSS version 29.0. Data were extracted from ADNI3 on April 21, 2023. Independent samples t-tests and 2x2 chi-square analyses were utilized for descriptive group comparisons of continuous and categorical variables, respectively. To determine the utility of a single OCL administration to predict conversion to aMCI or AD over a four-year follow-up period and baseline Aβ status compared to conventional tests, area under the curve (AUC) receiver operating characteristics were analyzed using multiple logistic regression models. Three logistic regression models were conducted separately for the OCL, MoCA, and RAVLT with the same covariates: age, sex, education, race/ethnicity, and ApoE ε4 status. Predicted probabilities of cognitive status (CN or aMCI/AD) calculated from the logistic regression models were entered into the AUC analyses to produce adjusted values. AUCs determine sensitivity and specificity levels for binary classification (CN or aMCI/AD) based on test cutoffs at various points (Hajian-Tilaki, 2013). Optimal cutoffs for each cognitive test (OCL, RAVLT, MoCA) that maximized sensitivity and specificity values were determined using Youden's index (Unal, 2017).

To evaluate the utility of repeated OCL administrations to predict conversion to aMCI or AD and associate with Aβ accumulation compared to conventional neuropsychological tests, separate generalized linear mixed models (GLMMs) were conducted for each instrument. Age, sex, education, race/ethnicity, ApoE ε4 status, test session, test score, and a test score*test session interaction were entered as fixed effects for each GLMM. The interaction variable accounted for difference in cognitive slopes over serial assessments. An additional covariate, test location, was

added to the OCL GLMM to account for in clinic versus remote test administration. No random effects were entered as doing so significantly reduced model fit. The GLMMs assumed a binomial logistic distribution and logit link as the two outcomes were coded categorically.

Results

Primary Objective: Baseline Assessments Predicting Conversion to aMCI or AD

Sample Characteristics

Demographics for the 339 participants included in this analysis are summarized in Table 1. Ages ranged from 56.0 to 92.1 years. The sample was mostly female, highly educated (10-20 years), and White. The ethnoracial breakdown of non-White individuals included Indigenous (*n*= 1, 0.3%), Asian (*n*= 3, 0.9%), Black (*n*= 17, 5.0%), and mixed (*n*= 9, 2.7%) participants. About one third of the sample was ApoE ε4+. Thirty (8.8%) individuals converted from cognitive health to aMCI (*n*= 27, 8.0%) or AD (*n*= 3, 0.9%). Average time to conversion was 3.2 years. There were no significant differences in demographics between conversion groups (Appendix Table A1).

Table 1

Sample Characteristics for Primary Analyses: Baseline Assessments Predicting aMCI or AD

N= 339	*M* (*SD*) / *N* (%)
Age (years)	73.4 (7.2)
Sex (female)	197 (58.1)
Education (years)	16.8 (2.3)
Race/ethnicity (White)	309 (91.2)
ApoE ε4+	110 (32.4)

Cognitive Assessment Measures

Descriptive cognitive scores for nonconverters and converters are shown in Table 2. Mean OCL accuracy scores are presented as raw arcsine square-root transformed values. Mean OCL scores

for CN participants were similar (*M*= 0.987, *SD*= 0.109) to previous research (*M*= 1.0 *SD*= 0.10) (White et al., 2021). While converters performed slightly worse than controls at baseline, this difference was non-significant and represented only approximately a -0.5 SD change from CN performance. Baseline conventional MoCA and RAVLT scores were statistically different between groups. At baseline, participants who remained CN had higher total MoCA scores by two points and recalled about two more words on the RAVLT 30-minute delay.

Table 2

Cognitive Scores for Primary Analyses: Baseline Assessments Predicting aMCI or AD

	Total Sample N= 339	*Cognitively Normal n= 309*	*Incident aMCI/AD n= 30*	p^a
OCL accuracy	0.984 (0.108)	0.987 (0.109)	0.947 (0.099)	.052
MoCA total score	26.5 (2.6)	26.6 (2.6)	24.6 (2.3)	<.001
RAVLT 30-minute delay	8.4 (4.1)	8.5 (4.1)	6.4 (3.6)	.007

[a] Indicates significant differences between cognitively normal and incident aMCI or AD groups.

Logistic Regression Analyses

The OCL logistic regression was assessed for goodness-of-fit using the Hosmer-Lemeshow test, $\chi^2(8)= 7.67$, $p= .466$, indicating adequate fit. The model was not statistically significant compared to the null model, $\chi^2(6)= 6.22$, $p= .399$, Nagelkerke's $R^2= .040$, indicating that predictors did not significantly improve model fit. Baseline OCL accuracy did not significantly predict conversion to aMCI or AD (Appendix Table B1). The MoCA logistic regression demonstrated adequate goodness-of-fit, $\chi^2(8)= 15.46$, $p= .051$. The model was statistically significant compared to the null model, $\chi^2(6)= 16.01$, $p= .014$, Nagelkerke's $R^2= .102$. Baseline total MoCA score was the only significant predictor of conversion (Appendix Table B2). The RAVLT logistic regression had adequate goodness-of-fit, $\chi^2(8)= 5.12$, $p= .745$, however the model was not statistically significant compared to the null model, $\chi^2(6)= 9.81$, $p= .133$, Nagelkerke's $R^2= .063$ (Appendix Table B3).

AUC Analyses

Unadjusted AUCs for each cognitive assessment are shown in Table 3. Baseline OCL scores did not significantly predict conversion above chance level, while baseline MoCA and RAVLT scores did. The MoCA provided the strongest predictive power (Appendix Figure C1). Maximized sensitivity and specificity values using Youden's index produced an optimal baseline OCL arcsine transformed accuracy score of 1.035 (sensitivity= 32.7%, specificity= 90.0%), MoCA total score of 25.5 (sensitivity= 68.6%, specificity= 73.3%), and RAVLT 30-minute delay score of 8.5 (sensitivity= 54.7%, specificity= 76.7%).

After adjusting for age, sex, education, ethnicity/race, and ApoE ε4 status, the three AUCs predicted conversion to aMCI or AD above chance as confidence intervals did not cover the 0.50 AUC threshold (Table 4). The MoCA AUC model had the strongest predictive power, followed by the RAVLT and then the OCL (Appendix Figure C2). However, logistic regression results indicated that neither the OCL nor RAVLT models improved prediction above the null model.

Table 3

Unadjusted Area Under the Receiver Operator Curves for Predicting Conversion to aMCI or AD

	AUC	Sensitivity	Specificity	*p*	95% CI	
					Lower	Upper
OCL accuracy	.603	32.7	90.0	.064	.507	.750
MoCA total score	.732	68.6	73.3	<.001	.646	.818
RAVLT 30-minute delay	.658	54.7	76.7	.004	.566	.750

Table 4

Adjusted Area Under the Receiver Operator Curves for Predicting Conversion to aMCI or AD

	AUC	*p*	95% CI	
			Lower	Upper
OCL accuracy	.650	.007	.561	.739
MoCA total score	.735	<.001	.652	.818
RAVLT 30-minute delay	.678	.001	.589	.766

Primary Objective: Repeated Assessments Predicting Conversion to aMCI or AD

Sample Characteristics

Demographic characteristics for the 301 participants included in this analysis are summarized in Table 5. Ages ranged from 56.0 to 91.4 years and participants were mostly female. The sample was highly educated (10 to 20 years) and mainly White. The ethnoracial breakdown of non-White individuals included Indigenous (*n*= 1, 0.3%), Asian (*n*= 2, 0.7%), Black (*n*= 14, 4.7%), and mixed (*n*= 9, 3.0%) participants. Approximately one third of the sample was ApoE ε4+. Twenty-five (8.3%) individuals included in the analysis converted from cognitive health to aMCI (*n*= 22, 7.3%) or AD (*n*= 3, 1.0%). The average time to conversion was 3.1 years. There were no significant differences in sample characteristics between conversion groups (Appendix Table A2).

Table 5

Sample Characteristics: Repeated Assessments Predicting Conversion to aMCI or AD

N= 301	*M* (*SD*) / *N* (%)
Age (years)	73.1 (7.0)
Sex (female)	175 (58.1)
Education (years)	16.9 (2.3)
Race/ethnicity (White)	275 (91.4)
ApoE ε4+	99 (32.9)

Cognitive Assessment Measures

Participants completed a mean number of 10.0 OCL sessions (*SD*= 6.6) while the median number was 8 and the range was 2 to 49. For testing location, 21.8% of OCL administrations were conducted in clinic and 78.2% were performed remotely. The sample completed an average of 2.8 MoCA sessions (*SD*= 0.8) with a median of 3 sessions and a range of 2 to 6. Mean RAVLT sessions was 2.8 (*SD*= 0.8) with a median of 3 sessions and a range of 2 to 6.

Generalized Linear Mixed Models

Model fit was assessed for the OCL GLMM using Akaike Information Criterion (AIC)= 1548.03. Results showed that the main effect of repeated OCL accuracy was a significant predictor of conversion, where higher scores predicted lower risk of conversion (Appendix Table D1). Being female, non-White, and remote administration were associated with reduced risk of aMCI or AD.

Model fit for the MoCA GLMM, AIC= 484.74, was stronger than the OCL GLMM. Results indicated that the main effect of repeated total MoCA score was a significant predictor of conversion, where higher MoCA scores were associated with lower risk of conversion (Appendix Table D2). Being non-White was also associated with reduced risk of incident aMCI or AD. Increased number of MoCA sessions was associated with lower risk of conversion. The interaction between MoCA session and MoCA total score was significant, whereby aMCI or AD converters demonstrated enhanced practice effects over serial MoCA administrations compared to CN peers.

Model fit for the RAVLT GLMM, AIC= 534.64, was stronger than the OCL model but weaker than the MoCA model. Results indicated that the main effect of repeated RAVLT scores significantly predicted conversion, where higher RAVLT scores were associated with lower risk of incident aMCI or AD (Appendix Table D3). No other variables predicted conversion.

Secondary Objective: Baseline Assessments Associated with Aβ Status

Sample Characteristics

Demographic characteristics for the 598 participants included in this analysis are summarized in Table 6. Participant's ages ranged from 56.0 to 97.5 years. The sample was mostly female, highly educated (range of 8 to 20 years), and White. The ethnoracial composition of non-White individuals included Indigenous (*n*= 2, 0.3%), Asian (*n*= 9, 1.5%), Black (*n*= 25, 4.2%), and mixed

(*n*= 12, 2.0%) participants. There were 210 (35.1%) ApoE ε4+ participants. Two-hundred-thirty-four (39.1%) individuals were Aβ+ at baseline. Participants who were Aβ+ at baseline were significantly older and more likely to be ApoE ε4+ (Appendix Table A3).

Table 6

Sample Characteristics of the Secondary Analyses: Baseline Assessments Predicting Aβ Status

N= 598	*M* (*SD*) / *N* (%)
Age (years)	73.7 (7.5)
Sex (female)	314 (52.5)
Education (years)	16.6 (2.4)
Race/ethnicity (White)	550 (92.0)
ApoE ε4+	210 (35.1)

Cognitive Assessment Measures

Descriptive cognitive scores for Aβ status groups at baseline are shown in Table 7. Mean OCL accuracy scores are presented as raw arcsine square-root transformed values. Baseline OCL scores were significantly different between Aβ status groups, where Aβ+ individuals performed worse than their Aβ- counterparts. Baseline conventional neuropsychological scores for the MoCA and RAVLT were statistically different between Aβ status groups. Participants who were Aβ- at baseline had higher total MoCA scores by approximately one point and recalled, on average, 1.5 more words on the RAVLT 30-minute delay.

Table 7

Baseline Cognitive Scores of the Secondary Analyses: Single Assessments Predicting Aβ Status

	Total Sample N= 598	*Aβ- at Baseline n= 364*	*Aβ+ at Baseline n= 234*	*p*[a]
OCL accuracy	0.956 (0.114)	0.965 (0.116)	0.943 (0.110)	.022
MoCA total score	25.3 (3.3)	25.7 (3.1)	24.6 (3.6)	<.001
RAVLT 30-minute delay	7.1 (4.4)	7.7 (4.3)	6.2 (4.5)	<.001

[a] Indicates significant differences between Aβ status groups.

Logistic Regression Analyses

The OCL logistic regression demonstrated adequate goodness-of-fit as measured by the Hosmer-Lemeshow test, $\chi^2(8)= 5.47$ $p= .708$. The model was statistically significant compared to the null model, $\chi^2(6)= 129.87$, $p< .001$, Nagelkerke's $R^2= .265$, indicating that selected predictors significantly improved model fit. Baseline OCL accuracy was not a significant predictor of Aβ status (Appendix Table B4). Both increased age and ApoE ε4+ were associated with Aβ+ status.

The MoCA logistic regression was statistically significant compared to the null model, $\chi^2(6)= 133.88$, $p< .001$, Nagelkerke's $R^2= .272$, and demonstrated adequate fit, $\chi^2(8)= 6.67$, $p= .572$. Baseline total MoCA scores significantly predicted Aβ, where higher MoCA scores were associated with decreased risk of Aβ+ at baseline (Appendix Table B5). Both increased age and ApoE ε4+ were also associated with Aβ+.

The RAVLT logistic regression was statistically significant compared to the null model, $\chi^2(6)= 135.38$, $p< .001$, Nagelkerke's $R^2= .275$, and demonstrated adequate fit, $\chi^2(8)= 3.84$, $p= .872$. Baseline RAVLT scores significantly predicted Aβ status, where higher recall was associated with decreased risk of Aβ+ (Appendix Table B6). Increased age, being female, and ApoE ε4+ status was associated with Aβ+ status at baseline.

AUC Analyses

Unadjusted AUCs for each cognitive assessment are shown in Table 8. Baseline OCL scores significantly predicted Aβ status above chance level. Baseline MoCA and RAVLT performance also significantly predicted Aβ status, with all three cognitive measures providing similar prediction (Appendix Figure C3). Maximized sensitivity and specificity values using Youden's index produced an optimal OCL raw arcsine transformed cutoff score of 0.940 (sensitivity= 55.8%,

specificity= 53.8%), MoCA total score cutoff of 23.5 (sensitivity= 78.8%, specificity= 36.8%), and RAVLT 30-minute delay threshold of 5.5 (sensitivity= 66.8%, specificity= 48.7%).

After adjusting for age, sex, education, ethnicity/race, and ApoE ε4 status, the three adjusted AUCs predicted Aβ status significantly above chance level (Table 9). From the AUC results, baseline OCL performance was similarly predictive of Aβ compared to the MoCA and RAVLT AUCs (Appendix Figure C4). However, from the logistic regression results, the OCL did not significantly predict Aβ status, while the MoCA and RAVLT did.

Table 8

Unadjusted AUCs for Baseline Cognitive Scores Predicting Aβ Status

	AUC	Sensitivity	Specificity	*p*	95% CI	
					Lower	Upper
OCL accuracy	.556	55.8%	53.8%	.020	.509	.603
MoCA total score	.588	78.8%	36.8%	<.001	.541	.636
RAVLT 30-minute delay	.596	66.8%	48.7%	<.001	.549	.643

Table 9

Adjusted AUCs for Baseline Cognitive Scores Predicting Aβ Status

	AUC	*p*	95% CI	
			Lower	Upper
OCL accuracy	.764	<.001	.725	.804
MoCA total score	.766	<.001	.726	.805
RAVLT 30-minute delay	.766	<.001	.726	.806

Secondary Objective: Repeated Assessments Associated with Aβ Accumulation

Sample Characteristics

Demographic characteristics for the 331 participants included in this analysis are summarized in Table 10. Ages ranged from 56.0 to 91.5 years and participants were mostly female. The sample

was highly educated (8 to 20 years) and mostly White. The ethnoracial breakdown of non-White individuals included Asian (*n*= 1, 0.3%), Black (*n*= 11, 3.3%), and mixed (*n*= 9, 2.7%) participants. There were 130 (39.3%) ApoE ε4+ participants. One-hundred-twenty-three (37.2%) individuals showed pathological Aβ accumulation over the study period. Participants who were Aβ+ at baseline were more likely to be ApoE ε4+ (Appendix Table A4).

Table 10

Sample Characteristics: Repeated Assessments Predicting Aβ Accumulation

N= 331	*M* (*SD*) / *N* (%)
Age (years)	73.1 (7.2)
Sex (female)	180 (54.4)
Education (years)	16.6 (2.5)
Race/ethnicity (White)	310 (93.7)
ApoE ε4+	130 (39.3)

Cognitive Assessment Measures

Participants completed a mean of 10.2 OCL sessions (*SD*= 6.7) while the median was 9. OCL sessions ranged from 2 to 49. For testing location, 24.4% of OCL administrations were conducted in clinic and 75.6% were performed remotely. The sample's mean completed MoCA sessions was 3.2 (*SD*= 0.9) with a median of 3 sessions. MoCA sessions ranged from 2 to 6. Similarly, mean RAVLT sessions was 3.2 (*SD*= 0.9) with a median of 3. RAVLT sessions ranged from 2 to 6.

Generalized Linear Mixed Models

Model fit for the OCL GLMM produced AIC= 3973.14. Results indicated that the main effect of repeated OCL accuracy was a significant predictor of Aβ accumulation, where higher scores indicated lower risk of Aβ accumulation (Appendix Table D4). Increased age and ApoE ε4+ were associated with increased risk of Aβ accumulation, while being non-White predicted lower risk of

pathological Aβ accumulation. The main effect of increased OCL sessions was associated greater Aβ accumulation and the interaction between OCL session and score was significant, whereby individuals with pathological Aβ accumulation over the study period demonstrated diminished practice effects over serial OCL administrations compared to Aβ- peers (Appendix Figure E1).

Model fit for the MoCA GLMM, AIC= 1230.85, was stronger than for the OCL model. Results indicated that the main effect of repeated total MoCA score significantly predicted Aβ accumulation, where higher MoCA scores predicted lower risk of Aβ accumulation (Appendix Table D5). Increased age, being female, and ApoE ε4+ were associated with higher risk of Aβ accumulation, while being non-White predicted lower risk of pathological Aβ accumulation.

Model fit for the RAVLT GLMM, AIC= 1243.89, was stronger than the OCL model, but weaker than the MoCA model. Results indicated that the main effect of repeated RAVLT 30-minute delay score was a significant predictor of Aβ accumulation, where higher RAVLT scores were associated with lower risk of Aβ accumulation (Appendix Table D6). Increased age, being female, and ApoE ε4+ were associated with higher risk of Aβ accumulation, while being non-White predicted lower risk of pathological Aβ accumulation. In addition, the main effect of RAVLT sessions was associated with lower risk of Aβ accumulation.

Discussion

The present study investigated the clinical utility of a computerized assessment, the OCL, to predict two markers of neurodegeneration in a community-dwelling older adult sample. To my knowledge, this is the first investigation of CCA performance to predict conversion to aMCI or AD from cognitive health and associate with Aβ imaging markers compared to a gold-standard conventional screening tool, the MoCA, using single and repeated assessments. Results showed

that baseline OCL scores did not provide comparable efficacy to the MoCA and RAVLT to predict aMCI or AD conversion and Aβ status. When assessed over repeated administrations, the OCL significantly associated with incident dementia and Aβ accumulation over four years, however computerized performance did not provide improved prediction of outcomes compared to conventional tests. In line with expectations, the OCL uniquely detected diminished practice effects associated with Aβ accumulation. Specific results for each objective are explored in detail.

For the primary objective, it was hypothesized that baseline OCL performance would provide equivalent utility to the MoCA and RAVLT to identify which CN individuals will go on to develop aMCI or AD. This hypothesis was not supported by the observed results. Unadjusted baseline OCL scores did not significantly predict conversion. In line with previous research, unadjusted MoCA and RAVLT scores identified risk of progression to aMCI or AD from cognitive health (Hassenstab et al., 2021; Qin, Zhao, Zhu, & Hu, 2020; Yue et al., 2021). Baseline MoCA performance produced stronger predictive power in the current study (AUC= .73) compared to Hassenstab et al. (2021) (AUC= .64), although these authors utilized a shorter follow-up period and normative cutoffs for prediction. The MoCA AUCs were the only models to provide adequate prognostic validity (AUC > .70) for identifying incident aMCI or AD over a four-year-period, indicating that baseline MoCA performance alone provides clinically valuable information (Swets, 1996). Adjusted AUC analyses for the OCL and RAVLT indicated that when accounting for demographic and clinical variables, both instruments can successfully predict conversion to aMCI or AD from cognitive health. However, the significant AUCs conflicted with non-significant omnibus results from the accompanying OCL and RAVLT logistic regression models. One potential explanation for discrepant results is that AUC analyses are less robust for testing

incremental prediction and are biased towards inflating estimation upwards compared to logistic regressions (Vickers, Cronin, & Begg, 2011). In this case, the regression results are preferred.

Poor prediction of conversion from baseline OCL scores is somewhat unexpected in the context of the current literature. The OCL produced less accurate prediction of future conversion to aMCI or AD than two previous studies drawn from the Mayo Clinic Study of Aging. Pudumjee et al. (2021) found that among a comparable sample of community-dwelling older adults, age-adjusted OCL scores at baseline significantly predicted conversion to MCI over a period of 4.8 years. Similarly, Stricker, Lundt, Albertson, et al. (2020) determined that baseline OCL showed predictive power to identify MCI incidence over 3.2 years that was comparable to the RAVLT when assessed at baseline. The discrepancy in results between the present study and previous research may be attributed to differences in MCI subtype for diagnosis. In the current study, participants were required to demonstrate memory impairments suggestive of an amnestic type disorder, while the above studies included alternative impairments sufficient for diagnosis such as executive function, language, and visuospatial domains. The OCL is theorized to be most sensitive to memory dysfunction as it is designed as a measure of visual learning and recognition (Hammers et al., 2012; Lim et al., 2012). However, the OCL may also tap into domains like executive functioning, making this task sensitive to early impairments seen in multidomain and naMCI, and thus more effective for predicting conversion as seen in Pudumjee et al. (2021) and Stricker, Lundt, Albertson, et al. (2020) (Kirova, Bays, & Lagalwar, 2015). This notion is supported by evidence that the OCL is moderately correlated with executive function tasks in participants who are cognitively healthy as well as those who are demented (Maruff et al., 2009; Racine et al., 2016).

In addition, the OCL's poorer prognostic validity compared to the RAVLT and MoCA may be attributed to the specific memory domains assessed in the cognitive instruments evaluated in

this study. As the OCL measures visual learning and recognition capacity, it may provide a less sensitive estimate of early neurodegenerative processes compared to the RAVLT and MoCA, as relevant specifically to the AD continuum. Previous research suggests that while memory impairment is the earliest detectable cognitive domain affected in AD, verbally mediated memory decline may precede visual memory impairments by up to four years (Mistridis et al., 2015). The RAVLT 30-minute delayed recall score, as a direct measure of verbal episodic memory recall, may be more sensitive to objective decline in the early transition from cognitive health to aMCI than the OCL, which is visually mediated and utilizes a recognition paradigm. While the MoCA broadly measures global cognition, it also contains a verbal memory recall subscale that may make this conventional tool more sensitive to the earliest objective cognitive impairments in aMCI.

For the primary objective, it was hypothesized that repeated OCL testing would better predict incident aMCI or AD compared to conventional neuropsychological measures. This hypothesis was not supported by the observed results. Model fit indicators suggested that the OCL repeated measures analysis was outperformed by the RAVLT and MoCA models, with the latter having the strongest fit overall. While the OCL model provided the least robust estimate of conversion, the main effect of OCL performance, accounting for average OCL accuracy scores across all visits, did significantly predict conversion. One explanation for why the OCL model less effectively predicted conversion is that task difficulty was inadequately calibrated for subtle cognitive decrements expected in a CN sample. Recent research has shown deficient mean performance in CN adults and floor effects in AD groups with the OCL version used in ADNI3 (White et al., 2021). The authors suggest that test difficulty may lead to restricted score ranges that inadequately distinguish between healthy and cognitively impaired individuals. A modified OCL version with reduced test difficulty demonstrated improved efficacy to differentiate CN and AD

groups (White et al., 2021). While, to this author's knowledge, there is no validation study for the modified OCL version in MCI, it is possible that reducing test difficulty may also improve sensitivity to detect future conversion. This is a question for future research to explore.

The interaction variable for change in OCL accuracy over continuous sessions was not a significant predictor of conversion to aMCI or AD, indicating that OCL performance trajectories did not differ between converters and non-converters. This finding contrasted with previous research that showed a significant effect of repeated OCL scores to identify incident MCI over a similar follow-up period (Pudumjee et al., 2021). Change in OCL scores may have provided poorer prediction of incident aMCI or AD over serial administrations due to contextual factors that were not accounted for in this study. ADNI3 participants completed a mixture of supervised in-clinic and unsupervised remote OCL sessions, with about four out of every five OCLs conducted remotely. Research on the effects of supervision and test-taking location on OCL psychometrics show that these contextual factors may affect computerized test performance: OCL within-subject performance differences as measured by intra class correlation have moderate reliability (ICC= .61) between supervised and unsupervised conditions, indicating that the use of both in ADNI3 may have introduced additional error variability when estimating longitudinal cognitive trajectories (Cromer et al., 2015). While one study showed that mean OCL accuracy scores were approximately 0.5 SDs lower when administered remotely, a second study demonstrated negligible mean differences between testing locations (Mielke et al., 2015; Stricker, Lundt, Alden, et al., 2020). Future longitudinal research on the use of computerized tools to track cognitive progression may benefit from consistent study designs regarding supervision and test taking location.

The MoCA was the only cognitive measure that provided significant prediction of aMCI or AD conversion from cognitive health based on change in performance trajectories over testing

sessions. Visual inspection of group scores revealed that, unexpectedly, converter's MoCA scores improved over serial administrations while nonconverter performance was relatively stable. It is surprising that improved MoCA trajectories predicted incident dementia while CN participants demonstrated flat performance slopes over time, however this finding may be an artefact of the study design. CN groups were assessed in clinic for conventional neuropsychological testing once every two years in ADNI3. If a clinical diagnosis of aMCI or AD was made, participants were then evaluated in clinic once annually, meaning that by design, aMCI/AD converters have more conventional testing sessions. Post hoc analyses confirmed that converters had significantly more MoCA sessions (M= 3.5, SD= 0.7) compared to nonconverters (M= 2.8, SD= 1.2), p= .006, which accounts for the main effect of session on the conversion outcome in this analysis. One explanation for the significant interaction is that converters benefited from additional MoCA trials with stronger practice effects while nonconverters received fewer opportunities to learn from practice. Indeed, previous research indicates that the magnitude of practice effects in cognitive testing is influenced by the number of testing sessions in both healthy and cognitively impaired adults (Bartels, Wegrzyn, Wiedl, Ackermann, & Ehrenreich, 2010; Sanderson-Cimino et al., 2023).

For the secondary objective, it was hypothesized that baseline OCL performance would provide equivalent utility to the MoCA and RAVLT to associate with Aβ. This hypothesis was partially supported by the observed results. In the present study, unadjusted AUCs for each instrument demonstrated similar and significant, albeit low, efficacy to identify Aβ status. While the AUCs between the three cognitive measures were similar, there were notable differences in maximized sensitivity and specificity values based on Youden's index. The OCL produced the highest optimized specificity value (53.8%) compared to the MoCA (36.8%) and RAVLT (48.7%) but had the lowest sensitivity levels (55.8%) in contrast to the MoCA (78.8%) and RAVLT

(66.8%). Relatively low sensitivity values for the OCL indicate that it is an inadequate tool for screening for Aβ status in community-dwelling older adults when used alone (Swets, 1996). Clinical practice guidelines for identifying dementia risk recommend that validated screening tools with optimal sensitivity values be used as a prelude to formal diagnostic assessment in older adults with suspected risk of neurodegenerative disease (Petersen et al., 2018)

Adjusted AUCs for the OCL, MoCA, and RAVLT were nearly identical to each other and above the recommended AUC threshold of .70. However, logistic regression results showed that OCL accuracy was not a significant predictor of Aβ status when adjusting for demographic and clinical variables. This suggests that in the OCL AUC, age and ApoE ε4 mainly contributed to the prognostic validity of the analysis. In contrast, the MoCA and RAVLT logistic regressions confirmed that both conventional measures maintained their predictive power when accounting for added demographic and clinical variables. Results help distinguish mixed findings from previous research on baseline OCL scores predicting Aβ. The CBB visual memory (OCL) and executive function (One Back) composite, used in two previous studies, may provide a better estimate of PET Aβ+ than OCL performance alone (Alden et al., 2021; Pudumjee et al., 2021). Apart from the OCL's correlation with conventional executive function measures and restricted score ranges discussed above, this finding may be related to enhanced measurement reliability with composite scores versus individual test scores. The use of composite cognitive scores has previously shown improved estimation of cognition by reducing error variability using mean scores, making significance testing more feasible (Jonaitis et al., 2019).

It was predicted that repeated OCL scores would better associate with Aβ accumulation compared to conventional neuropsychological measures. This hypothesis was partially supported. Model fit indices suggested that the OCL repeated measures analysis was outperformed by the

RAVLT and MoCA models, with the latter having the strongest fit overall. The main effects for all three cognitive measures, accounting for average scores across all visits, significantly predicted Aβ accumulation in the expected direction. These findings align with a recent systematic review that shows that conventional global cognition and episodic memory measures can effectively associate with Aβ accumulation (Parent, Rousseau, Predovan, Duchesne, & Hudon, 2023). For the OCL, it is likely that restricted score distributions and contextual factors, such as differences in test location and supervision, impacted prediction of Aβ accumulation similar to the primary repeated analyses. While the repeated measures analyses used in this study do not allow direct comparison of main effects between models, findings indicate that overall, conventional tools used in clinical practice provide better estimation of concurrent Aβ pathological processes.

Interestingly, change in OCL performance was the only cognitive measure to significantly predict Aβ accumulation based on slope estimation over serial assessments. Visual plot inspection revealed that the Aβ+ accumulation group had diminished practice effects on OCL accuracy across serial administrations compared to Aβ- peers. A post hoc analysis to evaluate whether the significant interaction was driven by differences in OCL sessions indicated that the number of completed trials was similar for Aβ+ (M= 11.0, SD= 7.1) and Aβ- (M= 10.1, SD= 6.8) accumulation groups, p= .289. While several studies have found that CCAs can successfully differentiate between individuals once they reach clinically significant thresholds that warrant biomarker positivity (Lim et al., 2014; Pudumjee et al., 2021; Young et al., 2023), fewer studies have examined the use of computerized tools to detect concurrent pathological biomarker accumulation (Lim et al., 2015). This is an important distinction as amyloid accumulation is estimated to be one of the earliest detectable biomarkers in AD, emerging 15-20 years prior to clinical diagnosis (Rowe et al., 2007). While this study used an amyloid-enriched older adult

sample, suggesting that AD processes were well-underway, future research on the OCL and associations with cognitive trajectories and biomarker accumulation will inform whether this feasible and cost-efficient tool can detect AD pathology earlier in the lifespan.

Clinical Implications

Findings from this study elucidate potential clinical applications of a novel computerized tool to detect incipient neurodegeneration and AD brain pathology in community-dwelling older adults. Practitioners are tasked with selecting appropriate screening measures for specific clinical questions from an array of available tools. By comparing the clinical utility of the OCL to existing conventional screening instruments, this research provides useful information to aid clinical decision-making on test selection. Results indicate that the OCL does not offer additional utility for predicting progression to aMCI or AD compared to the MoCA in CN older adults. A single MoCA assessment alone provided 68.8% sensitivity to detect dementia incidence over four years. As it stands, the potential practical utility of computerized screening through remote, unsupervised monitoring does not offset loss of prognostic accuracy in the OCL. This pattern was also replicated for serial OCL scores to predict aMCI or AD incidence, where global cognition as measured by the MoCA demonstrated superior outcome estimation. Routine conventional screening in primary care may currently provide clinicians with the best method for assessing risk of cognitive decline.

Examination of the comparative clinical utility of the OCL to associate with Aβ burden and accumulation provided more equivocal results. Cognitive testing with the OCL may be useful for clinicians and researchers seeking to identify AD as a biological entity. When considered in isolation, all three cognitive measures had similar, albeit low, prediction of Aβ status, signifying that a single OCL score provides equivalent predictive power to the MoCA and RAVLT. This

finding suggests that in circumstances when paper-and-pencil screening is not possible, such as in rural areas with low access to trained professionals, a computerized test may supplant conventional assessment methods to identify Aβ burden. While adjusted estimates showed that the MoCA provided the highest sensitivity (78.8%) to Aβ burden, the overall predictive power was strongly influenced by ApoE ε4 status which is a costly and relatively inaccessible data point compared to cognitive information. Repeated analyses showed that the MoCA provided the most accurate overall estimation of Aβ accumulation. However, OCL accuracy slopes uniquely detected subtle changes in visual learning and recognition performance associated with preclinical AD pathology compared to conventional screening tools. This finding suggests that serial OCL performance may provide a viable preclinical AD monitoring alternative where collection of costly biomarker data is not feasible. Further research is needed to determine whether repeated computerized assessments can detect practice effects associated with amyloid biomarkers in other neurodegenerative disorders, or if this cognitive marker is unique to AD-related processes.

Strengths and Limitations

There were several notable strengths of the present study. First, the relatively large sample size and longitudinal observational design increased the internal and external validity of the study. Longitudinal observation over a four-year-span allowed for detection of the earliest emerging cognitive and biological markers of neurodegeneration as they occur in community-dwelling older adults. Second, the selected outcomes permitted comparison of the clinical utility of various cognitive tools to associate with clinical and biological definitions of neurodegenerative disease. There is currently theoretical debate surrounding what constitutes preclinical AD, and thus what outcomes current screening tools ought to target in clinical practice and therapeutic research. This

study provided useful data for both AD definitions. Third, results have clear clinical implications for neuropsychological test selection as digitally mediated tools become more prevalent in research and clinical settings. Finally, this study provided an independent evaluation of the OCL as much previous research is linked to CogState Ltd. affiliates, the developers of the CBB.

This study is not without limitations. First, ADNI3 is subject to sampling bias given participants' high levels of education, homogenously White ethnoracial composition, and relative amyloid-enrichment. Due to these factors, results from the current study may lack widespread applicability to the general Canadian population. Second, the diagnostic criteria utilized to determine MCI status relied on episodic memory dysfunction over alternative cognitive domains, such as executive function or attention. Thus, the results from this study are specific to aMCI versus naMCI, which represents an estimated one out of every three MCI cases (Petersen, 2011). Findings should be interpreted with caution when applied to naMCI groups. In addition, the Petersen MCI diagnostic criteria utilized in this study require the presence of subjective cognitive impairment for diagnosis, however there is ongoing debate about the validity of subjective complaints as a true marker of neurodegeneration (Edmonds et al., 2014; Hessen et al., 2017). Results should be interpreted as specific to identifying aMCI with accompanying subjective impairment. Finally, there are limitations related to OCL testing procedures. As ADNI's main objectives are related to longitudinal aging and biomarkers of dementia, the OCL was not a required component for enrolment in the study. There is likely bias in the sample of participants who completed computerized testing versus those who did not. Confounding factors, such as low computer familiarity and presence of environmental distractions during remote testing, were not accounted for in this study and may have affected OCL performance trajectories.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., . . . 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 workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 270-279. doi:10.1016/j.jalz.2011.03.008

Alden, E. C., Pudumjee, S. B., Lundt, E. S., Albertson, S. M., Machulda, M. M., Kremers, W. K., . . . Stricker, N. H. (2021). Diagnostic accuracy of the Cogstate Brief Battery for prevalent MCI and prodromal AD (MCI A(+) T(+)) in a population-based sample. *Alzheimers Dement, 17*(4), 584-594. doi:10.1002/alz.12219

Alzheimer Society of Canada. (2016). Report summary Prevalence and monetary costs of dementia in Canada (2016): a report by the Alzheimer Society of Canada. *Health Promot Chronic Dis Prev Can, 36*(10), 231-232. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27768560

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5 ed.).

American Psychological Association. (2021). *Guidelines for the Evaluation of Dementia and Age-Related Cognitive Change*.

Aprahamian, I., Nunes, P. V., & Forlenza, O. V. (2013). Cognitive impairment and dementia in late-life bipolar disorder. *Curr Opin Psychiatry, 26*(1), 120-123. doi:10.1097/YCO.0b013e32835ac5f6

Bae, J. B., Han, J. W., Kwak, K. P., Kim, B. J., Kim, S. G., Kim, J. L., . . . Kim, K. W. (2018). Impact of Mild Cognitive Impairment on Mortality and Cause of Death in the Elderly. *J Alzheimers Dis, 64*(2), 607-616. doi:10.3233/JAD-171182

Baker, J. E., Lim, Y. Y., Pietrzak, R. H., Hassenstab, J., Snyder, P. J., Masters, C. L., & Maruff, P. (2017). Cognitive impairment and decline in cognitively normal older adults with high amyloid-beta: A meta-analysis. *Alzheimers Dement (Amst), 6*, 108-121. doi:10.1016/j.dadm.2016.09.002

Balthazar, M. L., Yasuda, C. L., Cendes, F., & Damasceno, B. P. (2010). Learning, retrieval, and recognition are compromised in aMCI and mild AD: are distinct episodic memory processes mediated by the same anatomical structures? *J Int Neuropsychol Soc, 16*(1), 205-209. doi:10.1017/S1355617709990956

Banh, T., Jin, C., Neuhaus, J., Mackin, R. S., Maruff, P., Stricker, N., . . . Nosheny, R. L. (2022). Unsupervised Performance of the CogState Brief Battery in the Brain Health Registry: Implications for Detecting Cognitive Decline. *J Prev Alzheimers Dis, 9*(2), 262-268. doi:10.14283/jpad.2021.68

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neurosci, 11*, 118. doi:10.1186/1471-2202-11-118

Bellone, J., Van Patten, R. (Hosts). (January 1, 2020). *Navigating Neuropsychology*. [Audio podcast episode]

Bilder, R. M. (2011). Neuropsychology 3.0: evidence-based science and practice. *J Int Neuropsychol Soc, 17*(1), 7-13. doi:10.1017/S1355617710001396

Binng, D., Splonskowski, M., & Jacova, C. (2020). Distance Assessment for Detecting Cognitive Impairment in Older Adults: A Systematic Review of Psychometric Evidence. *Dement Geriatr Cogn Disord, 49*(5), 456-470. doi:10.1159/000511945

Bondi, M. W., Jak, A. J., Delano-Wood, L., Jacobson, M. W., Delis, D. C., & Salmon, D. P. (2008). Neuropsychological contributions to the early identification of Alzheimer's disease. *Neuropsychol Rev, 18*(1), 73-90. doi:10.1007/s11065-008-9054-1

Briceno, E. M., Mehdipanah, R., Gonzales, X. F., Langa, K. M., Levine, D. A., Garcia, N. M., . . . Morgenstern, L. B. (2020). Neuropsychological assessment of mild cognitive impairment in Latinx adults: A scoping review. *Neuropsychology, 34*(5), 493-510. doi:10.1037/neu0000628

Brooks, B. L., Holdnack, J. A., & Iverson, G. L. (2011). Advanced clinical interpretation of the WAIS-IV and WMS-IV: prevalence of low scores varies by level of intelligence and years of education. *Assessment, 18*(2), 156-167. doi:10.1177/1073191110385316

Brooks, B. L., Iverson, G. L., Holdnack, J. A., & Feldman, H. H. (2008). Potential for misclassification of mild cognitive impairment: a study of memory scores on the Wechsler Memory Scale-III in healthy older adults. *J Int Neuropsychol Soc, 14*(3), 463-478. doi:10.1017/S1355617708080521

Brooks, B. L., Iverson, G. L., & White, T. (2007). Substantial risk of "Accidental MCI" in healthy older adults: base rates of low memory scores in neuropsychological assessment. *J Int Neuropsychol Soc, 13*(3), 490-500. doi:10.1017/S1355617707070531

Brooks, B. L., Iverson, G. L., & White, T. (2009). Advanced interpretation of the Neuropsychological Assessment Battery with older adults: base rate analyses, discrepancy

scores, and interpreting change. *Arch Clin Neuropsychol, 24*(7), 647-657. doi:10.1093/arclin/acp061

Bucci, M., Chiotis, K., Nordberg, A., & Alzheimer's Disease Neuroimaging, I. (2021). Alzheimer's disease profiled by fluid and imaging markers: tau PET best predicts cognitive decline. *Mol Psychiatry, 26*(10), 5888-5898. doi:10.1038/s41380-021-01263-2

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol, 26*(4), 543-570. doi:10.1080/13854046.2012.680913

Callahan, B. L., Ramirez, J., Berezuk, C., Duchesne, S., Black, S. E., & Alzheimer's Disease Neuroimaging, I. (2015). Predicting Alzheimer's disease development: a comparison of cognitive criteria and associated neuroimaging biomarkers. *Alzheimers Res Ther, 7*(1), 68. doi:10.1186/s13195-015-0152-z

Callahan, B. L., Shammi, P., Taylor, R., Ramakrishnan, N., & Black, S. E. (2021). Longitudinal Cognitive Performance of Older Adults With ADHD Presenting to a Cognitive Neurology Clinic: A Case Series of Change Up to 21 Years. *Front Aging Neurosci, 13*, 726374. doi:10.3389/fnagi.2021.726374

Castora-Binkley, M., Peronto, C. L., Edwards, J. D., & Small, B. J. (2015). A longitudinal analysis of the influence of race on cognitive performance. *J Gerontol B Psychol Sci Soc Sci, 70*(4), 512-518. doi:10.1093/geronb/gbt112

Cheng, Y. W., Chen, T. F., & Chiu, M. J. (2017). From mild cognitive impairment to subjective cognitive decline: conceptual and methodological evolution. *Neuropsychiatr Dis Treat, 13*, 491-498. doi:10.2147/NDT.S123428

Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., . . . Von Zglinicki, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society, 55*(10), 1630-1635. doi:10.1111/j.1532-5415.2007.01379.x

Cooley, S. A., Heaps, J. M., Bolzenius, J. D., Salminen, L. E., Baker, L. M., Scott, S. E., & Paul, R. H. (2015). Longitudinal Change in Performance on the Montreal Cognitive Assessment in Older Adults. *Clin Neuropsychol, 29*(6), 824-835. doi:10.1080/13854046.2015.1087596

Cromer, J. A., Harel, B. T., Yu, K., Valadka, J. S., Brunwin, J. W., Crawford, C. D., . . . Maruff, P. (2015). Comparison of Cognitive Performance on the Cogstate Brief Battery When Taken In-Clinic, In-Group, and Unsupervised. *Clin Neuropsychol, 29*(4), 542-558. doi:10.1080/13854046.2015.1054437

Crossen, J. R., & Wiens, A. N. (1994). Comparison of the Auditory-Verbal Learning Test (AVLT) and California Verbal Learning Test (CVLT) in a sample of normal subjects. *J Clin Exp Neuropsychol, 16*(2), 190-194. doi:10.1080/01688639408402630

Darby, D. G., Brodtmann, A., Pietrzak, R. H., Fredrickson, J., Woodward, M., Villemagne, V. L., . . . Rowe, C. (2011). Episodic memory decline predicts cortical amyloid status in community-dwelling older adults. *J Alzheimers Dis, 27*(3), 627-637. doi:10.3233/JAD-2011-110818

De Meijer, L., Merlo, D., Skibina, O., Grobbee, E. J., Gale, J., Haartsen, J., . . . Van der Walt, A. (2018). Monitoring cognitive change in multiple sclerosis using a computerized cognitive battery. *Mult Scler J Exp Transl Clin, 4*(4), 2055217318815513. doi:10.1177/2055217318815513

De Roeck, E. E., De Deyn, P. P., Dierckx, E., & Engelborghs, S. (2019). Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review. *Alzheimers Res Ther, 11*(1), 21. doi:10.1186/s13195-019-0474-3

Denboer, J. W., Nicholls, C., Corte, C., & Chestnut, K. (2014). National Institutes of Health Toolbox Cognition Battery. *Archives of Clinical Neuropsychology, 29*(7), 692-694. doi:DOI 10.1093/arclin/acu033

Ebenau, J. L., Timmers, T., Wesselman, L. M. P., Verberk, I. M. W., Verfaillie, S. C. J., Slot, R. E. R., . . . van der Flier, W. M. (2020). ATN classification and clinical progression in subjective cognitive decline: The SCIENCe project. *Neurology, 95*(1), e46-e58. doi:10.1212/WNL.0000000000009724

Edmonds, E. C., Delano-Wood, L., Galasko, D. R., Salmon, D. P., Bondi, M. W., & Alzheimer's Disease Neuroimaging, I. (2014). Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. *J Int Neuropsychol Soc, 20*(8), 836-847. doi:10.1017/S135561771400068X

Elman, J. A., Jak, A. J., Panizzon, M. S., Tu, X. M., Chen, T., Reynolds, C. A., . . . Kremen, W. S. (2018). Underdiagnosis of mild cognitive impairment: A consequence of ignoring practice effects. *Alzheimers Dement (Amst), 10*, 372-381. doi:10.1016/j.dadm.2018.04.003

Estevez-Gonzalez, A., Kulisevsky, J., Boltes, A., Otermin, P., & Garcia-Sanchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging. *Int J Geriatr Psychiatry, 18*(11), 1021-1028. doi:10.1002/gps.1010

Ferman, T. J., Smith, G. E., Kantarci, K., Boeve, B. F., Pankratz, V. S., Dickson, D. W., . . . Petersen, R. C. (2013). Nonamnestic mild cognitive impairment progresses to dementia

with Lewy bodies. *Neurology, 81*(23), 2032-2038. doi:10.1212/01.wnl.0000436942.55281.47

Fredrickson, J., Maruff, P., Woodward, M., Moore, L., Fredrickson, A., Sach, J., & Darby, D. (2010). Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies. *Neuroepidemiology, 34*(2), 65-75. doi:10.1159/000264823

Freitas, S., Simoes, M. R., Alves, L., & Santana, I. (2013). Montreal cognitive assessment: validation study for mild cognitive impairment and Alzheimer disease. *Alzheimer Dis Assoc Disord, 27*(1), 37-43. doi:10.1097/WAD.0b013e3182420bfe

Glymour, M. M., Brickman, A. M., Kivimaki, M., Mayeda, E. R., Chene, G., Dufouil, C., & Manly, J. J. (2018). Will biomarker-based diagnosis of Alzheimer's disease maximize scientific progress? Evaluating proposed diagnostic criteria. *Eur J Epidemiol, 33*(7), 607-612. doi:10.1007/s10654-018-0418-4

Goldberg, T. E., Harvey, P. D., Wesnes, K. A., Snyder, P. J., & Schneider, L. S. (2015). Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimers Dement (Amst), 1*(1), 103-111. doi:10.1016/j.dadm.2014.11.003

Grober, E., Hall, C. B., Lipton, R. B., Zonderman, A. B., Resnick, S. M., & Kawas, C. (2008). Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *J Int Neuropsychol Soc, 14*(2), 266-278. doi:10.1017/S1355617708080302

Gross, A. L., Mungas, D. M., Crane, P. K., Gibbons, L. E., MacKay-Brandt, A., Manly, J. J., . . . Jones, R. N. (2015). Effects of education and race on cognitive decline: An integrative

study of generalizability versus study-specific results. *Psychol Aging, 30*(4), 863-880. doi:10.1037/pag0000032

Haaksma, M. L., Vilela, L. R., Marengoni, A., Calderon-Larranaga, A., Leoutsakos, J. S., Olde Rikkert, M. G. M., & Melis, R. J. F. (2017). Comorbidity and progression of late onset Alzheimer's disease: A systematic review. *PLoS One, 12*(5), e0177044. doi:10.1371/journal.pone.0177044

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med, 4*(2), 627-635. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24009950

Hammers, D., Spurgeon, E., Ryan, K., Persad, C., Barbas, N., Heidebrink, J., . . . Giordani, B. (2012). Validity of a brief computerized cognitive screening test in dementia. *J Geriatr Psychiatry Neurol, 25*(2), 89-99. doi:10.1177/0891988712447894

Hassenstab, J., Nicosia, J., LaRose, M., Aschenbrenner, A. J., Gordon, B. A., Benzinger, T. L. S., . . . Morris, J. C. (2021). Is comprehensiveness critical? Comparing short and long format cognitive assessments in preclinical Alzheimer disease. *Alzheimers Res Ther, 13*(1), 153. doi:10.1186/s13195-021-00894-5

Hessen, E., Eckerstrom, M., Nordlund, A., Selseth Almdahl, I., Stalhammar, J., Bjerke, M., . . . Wallin, A. (2017). Subjective Cognitive Impairment Is a Predominantly Benign Condition in Memory Clinic Patients Followed for 6 Years: The Gothenburg-Oslo MCI Study. *Dement Geriatr Cogn Dis Extra, 7*(1), 1-14. doi:10.1159/000454676

Hobson, J. (2015). The Montreal Cognitive Assessment (MoCA). *Occup Med (Lond), 65*(9), 764-765. doi:10.1093/occmed/kqv078

Huebner, A. (2010). An Overview of Recent Developments in Cognitive Diagnostic Computer Adaptive Assessments. *Practical Assessment, Research, and Evaluation, 15*. doi:10.7275/7fdd-6897

Ihara, R., Iwata, A., Suzuki, K., Ikeuchi, T., Kuwano, R., Iwatsubo, T., & Japanese Alzheimer's Disease Neuroimaging, I. (2018). Clinical and cognitive characteristics of preclinical Alzheimer's disease in the Japanese Alzheimer's Disease Neuroimaging Initiative cohort. *Alzheimers Dement (N Y), 4*, 645-651. doi:10.1016/j.trci.2018.10.004

Iverson, G. L., Brooks, B. L., Ashton, V. L., Johnson, L. G., & Gualtieri, C. T. (2009). Does familiarity with computers affect computerized neuropsychological test performance? *J Clin Exp Neuropsychol, 31*(5), 594-604. doi:10.1080/13803390802372125

Iverson, G. L., & Karr, J. E. (2021). Improving the Methodology for Identifying Mild Cognitive Impairment in Intellectually High-Functioning Adults Using the NIH Toolbox Cognition Battery. *Front Psychol, 12*, 724888. doi:10.3389/fpsyg.2021.724888

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., . . . Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 535-562. doi:10.1016/j.jalz.2018.02.018

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. *Am J Geriatr Psychiatry, 17*(5), 368-375. doi:10.1097/JGP.0b013e31819431d5

Jonaitis, E. M., Koscik, R. L., Clark, L. R., Ma, Y., Betthauser, T. J., Berman, S. E., . . . Johnson, S. C. (2019). Measuring longitudinal cognition: Individual tests versus composites. *Alzheimers Dement (Amst), 11*, 74-84. doi:10.1016/j.dadm.2018.11.006

Jongsiriyanyong, S., & Limpawattana, P. (2018). Mild Cognitive Impairment in Clinical Practice: A Review Article. *Am J Alzheimers Dis Other Demen, 33*(8), 500-507. doi:10.1177/1533317518791401

Jungwirth, S., Zehetmayer, S., Hinterberger, M., Tragl, K. H., & Fischer, P. (2012). The validity of amnestic MCI and non-amnestic MCI at age 75 in the prediction of Alzheimer's dementia and vascular dementia. *Int Psychogeriatr, 24*(6), 959-966. doi:10.1017/S1041610211002870

Jutten, R. J., Grandoit, E., Foldi, N. S., Sikkes, S. A. M., Jones, R. N., Choi, S. E., . . . Rabin, L. A. (2020). Lower practice effects as a marker of cognitive performance and dementia risk: A literature review. *Alzheimers Dement (Amst), 12*(1), e12055. doi:10.1002/dad2.12055

Kasper, S., Bancher, C., Eckert, A., Forstl, H., Frolich, L., Hort, J., . . . Palomo, M. S. M. (2020). Management of mild cognitive impairment (MCI): The need for national and international guidelines. *World J Biol Psychiatry, 21*(8), 579-594. doi:10.1080/15622975.2019.1696473

Kirova, A. M., Bays, R. B., & Lagalwar, S. (2015). Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *Biomed Res Int, 2015*, 748212. doi:10.1155/2015/748212

Koo, B. M., & Vizer, L. M. (2019). Mobile Technology for Cognitive Assessment of Older Adults: A Scoping Review. *Innov Aging, 3*(1), igy038. doi:10.1093/geroni/igy038

Koski, L. (2013). Validity and applications of the Montreal cognitive assessment for the assessment of vascular cognitive impairment. *Cerebrovasc Dis, 36*(1), 6-18. doi:10.1159/000352051

Landau, S. M., Harvey, D., Madison, C. M., Reiman, E. M., Foster, N. L., Aisen, P. S., . . . Alzheimer's Disease Neuroimaging, I. (2010). Comparing predictors of conversion and

decline in mild cognitive impairment. *Neurology, 75*(3), 230-238. doi:10.1212/WNL.0b013e3181e8e8b8

Langa, K. M., & Levine, D. A. (2014). The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA, 312*(23), 2551-2561. doi:10.1001/jama.2014.13806

Lanza, C. E., Sejunaite, K., Steindel, C., Scholz, I., & Riepe, M. W. (2020). On the conundrum of cognitive impairment due to depressive disorder in older patients. *PLoS One, 15*(4), e0231111. doi:10.1371/journal.pone.0231111

Lapshin, H., O'Connor, P., Lanctot, K. L., & Feinstein, A. (2012). Computerized cognitive testing for patients with multiple sclerosis. *Mult Scler Relat Disord, 1*(4), 196-201. doi:10.1016/j.msard.2012.05.001

Lee, Y. C., Lin, Y. T., & Chiu, E. C. (2021). A comparison of test-retest reliability of four cognitive screening tools in people with dementia. *Disabil Rehabil*, 1-6. doi:10.1080/09638288.2021.1891466

Lezak, M. D., & Lezak, M. D. (2004). *Neuropsychological assessment* (4th ed.). Oxford ; New York: Oxford University Press.

Lim, Y. Y., Ellis, K. A., Harrington, K., Ames, D., Martins, R. N., Masters, C. L., . . . The Aibl Research, G. (2012). Use of the CogState Brief Battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. *J Clin Exp Neuropsychol, 34*(4), 345-358. doi:10.1080/13803395.2011.643227

Lim, Y. Y., Ellis, K. A., Harrington, K., Pietrzak, R. H., Gale, J., Ames, D., . . . Maruff, P. (2013). Cognitive decline in adults with amnestic mild cognitive impairment and high amyloid-

beta: prodromal Alzheimer's disease? *J Alzheimers Dis, 33*(4), 1167-1176. doi:10.3233/JAD-121771

Lim, Y. Y., Jaeger, J., Harrington, K., Ashwood, T., Ellis, K. A., Stoffler, A., . . . Maruff, P. (2013). Three-month stability of the CogState brief battery in healthy older adults, mild cognitive impairment, and Alzheimer's disease: results from the Australian Imaging, Biomarkers, and Lifestyle-rate of change substudy (AIBL-ROCS). *Arch Clin Neuropsychol, 28*(4), 320-330. doi:10.1093/arclin/act021

Lim, Y. Y., Maruff, P., Pietrzak, R. H., Ellis, K. A., Darby, D., Ames, D., . . . Group, A. R. (2014). Abeta and cognitive change: examining the preclinical and prodromal stages of Alzheimer's disease. *Alzheimers Dement, 10*(6), 743-751 e741. doi:10.1016/j.jalz.2013.11.005

Lim, Y. Y., Pietrzak, R. H., Bourgeat, P., Ames, D., Ellis, K. A., Rembach, A., . . . Maruff, P. (2015). Relationships between performance on the Cogstate Brief Battery, neurodegeneration, and Abeta accumulation in cognitively normal older adults and adults with MCI. *Arch Clin Neuropsychol, 30*(1), 49-58. doi:10.1093/arclin/acu068

Lim, Y. Y., Villemagne, V. L., Laws, S. M., Pietrzak, R. H., Ames, D., Fowler, C., . . . Group, A. R. (2016). Performance on the Cogstate Brief Battery Is Related to Amyloid Levels and Hippocampal Volume in Very Mild Dementia. *J Mol Neurosci, 60*(3), 362-370. doi:10.1007/s12031-016-0822-8

Lissek, V., & Suchan, B. (2021). Preventing dementia? Interventional approaches in mild cognitive impairment. *Neurosci Biobehav Rev, 122*, 143-164. doi:10.1016/j.neubiorev.2020.12.022

Liu, C. C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol, 9*(2), 106-118. doi:10.1038/nrneurol.2012.263

Machulda, M. M., Hagen, C. E., Wiste, H. J., Mielke, M. M., Knopman, D. S., Roberts, R. O., . . . Petersen, R. C. (2017). [Formula: see text]Practice effects and longitudinal cognitive change in clinically normal older adults differ by Alzheimer imaging biomarker status. *Clin Neuropsychol, 31*(1), 99-117. doi:10.1080/13854046.2016.1241303

Machulda, M. M., Pankratz, V. S., Christianson, T. J., Ivnik, R. J., Mielke, M. M., Roberts, R. O., . . . Petersen, R. C. (2013). Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. *Clin Neuropsychol, 27*(8), 1247-1264. doi:10.1080/13854046.2013.836567

Mackin, R. S., Insel, P. S., Truran, D., Finley, S., Flenniken, D., Nosheny, R., . . . Weiner, M. W. (2018). Unsupervised online neuropsychological test performance for individuals with mild cognitive impairment and dementia: Results from the Brain Health Registry. *Alzheimers Dement (Amst), 10*, 573-582. doi:10.1016/j.dadm.2018.05.005

Maruff, P., Lim, Y. Y., Darby, D., Ellis, K. A., Pietrzak, R. H., Snyder, P. J., . . . Group, A. R. (2013). Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild cognitive impairment and Alzheimer's disease. *BMC Psychol, 1*(1), 30. doi:10.1186/2050-7283-1-30

Maruff, P., Thomas, E., Cysique, L., Brew, B., Collie, A., Snyder, P., & Pietrzak, R. H. (2009). Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Arch Clin Neuropsychol, 24*(2), 165-178. doi:10.1093/arclin/acp010

McGrattan, A. M., Pakpahan, E., Siervo, M., Mohan, D., Reidpath, D. D., Prina, M., . . . De, P. E. C. t. (2022). Risk of conversion from mild cognitive impairment to dementia in low- and middle-income countries: A systematic review and meta-analysis. *Alzheimers Dement (N Y), 8*(1), e12267. doi:10.1002/trc2.12267

Miao, R., Chen, H. Y., Gill, S., Naude, J., Smith, E. E., Ismail, Z., & Alzheimer's Disease Neuroimaging, I. (2022). Plasma beta-Amyloid in Mild Behavioural Impairment - Neuropsychiatric Symptoms on the Alzheimer's Continuum. *J Geriatr Psychiatry Neurol, 35*(3), 434-441. doi:10.1177/08919887211016068

Michaud, T. L., Su, D., Siahpush, M., & Murman, D. L. (2017). The Risk of Incident Mild Cognitive Impairment and Progression to Dementia Considering Mild Cognitive Impairment Subtypes. *Dement Geriatr Cogn Dis Extra, 7*(1), 15-29. doi:10.1159/000452486

Mickes, L., Wixted, J. T., Fennema-Notestine, C., Galasko, D., Bondi, M. W., Thal, L. J., & Salmon, D. P. (2007). Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology, 21*(6), 696-705. doi:10.1037/0894-4105.21.6.696

Mielke, M. M., Machulda, M. M., Hagen, C. E., Edwards, K. K., Roberts, R. O., Pankratz, V. S., . . . Petersen, R. C. (2015). Performance of the CogState computerized battery in the Mayo Clinic Study on Aging. *Alzheimers Dement, 11*(11), 1367-1376. doi:10.1016/j.jalz.2015.01.008

Mielke, M. M., Weigand, S. D., Wiste, H. J., Vemuri, P., Machulda, M. M., Knopman, D. S., . . . Petersen, R. C. (2014). Independent comparison of CogState computerized testing and a

standard cognitive battery with neuroimaging. *Alzheimers Dement, 10*(6), 779-789. doi:10.1016/j.jalz.2014.09.001

Milani, S. A., Marsiske, M., Cottler, L. B., Chen, X., & Striley, C. W. (2018). Optimal cutoffs for the Montreal Cognitive Assessment vary by race and ethnicity. *Alzheimers Dement (Amst), 10*, 773-781. doi:10.1016/j.dadm.2018.09.003

Minnesota 2022 Update Conference Planning Commission. (2022). Minnesota 2022 Update Conference: Objectives. Retrieved from https://minnesotaconference.org/conference-objectives/

Mistridis, P., Krumm, S., Monsch, A. U., Berres, M., & Taylor, K. I. (2015). The 12 Years Preceding Mild Cognitive Impairment Due to Alzheimer's Disease: The Temporal Emergence of Cognitive Decline. *J Alzheimers Dis, 48*(4), 1095-1107. doi:10.3233/JAD-150137

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand, 119*(4), 252-265. doi:10.1111/j.1600-0447.2008.01326.x

Molano, J., Boeve, B., Ferman, T., Smith, G., Parisi, J., Dickson, D., . . . Petersen, R. (2010). Mild cognitive impairment associated with limbic and neocortical Lewy body disease: a clinicopathological study. *Brain, 133*(Pt 2), 540-556. doi:10.1093/brain/awp280

Morimoto, S. S., & Alexopoulos, G. S. (2013). Cognitive deficits in geriatric depression: clinical correlates and implications for current and future treatment. *Psychiatr Clin North Am, 36*(4), 517-531. doi:10.1016/j.psc.2013.08.002

Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., . . . Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimers Dement, 13*(4), 468-492. doi:10.1016/j.jalz.2016.06.2365

Mukku, S. S. R., Dahale, A. B., Muniswamy, N. R., Muliyala, K. P., Sivakumar, P. T., & Varghese, M. (2021). Geriatric Depression and Cognitive Impairment-An Update. *Indian J Psychol Med, 43*(4), 286-293. doi:10.1177/0253717620981556

Murante, T., & Cohen, C. I. (2017). Cognitive Functioning in Older Adults With Schizophrenia. *Focus (Am Psychiatr Publ), 15*(1), 26-34. doi:10.1176/appi.focus.20160032

Murphy, M. P., & LeVine, H., 3rd. (2010). Alzheimer's disease and the amyloid-beta peptide. *J Alzheimers Dis, 19*(1), 311-323. doi:10.3233/JAD-2010-1221

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc, 53*(4), 695-699. doi:10.1111/j.1532-5415.2005.53221.x

Nation, D. A., Ho, J. K., Dutt, S., Han, S. D., Lai, M. H. C., & Alzheimer's Disease Neuroimaging, I. (2019). Neuropsychological Decline Improves Prediction of Dementia Beyond Alzheimer's Disease Biomarker and Mild Cognitive Impairment Diagnoses. *J Alzheimers Dis, 69*(4), 1171-1182. doi:10.3233/JAD-180525

O'Connell, M. E., Vellani, S., Robertson, S., O'Rourke, H. M., & McGilton, K. S. (2021). Going From Zero to 100 in Remote Dementia Research: A Practical Guide. *J Med Internet Res, 23*(1), e24098. doi:10.2196/24098

Ohman, F., Hassenstab, J., Berron, D., Scholl, M., & Papp, K. V. (2021). Current advances in digital cognitive assessment for preclinical Alzheimer's disease. *Alzheimers Dement (Amst), 13*(1), e12217. doi:10.1002/dad2.12217

Parent, C., Rousseau, L. S., Predovan, D., Duchesne, S., & Hudon, C. (2023). Longitudinal association between ss-amyloid accumulation and cognitive decline in cognitively healthy older adults: A systematic review. *Aging Brain, 3*, 100074. doi:10.1016/j.nbas.2023.100074

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *J Intern Med, 256*(3), 183-194. doi:10.1111/j.1365-2796.2004.01388.x

Petersen, R. C. (2011). Clinical practice. Mild cognitive impairment. *N Engl J Med, 364*(23), 2227-2234. doi:10.1056/NEJMcp0910237

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: a concept in evolution. *J Intern Med, 275*(3), 214-228. doi:10.1111/joim.12190

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., . . . 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. doi:10.1212/WNL.0000000000004826

Pudumjee, S. B., Lundt, E. S., Albertson, S. M., Machulda, M. M., Kremers, W. K., Jack, C. R., . . . Stricker, N. H. (2021). A Comparison of Cross-Sectional and Longitudinal Methods of Defining Objective Subtle Cognitive Decline in Preclinical Alzheimer's Disease Based on

Cogstate One Card Learning Accuracy Performance. *J Alzheimers Dis, 83*(2), 861-877. doi:10.3233/JAD-210251

Qin, H. Y., Zhao, X. D., Zhu, B. G., & Hu, C. P. (2020). Demographic Factors and Cognitive Function Assessments Associated with Mild Cognitive Impairment Progression for the Elderly. *Biomed Res Int, 2020*, 3054373. doi:10.1155/2020/3054373

Racine, A. M., Clark, L. R., Berman, S. E., Koscik, R. L., Mueller, K. D., Norton, D., . . . Johnson, S. C. (2016). 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, 54*(4), 1395-1408. doi:10.3233/JAD-160528

Rajji, T. K., & Mulsant, B. H. (2008). Nature and course of cognitive function in late-life schizophrenia: a systematic review. *Schizophr Res, 102*(1-3), 122-140. doi:10.1016/j.schres.2008.03.015

Reitmanova, S., & Gustafson, D. L. (2009). Primary mental health care information and services for St. John's visible minority immigrants: gaps and opportunities. *Issues Ment Health Nurs, 30*(10), 615-623. doi:10.1080/01612840903033733

Rowe, C. C., Ng, S., Ackermann, U., Gong, S. J., Pike, K., Savage, G., . . . Villemagne, V. L. (2007). Imaging beta-amyloid burden in aging and dementia. *Neurology, 68*(20), 1718-1725. doi:10.1212/01.wnl.0000261919.22630.ea

Royse, S. K., Minhas, D. S., Lopresti, B. J., Murphy, A., Ward, T., Koeppe, R. A., . . . Alzheimer's Disease Neuroimaging, I. (2021). Validation of amyloid PET positivity thresholds in centiloids: a multisite PET study approach. *Alzheimers Res Ther, 13*(1), 99. doi:10.1186/s13195-021-00836-1

Sanderson-Cimino, M., Chen, R., Tu, X. M., Elman, J. A., Jak, A. J., & Kremen, W. S. (2023). Misinterpreting cognitive change over multiple timepoints: When practice effects meet age-related decline. *Neuropsychology*. doi:10.1037/neu0000903

Sanderson-Cimino, M., Elman, J. A., Tu, X. M., Gross, A. L., Panizzon, M. S., Gustavson, D. E., . . . Kremen, W. S. (2020). Cognitive Practice Effects Delay Diagnosis; Implications for Clinical Trials. *medRxiv*. doi:10.1101/2020.11.03.20224808

Schouws, S. N., Comijs, H. C., Dols, A., Beekman, A. T., & Stek, M. L. (2016). Five-year follow-up of cognitive impairment in older adults with bipolar disorder. *Bipolar Disord, 18*(2), 148-154. doi:10.1111/bdi.12374

Schraen-Maschke, S., Sergeant, N., Dhaenens, C. M., Bombois, S., Deramecourt, V., Caillet-Boudin, M. L., . . . Buee, L. (2008). Tau as a biomarker of neurodegenerative diseases. *Biomark Med, 2*(4), 363-384. doi:10.2217/17520363.2.4.363

Silverberg, N., Elliott, C., Ryan, L., Masliah, E., & Hodes, R. (2018). NIA commentary on the NIA-AA Research Framework: Towards a biological definition of Alzheimer's disease. *Alzheimers Dement, 14*(4), 576-578. doi:10.1016/j.jalz.2018.03.004

Staffaroni, A. M., Tsoy, E., Taylor, J., Boxer, A. L., & Possin, K. L. (2020). Digital Cognitive Assessments for Dementia: Digital assessments may enhance the efficiency of evaluations in neurology and other clinics. *Pract Neurol (Fort Wash Pa), 2020*, 24-45. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/33927583

Stallings, G., Boake, C., & Sherer, M. (1995). Comparison of the California Verbal Learning Test and the Rey Auditory Verbal Learning Test in head-injured patients. *J Clin Exp Neuropsychol, 17*(5), 706-712. doi:10.1080/01688639508405160

Stephan, B. C., Brayne, C., Savva, G. M., Matthews, F. E., Medical Research Council Cognitive, F., & Ageing, S. (2011). Occurrence of medical co-morbidity in mild cognitive impairment: implications for generalisation of MCI research. *Age Ageing, 40*(4), 501-507. doi:10.1093/ageing/afr057

Sternin, A., Burns, A., & Owen, A. M. (2019). Thirty-Five Years of Computerized Cognitive Assessment of Aging-Where Are We Now? *Diagnostics (Basel), 9*(3). doi:10.3390/diagnostics9030114

Strauss, E., Sherman, E. M. S., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests : administration, norms, and commentary* (3rd ed.). Oxford ; New York: Oxford University Press.

Stricker, N. H., Lundt, E. S., Albertson, S. M., Machulda, M. M., Pudumjee, S. B., Kremers, W. K., . . . Mielke, M. M. (2020). Diagnostic and Prognostic Accuracy of the Cogstate Brief Battery and Auditory Verbal Learning Test in Preclinical Alzheimer's Disease and Incident Mild Cognitive Impairment: Implications for Defining Subtle Objective Cognitive Impairment. *J Alzheimers Dis, 76*(1), 261-274. doi:10.3233/JAD-200087

Stricker, N. H., Lundt, E. S., Alden, E. C., Albertson, S. M., Machulda, M. M., Kremers, W. K., . . . Mielke, M. M. (2020). Longitudinal Comparison of in Clinic and at Home Administration of the Cogstate Brief Battery and Demonstrated Practice Effects in the Mayo Clinic Study of Aging. *J Prev Alzheimers Dis, 7*(1), 21-28. doi:10.14283/jpad.2019.35

Stricker, N. H., Lundt, E. S., Edwards, K. K., Machulda, M. M., Kremers, W. K., Roberts, R. O., . . . Mielke, M. M. (2019). Comparison of PC and iPad administrations of the Cogstate Brief Battery in the Mayo Clinic Study of Aging: Assessing cross-modality equivalence of

computerized neuropsychological tests. *Clin Neuropsychol, 33*(6), 1102-1126. doi:10.1080/13854046.2018.1519085

Suppiah, S., Didier, M. A., & Vinjamuri, S. (2019). The Who, When, Why, and How of PET Amyloid Imaging in Management of Alzheimer's Disease-Review of Literature and Interesting Images. *Diagnostics (Basel), 9*(2). doi:10.3390/diagnostics9020065

Suzuki, T., Shimada, H., Makizako, H., Doi, T., Yoshida, D., Ito, K., . . . Kato, T. (2013). A randomized controlled trial of multicomponent exercise in older adults with mild cognitive impairment. *PLoS One, 8*(4), e61483. doi:10.1371/journal.pone.0061483

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics : collected papers*. Mahwah, N.J.: L. Erlbaum Associates.

Thomson, M. S., Chaze, F., George, U., & Guruge, S. (2015). Improving Immigrant Populations' Access to Mental Health Services in Canada: A Review of Barriers and Recommendations. *J Immigr Minor Health, 17*(6), 1895-1905. doi:10.1007/s10903-015-0175-3

Tierney, M. C., Naglie, G., Upshur, R., Moineddin, R., Charles, J., & Jaakkimainen, R. L. (2014). Feasibility and validity of the self-administered computerized assessment of mild cognitive impairment with older primary care patients. *Alzheimer Dis Assoc Disord, 28*(4), 311-319. doi:10.1097/wad.0000000000000036

Tsoy, E., Erlhoff, S. J., Goode, C. A., Dorsman, K. A., Kanjanapong, S., Lindbergh, C. A., . . . Possin, K. L. (2020). BHA-CS: A novel cognitive composite for Alzheimer's disease and related disorders. *Alzheimers Dement (Amst), 12*(1), e12042. doi:10.1002/dad2.12042

Tsoy, E., Zygouris, S., & Possin, K. L. (2021). Current State of Self-Administered Brief Computerized Cognitive Assessments for Detection of Cognitive Disorders in Older

Adults: A Systematic Review. *J Prev Alzheimers Dis, 8*(3), 267-276. doi:10.14283/jpad.2021.11

Unal, I. (2017). Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. *Comput Math Methods Med, 2017*, 3762651. doi:10.1155/2017/3762651

Valdes, E. G., Sadeq, N. A., Harrison Bush, A. L., Morgan, D., & Andel, R. (2016). Regular cognitive self-monitoring in community-dwelling older adults using an internet-based tool. *J Clin Exp Neuropsychol, 38*(9), 1026-1037. doi:10.1080/13803395.2016.1186155

Vassilaki, M., Cha, R. H., Aakre, J. A., Therneau, T. M., Geda, Y. E., Mielke, M. M., . . . Roberts, R. O. (2015). Mortality in mild cognitive impairment varies by subtype, sex, and lifestyle factors: the Mayo Clinic Study of Aging. *J Alzheimers Dis, 45*(4), 1237-1245. doi:10.3233/JAD-143078

Vickers, A. J., Cronin, A. M., & Begg, C. B. (2011). One statistical test is sufficient for assessing new predictive markers. *BMC Med Res Methodol, 11*, 13. doi:10.1186/1471-2288-11-13

Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., . . . Alzheimer's Disease Neuroimaging, I. (2017). The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. *Alzheimers Dement, 13*(5), 561-571. doi:10.1016/j.jalz.2016.10.006

White, J. P., Schembri, A., Edgar, C. J., Lim, Y. Y., Masters, C. L., & Maruff, P. (2021). A Paradox in Digital Memory Assessment: Increased Sensitivity With Reduced Difficulty. *Front Digit Health, 3*, 780303. doi:10.3389/fdgth.2021.780303

Williams, C. C. (2001). Increasing access and building equity into mental health services: an examination of the potential for change. *Can J Commun Ment Health, 20*(1), 37-51. doi:10.7870/cjcmh-2001-0003

Wojcik, C. M., Beier, M., Costello, K., DeLuca, J., Feinstein, A., Goverover, Y., . . . National, M. S. S. C. W. T. (2019). Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review. *Mult Scler, 25*(14), 1848-1869. doi:10.1177/1352458519879094

Woodhouse, J., Heyanka, D. J., Scott, J., Vincent, A., Roebuck-Spencer, T., Domboski-Davidson, K., . . . Adams, R. (2013). Efficacy of the ANAM General Neuropsychological Screening Battery (ANAM GNS) for detecting neurocognitive impairment in a mixed clinical sample. *Clin Neuropsychol, 27*(3), 376-385. doi:10.1080/13854046.2012.762427

Young, C. B., Mormino, E. C., Poston, K. L., Johnson, K. A., Rentz, D. M., Sperling, R. A., & Papp, K. V. (2023). Computerized cognitive practice effects in relation to amyloid and tau in preclinical Alzheimer's disease: Results from a multi-site cohort. *Alzheimers Dement (Amst), 15*(1), e12414. doi:10.1002/dad2.12414

Yue, L., Hu, D., Zhang, H., Wen, J., Wu, Y., Li, W., . . . Xiao, S. (2021). Prediction of 7-year's conversion from subjective cognitive decline to mild cognitive impairment. *Hum Brain Mapp, 42*(1), 192-203. doi:10.1002/hbm.25216

Zarshenas, S., & Cullen, N. (2018). Association between cognitive ability and functional outcomes following traumatic brain injury- using a computer-based neurocognitive battery. *Brain Inj, 32*(13-14), 1678-1683. doi:10.1080/02699052.2018.1517223

Appendix A

Sample Characteristics for Group Comparisons

Table A1

Group Differences for the Primary Analyses: Baseline Assessments Predicting Conversion to aMCI or AD

	Cognitively Normal n= 309	*Incident aMCI/AD n= 30*	*p*
Age (years)	73.2 (7.2)	75.5 (6.8)	.093
Sex (female)	180 (58.3)	17 (56.7)	.867
Education (years)	16.9 (2.3)	16.6 (2.2)	.494
Race/ethnicity (White)	26 (8.4)	4 (13.3)	.322
ApoE ε4+	100 (32.4)	10 (33.3)	.914

Table A2

Group Differences for the Primary Analyses: Repeated Assessments Predicting Conversion to aMCI or AD

	Cognitively Normal n= 276	*Incident aMCI/AD n= 25*	*p*
Age (years)	72.9 (7.0)	74.6 (6.5)	.261
Sex (female)	162 (58.7)	13 (52.0)	.516
Education (years)	16.9 (2.3)	16.5 (2.2)	.388
Race/ethnicity (White)	24 (8.7)	2 (7.7)	1.000
ApoE ε4+	90 (29.9)	9 (36.0)	.730

Table A3

Group Differences for the Secondary Analyses: Baseline Assessments Predicting Aβ Status

	Aβ- at Baseline *n= 364*	*Aβ+ at Baseline* *n= 234*	*p*
Age (years)	72.7 (7.5)	75.1 (7.3)	<.001
Sex (female)	191 (52.5)	123 (52.6)	.983
Education (years)	16.7 (2.5)	16.5 (2.4)	.247
Race/ethnicity (White)	33 (9.1)	15 (6.4)	.243
ApoE ε4+	73 (20.1)	137 (58.5)	<.001

Table A4

Group Differences for the Secondary Analyses: Repeated Assessments Predicting Aβ Accumulation

	Aβ- Accumulation n= 123	*Aβ+ Accumulation n= 208*	*p*
Age (years)	72.6 (7.3)	74.0 (7.1)	.099
Sex (female)	110 (52.9)	70 (56.9)	.477
Education (years)	16.7 (2.6)	16.6 (2.4)	.811
Race/ethnicity (White)	191 (91.8)	119 (96.7)	.076
ApoE ε4+	54 (26.0)	76 (61.8)	<.001

Appendix B

Logistic Regressions for Baseline Adjusted Area Under the Receiver Operator Curves

Table B1

Logistic Regression for Baseline OCL Performance Predicting Conversion To aMCI or AD

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.036	.027	1.749	1	.186	1.037	.983	1.094
Sex	.005	.400	.000	1	.990	1.005	.459	2.204
Education	-.053	.081	.433	1	.511	.948	.808	1.112
Race/ethnicity	.395	.592	.444	1	.505	1.484	.465	4.739
ApoE ε4+	.154	.417	.137	1	.711	1.167	.515	2.645
OCL accuracy	-2.614	1.925	1.843	1	.175	.073	.002	3.188
Constant	-1.699	3.198	.282	1	.595	.183		

Table B2

Logistic Regression for Baseline MoCA Performance Predicting Conversion to aMCI or AD

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.023	.028	.659	1	.417	1.023	.969	1.080
Sex	.238	.416	.327	1	.567	1.268	.562	2.865
Education	.028	.087	.103	1	.748	1.028	.867	1.220
Race/ethnicity	.165	.622	.071	1	.790	1.180	.349	3.991
ApoE ε4+	.124	.422	.087	1	.768	1.132	.495	2.592
MoCA total score	-.258	.076	11.581	1	<.001	.772	.665	.896
Constant	2.204	3.030	.529	1	.467	9.062		

Table B3

Logistic Regression for Baseline RAVLT Performance Predicting Conversion to aMCI or AD

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.037	.027	1.931	1	.165	1.038	.985	1.094
Sex	.202	.410	.244	1	.621	1.224	.549	2.732
Education	-.036	.081	.195	1	.659	.965	.823	1.131
Race/ethnicity	.383	.592	.418	1	.518	1.466	.459	4.680
ApoE ε4+	.126	.418	.090	1	.764	1.134	.500	2.573
RAVLT 30-minute delay	-.116	.050	5.457	1	.019	.890	.807	.981
Constant	-3.632	2.385	2.319	1	.128	.026		

Table B4

Logistic Regression for Baseline OCL Performance Predicting Aβ Status

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.074	.014	28.757	1	<.001	1.077	1.048	1.107
Sex	-.278	.197	1.981	1	.159	.757	.514	1.115
Education	-.034	.039	.760	1	.383	.967	.896	1.043
Race/ethnicity	-.308	.362	.723	1	.395	.735	.362	1.494
ApoE ε4+	2.003	.205	95.324	1	<.001	7.408	4.956	11.074
OCL accuracy	-.860	.859	1.000	1	.317	.423	.079	2.282
Constant	-5.154	1.502	11.772	1	<.001	.006		

Table B5

Logistic Regression for Baseline MoCA Performance Predicting Aβ Status

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.070	.014	25.866	1	<.001	1.073	1.044	1.102
Sex	-.343	.201	2.915	1	.088	.709	.478	1.052
Education	-.010	.041	.060	1	.807	.990	.914	1.072
Race/ethnicity	-.377	.356	1.068	1	.301	.686	.336	1.402
ApoE ε4+	1.989	.206	93.299	1	<.001	7.308	4.881	10.942
MoCA total score	-.070	.031	4.943	1	.026	.933	.877	.992
Constant	-4.290	1.416	9.177	1	.002	.014		

Table B6

Logistic Regression for Baseline RAVLT Performance Predicting Aβ Status

	B	*SE*	Wald	*df*	*p*	OR	OR 95% CI	
							Lower	Upper
Age	.071	.014	26.507	1	<.001	1.073	1.045	1.103
Sex	.423	.208	4.142	1	.042	1.527	1.016	2.296
Education	-.018	.039	.198	1	.657	.983	.910	1.062
Race/ethnicity	-.281	.363	.600	1	.438	.755	.371	1.537
ApoE ε4+	1.979	.206	92.318	1	<.001	7.239	4.834	10.840
RAVLT 30-minute delay	-.060	.024	6.454	1	.011	.942	.899	.986
Constant	-5.925	1.240	22.816	1	<.001	.003		

Appendix C

Area Under the Receiver Operator Curves for Baseline Analyses

Figure C1

Unadjusted Area Under the Receiver Operator Curves for Baseline Cognitive Scores Predicting Conversion to aMCI or AD

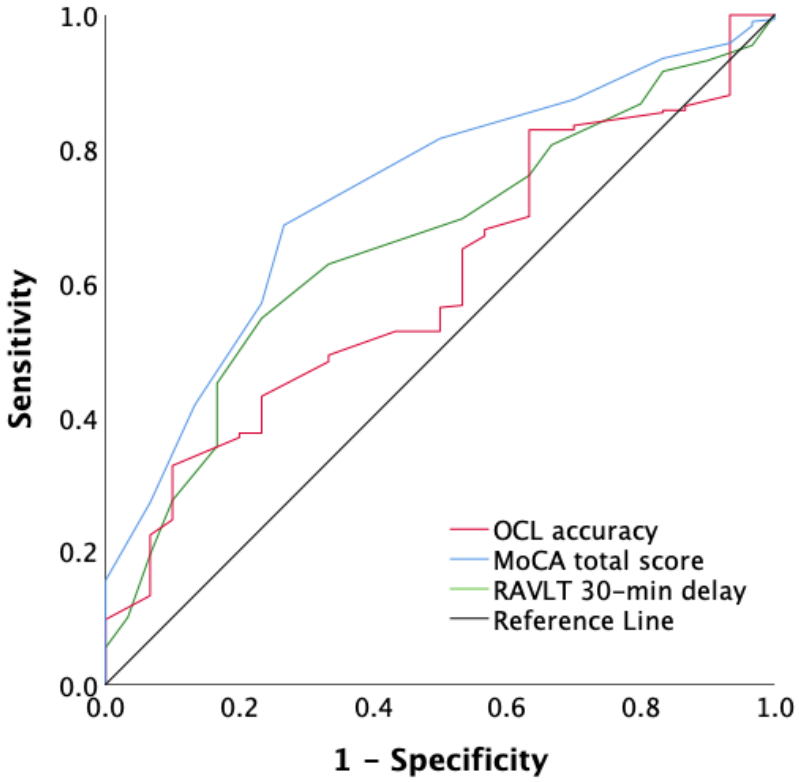


Figure C2

Adjusted Area Under the Receiver Operator Curves for Baseline Cognitive Scores Predicting Conversion to aMCI or AD

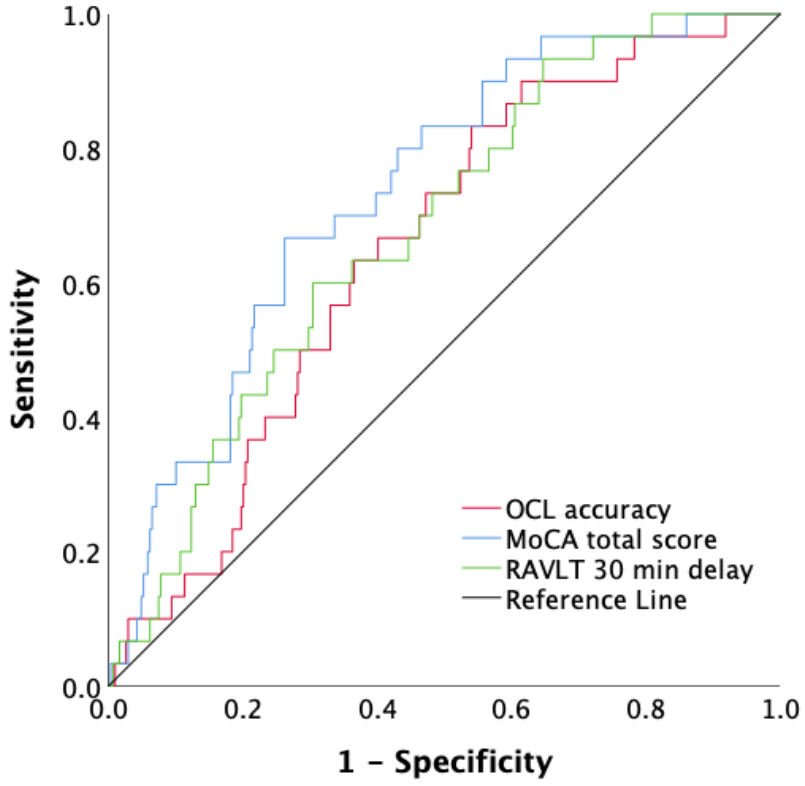


Figure C3

Unadjusted Area Under the Receiver Operator Curves for Baseline Cognitive Scores Predicting Aβ Status

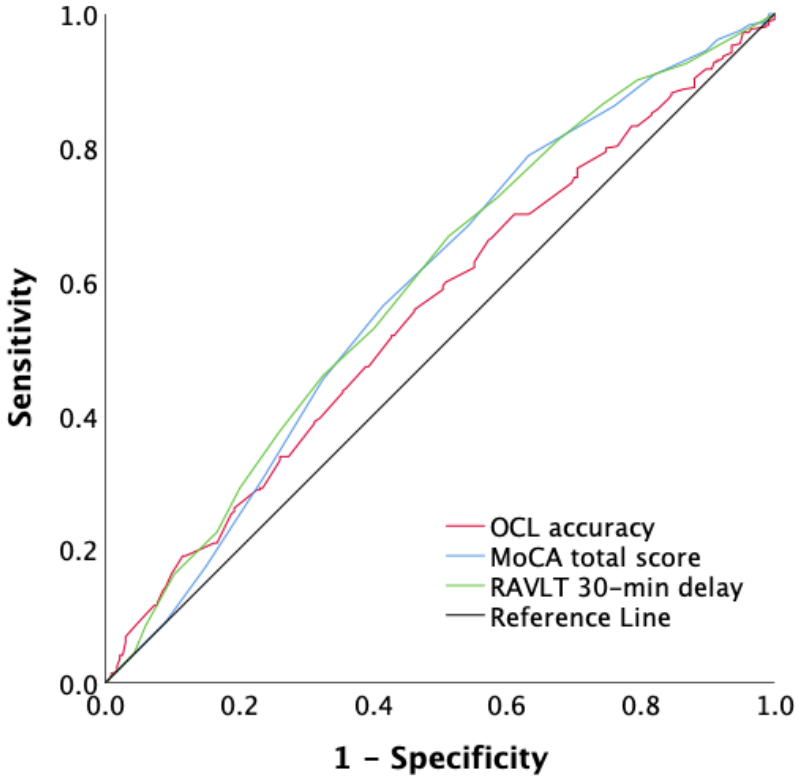
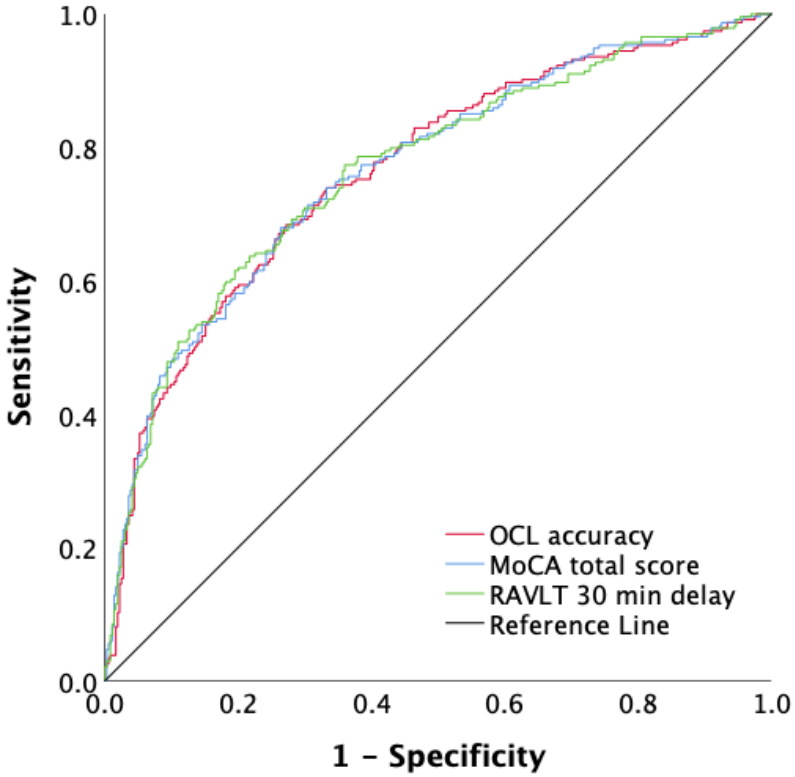


Figure C4

Adjusted Area Under the Receiver Operator Curves for Baseline Cognitive Scores Predicting Aβ Status



Appendix D

Generalized Linear Mixed Models for Repeated Measures Analyses

Table D1

Generalized Linear Mixed Model for Repeated OCL Performance Predicting Conversion to aMCI or AD

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	2.025	1.320	1.534	.125			
Age	.005	.010	.485	.628	1.005	.984	1.027
Sex	-.419	.141	-2.967	.003	.657	.498	.867
Education	.025	.030	.840	.401	1.026	.967	1.088
Race/ethnicity	-1.158	.460	-2.517	.012	.314	.128	.774
ApoE ε4+	-.102	.149	-.682	.495	.903	.675	1.210
Test location	-.499	.161	-3.101	.002	.607	.442	.832
OCL accuracy	-5.115	.874	-5.851	<.001	.006	.001	.033
Session number	.036	.066	.541	.588	1.036	.911	1.178
OCL accuracy *session number	.030	.060	.505	.614	1.030	.917	1.157

Table D2

Generalized Linear Mixed Model for Repeated MoCA Performance Predicting Conversion to aMCI or AD

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	11.162	2.480	4.502	<.001			
Age	-.012	.020	-.631	.528	.988	.950	1.027
Sex	-.029	.245	-.199	.905	.971	.601	1.570
Education	-.009	.051	-.184	.854	.991	.896	1.095
Race/ethnicity	-1.140	.513	-2.225	.026	.320	.117	.874
ApoE ε4+	.388	.243	1.597	.111	1.475	.915	2.377
MoCA total score	-.524	.084	-6.241	<.001	.592	.502	.698
Session number	-.759	.353	-2.149	.032	.468	.234	.936
MoCA total score *session number	.038	.014	2.706	.007	1.039	1.011	1.068

Table D3

Generalized Linear Mixed Model for Repeated RAVLT Performance Predicting Conversion to aMCI or AD

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	-1.179	1.610	-.732	.464			
Age	.004	.020	.228	.820	1.004	.967	1.044
Sex	-.009	.250	-.035	.972	.991	.607	1.619
Education	-.037	.050	-.751	.453	.964	.875	1.061
Race/ethnicity	-.703	.491	-1.431	.153	.495	.189	1.299
ApoE ε4+	.273	.243	1.123	.262	1.314	.815	2.118
RAVLT 30-minute delay	-.159	.051	-3.129	.002	.853	.773	.943
Session number	.086	.064	1.357	.175	1.090	.962	1.235
RAVLT score* session number	-.001	.008	-.117	.907	.999	.983	1.015

Table D4

Generalized Linear Mixed Model for Repeated OCL Performance Predicting Aβ Accumulation

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	-1.802	.748	-2.410	.016			
Age	.021	.006	3.662	<.001	1.022	1.010	1.033
Sex	.151	.080	1.884	.060	1.163	.994	1.362
Education	.004	.016	.251	.802	1.004	.973	1.035
Race/ethnicity	-.654	.207	-3.155	.002	.520	.347	.781
ApoE ε4+	1.507	.078	19.231	<.001	4.514	3.871	5.263
Test location	.017	.093	.184	.854	1.017	.847	1.222
OCL accuracy	-1.180	.549	-2.149	.032	.307	.105	.902
Session number	.201	.054	3.695	<.001	1.222	1.099	1.360
OCL accuracy *session number	-.176	.052	-3.423	<.001	.838	.758	.927

Table D5

Generalized Linear Mixed Model for Repeated MoCA Performance Predicting Aβ Accumulation

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	-2.696	1.328	-2.030	.043			
Age	.037	.011	3.393	<.001	1.037	1.016	1.059
Sex	.366	.148	2.468	.014	1.441	1.078	1.928
Education	.052	.030	1.773	.076	1.054	.994	1.117
Race/ethnicity	-1.122	.335	-3.354	<.001	.326	.169	.628
ApoE ε4+	1.787	.146	12.269	<.001	5.970	4.486	7.944
MoCA total score	-.079	.038	-2.074	.038	.924	.858	.996
Session number	.197	.170	1.161	.246	1.217	.873	1.698
MoCA total score *session number	-.010	.007	-1.505	.133	.990	.977	1.003

Table D6

Generalized Linear Mixed Model for Repeated RAVLT Scores Predicting Aβ Accumulation

	B	*SE*	*t*	*p*	OR	OR 95% CI	
						Lower	Upper
Intercept	-3.996	.931	-4.292	<.001			
Age	.046	.011	4.173	<.001	1.047	1.025	1.070
Sex	.407	.151	2.699	.007	1.502	1.117	2.019
Education	.021	.028	.725	.469	1.021	.965	1.079
Race/ethnicity	-1.121	.333	-3.372	<.001	.326	.170	.626
ApoE ε4+	1.743	.145	12.035	<.001	5.712	4.299	7.589
RAVLT 30-minute delay	-.103	.028	-3.718	<.001	.902	.854	.952
Session number	-.135	.036	-3.799	<.001	.873	.814	.937
RAVLT score *session number	.008	.005	1.601	.110	1.008	.998	1.017

Appendix E

Figure E1

Interaction Between OCL Accuracy and OCL Sessions for GLMM Predicting Aβ Accumulation