




Research Article

Comparing the predictive validity of four MCI definitions for incident dementia in demographically diverse community-dwelling individuals: Results from the Einstein Aging Study (EAS)

Katherine H. Chang^{1,2,3} , Cuiling Wang^{3,4}, Jiyue Qin⁵, Mindy J. Katz³, Desiree A. Byrd^{1,2}, Richard B. Lipton^{3,4,6} and Laura A. Rabin^{1,2,3,7}

¹Department of Psychology, Queens College, City University of New York (CUNY), Flushing, NY, USA, ²Department of Psychology, The Graduate Center, City University of New York (CUNY), New York, NY, USA, ³Department of Neurology, Albert Einstein College of Medicine, Bronx, NY, USA, ⁴Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA, ⁵Division of Biostatistics and Bioinformatics, Herbert Wertheim School of Public Health and Human Longevity Science, UC San Diego, La Jolla, CA, USA, ⁶Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, Bronx, NY, USA and ⁷Department of Psychology, Brooklyn College, City University of New York (CUNY), Brooklyn, NY, USA

Abstract

Objective: Research examining (MCI) criteria in diverse and/or health-disparate populations is limited. There is a critical need to investigate the predictive validity for incident dementia of widely used MCI definitions in diverse populations. **Method:** Eligible participants were non-Hispanic White or Black Bronx community residents, free of dementia at enrollment, with at least one annual follow-up visit after baseline. Participants completed annual neurological and neuropsychological evaluations to determine cognitive status. Dementia was defined based on DSM-IV criteria using case conferences. Cox proportional hazard models assessed predictive validity for incident dementia of four specific MCI definitions (Petersen, Jak/Bondi, number of impaired tests, Global Clinical Ratings) at baseline, controlling for age, sex, education, and race/ethnicity. Time-dependent sensitivity and specificity at 2–7 years for each definition, and Youden's index were calculated as accuracy measures. **Results:** Participants (N = 1073) ranged in age from 70 to 100 (mean = 78.4 ± 5.3) years at baseline. The sample was 62.5% female, and educational achievement averaged 13.9 ± 3.5 years. Most participants identified as White (70.0%), though Black participants were well-represented (30.0%). In general, MCI definitions differed in sensitivity and specificity for incident dementia. However, there were no significant differences in Youden's index for any definition, across all years of follow-up. **Conclusions:** This work provides an important step toward improving the generalizability of the MCI diagnosis to underrepresented/health-disparate populations. While our findings suggest the studied MCI classifications are comparable, researchers and clinicians may choose to consider one method over another depending on the rationale for evaluation or question of interest.

Keywords: Mild cognitive impairment; cross-cultural neuropsychology; cross-cultural issues; diversity; longitudinal; aging

(Received 18 March 2024; final revision 15 November 2024; accepted 19 November 2024)

Introduction

Mild cognitive impairment (MCI) represents an intermediate stage of cognitive impairment that is often, but not always, a transitional phase between normal aging and dementia (Petersen et al., 1999, 2014; Scharre, 2019; Winblad et al., 2004). The original construct of MCI was intended to identify and evaluate individuals for suspected Alzheimer's disease (AD) (Petersen et al., 1999; Petersen, 2004), and indeed those who carry a diagnosis of MCI are at increased risk for further cognitive and functional decline with the rate of conversion to dementia ranging from 5 to 15% annually (Angevaere et al., 2022; Dunne et al., 2021, Thomas et al., 2020). However, not everyone who carries a diagnosis of MCI progresses to AD or other subtypes of dementia, and longitudinal studies show that 5–53% of individuals with MCI at baseline no longer

meet diagnostic criteria at their subsequent study visit. Higher rates of reversion occur in community-based settings (Ganguli et al., 2011; Manly et al., 2008; Pandya, Lacritz, Weiner, Deschner, & Woon, 2017; Angevaere et al., 2022). Importantly, it is now widely recognized that the underlying causes of MCI are heterogeneous (e.g., stroke, traumatic brain injury, medication side effects, low vitamin B12 levels, etc.; Al-Qazzaz, Ali, Ahmad, Islam, & Mohamad., 2014; Calvillo & Irimia, 2020; Marvanova, 2016; Moore et al., 2012), reversible (in the case of medication side effects or recovery from neurological insult), and not necessarily due to a neurodegenerative disorder (Albert et al., 2011; Dunne et al., 2021; Petersen et al., 2001). In the absence of methods to slow or halt many causes of neurodegeneration, a diagnosis of MCI without a clear understanding about the likely prognosis carries the risk of causing

Corresponding author: Katherine H. Chang; Email: kchang@gradcenter.cuny.edu

Cite this article: Chang K.H., Wang C., Qin J., Katz M.J., Byrd D.A., Lipton R.B., & Rabin Laura A. Comparing the predictive validity of four MCI definitions for incident dementia in demographically diverse community-dwelling individuals: Results from the Einstein Aging Study (EAS). *Journal of the International Neuropsychological Society*, 1–10, <https://doi.org/10.1017/S1355617724000729>

© The Author(s), 2024. Published by Cambridge University Press on behalf of International Neuropsychological Society

confusion, concern, and potential overtreatment of individuals who may never progress to dementia (Kaduszkiewicz *et al.*, 2014; Gomersall *et al.*, 2015; Visser *et al.*, 2020).

Although diagnostic criteria for “MCI” and international consensus guidelines have been developed and refined over decades beginning in the 1990s, (Albert *et al.*, 2011; Artero *et al.*, 2006; Chen *et al.*, 2021; Dunne *et al.*, 2021; Petersen *et al.*, 1999; Winblad *et al.*, 2004) several definitions of MCI are still widely used (Bondi *et al.*, 2014; Edmonds *et al.*, 2021; Graves *et al.*, 2020; Vuoksimaa *et al.*, 2020; Wong *et al.*, 2018). Fluid biomarkers, alone or in combination with cognitive assessments, are now recommended for earlier identification of individuals in the preclinical or prodromal stages of AD (Dubois *et al.*, 2021; Hansson *et al.*, 2022; Jack *et al.*, 2018; Scharre, 2019), although their use outside research settings or in low-income countries is limited. There are no specific biomarkers to confirm the presence of MCI (Giau *et al.*, 2019), nor does an abnormal result provide certainty of decline (Dubois *et al.*, 2021; Mormino & Papp, 2018). Rather, MCI diagnosis continues to rely largely on phenotypic presentation and clinical judgment (Donders, 2020; Dubois *et al.*, 2021; Hansson *et al.*, 2022; Scharre, 2019). Current conventional clinical criteria are largely iterations of criteria first proposed by Petersen *et al.* (1999) and updated by Winblad *et al.* (2004) at the First Key Symposium on MCI. Notably, these criteria were never explicitly operationally defined, leaving them open for divergent interpretations and applications. There is no consensus, for example, for how objective impairment should be operationalized. This results in wide variation in the application of the definition, including which neurocognitive domains to include, the number of tests within a domain required to demonstrate impairment, the specific tests to use, and the level of impairment required (Ritchie *et al.*, 2001; de Vent *et al.*, 2020). Differences in the operationalization of various case definitions contribute to variations in prevalence estimates, reversion rates, and predictive validity for subsequent dementia (Bondi & Smith, 2014).

With no guidance from the conventional Petersen/Winblad definition on the number of tests that must show impairment within a cognitive domain, the minimum requirement is often assumed (*i.e.*, one impaired score). However, requiring only one impaired score has been criticized (Jak *et al.*, 2016; Loewenstein *et al.*, 2009; Trittschuh *et al.*, 2011), as upwards of 55% of cognitively normal adults have one or more impaired scores when given a battery of neurocognitive tests (Binder *et al.*, 2009; Blackford & LaRue, 1989; Brooks *et al.*, 2007; Brooks *et al.*, 2008). Having a false-positive diagnosis increases with the number of tests administered. In addition, the number of false positives followed by reversion instead of progression, might also increase due to day-to-day variation in cognitive performance if testing occurs on a bad day (Brose *et al.*, 2012; von Stumm, 2016; Facer-Childs, Boiling, & Balanos, 2018).

To improve diagnostic stability and reduce the number of false-positive diagnostic errors, Jak/Bondi proposed an MCI classification using comprehensive neuropsychological criteria requiring two impaired scores to establish reliable cognitive impairment (Jak, Bondi *et al.*, 2009). Using this comprehensive neuropsychological approach, Jak/Bondi’s definition increased diagnostic accuracy (*i.e.*, had a similar strength of association with incident dementia to Petersen/Winblad criteria despite diagnosing 30% fewer participants with MCI) and correlated more strongly with AD biomarkers when compared to Petersen/Winblad criteria in the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (Bondi *et al.*, 2014; Jak *et al.*, 2016).

Despite progress in refining MCI criteria, there is still no consensus as to (1) which specific cognitive tests (*e.g.*, list learning, story memory, or both, within the memory domain) should be used; (2) how many tests across cognitive domains should comprise a test battery; or (3) what norms or thresholds should indicate cognitive impairment (Vuoksimaa *et al.*, 2018; Oltra-Cucarella *et al.*, 2018). Moreover, sole reliance on quantitative data may omit important contextual information for interpreting whether an abnormal result truly reflects a decline (Ganguli *et al.*, 2011). For example, it is well-documented that racially/ethnically diverse individuals, on average, tend to obtain lower scores on cognitive tests than their non-Hispanic White counterparts due to a variety of socioeconomic factors such as quality of early education, acculturation, and health environments (Boone *et al.*, 2007; Byrd *et al.*, 2006; Manly & Echemendia, 2007; Zahodne *et al.*, 2021). In the oldest-old (aged 80+), there may be several reasons beyond incipient dementia for low cognitive scores (Hong *et al.*, 2003; Kravitz *et al.*, 2012; Wong *et al.*, 2018), such as age-related accumulation of comorbidities, fatigue, or sensory impairment (Arosio *et al.*, 2017). Furthermore, Brooks and colleagues (2007) demonstrated that low scores on cognitive tests occur in 44% of individuals with high average or extremely high intellectual functioning and 80% of older adults with low average intellectual abilities using a 1 SD cutoff. The risk of overinterpreting low test scores because of high base rates and normal variability in these various populations is high, and as such, all definitions, including the Jak/Bondi definition, may be subject to false-positive errors.

To account for the effects of normal variability in cognitive test performance and to address certain criticisms of Petersen and Jak/Bondi classification criteria, Oltra-Cucarella and colleagues (2018) proposed to classify individuals with MCI based on the number of impaired tests (NIT). The researchers first defined a base rate of low scores as the average number of low scores in the worst-performing 10% of cognitively normal individuals. They then theorized that individuals with a greater number of low scores than expected would be more likely to demonstrate true impairment (thereby minimizing the number of false positives). When applied to the ADNI dataset, the use of this classification led to better prediction of progression from MCI to AD (34%) than both the Petersen/Winblad (29%) and the Jak/Bondi criteria (31%) (Oltra-Cucarella *et al.*, 2018). However, the following concerns were raised regarding wide utility of this approach: (1) cutoff for impairment is based on the ADNI sample rather than external or nationwide norms, and it may not be as easy to calculate a base rate of low scores in practice; (2) the number of impaired scores in the lowest 10% will vary depending on the number of measures and scores available; and (3) this study used multiple scores from the same test, possibly inflating the number of impaired scores (*i.e.*, individuals with impaired list-learning recognition will likely have impaired recall as well) (Vuoksimaa *et al.*, 2020).

Most recently, Alfano and colleagues (2022) suggested using a clinical ratings (Global CR) approach, which has been validated and is considered a “gold standard” for detecting cognitive impairment in the human immunodeficiency virus literature (Heaton *et al.*, 1995; Malaspina *et al.*, 2011; Blackstone *et al.*, 2012) and closely parallels the profile analysis strategy seen in the clinical decision-making process (Alfano *et al.*, 2022). Using this classification method, 26% of the sample was diagnosed with MCI and there was substantial agreement between Petersen/Winblad and Jak/Bondi definitions (Alfano *et al.*, 2022). Notably, this study was cross-sectional so there is no information regarding the predictive validity of the Global CR approach for incident

dementia. Moreover, Alfano and colleagues (2022) acknowledged the importance of further validating the approach in individuals with known or suspected MCI and of incorporating demographic variables such as race/ethnicity.

Much of the work on defining MCI criteria has been carried out at the Mayo Clinic in Rochester, Minnesota, and in ADNI where over 90% of participants are non-Hispanic White and highly educated with college degrees (Bondi et al., 2014; Oltra-Cucarella et al., 2018; Petersen et al., 2010; Vuoksima et al., 2020). Comparisons between conventional and neuropsychological criteria have similarly been attempted in demographically homogeneous samples (Bondi et al., 2014; Edmonds et al., 2021; Graves et al., 2020; Jak et al., 2016; Oltra-Cucarella et al., 2018; Thomas et al., 2019; Wong et al., 2018). However, diverse older adults tend to obtain lower test scores across cognitive domains (Díaz-Venegas et al., 2016; Gasquoin, 2009; Zahodne, Manly, Azar, Brickman, & Glymour, 2016; Werry et al., 2019) due to non-neurological factors including, but not exclusive to, quality of education, acculturation, or bias (Boone et al., 2007; Byrd et al., 2006; Manly & Echemendia, 2007; Manly, 2005). Thus far, two studies have compared the diagnostic agreement of MCI definitions in diverse older adults. However, these studies utilized individuals from a specialized memory clinic and measured Jak/Bondi criteria against the consensus approach and not Petersen/Winblad criteria alone (Devlin et al., 2022; Graves et al., 2022); furthermore, neither study included NIT or Global CR approaches. Thus, it remains unclear (1) how Petersen/Winblad (not a consensus definition) and Jak/Bondi criteria compare in predicting incident dementia in a community-based sample and (2) how other methods of MCI measurement that may account for normal variability (i.e., NIT or Global CR) would fare relative to these two-widely used definitions.

In the current study, we compared the four definitions of MCI described above: conventional Petersen/Winblad criteria, Jak/Bondi's comprehensive neuropsychological criteria, NIT, and Global CR within the Einstein Aging Study (EAS), a community-residing, racially/ethnically and educationally diverse cohort of older adults from Bronx, NY. Specifically, we sought to address gaps in the literature by comparing the prevalence of MCI diagnosis and predictive validity for incident dementia to determine which MCI definition approach offers the greatest diagnostic and prognostic accuracy, particularly in diverse, community-based samples. These results will inform how MCI will be operationalized in the EAS sample and may be relevant to other samples with similar demographic characteristics.

Methods

Participants

The EAS is a longitudinal study of community-residing individuals in a racially and ethnically diverse urban setting (Katz et al., 2012). Details about study recruitment have been described elsewhere (Katz et al., 2012). In brief, participants were systematically recruited using Bronx County Voter Registration lists, mailed introductory letters, and given a telephone screen to determine study eligibility. Those who met preliminary eligibility criteria were invited for further in-person evaluations. In-person assessments were then conducted annually and included comprehensive neurological, medical, psychosocial, and neuropsychological evaluations. All protocols were approved by the Einstein Institutional Review Board and written informed consent was obtained at the initial clinic visit. This research was completed in accordance with the Helsinki Declaration. Inclusion criteria were

age 70 and above, resident of Bronx, NY, community-dwelling, and English speaking. Exclusion criteria at enrollment included severe audiovisual or physical impairments, or active psychiatric symptomatology, which may interfere with the ability to complete assessments (Katz et al., 2012). Participants eligible for the analysis were enrolled between October 1993 and June 2016, free of dementia at baseline, and had at least one annual follow-up. Due to the limited number of participants who were Hispanic or of other racial/ethnic groups, the study was further restricted to those who self-reported as non-Hispanic White or non-Hispanic Black.

Neuropsychological assessment

Participants completed standardized neuropsychological testing at baseline and all follow-up visits, which were conducted annually (Katz et al., 2012). Five cognitive domains were used for MCI diagnosis: memory, attention, executive functioning, language, and visuospatial functioning, with two tests included in each domain. The memory domain included the free recall measure from the Free and Cued Selective Reminding Test (Buschke, 1984) and the Wechsler Memory Scale-Revised Logical Memory I subtest (WMS-R-LMI) (Wechsler, 1987). Attention/processing speed was measured using the Digit Span subtest of the Wechsler Adult Intelligence Scale-III (Wechsler, 1997) and the Trail Making Test, part A (Reitan, 1958). Executive functioning tests included the Trail Making Test, part B (Reitan, 1958), and the Letter Fluency "FAS" task (Spree & Strauss, 2006). Language was measured with the Category Fluency task (animals, vegetables, fruits) (Rosen, 1980) and the Boston Naming Test (Kaplan et al., 1983). Visuospatial functioning tests included the Block Design and Digit Symbol subtest from the WAIS-III (Wechsler, 1997).

Group/diagnostic MCI classification

MCI classifications were made using four sets of criteria: Petersen/Winblad standard criteria (Petersen et al., 1999; Winblad et al., 2004), Jak/Bondi's comprehensive neuropsychological criteria (Jak, Bondi et al., 2009; Bondi et al., 2014), NIT (Oltra-Cucarella et al., 2018), and Global CR (Alfano et al., 2022). Normative data were calculated using local norms derived for cognitively unimpaired (CU) individuals in the sample (described in Supplementary Methods), and scores for all cognitive tests used were adjusted for age, sex, education, and race/ethnicity.

Petersen/Winblad

Updated Petersen criteria (Artero et al., 2006; Winblad et al., 2004) were used and to receive a diagnosis of MCI, it required (1) objective cognitive impairment (>1.5 SD below the age-, sex-, education-, and race/ethnicity-adjusted mean) in at least one test; (2) subjective cognitive concern operationalized as any SCC indicated by self- or informant-report; (3) absence of functional decline as measured by the IADL Lawton Brody scale (Lawton & Brody, 1969); and (4) no diagnosis of dementia.

Jak/Bondi

The most widely used Jak/Bondi definition (Bondi et al., 2014) for MCI was adapted for use in our study: (1) one low score (>1 SD below the age-, sex-, education-, and race/ethnicity-adjusted mean) on both measures within at least one cognitive domain or (2) at least one low score (>1 SD below the age-, sex-, education-, and race/ethnicity-adjusted mean) across at least two cognitive domains. The three-domain criterion that was previously used in a study conducted in the ADNI sample (Bondi et al., 2014) was modified to

account for a larger battery of five cognitive domains and follows a similar approach that Graves and colleagues (2020, 2022) used for application of the Jak/Bondi criteria in the National Alzheimer's Coordinating Center cohort. Although Functional Activities Questionnaire (FAQ; Pfeffer *et al.*, 1982) scores were included in the Jak/Bondi definition from this previous ADNI study (Bondi *et al.*, 2014), this was a standalone criterion used to classify MCI (i.e., individuals had to meet one of three criteria; the third being FAQ scores), and most recent studies using Jak/Bondi do not include it in their operationalizations. Moreover, there are no guidelines on cutoff scores (i.e., 9 as a standalone criterion in the original definition but more recent studies incorporate FAQ + objective impairment, but only for the multiple domain criteria). Lastly, FAQ is an informant-reported measure that may be affected by demographic and relational characteristics (Hackett *et al.*, 2020; Jessen *et al.*, 2014). In a recent study examining sources of discrepancy on FAQ in a diverse sample, informants who identify as Black/African American endorsed lower ratings (less impairment) (Hackett *et al.*, 2020), suggesting there may be some cross-cultural differences that reduce the validity of this measure. As such, FAQ scores were not included for the Jak/Bondi criteria. Further, SCC is not used in this definition.

Number of impaired tests (NIT)

For the NIT definition (Oltra-Cucarella *et al.*, 2018), MCI was diagnosed when the NIT (>1.5 SD below the age-, sex-, education-, and race/ethnicity-adjusted mean) equaled or exceeded the NIT obtained by the worst performing 10% of the CU group (inclusion criteria described in Supplementary Methods). To diagnose MCI, three or more impaired scores were required. Functional impairment and SCC are not included in this definition.

Clinical ratings (Global CR)

For the Global CR approach, MCI was diagnosed according to Alfano and colleagues (2022). Individual scores were converted to demographically (age-, sex-, education-, race/ethnicity-) corrected T-scores and assigned a CR ranging from 1 (above average; $T \geq 55$) to 9 (severe impairment; $T \leq 19$), with a CR of 5 indicating "definite mild impairment" ($35 < T < 39$). Domain CRs were derived from the highest individual test CR in a given domain. If both tests in a given domain had different CRs, then domain CR was defined as the higher CR minus one. If not (i.e., both tests had the same CRs), that rating was the domain CR. Global CR was derived from the highest domain CR such that if only one domain has the highest CR, then Global CR was the highest domain CR minus one. If two or more domains had the same highest CR, that rating was the Global CR. MCI was diagnosed if an individual's Global CR equaled or exceeded a rating of five. Functional impairment and SCC are not included in this definition.

Dementia

Incident dementia of any etiology at follow-up was the main study outcome. DSM-IV criteria were used for dementia diagnosis (APA, 1994), and participants were classified as having incident dementia if (1) there was substantial cognitive impairment – that is, scores at least 1.5 standard deviations below the age-adjusted mean across multiple cognitive domains; (2) the participant or study informant reported changes in cognitive function; (3) there was functional decline determined at a case conference based on information from the IADL Lawton Brody Scale (Lawton & Brody, 1969) and clinical evaluation; and (4) cognitive impairment was not better explained by the effects of a substance or medication.

Demographic and clinical characteristics

Demographic information from the EAS included self-reported race/ethnicity as defined by the U.S. Census Bureau in 1994 (recategorized to non-Hispanic White, non-Hispanic Black), number of years of education, sex, and age. Subclinical symptoms of depression were assessed using the GDS short form, excluding the memory item (Sheikh & Yesavage, 1986).

Data analysis

Demographic information, neuropsychological test scores, and all definitions of MCI status at baseline were summarized using descriptive statistics by incident dementia status during follow-up. Preliminary comparisons of these characteristics were performed using the Wilcoxon rank sum test and chi-square test for continuous and categorical variables, respectively. Time to event was defined as the time in years from the baseline clinic visit to the visit when dementia was diagnosed or to the last follow-up visit. Cox proportional hazard models were used to assess the associations of various definitions of MCI at baseline with the risk of incident dementia. Additional models controlling for covariates including age, gender, education, race/ethnicity, and depressive symptoms at baseline were also applied. The modification effects of race/ethnicity on the associations of each MCI definition with the risk of dementia were examined by testing the interaction between each MCI definition and race/ethnicity in the Cox models. To evaluate the accuracy of discriminating those at risk of developing dementia within specific time periods using each MCI definition at baseline, time-dependent sensitivity and specificity were examined at 2, 3, 5, and 7 years, and Youden's index was calculated. Variations and confidence intervals (CIs) of these statistics were obtained using the bootstrap method (Tibshirani & Efron, 1993) with 1000 resamples and 2.5% and 97.5% percentiles used as limits of the 95% CI. Comparisons of sensitivity, specificity, and Youden's index among different MCI definitions were performed by obtaining bootstrap confidence intervals of the differences and checking if zero was covered in the intervals. The analysis was further stratified by race group, and comparisons of sensitivity, specificity, and Youden's index for each MCI definition between racial groups were performed using Z-tests based on the estimates and their variances from the bootstrap method. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC).

Results

At baseline ($N = 1073$), participants' ages ranged from 70 to 100 (mean = 78.4 ± 5.3) years. The sample was 62.5% female, and educational achievement averaged 13.9 ± 3.5 years (range 2–24 years), with 42.6% obtaining 12 years or fewer years of education. Notably, nearly one-fifth of participants (17%) did not complete high school, highlighting the rich educational diversity within the sample. Most participants identified as White (70.0%), though Black participants were well-represented (30.0%). The total number and proportion of participants diagnosed with MCI using each classification method are shown in Table 1. During a mean of 4.5 (median 3.3, maximum 19) years of follow-up, 118 individuals developed incident dementia. As shown in Table 1, those who developed dementia were older at baseline. The groups did not differ in number of depressive symptoms, sex, years of education, race/ethnicity, or follow-up time.

Table 1. Baseline descriptive characteristics of the sample by dementia status

	All ^a	No dementia during follow-up	Incident dementia	<i>p</i>
Sample size	1073	955	118	
Age at baseline: <i>M</i> (<i>SD</i>)	78.4 (5.3)	78.1 (5.2)	81.0 (5.4)	< .0001
Sex (#, % Female)	671, 62.5%	594, 62.2%	77, 65.3%	0.518
Education: <i>M</i> (<i>SD</i>)	13.9 (3.5)	14.0 (3.5)	13.4 (3.5)	0.058
Race/ethnicity				
#, % non-Hispanic White	751, 70.0%	674, 70.6%	77, 65.3%	0.234
#, % non-Hispanic Black	322, 30.0%	281, 29.4%	41, 34.7%	
GDS score ⁺ <i>M</i> (<i>SD</i>)	2.2 (2.3)	2.2 (2.3)	2.5 (2.3)	0.085
MCI diagnosis (#, %)				
Petersen/Winblad	463, 43.2%			
Jak/Bondi	386, 36.0%			
NIT	169, 15.8%			
Global CR	516, 48.1%			
Follow-up time, years median (IQR)	3.3 (2.0–6.3)	3.4 (2.0–6.4)	3.1 (1.7–6.1)	0.485

Note. NIT = number of impaired test; CR = clinical ratings.

^a Participants free of prevalent dementia at baseline who had at least one follow-up evaluation, completed neuropsychological tests and self-reported subjective cognitive concern at enrollment and self-reported as non-Hispanic White or non-Hispanic Black.

⁺ Excluding the dichotomous memory item from the short form of the GDS.

Mean time between baseline and conversion to dementia or end of study participation was 4.5 ± 3.5 years.

Table 2. Hazard ratios using diagnostic classifications of MCI to predict incident dementia

MCI classification	Model 1: unadjusted		Model 2: + age, sex, education, race/ethnicity		Model 3: + age, sex, education, race/ethnicity, GDS score	
	HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>	HR (95% CI)	<i>p</i>
Petersen/Winblad	4.6 (3.1–7.0)	< .001	4.3 (2.9–6.5)	< .001	4.8 (3.1–7.7)	< .001
Jak/Bondi	5.3 (3.6–7.9)	< .001	5.2 (3.5–7.8)	< .001	5.2 (3.4–8.0)	< .001
NIT	6.5 (4.5–9.5)	< .001	5.9 (4.0–8.6)	< .001	6.3 (4.2–9.6)	< .001
Global CR	6.2 (3.9–9.9)	< .001	6.1 (3.8–9.8)	< .001	6.4 (3.9–10.7)	< .001

Note. HR = hazard ratio; CI = confidence interval; NIT = number of impaired test; CR = clinical ratings.

Hazard ratios from unadjusted models were similar among the four definitions (Table 2) and all classifications significantly predicted incident dementia (all $p < 0.001$). Results were similar when adjusted for age, sex, education, race/ethnicity (Model 2), and further adjusted for depressive symptoms at baseline (Model 3). Results of time-dependent Receiver Operating Characteristic (ROC) statistics for MCI classifications, at 2-, 3-, 5- and 7-years of follow-up, and the number of participants at risk at the time are shown in Table 3. The tests of interactions between race/ethnicity and MCI on the risk of dementia in the aforementioned Cox models were not statistically significant for any MCI definition. Nonetheless, we further evaluated the time-dependent ROC statistics within each racial/ethnic group in a secondary analysis due to a specific interest in exploring any disparities between groups.

The bootstrap method comparing sensitivity and specificity among the MCI classifications demonstrated varying statistical significance in sensitivity and specificity across definitions. Between Jak/Bondi and Petersen definitions, Jak/Bondi classification significantly increased specificity ($p < 0.05$) while there was no significant difference between levels of sensitivity. All other definitions significantly differed in sensitivity and specificity ($p < 0.05$), with Jak/Bondi and Petersen classifications demonstrating more intermediate values of sensitivity and specificity relative to NIT and Global CR classifications.

In the overall sample, Youden's index, which is determined by the sum of sensitivity and specificity minus 1, was highest for the Jak/Bondi classification for dementia incidence within 2-, 3-, 5-, and 7-years of follow-up, followed by the Global CR classification (Table 3). While NIT had significantly higher specificity than all other definitions, these gains were more than offset by a substantial loss of sensitivity that was significantly lower than all the other definitions. The most extreme definition in the opposite direction, the Global CR classification, had significantly higher sensitivity at all follow-up; however, this was achieved through a significant sacrifice to specificity (e.g., specificity was significantly lower than in all other definitions). Between Jak/Bondi and Petersen/Winblad classification methods, Jak/Bondi classification demonstrated significantly higher specificity than Petersen/Winblad; sensitivity was slightly lower but this was not a significant finding. Importantly, differences in Youden's index were not statistically significant across definitions as measured by the bootstrap confidence intervals. Specificity, sensitivity, and Youden's index across MCI classifications and follow-up years for non-Hispanic Blacks and non-Hispanic Whites are shown in Appendix Tables S1a and S1b. Within the non-Hispanic Black group, the Global CR classification showed the highest Youden's index, though this was not significantly different from other MCI classifications. Comparisons of sensitivity,

Table 3. Estimates (95% CI) of sensitivity, specificity, and Youden's index for diagnostic definitions of MCI at baseline for cumulative dementia after 2 to 7 years of follow-up

Follow-up year (# cumulative incidence; # of participants at risk)	MCI classification	Sensitivity (95% CI)	Specificity (95% CI)	Youden's index (95% CI)
2 (34; 787)	Petersen/Winblad	0.77 (0.72–0.83)	0.58 (0.56–0.60)	0.36 (0.30–0.41)
	Jak/Bondi	0.74 (0.68–0.80)	0.66 (0.63–0.68)	0.40 (0.34–0.46)
	NIT	0.54 (0.46–0.61)	0.86 (0.84–0.87)	0.39 (0.32–0.47)
	Global CR	0.85 (0.80–0.90)	0.53 (0.51–0.56)	0.38 (0.33–0.43)
3 (56; 608)	Petersen/Winblad	0.77 (0.72–0.83)	0.59 (0.57–0.62)	0.36 (0.31–0.42)
	Jak/Bondi	0.74 (0.68–0.80)	0.67 (0.64–0.69)	0.41 (0.35–0.47)
	NIT	0.52 (0.45–0.59)	0.87 (0.85–0.88)	0.39 (0.32–0.47)
	Global CR	0.85 (0.80–0.89)	0.55 (0.52–0.57)	0.39 (0.34–0.44)
5 (78; 381)	Petersen/Winblad	0.76 (0.71–0.82)	0.61 (0.59–0.63)	0.37 (0.31–0.44)
	Jak/Bondi	0.73 (0.67–0.79)	0.69 (0.66–0.71)	0.42 (0.35–0.48)
	NIT	0.51 (0.44–0.58)	0.89 (0.87–0.90)	0.39 (0.32–0.46)
	Global CR	0.84 (0.80–0.89)	0.56 (0.54–0.59)	0.41 (0.35–0.46)
7 (95; 232)	Petersen/Winblad	0.76 (0.70–0.81)	0.63 (0.61–0.66)	0.39 (0.32–0.45)
	Jak/Bondi	0.72 (0.66–0.78)	0.71 (0.68–0.74)	0.43 (0.36–0.50)
	NIT	0.48 (0.42–0.55)	0.91 (0.89–0.92)	0.39 (0.32–0.46)
	Global CR	0.83 (0.79–0.88)	0.59 (0.56–0.62)	0.42 (0.37–0.48)

Note. MCI = mild cognitive impairment; CI = confidence interval; NIT = number of impaired test; CR = clinical ratings.

specificity, and Youden's index among the MCI definitions within each racial group were overall similar to those of the whole sample. Between the two racial/ethnic groups, MCI based on the Global CR classification had significantly higher sensitivities at 2, 3, 5, and 7 years and lower specificities at 2, 3, and 5 years in the non-Hispanic Black group when compared with the non-Hispanic White group. MCI based on Petersen/Winblad had significantly lower specificities at 2, 3, and 5 years in the non-Hispanic Black group, but differences in sensitivities and specificities for other definitions, and differences in Youden's index for all definitions, were not statistically significant.

Discussion

The present study investigated four specific approaches to MCI classifications (Petersen, Jak/Bondi, NIT, and Global CR) in an educationally and racially/ethnically diverse cohort of older adults. Overall, we found that the MCI classifications varied significantly in sensitivity and specificity for incident dementia, with Jak/Bondi and Petersen classifications demonstrating more intermediate values of sensitivity and specificity relative to NIT and Global CR classifications, which demonstrated the greatest specificity and greatest sensitivity values, respectively. However, when examining Youden's index, reflecting the balance between sensitivity and specificity, all four classifications performed comparably. Jak/Bondi MCI classification demonstrated the highest Youden's index, but this was not significantly different from the other classification methods. Comparisons of sensitivity, specificity, and

Youden's index among the MCI definitions within each racial group were overall similar to those of the whole sample.

This is an important finding as researchers or clinicians could have increased confidence in choosing one classification method over another, depending on their referral or research question of interest. For example, if a group was interested in examining SCC, they might select the Jak/Bondi classification to avoid circularity, as other criteria, such as Petersen/Winblad include SCC in the definition of MCI. The optimal balance between sensitivity and specificity depends upon the research, clinical, or public health setting of intended use. Sensitivity might be a priority if those who screen positive receive additional cognitive testing or low-cost, blood-based biomarkers. Specificity might be a priority if those who meet the MCI definition will receive invasive or expensive follow-up and resources are constrained. Professionals could consider using the cost-benefit matrix as proposed by Manly and Echemendia (2007) when considering whether they require increased sensitivity or specificity and then select their classification method accordingly. Due to the benefits of early detection and treatment for dementia, increased diagnostic sensitivity and consideration of Global CR may be warranted. Moreover, when evaluating non-Hispanic Black or non-Hispanic White individuals, professionals might choose to utilize one definition over another in terms of differing sensitivity or specificity within these groups.

These results also replicate and extend previous findings comparing Jak/Bondi and Petersen/Winblad classifications in non-Hispanic White samples (Bondi *et al.*, 2014; Edmonds *et al.*, 2021; Jak *et al.*, 2016; Thomas *et al.*, 2019; Wong *et al.*, 2018). In our

overall sample, the Jak/Bondi definition significantly increased specificity (decreased false positives) while maintaining comparable sensitivity when compared to Petersen/Winblad. These findings suggest Petersen/Winblad may result in more false positive MCI diagnoses. This is consistent with literature demonstrating that (1) defining cognitive impairment based on a single impaired score does not account for normal variability and may inflate MCI prevalence rates (Loewenstein et al., 2009; Trittschuh et al., 2011), and (2) the subjective cognitive concern criterion in Petersen/Winblad is not additive to predicting incident dementia and may contribute to misdiagnosis of MCI (Chang et al., 2023 in press; Edmonds et al., 2015, 2018; Ilardi et al., 2021).

Although the NIT classification (Oltra-Cucarella et al., 2018) performed well in the ADNI sample, its criteria appeared to be too stringent for our participants. In our study, the specificity for incident dementia was highest using the NIT approach, ranging from 0.86 to 0.91; however, sensitivity for incident dementia was substantially reduced (0.48–0.54), resulting in a failure to identify almost half of the individuals who ultimately progressed to dementia. It is important to note certain methodological differences between the studies of interest. Oltra-Cucarella and colleagues' original approach was previously criticized for possibly inflating the number of impaired scores by using multiple scores from the same test (Vuoksimaa et al., 2020). For example, if delayed recall of a list-learning test is impaired, it is possible that recognition may also be impaired. To avoid this possibility, the current study used 10 separate measures and their corresponding scores. To diagnose MCI, three or more impaired scores were required – and notably a higher cutoff than the two impaired tests required for Jak/Bondi classification. Oltra-Cucarella and colleagues (2018) also only adjusted for age, sex, and education, while the current study additionally adjusted for race/ethnicity. Lastly, compared to our sample of demographically diverse adults, over 90% of ADNI participants are non-Hispanic White and have a mean education corresponding to a four-year university degree. Further study using non-overlapping measures and diverse populations is needed to fully understand the possible utility of this approach.

The Global CR approach (Alfano et al., 2022) demonstrated good sensitivity for incident dementia, possibly supporting the authors' claim that this approach may parallel the profile analysis strategies in the clinical decision-making process (Alfano et al., 2022). However, there was a substantial loss of specificity, resulting in false positives. The authors acknowledged the risk of Type I error due to the relatively high number of comparisons that occur throughout the classification's formulation (Alfano et al., 2022). Important to consider is the scale used by Alfano and colleagues to convert T-scores to their respective clinical ratings – wherein a T-score between 35 and 39 (equivalent to -1.0 to -1.5 SD) is converted to a clinical rating of 5 and labeled as “definite mild impairment,” despite some of these T-scores falling within the “low average” range. Because the Global CR approach places greater importance on CRs that are in the “impaired” range (Alfano et al., 2022), it is possible the current conversion scale may misdiagnose individuals who at a cognitive unimpaired baseline may fall within the “low average” range with “definite mild impairment.” Within the clinical decision-making process are behavioral observations and the interpretation of a low score. In this case, sole reliance on quantitative data and labeling may omit important contextual information for interpreting whether an “abnormal” result truly reflects impairment (Ganguli et al., 2011). That said, because of its high sensitivity to incident dementia, clinicians may still choose to consider using the Global CR approach if the referral question

requires greater sensitivity. Future work could improve this classification by revising the conversion scale to flag performance more appropriately as “impaired.”

To our knowledge, this was the first longitudinal study directly comparing four specific MCI criteria in a diverse community-based sample. Unfortunately, although non-Hispanic Blacks were well-represented, the population of Hispanic older adults or those of other racial/ethnic groups was insufficient to include in these analyses. There was a relatively low number of participants with incident dementia at follow-up and a lower proportion of non-Hispanic Blacks who converted compared to non-Hispanic Whites. Also, we used normative data derived from our study. We previously showed that the EAS sample performs more poorly relative to national normative data even after adjusting for demographic variables (Wang et al., 2021). It will be important to replicate these results in larger, demographically diverse samples.

Study limitations related to our test battery also warrant mention. For example, because we did not have Logical Memory II data available, we included Logical Memory I, which is not a pure measure of episodic memory and can be impacted by attentional and working memory abilities. In addition, to enable us to include a visuospatial domain for the Jak/Bondi criteria, we utilized the Digit Symbol task as a task of visuospatial functioning. While it is typically considered a measure of psychomotor processing speed, previous research has demonstrated that the Digit Symbol test is sensitive to visuospatial and spatial rotation variables (Bigelow et al., 2015; Crowe et al., 1999; Glosser et al., 1977; Shao et al., 2023), and is most strongly related to spatial navigation when compared to other cognitive measures (Wei et al., 2020). However, performance on this task may still be impacted by other domains (e.g., working memory, processing speed, motor function) making it suboptimal as a pure measure of visuospatial functioning.

In sum, although not statistically significant, the Jak/Bondi MCI classification provided the highest balance of sensitivity and specificity for predicting incident dementia across all years of follow-up considered. This classification method significantly decreased false positives while maintaining comparable sensitivity to the Petersen classification. Clinicians and researchers should consider designing their neuropsychological batteries to allow for the utilization of this classification method. Beyond this, clinicians and researchers may have increased confidence in using one classification method over another as they were comparable to Youden's index for predicting incident dementia. This work provides an important step toward improving the generalizability of the MCI diagnosis to underrepresented or health-disparate populations and can inform researchers and clinicians on the most appropriate MCI classification method for their populations or questions of interest. Importantly, this research can lead to meaningful discussions among experts to arrive at an optimal MCI definition that promotes diagnostic accuracy for incident dementia in diverse older adult populations.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1355617724000729>.

Acknowledgments. We thank the EAS participants, investigators, and staff for their contribution to this study.

Financial support. This work was supported by the National Institute on Aging at the National Institutes of Health (grant numbers P01 AG03949, R21 AG056920), the Leonard and Sylvia Marx Foundation, and the Czup Foundation.

Competing interests. The authors report there are no conflict-of-interest disclosures.

References

- Al-Qazzaz, N. K., Ali, S. H., Ahmad, S. A., Islam, S., & Mohamad, K. (2014). Cognitive impairment and memory dysfunction after a stroke diagnosis: A post-stroke memory assessment. *Neuropsychiatric Disease and Treatment, 10*, 1677–1691.
- Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the national institute on aging—Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia, 7*, 270–279.
- Alfano, D. P., Grummisch, J. A., Gordon, J. L., & Hadjistavropoulos, T. (2022). A neuropsychological approach to mild cognitive impairment. *Archives of Clinical Neuropsychology, 37*, 873–890.
- American Psychiatric Association (APA). (1994). Diagnostic and statistical manual of mental disorders (4th ed.). American Psychiatric Publishing, Inc.
- Angevaere, M. J., Vonk, J. M. J., Bertola, L., Zahodne, L., Watson, C. W.-M., Boehme, A., Schupf, N., Mayeux, R., Geerlings, M. I., & Manly, J. J. (2022). Predictors of incident mild cognitive impairment and its course in a diverse community-based population. *Neurology, 98*, e15–e26.
- Arosio, B., Ostan, R., Mari, D., Damanti, S., Ronchetti, F., Arcudi, S., Scurti, M., Franceschi, C., & Monti, D. (2017). Cognitive status in the oldest old and centenarians: A condition crucial for quality of life methodologically difficult to assess. *Mechanisms of Ageing and Development, 165*, 185–194.
- Artero, S., Petersen, R. C., Touchon, J., & Ritchie, K. (2006). Revised criteria for mild cognitive impairment: Validation within a longitudinal population study. *Dementia and Geriatric Cognitive Disorders, 22*, 465–470.
- Bigelow, R. T., Semenov, Y. R., Trevino, C., Ferrucci, L., Resnick, S. M., Simonsick, E. M., Xue, Q. L., & Agrawal, Y. (2015). Association between visuospatial ability and vestibular function in the Baltimore longitudinal study of aging. *Journal of the American Geriatrics Society, 63*, 1837–1844.
- Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology, 24*, 31–46.
- Blackford, R. C., & LaRue, A. (1989). Criteria for diagnosing age-associated memory impairment: Proposed improvements from the field. *Developmental Neuropsychology, 5*, 295–306.
- Blackstone, K., Moore, D. J., Franklin, D. R., Clifford, D. B., Collier, A. C., Marra, C. M., Gelman, B. B., McArthur, J. C., Morgello, S., Simpson, D. M., Ellis, R. J., Atkinson, J. H., Grant, I., & Heaton, R. K. (2012). Defining neurocognitive impairment in HIV: Deficit scores versus clinical ratings. *The Clinical Neuropsychologist, 26*, 894–908.
- Bondi, M. W., Edmonds, E. C., Jak, A. J., Clark, L. R., Delano-Wood, L., McDonald, C. R., Nation, D. A., Libon, D. J., Au, R., Galasko, D., & Salmon, D. P. (2014). Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *Journal of Alzheimer's Disease, 42*, 275–289.
- Bondi, M. W., & Smith, G. E. (2014). Mild cognitive impairment: A concept and diagnostic entity in need of input from neuropsychology. *Journal of the International Neuropsychological Society, 20*, 129–134.
- Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology, 22*, 355–365.
- 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. *Journal of the International Neuropsychological Society, 14*, 463–478.
- 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. *Journal of the International Neuropsychological Society, 13*, 490–500.
- Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: The role of attention and motivation. *Emotion, 12*, 605–617.
- Buschke, H. (1984). Cued recall in Amnesia. *Journal of Clinical Neuropsychology, 6*, 433–440.
- Byrd, D. A., Walden Miller, S., Reilly, J., Weber, S., Wall, T. L., & Heaton, R. K. (2006). Early environmental factors, ethnicity, and adult cognitive test performance. *The Clinical Neuropsychologist, 20*, 243–260.
- Calvillo, M., & Irimia, A. (2020). Neuroimaging and psychometric assessment of mild cognitive impairment after traumatic brain injury. *Frontiers in Psychology, 11*, 1423.
- Chang, K. H., Wang, C., Nester, C. O., Katz, M. J., Byrd, D. A., Lipton, R. B., & Rabin, L. A. (2023). Examining the role of participant and study partner report in widely-used classification approaches of mild cognitive impairment in demographically-diverse community dwelling individuals: Results from the Einstein Aging Study (EAS). *Frontiers in Aging Neuroscience, 15*, 1221768. doi: 10.3389/fnagi.2023.1221768
- Chen, Y., Liang, N., Li, X., Yang, S., Wang, Y., & Shi, N. (2021). Diagnosis and treatment for mild cognitive impairment: A systematic review of clinical practice guidelines and consensus statements. *Frontiers in Neurology, 12*, 719849.
- Crowe, S. F., Benedict, T., Enrico, J., Mancuso, N., Matthews, C., & Wallace, J. (1999). Cognitive determinants of performance on the digit-symbol coding test, and the symbol search test of the WAIS-III and the symbol digit modalities test: An analysis in a healthy sample. *Australian Psychologist, 34*, 204–210.
- de Vent, N. R., Agelink van Rentergem, J. A., Huizenga, H. M., van der Flier, W. M., Sikkes, S. A. M., Murre, J. M. J., van den Bosch, K. A., Scheltens, P., & Schmand, B. A. (2020). An operational definition of abnormal cognition to optimize the prediction of progression to dementia: What are optimal cut-off points for univariate and multivariate normative comparisons? *Journal of Alzheimer's Disease, 77*, 1693–1703.
- Devlin, K. N., Brennan, L., Saad, L., Giovannetti, T., Hamilton, R. H., Wolk, D. A., Xie, S. X., Mechanic-Hamilton, D., & Okonkwo, O. (2022). Diagnosing mild cognitive impairment among racially diverse older adults: Comparison of consensus, actuarial, and statistical methods. *Journal of Alzheimer's Disease, 85*, 627–644.
- Díaz-Venegas, C., Downer, B., Langa, K. M., & Wong, R. (2016). Racial and ethnic differences in cognitive function among older adults in the USA. *International Journal of Geriatric Psychiatry, 31*, 1004–1012.
- Donders, J. (2020). The incremental value of neuropsychological assessment: A critical review. *The Clinical Neuropsychologist, 34*, 56–87.
- Dubois, B., Villain, N., Frisoni, G. B., Rabinovici, G. D., Sabbagh, M., Cappa, S., Bejanin, A., Bombois, S., Epelbaum, S., Teichmann, M., Habert, M.-O., Nordberg, A., Blennow, K., Galasko, D., Stern, Y., Rowe, C. C., Salloway, S., Schneider, L. S., Cummings, J. L., & Feldman, H. H. (2021). Clinical diagnosis of Alzheimer's disease: Recommendations of the international working group. *The Lancet Neurology, 20*, 484–496.
- Dunne, R. A., Aarsland, D., O'Brien, J. T., Ballard, C., Banerjee, S., Fox, N. C., Isaacs, J. D., Underwood, B. R., Perry, R. J., Chan, D., Denning, T., Thomas, A. J., Schryer, J., Jones, A.-M., Evans, A. R., Alessi, C., Coulthard, E. J., Pickett, J., Elton, P., . . . Burns, A. (2021). Mild cognitive impairment: The Manchester consensus. *Age and Ageing, 50*, 72–80.
- Edmonds, E. C., Delano-Wood, L., Clark, L. R., Jak, A. J., Nation, D. A., McDonald, C. R., Libon, D. J., Au, R., Galasko, D., Salmon, D. P., Bondi, M. W., & Alzheimer's Disease Neuroimaging Initiative (2015a). Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia, 11*, 415–424.
- Edmonds, E. C., Smirnov, D. S., Thomas, K. R., Graves, L. V., Bangen, K. J., Delano-Wood, L., Galasko, D. R., Salmon, D. P., & Bondi, M. W. (2021). Data-driven vs consensus diagnosis of MCI: Enhanced sensitivity for detection of clinical, biomarker, and neuropathologic outcomes. *Neurology, 97*, e1288–e1299.
- Edmonds, E. C., Weigand, A. J., Thomas, K. R., Eppig, J., Delano-Wood, L., Galasko, D. R., Salmon, D. P., & Bondi, M. W. (2018). Increasing inaccuracy of self-reported subjective cognitive complaints over 24 months in empirically derived subtypes of mild cognitive impairment. *Journal of the International Neuropsychological Society, 24*, 842–853.
- Facer-Childs, E., Boiling, S., & Balanos, G. (2018). The effects of time of day and chronotype on cognitive and physical performance in healthy volunteers. *Sports Medicine, 48*, 47. doi: 10.1186/s40798-018-0162-z

- Ganguli, M., Snitz, B. E., Saxton, J. A., Chang, C.-C. H., Lee, C.-W., Vander Bilt, J., Hughes, T. F., Loewenstein, D. A., Unverzagt, F. W., & Petersen, R. C. (2011). Outcomes of mild cognitive impairment by definition: A population study. *Archives of Neurology*, *68*, 761–767.
- Gasquoine, P. G. (2009). Race-norming of neuropsychological tests. *Neuropsychology Review*, *19*, 250–262.
- Giau, V. V., Bagyinszky, E., & An, S. S. (2019). Potential fluid biomarkers for the diagnosis of mild cognitive impairment. *International Journal of Molecular Sciences*, *20*, 4149.
- Glosser, G., Butters, N., & Kaplan, E. (1977). Visuo-perceptual processes in brain damaged patients on the digit symbol substitution test. *International Journal of Neuroscience*, *7*, 59–66.
- Gomersall, T., Astell, A., Nygard, L., Sixsmith, A., Milhailidis, A., & Hwang, A. (2015). Living with ambiguity: A metasynthesis of qualitative research on mild cognitive impairment. *The Gerontologist*, *55*, 892–912.
- Graves, L. V., Edmonds, E. C., Thomas, K. R., Weigand, A. J., Cooper, S., & Bondi, M. W. (2020). Evidence for the utility of actuarial neuropsychological criteria across the continuum of normal aging, mild cognitive impairment, and dementia. *Journal of Alzheimer's Disease*, *37*, 1–385.
- Graves, L. V., Edmonds, E. C., Thomas, K. R., Weigand, A. J., Cooper, S., Stickel, A. M., Zlatar, Z. Z., Clark, A. L., & Bondi, M. W. (2022). Diagnostic accuracy and differential associations between ratings of functioning and neuropsychological performance in non-hispanic Black and White older adults. *The Clinical Neuropsychologist*, *36*, 287–310.
- Hackett, K., Mis, R., Drabick, D. A., & Giovannetti, T. (2020). Informant reporting in mild cognitive impairment: Sources of discrepancy on the functional activities questionnaire. *Journal of the International Neuropsychological Society*, *26*, 503–514.
- Hansson, O., Edelmayer, R. M., Boxer, A. L., Carrillo, M. C., Mielke, M. M., Rabinovici, G. D., Salloway, S., Sperling, R., Zetterberg, H., & Teunissen, C. E. (2022). The Alzheimer's association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimer's & Dementia*, *18*, 2669–2686.
- Heaton, R. K., Grant, I., Butters, N., White, D. A., Kirson, D., Atkinson, J. H., McCutchan, J. A., Taylor, M. J., Kelly, M. D., Ellis, R. J., Wolfson, T., Velin, R., Marcotte, T. D., Hesselink, J. R., Jernigan, T. L., Chandler, J., Wallace, M., Abramson, I., & The HNRC group (1995). The HNRC 500-Neuropsychology of HIV infection at different disease stages. *Journal of the International Neuropsychological Society*, *1*, 231–251.
- Hong, T. B., Zarit, S. H., & Johansson, B. (2003). Mild cognitive impairment in the oldest old: A comparison of two approaches. *Aging & Mental Health*, *7*, 271–276.
- Ilardi, C. R., Chieffo, S., Iachini, T., & Iavarone, A. (2021). Neuropsychology of posteromedial parietal cortex and conversion factors from mild cognitive impairment to Alzheimer's disease: Systematic search and state-of-the-art review. *Aging Clinical and Experimental Research*, *34*, 289–307.
- Jack C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., . . . , Silverberg, N. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*, 535–562.
- 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. *The American Journal of Geriatric Psychiatry*, *17*, 368–375.
- Jak, A. J., Preis, S. R., Beiser, A. S., Seshadri, S., Wolf, P. A., Bondi, M. W., & Au, R. (2016). Neuropsychological criteria for mild cognitive impairment and dementia risk in the Framingham heart study. *Journal of the International Neuropsychological Society*, *22*, 937–943.
- Jessen, F., Amariglio, R. E., van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., van der Flier, W. M., Glodzik, L., van Harten, A. C., de Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., . . . , Subjective Cognitive Decline Initiative (SCD-I) Working Group (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, *10*, 844–852.
- Kaduskiewicz, H., Eisele, M., Wiese, B., Prokein, J., Luppa, M., Luck, T., Jessen, F., Bickel, H., Mosch, E., Pentzek, M., Fuchs, A., Eifflaender-Gorfer, S., Weyerer, S., König, H.-H., Bretschneider, C., van den Bussche, H., Maier, W., Scherer, M., Riedel-Heller, S. G., & Study on Aging, Cognition, and Dementia in Primary Care Patients (AgeCoDe) Study Group (2014). Prognosis of mild cognitive impairment in general practice: Results of the German AgeCoDe study. *The Annals of Family Medicine*, *12*, 158–165.
- Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Lea & Febiger.
- Katz, M. J., Lipton, R. B., Hall, C. B., Zimmerman, M. E., Sanders, A. E., Verghese, J., Dickson, D. W., & Derby, C. A. (2012). Age-specific and sex-specific prevalence and incidence of mild cognitive impairment, dementia, and Alzheimer dementia in Blacks and Whites. *Alzheimer Disease & Associated Disorders*, *26*, 335–343.
- Kravitz, E., Schmeidler, J., & Schnaider Beeri, M. (2012). Cognitive decline and dementia in the oldest-old. *Rambam Maimonides Medical Journal*, *3*, e0026.
- Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*, 179–186.
- Loewenstein, D. A., Acevedo, A., Small, B. J., Agron, J., Crocco, E., & Duara, R. (2009). Stability of different subtypes of mild cognitive impairment among the elderly over a 2- to 3-year follow-up period. *Dementia and Geriatric Cognitive Disorders*, *27*, 418–423.
- Malaspina, L., Woods, S. P., Moore, D. J., Depp, C., Letendre, S. L., Jeste, D., Grant, I., & HIV Neurobehavioral Research Programs (HNRP) Group. (2011). Successful cognitive aging in persons living with HIV infection. *Journal of NeuroVirology*, *17*, 110–119.
- Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist*, *19*, 270–275.
- Manly, J. J., & Echemendia, R. J. (2007). Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Archives of Clinical Neuropsychology*, *22*, 319–325.
- Manly, J. J., Tang, M. X., Schupf, N., Stern, Y., Vonsattel, J. G., & Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Annals of Neurology*, *63*, 494–506.
- Marvanova, M. (2016). Drug-induced cognitive impairment: Effect of cardiovascular agents. *Mental Health Clinician*, *6*, 201–206.
- Moore, E., Mander, A., Ames, D., Carne, R., Sanders, K., & Watters, D. (2012). Cognitive impairment and vitamin B12: A review. *International Psychogeriatrics*, *24*, 541–556.
- Mormino, E. C., Papp, K. V., Perry, G., Avila, J., Moreira, P. I., Sorensen, A. A., & Tabaton, M. (2018). Amyloid accumulation and cognitive decline in clinically normal older adults: Implications for aging and early Alzheimer's disease. *Journal of Alzheimer's Disease*, *64*, S633–S646.
- Oltra-Cucarella, J., Sánchez-SanSegundo, M., Lipnicki, D. M., Sachdev, P. S., Crawford, J. D., Pérez-Vicente, J. A., Cabello-Rodríguez, L., Ferrer-Cascales, R., & Alzheimer's Disease Neuroimaging Initiative (2018). Using base rate of low scores to identify progression from amnesic mild cognitive impairment to Alzheimer's disease. *Journal of the American Geriatrics Society*, *66*, 1360–1366.
- Pandya, S. Y., Lactriz, L. H., Weiner, M. F., Deschner, M., & Woon, F. L. (2017). Predictors of reversion from mild cognitive impairment to normal cognition. *Dementia and Geriatric Cognitive Disorders*, *43*, 204–214.
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*, 183–194.
- Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., Jack C. R., Jagust, W. J., Shaw, L. M., Toga, A. W., Trojanowski, J. Q., & Weiner, M. W. (2010). Alzheimer's disease neuroimaging initiative (ADNI): Clinical characterization. *Neurology*, *74*, 201–209.
- Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. *Journal of Internal Medicine*, *275*, 214–228.
- Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., Ritchie, K., Rosser, M., Thal, L., & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*, 1985–1992.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*, 303–308.

- Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, 37, 323–329.
- Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, 8, 271–276.
- Ritchie, K., Artero, S., & Touchon, J. (2001). Classification criteria for mild cognitive impairment. *Neurology*, 56, 37–42.
- Rosen, W. G. (1980). Verbal fluency in aging and dementia. *Journal of Clinical Neuropsychology*, 2, 135–146.
- Scharre, D. W. (2019). Preclinical, prodromal, and dementia stages of Alzheimer's Disease. *Practical Neurology*, 36–47.
- Shao, X., Wang, K., Zhang, Y., Zhen, X., Dong, F., Tian, H., & Yu, Y. (2023). Outcome of visuospatial dysfunction assessment in patients with Parkinson's disease using mobile application software. *Frontiers in Aging Neuroscience*, 15, 1108166.
- Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric depression scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontologist*, 5, 165–173.
- Spreen, O., & Strauss, E. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press.
- Thomas, K. R., Cook, S. E., Bondi, M. W., Unverzagt, F. W., Gross, A. L., Willis, S. L., & Marsiske, M. (2020). Application of neuropsychological criteria to classify mild cognitive impairment in the ACTIVE study. *Neuropsychology*, 34, 862–873.
- Thomas, K. R., Eppig, J. S., Weigand, A. J., Edmonds, E. C., Wong, C. G., Jak, A. J., Delano-Wood, L., Galasko, D. R., Salmon, D. P., Edland, S. D., Bondi, M. W., & Alzheimer's Disease Neuroimaging Initiative. (2019). MCI-to-normal reversion using neuropsychological criteria in the Alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia*, 15, 561–569.
- Tibhirani, R. J., & Efron, B. (1993). *An introduction to the bootstrap*. Chapman and Hall/CRC.
- Trittschuh, E. H., Crane, P. K., Larson, E. B., Cholerton, B., McCormick, W. C., McCurry, S. M., Bowen, J. D., Baker, L. D., & Craft, S. (2011). Effects of varying diagnostic criteria on prevalence of mild cognitive impairment in a community sample. *Journal of Alzheimer's Disease*, 25, 163–173.
- Visser, L. N. C., van Maurik, I. S., Bouwman, F. H., Staekenborg, S., Vreeswijk, R., Hempenius, L., de Beer, M. H., Roks, G., Boelaarts, L., Kleijer, M., van der Flier, W. M., Smets, E. M. A., & Ginsberg, S. D. (2020). Clinicians' communication with patients receiving a MCI diagnosis: The ABIDE project. *PLOS ONE*, 15, e0227282.
- von Stumm, S. (2016). Is day-to-day variability in cognitive function coupled with day-to-day variability in affect? *Intelligence*, 55, 1–6.
- Vuoksimaa, E., McEvoy, L. K., Holland, D., Franz, C. E., Kremen, W. S., & Alzheimer's Disease Neuroimaging Initiative (2020). Modifying the minimum criteria for diagnosing amnesic MCI to improve prediction of brain atrophy and progression to Alzheimer's disease. *Brain Imaging and Behavior*, 14, 787–796.
- Wang, C., Katz, M. J., Chang, K. H., Qin, J., Lipton, R. B., Zwerling, J. L., Sliwinski, M. J., Derby, C. A., Rabin, L. A., & Abner, E. (2021). UDSNB 3.0 neuropsychological test norms in older adults from a diverse community: Results from the Einstein Aging Study (EAS). *Journal of Alzheimer's Disease*, 83, 1665–1678.
- Wechsler, D. (1987). *WMS-R: Wechsler memory scale-revised*. Psychological Corporation.
- Wechsler, D. (1997). *WAIS-III: Administration and scoring manual: Wechsler Adult Intelligence Scale*. Psychological Corporation.
- Wei, E. X., Anson, E. R., Resnick, S. M., & Agrawal, Y. (2020). Psychometric tests and spatial navigation: Data from the Baltimore longitudinal study of aging. *Frontiers in Neurology*, 11, 484.
- Werry, A. E., Daniel, M., & Bergström, B. (2019). Group differences in normal neuropsychological test performance for older non-Hispanic White and Black/African American adults. *Neuropsychology*, 33, 1089–1100.
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., Nordberg, A., Bäckman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., De Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., & Petersen, R. C. (2004). Mild cognitive impairment-beyond controversies, towards a consensus: Report of the international working group on mild cognitive impairment. *Journal of Internal Medicine*, 256, 240–246.
- Wong, C. G., Thomas, K. R., Edmonds, E. C., Weigand, A. J., Bangen, K. J., Eppig, J. S., Jak, A. J., Devine, S. A., Delano-Wood, L., Libon, D. J., Edland, S. D., Au, R., & Bondi, M. W. (2018). Neuropsychological criteria for mild cognitive impairment in the Framingham heart study's old-old. *Dementia and Geriatric Cognitive Disorders*, 46, 253–265.
- Zahodne, L. B., Manly, J. J., Azar, M., Brickman, A. M., & Glymour, M. M. (2016). Racial disparities in cognitive performance in mid- and late adulthood: Analyses of two cohort studies. *Journal of the American Geriatrics Society*, 64, 959–964.
- Zahodne, L. B., Sharifian, N., Kraal, A. Z., Zaheed, A. B., Sol, K., Morris, E. P., Schupf, N., Manly, J. J., & Brickman, A. M. (2021). Socioeconomic and psychosocial mechanisms underlying racial/ethnic disparities in cognition among older adults. *Neuropsychology*, 35, 265–275.