Title

Data-driven classification of cognitively normal and mild cognitive impairment subtypes predicts progression in the NACC dataset

Journal

Alzheimer's & Dementia, 20(5)

ISSN

1552-5260

Authors

Edmonds, Emily C

Thomas, Kelsey R

Rapcsak, Steven Z

et al.

Publication Date

2024-05-01

DOI

10.1002/alz.13793

Peer reviewed

Received: 14 June 2023 | Revised: 20 October 2023 | Accepted: 23 February 2024

DOI: 10.1002/alz.13793

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Data-driven classification of cognitively normal and mild cognitive impairment subtypes predicts progression in the NACC dataset

Emily C. Edmonds[1,2] | **Kelsey R. Thomas**[3,4] | **Steven Z. Rapcsak**[1,2,5] | **Shannon L. Lindemer**[1] | **Lisa Delano-Wood**[4,6] | **David P. Salmon**[7] | **Mark W. Bondi**[4,6]

[1]Banner Alzheimer's Institute, Tucson, Arizona, USA

[2]Departments of Neurology and Psychology, University of Arizona, Tucson, Arizona, USA

[3]Research Service, Veterans Affairs San Diego Healthcare System, San Diego, California, USA

[4]Department of Psychiatry, University of California, San Diego, La Jolla, California, USA

[5]Department of Speech, Language, & Hearing Sciences, University of Arizona, Tucson, Arizona, USA

[6]Psychology Service, Veterans Affairs San Diego Healthcare System, San Diego, California, USA

[7]Department of Neurosciences, University of California, San Diego, La Jolla, California, USA

Correspondence
Emily C. Edmonds, Banner Alzheimer's Institute, 2626 E River Rd., Tucson, AZ 85718, USA.
Email: eedmonds@arizona.edu

Funding information
NIA/NIH, Grant/Award Numbers: P30 AG062429, P30 AG066468, P30 AG062421, P30 AG066509, P30 AG066514, P30 AG066530, P30 AG066507, P30 AG066444, P30 AG066518, P30 AG066512, P30 AG066462, P30 AG072979, P30 AG072972, P30 AG072976, P30 AG072975, P30 AG072978, P30 AG072977, P30 AG066519, P30 AG062677, P30 AG079280, P30 AG062422, P30 AG066511, P30 AG072946, P30 AG062715, P30 AG072973, P30 AG066506, P30 AG066508, P30 AG066515, P30 AG072947, P30 AG072931, P30 AG066546, P20 AG068024, P20 AG068053, P20 AG068077, P20 AG068082, P30 AG072958, P30 AG072959, P30 AG062429, R03 AG070435, RF1 AG082726; Arizona Alzheimer's Consortium and Arizona DHS; U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service, Grant/Award Number: 1IK2CX001865

Abstract

INTRODUCTION: Data-driven neuropsychological methods can identify mild cognitive impairment (MCI) subtypes with stronger associations to dementia risk factors than conventional diagnostic methods.

METHODS: Cluster analysis used neuropsychological data from participants without dementia (mean age = 71.6 years) in the National Alzheimer's Coordinating Center (NACC) Uniform Data Set (n = 26,255) and the "normal cognition" subsample (n = 16,005). Survival analyses examined MCI or dementia progression.

RESULTS: Five clusters were identified: "Optimal" cognitively normal (oCN; 13.2%), "Typical" CN (tCN; 28.0%), Amnestic MCI (aMCI; 25.3%), Mixed MCI-Mild (mMCI-Mild; 20.4%), and Mixed MCI-Severe (mMCI-Severe; 13.0%). Progression to dementia differed across clusters (oCN < tCN < aMCI < mMCI-Mild < mMCI-Severe). Cluster analysis identified more MCI cases than consensus diagnosis. In the "normal cognition" subsample, five clusters emerged: High-All Domains (High-All; 16.7%), Low-Attention/Working Memory (Low-WM; 22.1%), Low-Memory (36.3%), Amnestic MCI (16.7%), and Non-amnestic MCI (naMCI; 8.3%), with differing progression rates (High-All < Low-WM = Low-Memory < aMCI < naMCI).

DISCUSSION: Our data-driven methods outperformed consensus diagnosis by providing more precise information about progression risk and revealing heterogeneity in cognition and progression risk within the NACC "normal cognition" group.

KEYWORDS
Alzheimer's disease, cluster analysis, cognitive subtypes, dementia, mild cognitive impairment, neuropsychology, preclinical Alzheimer's disease, subtle cognitive decline

1 | BACKGROUND

Mild cognitive impairment (MCI) is typically diagnosed based on the presence of a subjective cognitive complaint, objective impairment on a cognitive test, and essentially normal day-to-day functioning.[1,2] A "consensus diagnosis" approach is often applied in which several experts use subjective and objective assessments to arrive at a diagnostic impression based on the above criteria in the context of information about a participant's background. The consensus approach is considered the "gold standard" method for MCI diagnosis and is widely used by MCI research studies, including the National Institute on Aging (NIA)-funded Alzheimer's Disease Research Centers (ADRC) across the United States.[3–5] However, this method has limitations such as reliance on clinical judgment, which can vary across clinicians, time points, and sites.

Previous work has shown that using objective, data-driven statistical methods for identifying MCI based on comprehensive neuropsychological test data is an alternative approach that can reliably identify subtypes of MCI. This data-driven approach identifies MCI groups that show stronger associations among cognition, Alzheimer's disease (AD) biomarkers, and risk for dementia than do groups based on conventional diagnostic methods.[6–13] It has also been shown that objective methods, including those that combine cognitive test scores with AD biomarkers, outperform the clinical judgment of memory clinic physicians in predicting risk of developing AD dementia.[14]

We recently applied a data-driven cluster analysis approach to baseline neuropsychological test data from > 700 non-demented older adult participants from the University of California San Diego (UCSD) ADRC, and identified five cognitive subgroups: "optimal" cognitively normal with above-average cognition, "typical" cognitively normal with average cognition, non-amnestic MCI, amnestic MCI, and mixed MCI.[15] Progression to dementia within the next 6 years (on average) differed across the three MCI subtypes with mixed MCI showing the highest rate of progression (mixed > amnestic > non-amnestic). Mixed MCI also had the highest prevalence of cerebrospinal fluid (CSF) biomarker positivity during life and AD pathology at autopsy. Our data-driven approach to identifying MCI outperformed consensus diagnoses in capturing those who had abnormal biomarkers, progressed to dementia, or had AD pathology at autopsy.[15]

To examine cognitive heterogeneity within pre-MCI, we also applied our objective methods to comprehensive neuropsychological test data from 365 participants in the UCSD ADRC sample who were determined to have "normal cognition" based on consensus diagnosis. Cluster analysis of neuropsychological test scores identified four subgroups of participants, including three with subtle cognitive decline who had low scores in memory/language, executive, and/or visuospatial domains, and a cognitively normal group with average performance across all cognitive domains assessed (i.e., an All-Average group). Rates of cognitive decline and progression to MCI/dementia were steeper in the subtle cognitive decline groups (Low-All Domains > Low Memory/Language ≥ Low-Visuospatial and Low-Executive) than the All-Average group.[16]

The present study extends this work from a single ADRC to the larger National Alzheimer's Coordinating Center (NACC) sample by applying cluster analysis to baseline neuropsychological test data for all NACC participants without dementia, and to a subsample limited to those classified as "normal cognition." Based on previous findings,[15,16] we hypothesized that cognitive subtypes of MCI and subtle cognitive impairment would emerge that would be predictive of subsequent risk of progression to MCI/dementia.

2 | METHODS

2.1 | Participants

Participants were 26,255 individuals aged ≥ 50 (mean age = 71.6 [standard deviation (SD) = 8.9]; mean education = 15.7 [SD = 3.1]; 60.0% female; 77.7% White; 92.4% non-Hispanic) with neuropsychological test scores in the NACC Uniform Data Set (UDS).[17,18] Baseline data were collected from 2005 to 2022 across 46 ADRC study sites. Written informed consent to participate in the study was obtained from all participants or their caregivers at each individual ADRC, as approved by individual institutional review boards (IRBs) at each site; the current study was approved by the Banner Health IRB. Inclusion criteria for enrollment in an ADRC include stable health status with no history of major stroke, neurologic disorders, severe psychiatric illness, substance abuse, or learning disability. For the current study, we excluded participants with a diagnosis of dementia at baseline, as determined by NACC via consensus diagnosis and National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria.[19,20]

2.2 | Diagnostic and neuropsychological procedures

Participants completed annual clinical, neurological, and neuropsychological evaluations as part of the ADRC/NACC research protocol. A diagnosis of normal cognition, impaired-not-MCI (i.e., used for participants whose presentation did not clearly fit into the normal cognition or MCI categories), MCI, or dementia was determined at baseline and at each subsequent annual visit by the consensus of a multidisciplinary team at each site. (One caveat is that some data in NACC may

include diagnostic decisions that did not result from a consensus conference process. Per the NACC website: Depending on a given ADRC's protocol, diagnosis is made by either a consensus team or a single physician [the one who conducted the examination]). For participants diagnosed with a cognitive disorder, the presumptive primary etiologic diagnosis and any contributing conditions are specified in the NACC database.

The UDS neuropsychological tests examined in the current study included measures of memory (Immediate and Delayed Recall from Logical Memory or Craft Story), attention/working memory (Forward and Backward Digit Span or Number Span), processing speed/executive functioning (Trail Making Test, Parts A and B), and language (Category Fluency [animals, vegetables], Boston Naming Test [BNT] or Multilingual Naming Test [MINT]).

Raw scores on each of these measures were converted into demographical adjusted (age, education, sex) *z* scores based on regression coefficients derived from performance of a subset of the NACC sample that we identified as "robust CN" participants ($n = 9742$). The robust CN group was defined as participants who had at least 2 years of data available and who remained classified as "normal cognition" by consensus diagnosis for the duration of their participation in the longitudinal study (mean follow-up = 5.4 years [SD = 3.4]). The robust CN sample was generally well matched at baseline (mean age = 70.3 [SD = 8.8]; mean education = 16.0 [SD = 2.9]; 66.7% female; 78.3% White; 93.7% non-Hispanic) to the full study sample.

The NACC UDS Neuropsychological Battery Crosswalk Study[21] found good correlation between different versions of tests that changed from UDS version 2 to version 3. Thus, for the current study, *z* scores were generated separately for each test and then combined for corresponding tests (i.e., Logical Memory/Craft Story, Digit Span/Number Span, and BNT/MINT). The other tests (Trail Making Test and Category Fluency) were administrated across UDS versions.

2.3 | Statistical analysis

Cluster analysis of baseline neuropsychological data was conducted by entering *z* scores into a hierarchical cluster analysis using the Ward method, consistent with our previous work.[6,10,15,16] Analysis of variance, Kruskal–Wallis tests, or chi-square tests compared cluster-derived groups on demographic characteristics, clinical and cognitive measures, apolipoprotein E (*APOE*) ε4 status, rate of progression, and year of dementia diagnosis. Bonferroni correction was used to account for multiple comparisons between cluster groups. A proportional hazards model for progression to a diagnosis of dementia or MCI/dementia was carried out using a Cox regression model that adjusted for demographics (age, education, sex, race, ethnicity). Kaplan–Meier curves were used to depict the rate of progression over time by cluster group, and survival curves were compared using a log-rank test with pairwise comparisons. Participants who did not progress during their follow-up period were censored at their last visit. Chi-square analysis was used to examine the presumptive primary etiology in those who progressed to dementia.

RESEARCH IN CONTEXT

1. **Systematic review**: The authors searched PubMed for studies related to cognitive subtypes in the National Alzheimer's Coordinating Center (NACC) Uniform Data Set (UDS). Results revealed that previous studies have identified subgroups of mild cognitive impairment (MCI) and cognitively normal in either single Alzheimer's Disease Research Center samples or in smaller subsets of the NACC UDS sample.
2. **Interpretation**: We extend this work by applying cluster analysis to baseline neuropsychological data from all NACC participants without dementia, and to those classified as "normal cognition." Our data-driven method outperformed the consensus diagnostic approach by providing more precise information about risk for future MCI/dementia, and by revealing heterogeneity within the NACC normal cognition group.
3. **Future directions**: Results have implications for future research by demonstrating a method to identify empirically derived subtypes of MCI and subtle cognitive decline that optimize prediction of progression. Continued research incorporating Alzheimer's disease biomarkers is needed to further determine the utility of data-driven diagnoses in diverse samples.

3 | RESULTS

3.1 | Neuropsychological cluster-derived groups

Results of the cluster analysis on the full sample of 26,255 participants revealed five cognitive subgroups: (1) Optimal CN (oCN; $n = 3465$; 13.2%) with above-average to average cognition in all domains examined; (2) Typical CN (tCN; $n = 7358$; 28.0%) with average cognition across domains; (3) Amnestic MCI (aMCI; $n = 6649$; 25.3%) with isolated low memory performance; (4) Mixed MCI-Mild (mMCI-Mild; $n = 5363$; 20.4%) with low performance across domains; and (5) Mixed MCI-Severe (mMCI-Severe; $n = 3420$; 13.0%) with more severe multidomain impairment including significant executive dysfunction; see Figure 1.

There were demographic differences among the five cluster groups; see Table 1. The oCN and tCN groups were significantly younger than the three MCI groups, and the mMCI-Severe group was the oldest of all five groups. The groups also differed in years of education, as the tCN group had a significantly higher level of education relative to all groups except mMCI-Mild, and the mMCI-Severe group the lowest level of education relative to all groups. The oCN group had more women relative to all other groups, and the tCN group had more women relative to the mMCI-Mild group. The proportion of non-White participants was lowest in the oCN group and increased across the declining cognitive

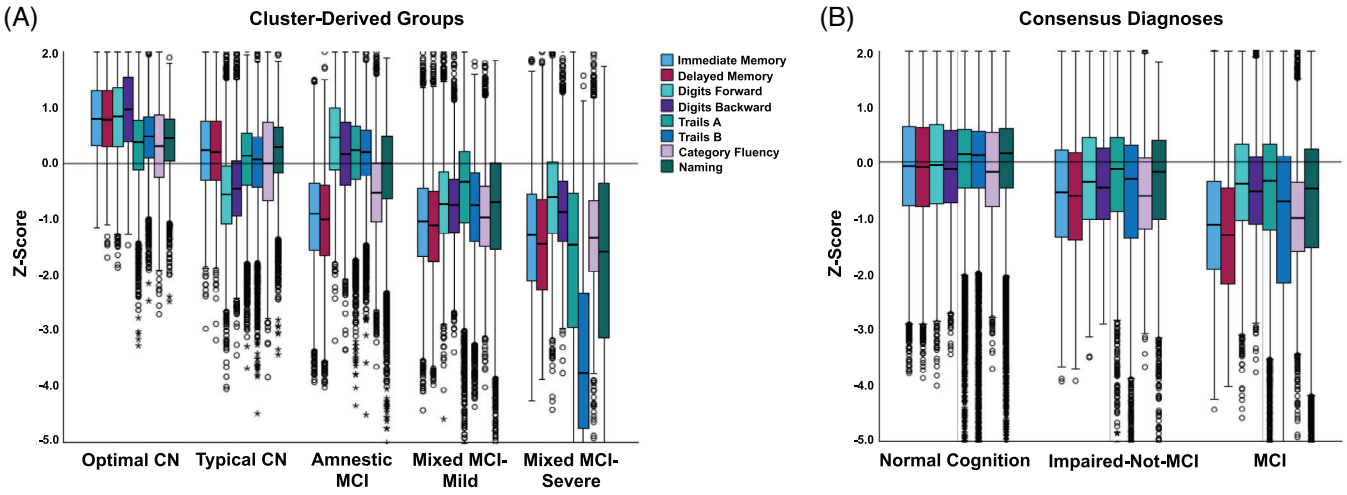

FIGURE 1 Baseline neuropsychological performance of the (A) cluster-derived groups, and (B) NACC consensus diagnostic groups. Error bars denote standard error of the mean. CN, cognitively normal; MCI, mild cognitive impairment; NACC, National Alzheimer's Coordinating Center.

continuum, with the exception of similar proportions in the tCN and aMCI groups. The proportion of Hispanic participants also increased across the groups.

APOE genetic biomarkers were available for 77.7% of the sample. The oCN and tCN groups had the fewest number of *APOE* ε4 carriers relative to the three MCI groups which did not differ from one another. Baseline Clinical Dementia Rating (CDR) scores (available for 100% of the sample) differed significantly among all groups and were higher (worse) across the declining cognitive continuum. Scores on the Mini-Mental State Examination (MMSE; available for 68.6% of the sample) and Montreal Cognitive Assessment (MoCA; available for 31.2% of the sample) declined across the groups, while depressive symptoms on the Geriatric Depression Scale (GDS-15 item; available for 98.7% of the sample) increased. Scores on the Functional Assessment Questionnaire (FAQ; available for 77.0% of the sample) increased (indicating greater functional difficulty) across the groups.

3.2 | Progression to dementia

Of the 26,255 participants at baseline, 3784 (14.4%) progressed to a consensus diagnosis of dementia after an average of 4.3 years of follow-up (SD = 2.8, range 2–17). Cox regression using tCN as the reference group and adjusting for demographics (age, education, sex, race, ethnicity) showed an increased risk of progression to dementia in the aMCI (hazard ratio [HR] = 2.42, 95% confidence interval [CI; 2.18, 2.68], $P < 0.001$), mMCI-Mild (HR = 4.04, 95% CI [3.63, 4.48], $P < 0.001$), and mMCI-Severe (HR = 9.14, 95% CI [8.18, 10.21], $P < 0.001$) groups. There was a decreased risk of progression in the oCN group (HR = 0.61, 95% CI [0.52, 0.72], $P < 0.001$) relative to tCN.

Kaplan–Meier curves depicting rate of progression to dementia over time by group are shown in Figure 2. A log-rank test revealed significant group differences in survival curves ($\chi^2[4] = 2677.94$; $P < 0.001$), with pairwise comparisons showing that all five groups differed significantly from one another (*Ps* < 0.001). The mMCI-Severe group had the steepest survival curve (i.e., fastest rate of progression), followed by mMCI-Mild, aMCI, tCN, and oCN.

Regarding type of dementia, 79.6% ($n = 3012$) of the participants who progressed were presumed to have AD as the primary etiology. The aMCI group had the highest rate of progression to AD dementia (significantly higher than the tCN and mMCI-Severe groups); see Table 1.

The remaining 20.4% ($n = 772$) of participants who progressed were presumed to have a primary etiology of non-AD. The most common was Lewy body disease (LBD; 6.2%), followed by vascular disease (4.8%); frontotemporal dementia (FTD; 3.2%); and other neurological, medical, psychiatric, or undetermined causes (6.6%). The mMCI-Severe group had a higher rate of dementia due to LBD relative to the aMCI and mMCI-Mild groups, and a higher rate of FTD relative to the aMCI group. The tCN group had a higher rate of progression to vascular dementia relative to the aMCI and mMCI-Severe groups.

Mixed etiologies were common. In participants with a primary etiology of AD ($n = 3012$), one or more secondary etiologies were present in 29.0% of cases, the most common being psychiatric conditions (14.8%), vascular disease (7.3%), systemic/medical illness (2.9%), and LBD (2.6%). In participants with a primary non-AD etiology ($n = 772$), one or more secondary etiologies were present in 46.6% of cases, the most common being AD (21.9%), psychiatric conditions (20.3%), systemic/medical illness (4.0%), and vascular disease (3.5%).

3.3 | Progression to dementia in non-White participants

Given that race differed across the cluster-derived groups, with a higher proportion of Black participants in particular in the MCI subgroups, we conducted subanalyses with only non-White participants ($n = 5558$; 69.9% Black; 11.3% Asian; 15.7% multiracial; 3.0% other). In this subsample, 505 (9.1%) progressed to a consensus diagnosis of dementia after an average of 4.2 years post-baseline (SD = 2.7,

TABLE 1 Baseline demographic and clinical characteristics of the cluster-derived groups in the full sample.

	Optimal CN ($n = 3465$)	Typical CN ($n = 7358$)	Amnestic MCI ($n = 6649$)	Mixed MCI-Mild ($n = 5363$)	Mixed MCI-Severe ($n = 3420$)	F, H, or X^2	Effect size	*P*
Demographic variables								
Age, years	70.96 (8.62)	70.84 (8.61)	71.50 (8.82)	71.77 (8.92)	74.02 (9.20)	F = 83.44	$\eta_p^2 = 0.01$	$P < 0.001$
Education, years	15.74 (2.74)	15.96 (2.79)	15.78 (2.78)	15.85 (3.17)	14.33 (3.89)	F = 193.28	$\eta_p^2 = 0.03$	$P < 0.001$
Sex: female, %	65.4%	60.5%	59.5%	57.3%	58.4%	$X^2 = 63.21$	$\varphi_c = 0.05$	$P < 0.001$
Race, %:						$X^2 = 1762.52$	$\varphi_c = 0.13$	$P < 0.001$
White	89.4%	83.4%	83.5%	68.3%	57.3%			
Black or African American	6.7%	10.4%	11.1%	21.3%	29.4%			
Amer. Indian or Alaska Native	0.1%	0.6%	0.5%	0.9%	0.6%			
Native Hawaiian or Pacific Islander	0.1%	0.0%	0.0%	0.2%	0.1%			
Asian	1.1%	1.7%	1.9%	3.9%	4.0%			
Multiracial	2.3%	3.0%	2.5%	4.2%	5.3%			
Unknown	0.3%	0.9%	0.5%	1.3%	3.2%			
Ethnicity: Hispanic, %	2.0%	6.0%	3.9%	10.7%	16.1%	$X^2 = 757.37$	$\varphi_c = 0.17$	$P < 0.001$
Clinical variables[a]								
APOE ε4 carrier, %	28.9%	31.2%	38.0%	39.7%	38.4%	$X^2 = 154.94$	$\varphi_c = 0.09$	$P < 0.001$
CDR Global	0.08 (0.18)	0.12 (0.22)	0.21 (0.25)	0.27 (0.26)	0.36 (0.24)	H = 3453.98	$\eta^2 = 0.13$	$P < 0.001$
CDR Sum of Boxes	0.17 (0.51)	0.29 (0.63)	0.56 (0.91)	0.79 (1.05)	1.18 (0.93)	H = 4000.93	$\eta^2 = 0.15$	$P < 0.001$
MMSE	29.35 (0.94)	28.84 (1.37)	28.49 (1.67)	27.80 (1.96)	26.33 (2.68)	H = 3419.73	$\eta^2 = 0.19$	$P < 0.001$
MoCA	27.38 (2.06)	26.21 (2.51)	25.30 (2.75)	23.44 (3.24)	20.89 (3.64)	H = 2339.15	$\eta^2 = 0.29$	$P < 0.001$
GDS	1.31 (1.95)	1.50 (2.12)	1.70 (2.26)	2.07 (2.58)	2.69 (2.91)	H = 829.61	$\eta^2 = 0.03$	$P < 0.001$
FAQ	0.32 (1.31)	0.53 (1.87)	1.09 (2.84)	1.60 (3.47)	2.92 (4.86)	H = 1586.04	$\eta^2 = 0.08$	$P < 0.001$
Clinical outcome								
Progression to dementia, %	5.7%	7.4%	15.7%	18.8%	28.8%	$X^2 = 1171.83$	$\varphi_c = 0.21$	$P < 0.001$
Year of dementia diagnosis	7.82 (3.47)	6.05 (3.33)	4.41 (2.70)	3.67 (2.06)	3.01 (1.67)	H = 714.89	$\eta^2 = 0.19$	$P < 0.001$
Primary dementia etiology:						$X^2 = 123.70$	$\varphi_c = 0.09$	$P < 0.001$
AD	77.7%	73.2%	85.5%	80.9%	76.0%			
LBD	6.1%	6.8%	5.0%	4.9%	8.7%			
Vascular	6.1%	7.8%	3.3%	5.1%	4.4%			
FTD	2.5%	3.8%	1.0%	3.7%	4.9%			
TBI	1.0%	0.4%	0.6%	0.2%	0.2%			
Systemic/medical illness	1.5%	0.5%	0.2%	0.4%	0.2%			
PSP	0.5%	0.4%	0.0%	0.3%	1.4%			
CBD	0.5%	0.5%	0.4%	0.5%	0.7%			
Psychiatric	0.0%	1.5%	0.4%	0.6%	0.3%			
Other[b]	2.0%	3.6%	2.3%	1.8%	2.1%			
Undetermined	2.0%	1.5%	1.5%	1.7%	1.1%			

Note: Values represent mean (standard deviation), unless otherwise indicated.

Abbreviations: AD, Alzheimer's disease; *APOE*, apolipoprotein E; CBD, corticobasal degeneration; CDR, Clinical Dementia Rating; CN, cognitively normal; CNS, central nervous system; FAQ, Functional Assessment Questionnaire; FTD, frontotemporal dementia; GDS, Geriatric Depression Scale; LBD, Lewy body disease; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment (education corrected); PSP, progressive supranuclear palsy; TBI, traumatic brain injury.

[a]*APOE* available for 77.7% of sample; CDR available for 100% of the sample; MMSE available for 68.6% of sample; MoCA available for 31.2% of sample; GDS available for 98.7% of the sample; FAQ available for 77.0% of sample.

[b]Other: Normal-pressure hydrocephalus ($n = 8$), prion disease ($n = 3$), CNS neoplasm ($n = 2$), epilepsy ($n = 1$), cognitive impairment due to alcohol abuse ($n = 4$), cognitive impairment due to medications ($n = 5$), other unspecified reason ($n = 64$).

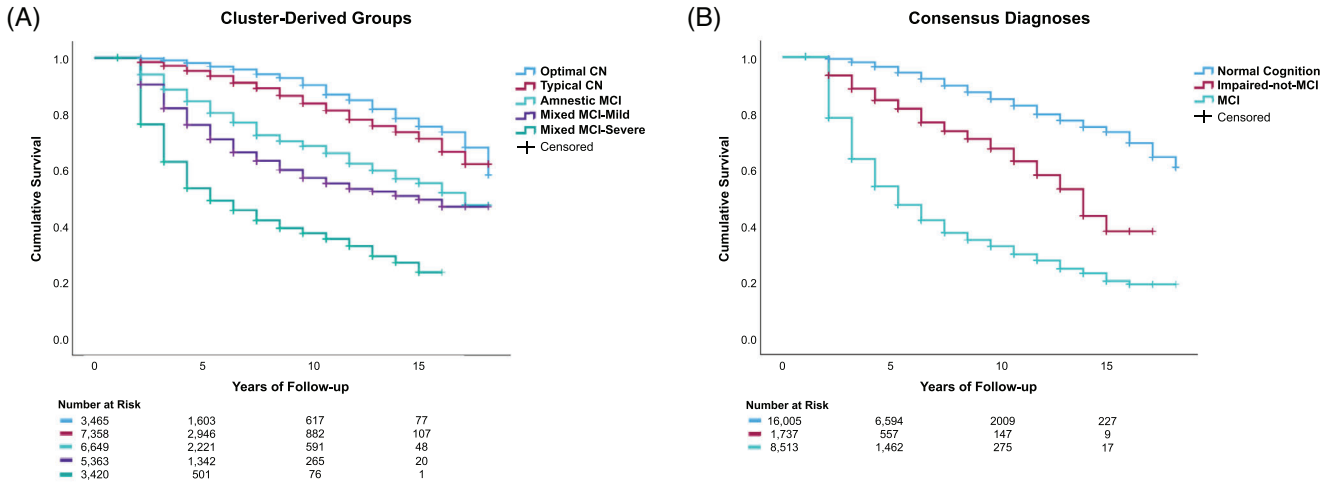

FIGURE 2 Kaplan–Meier survival curves showing progression to a consensus diagnosis of dementia in the (A) cluster-derived groups, and (B) NACC consensus diagnostic groups. An event was defined as the visit at which a participant first received a diagnosis of dementia. Participants who did not progress to dementia during their follow-up period were censored at their last visit. All groups in both analyses differed significantly from one another ($Ps < 0.001$). CN, cognitively normal; MCI, mild cognitive impairment; NACC, National Alzheimer's Coordinating Center.

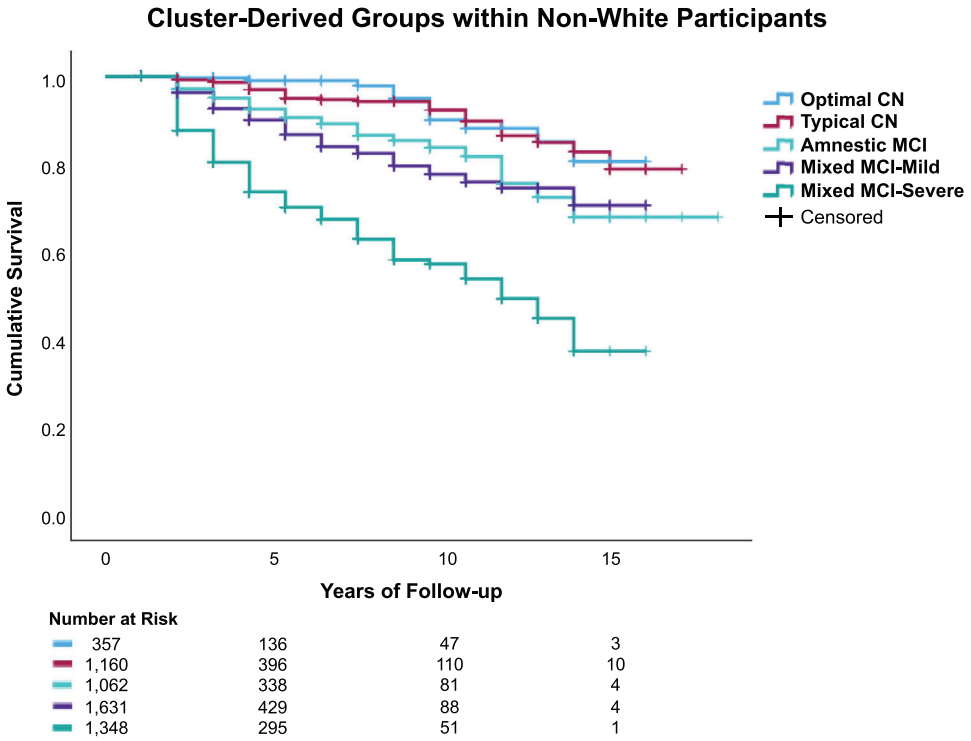

FIGURE 3 Kaplan–Meier survival curves showing progression to a consensus diagnosis of dementia in the cluster-derived groups for non-White participants ($n = 5558$ at baseline). An event was defined as the visit at which a participant first received a diagnosis of dementia. Participants who did not progress to dementia during their follow-up period were censored at their last visit. The Optimal CN and Typical CN groups did not differ from one another, but both showed lower rates of progression than the amnestic MCI and mixed MCI-Mild groups, which did not differ; the mixed MCI-Severe group had the highest rate of progression ($Ps < 0.001$). CN, cognitively normal; MCI, mild cognitive impairment.

range 2–14). Kaplan–Meier curves depicting rate of progression to dementia over time by group in non-White participants are shown in Figure 3. A log-rank test revealed significant group differences in survival curves ($\chi^2[4] = 305.25$; $P < 0.001$). Although all five cluster groups differed significantly in the full sample, pairwise comparisons showed that there were only three levels of risk within the non-White participants. Specifically, the oCN and tCN groups did not differ from one another, but both showed lower rates of progression than the aMCI and mMCI-Mild groups, which did not differ. Similar to results in the full sample, non-White participants in the mMCI-Severe group

TABLE 2 Number of participants in each cluster-derived group as a function of consensus diagnostic group at baseline.

	Consensus diagnosis		
	Normal cognition	Impaired-not-MCI	MCI
Cluster-derived group			
oCN	3081 (19.3%)	152 (8.8%)	232 (2.7%)
tCN	5797 (36.2%)	446 (25.7%)	1115 (13.1%)
aMCI	3937 (24.6%)	429 (24.7%)	2283 (26.8%)
mMCI-Mild	2354 (14.7%)	423 (24.4%)	2586 (30.4%)
mMCI-Severe	836 (5.2%)	287 (16.5%)	2297 (27.0%)
Total	16,005 (100%)	1737 (100%)	8513 (100%)

Abbreviations: aMCI, amnestic mild cognitive impairment; MCI; MCI, mild cognitive impairment; mMCI, mixed mild cognitive impairment; oCN, Optimal cognitively normal; tCN, Typical cognitively normal.

had a higher rate of progression to dementia than all other groups (oCN = tCN < aMCI = mMCI-Mild < mMCI-Severe). Three levels of risk were also found in a subanalysis limited to only Hispanic participants (n = 175 of 1892 Hispanic participants progressed; $\chi^2[4] = 107.00$; $P < 0.001$; oCN = tCN = aMCI < mMCI-Mild < mMCI-Severe).

Regarding the type of dementia, 81.6% (n = 412) of the non-White participants who progressed were thought to have AD as the primary etiology. The most common primary non-AD etiology was vascular disease (7.5%), followed by LBD (3.0%). Chi-square analysis showed no significant difference in primary etiology across the cluster-derived groups in the non-White participants who progressed ($\chi^2[36] = 30.79$; $P = 0.72$).

In those with a primary etiology of AD (n = 412), one or more secondary etiologies were present in 26.5% of cases, the most common being psychiatric conditions (12.9%), vascular disease (9.2%), and systemic or medical illness (3.2%). In participants with a primary non-AD etiology (n = 93), one or more secondary etiologies were present in 54.8% of cases, the most common being AD (28.0%), psychiatric conditions (18.3%), and systemic or medical illness (5.4%).

3.4 | Comparisons to consensus diagnosis

The number of participants in each cluster-derived group is shown as a function of consensus diagnostic groups in Table 2. Overall, the cluster analysis classified a greater number of individuals as having MCI (58.8% of the sample) than did the consensus method (32.4% of the sample with MCI, plus another 6.6% impaired-not-MCI). Within NACC's MCI cohort, 84.2% were also classified into one of the cluster-derived MCI groups. However, only 55.5% of NACC's "normal cognition" cohort were classified into one of the cluster-derived CN groups, while 24.6% were classified into our amnestic MCI group and 19.9% into one of the mixed MCI groups. The majority of the NACC "impaired-not-MCI" group was split fairly evenly across the tCN, aMCI, and mMCI-Mild groups. In contrast to the cluster-derived groups that were based on comprehensive neuropsychological test performance, the consensus diagnoses appeared to be heavily driven by CDR scores, as roughly 90% of consensus "normal cognition" participants had a global CDR of 0.0 (mean CDR = 0.05, SD = 0.15), and ≈ 90% of consensus MCI participants had a global CDR of 0.5 or above (mean CDR = 0.46, SD = 0.17).

Of participants who progressed to a diagnosis of dementia, 68.2% (n = 2579) were classified as having MCI at baseline by the consensus diagnosis, while 80.3% (n = 3039) were classified as having MCI at baseline by the cluster analysis (1046 aMCI, 1007 mMCI-Mild, 986 mMCI-Severe), suggesting that the data-driven method was more sensitive for detecting at-risk participants.

3.5 | Neuropsychological cluster-derived groups in the NACC UDS "normal cognition" subsample

Cluster analysis of baseline neuropsychological data from *only* those participants classified as "normal cognition" by consensus diagnosis in the NACC UDS (n = 16,005) revealed five cognitive subgroups: (1) High-All Domains (High-All n = 2672; 16.7%) with above average performance across domains; (2) Low-Attention/Working Memory (Low-WM; n = 3532; 22.1%) with low scores (approximately half a SD below the mean) on measures of verbal attention and working memory; (3) Low-Memory (n = 5811; 36.3%) with low immediate and delayed verbal memory scores; (4) Amnestic MCI (aMCI; n = 2669; 16.7%) with impaired performance (1 SD below the mean) on memory measures and low scores across other domains; and (5) Non-amnestic MCI (naMCI; n = 1321; 8.3%) with impaired performance on measures of processing speed, executive functioning, and language, as well as low scores across other domains; see Figure 4.

The naMCI group was significantly older than the other groups, and the Low-Memory group was older than the High-All and Low-WM groups; see Table 3. The naMCI group also had less education than all other groups, while the Low-WM group had higher education then the High-All and Low-Memory groups. The naMCI group had more women than the other groups except High-All. The proportion of non-White and Hispanic participants was lowest in the High-All group and increased across the five groups.

APOE genotype was available for 79.4% of the sample. The aMCI group had a higher proportion of *APOE* ε4 carriers than the High-All group. Baseline CDR scores (available for 100% of the sample) were near zero across all the groups, but were lowest for the High-All group; followed by the Low-WM and Low-Memory groups, which did not differ; followed by the aMCI and naMCI groups, which did not differ. Scores on the MMSE (available for 67.9% of the sample) and the MoCA (available for 31.9% of the sample) were similar for the Low-WM and Low-Memory groups, but otherwise differed significantly across groups. Scores on GDS and FAQ (available for 98.9% and 84.0% of the sample, respectively) were minimal overall, but were highest in naMCI, followed by aMCI, followed by the other groups, which did not differ.

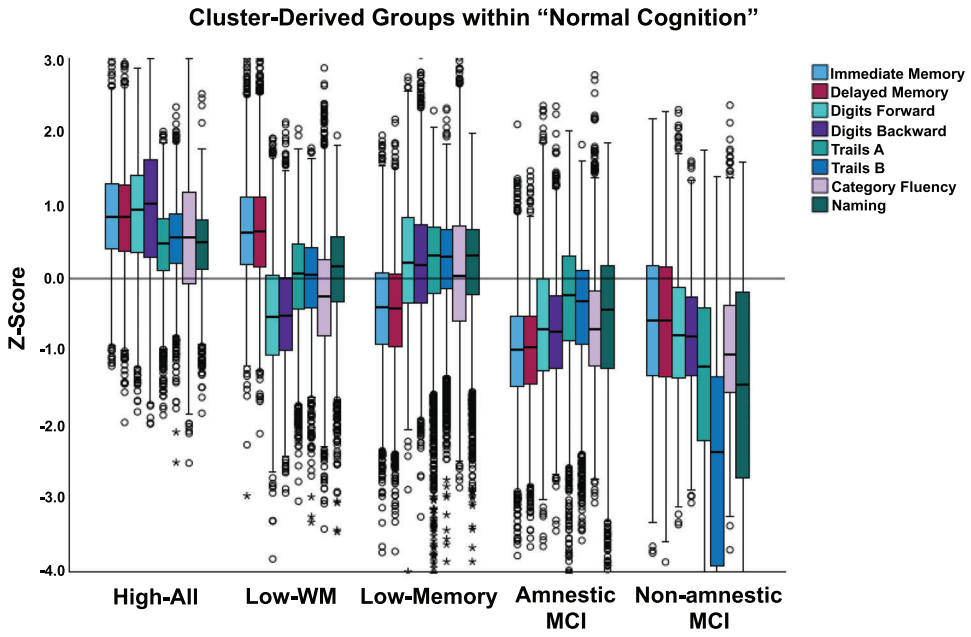

FIGURE 4 Baseline neuropsychological performance of the cluster-derived groups within NACC's "normal cognition" sample. CN, cognitively normal; MCI, mild cognitive impairment; NACC, National Alzheimer's Coordinating Center; WM, working memory.

3.6 | Progression to MCI/dementia in the NACC UDS "normal cognition" sample

Of the 16,005 "normal cognition" participants, 2810 (17.6%) progressed to a consensus diagnosis of either MCI ($n = 1846$) or dementia ($n = 964$) after an average of 4.8 years post-baseline (SD = 3.1, range 2–16). Cox regression using High-All as the reference group and adjusting for demographics (age, education, sex, race, ethnicity) showed an increased risk of progression to MCI/dementia in the Low-WM (HR = 1.79, 95% CI [1.55, 2.06], $P < 0.001$), Low-Memory (HR = 1.79, 95% CI [1.57, 2.04], $P < 0.001$), aMCI (HR = 3.26, 95% CI [2.82, 3.76], $P < 0.001$), and naMCI groups (HR = 3.44, 95% CI [2.92, 4.06], $P < 0.001$).

Kaplan–Meier curves depicting rate of progression to dementia over time by group are shown in Figure 5. A log-rank test revealed significant group differences in survival curves ($\chi^2[4] = 507.28$; $P < 0.001$). Pairwise comparisons showed that the naMCI group had the steepest survival curve (i.e., fastest rate of progression); followed by the aMCI group; followed by the Low-WM and Low-Memory groups, which did not differ from one another ($P = 0.57$); followed by the High-All group, which had slower progression than all other groups ($Ps < 0.001$).

For those participants who progressed to a consensus diagnosis of dementia specifically ($n = 964$), 78.8% ($n = 760$) were presumed to have AD as the primary etiology. The most common primary non-AD etiology was vascular disease (7.1%), followed by LBD (5.1%); see Table 3. Chi-square analysis showed no significant difference in primary etiology across the "normal cognition" cluster-derived groups ($\chi^2[24] = 26.07$; $P = 0.35$).

4 | DISCUSSION

Cluster analytic techniques that group individuals with similar cognitive profiles identified five distinct neuropsychological subgroups in 26,255 older adults without dementia within the NACC dataset, including two cognitively normal subtypes (oCN, tCN) and three MCI subtypes (aMCI, mMCI-Mild, and mMCI-Severe). The extent of cognitive impairment across the five groups was related to risk of progression to a diagnosis of dementia. The most impaired subtype, mMCI-Severe, was the oldest and had the fewest years of education, and the most functional difficulty (although still independent). Our MCI subgroups likely reflect a continuum of AD pathology, with mMCI-Severe representing a more advanced stage and aMCI an earlier stage. This is based on our previous work showing that a cluster-derived mixed MCI group had widespread cortical atrophy (corresponding to Braak stages V–VI), whereas an aMCI group had atrophy largely limited to medial and lateral temporal lobe regions (Braak stages III–IV).[10,13]

Most participants who progressed to dementia were presumed to have AD as the primary etiology (80%), although mixed etiologies were common. The aMCI group was most likely to progress to a primary etiology of AD dementia, while mMCI-Severe had the highest rate of non-AD pathologies (e.g., LBD, FTD), and tCN had a higher rate of progression to a primary etiology of vascular dementia. These findings suggest that our identified subtypes may help to guide individualized treatments, although future studies examining neuropathological diagnoses are needed.

We observed higher rates of racial/ethnic diversity in our more cognitively impaired groups, and subanalyses within non-White participants showed that participants in the more impaired groups had

TABLE 3 Baseline demographic and clinical characteristics of the cluster-derived groups in the NACC "normal cognition" sample only.

	High-All ($n = 2672$)	Low-WM ($n = 3532$)	Low-memory (5811)	Amnestic MCI ($n = 2669$)	Non-amnestic MCI ($n = 1321$)	F, H, or X^2	Effect Size	P
Demographic variables								
Age, years	70.33 (8.53)	70.38 (8.67)	71.18 (8.84)	70.91 (9.07)	73.44 (9.61)	F = 33.96	$\eta_p^2 = 0.01$	$P < 0.001$
Education, years	15.82 (2.61)	16.24 (2.66)	15.86 (2.75)	16.05 (3.03)	14.45 (3.94)	F = 97.72	$\eta_p^2 = 0.02$	$P < 0.001$
Sex: Female, %	68.1%	64.6%	65.2%	64.9%	69.6%	$X^2 = 18.89$	$\varphi_c = 0.03$	$P < 0.001$
Race, %:						$X^2 = 1744.20$	$\varphi_c = 0.17$	$P < 0.001$
White	91.2%	81.3%	84.0%	63.7%	42.9%			
Black or African American	5.6%	12.2%	10.5%	25.3%	40.7%			
American Indian or Alaska Native	0.1%	0.6%	0.4%	1.2%	0.9%			
Native Hawaiian or Pacific Islander	0.0%	0.1%	0.1%	0.1%	0.1%			
Asian	0.9%	2.0%	1.7%	3.9%	5.8%			
Multiracial	2.0%	2.9%	2.9%	3.9%	6.4%			
Unknown	0.2%	0.9%	0.4%	1.8%	3.2%			
Ethnicity: Hispanic, %	2.0%	6.0%	3.9%	11.8%	17.8%	$X^2 = 557.73$	$\varphi_c = 0.19$	$P < 0.001$
Clinical variables[a]								
APOE ε4 carrier, %	29.3%	30.8%	30.6%	33.7%	31.0%	$X^2 = 10.14$	$\varphi_c = 0.03$	$P = 0.04$
CDR: Global	0.03 (0.13)	0.05 (0.15)	0.05 (0.15)	0.06 (0.17)	0.07 (0.18)	H = 74.74	$\eta^2 = 0.004$	$P < 0.001$
CDR: Sum of Boxes	0.07 (0.26)	0.10 (0.33)	0.10 (0.33)	0.15 (0.42)	0.18 (0.51)	H = 146.56	$\eta^2 = 0.01$	$P < 0.001$
MMSE	29.40 (0.89)	29.07 (1.17)	29.03 (1.20)	28.50 (1.55)	27.54 (2.11)	H = 1047.97	$\eta^2 = 0.10$	$P < 0.001$
MoCA	27.60 (1.88)	26.61 (2.28)	26.59 (2.22)	24.80 (2.84)	23.32 (3.46)	H = 787.32	$\eta^2 = 0.15$	$P < 0.001$
GDS	1.17 (1.84)	1.24 (1.84)	1.29 (1.95)	1.53 (2.21)	1.81 (2.41)	H = 98.73	$\eta^2 = 0.01$	$P < 0.001$
FAQ	0.16 (0.76)	0.21 (0.98)	0.23 (1.05)	0.43 (1.79)	0.70 (2.53)	H = 110.09	$\eta^2 = 0.01$	$P < 0.001$
Clinical outcome								
Progression to MCI/dementia, %	11.2%	16.3%	16.8%	22.4%	27.4%	$X^2 = 214.60$	$\varphi_c = 0.12$	$P < 0.001$
Year of MCI/dementia diagnosis	6.19 (3.49)	5.22 (3.23)	4.90 (3.06)	4.16 (2.64)	3.90 (2.41)	H = 124.30	$\eta^2 = 0.04$	$P < 0.001$
Progression to dementia, %	4.0%	4.9%	6.2%	6.7%	10.7%	$X^2 = 80.57$	$\varphi_c = 0.07$	$P < 0.001$
Year of dementia diagnosis	8.71 (3.57)	7.19 (3.18)	6.75 (3.12)	5.47 (2.72)	4.99 (2.59)	H = 106.64	$\eta^2 = 0.11$	$P < 0.001$
Primary dementia etiology:						$X^2 = 26.07$	$\varphi_c = 0.08$	$P = 0.35$
AD	80.4%	77.5%	80.4%	78.3%	75.9%			
LBD	5.6%	6.9%	3.9%	4.4%	6.4%			
Vascular	6.5%	4.6%	8.8%	7.2%	5.7%			
FTD	1.9%	4.0%	0.6%	1.1%	1.4%			
Psychiatric	0.9%	1.2%	1.1%	1.7%	0.7%			
Other[b]	1.9%	5.2%	4.4%	6.1%	7.8%			
Undetermined	2.8%	0.6%	0.8%	1.1%	2.1%			

Note: Data are summarized as mean (standard deviation), unless otherwise indicated.

Abbreviations: AD, Alzheimer's disease; *APOE*, apolipoprotein E; CDR, Clinical Dementia Rating; CN, cognitively normal; CNS, central nervous system; FAQ, Functional Assessment Questionnaire; FTD, frontotemporal dementia; GDS, Geriatric Depression Scale; LBD, Lewy body disease; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment (education corrected).

[a]*APOE* available for 79.4% of sample; CDR available for 100% of the sample; MMSE available for 67.9% of sample; MoCA available for 31.9% of sample; GDS available for 98.9% of the sample; FAQ available for 84.0% of sample.

[b]Other: Progressive supranuclear palsy ($n = 1$), corticobasal degeneration ($n = 4$), traumatic brain injury ($n = 8$), systemic/medical illness ($n = 6$), normal-pressure hydrocephalus ($n = 5$), CNS neoplasm ($n = 2$), epilepsy ($n = 1$), cognitive impairment due to alcohol abuse ($n = 2$), cognitive impairment due to medications ($n = 3$), other unspecified reason ($n = 17$).

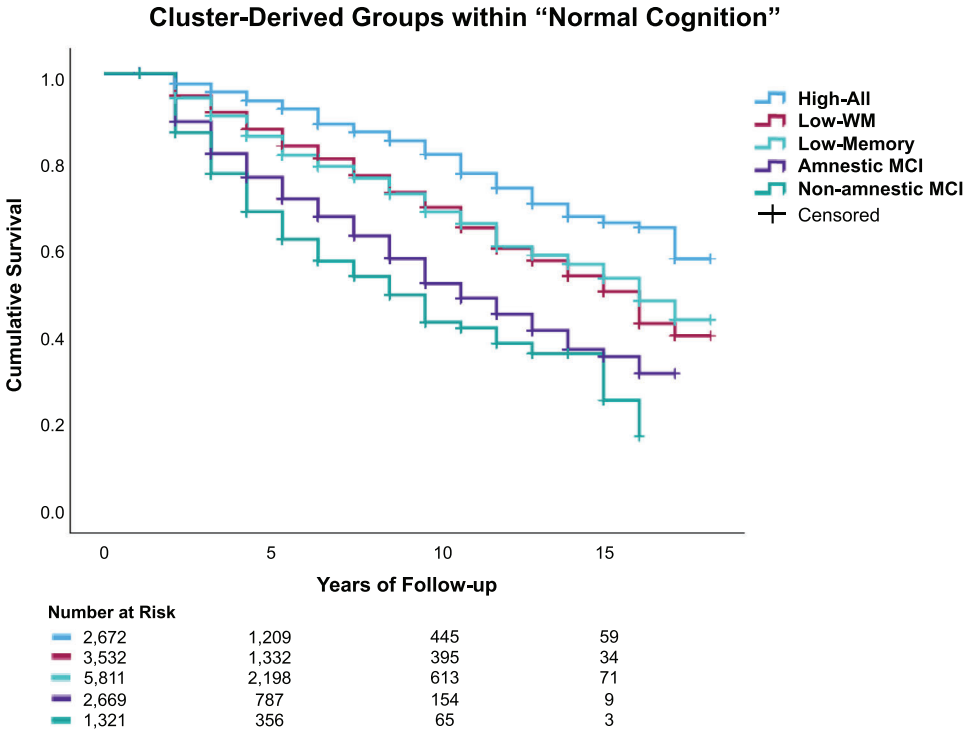

FIGURE 5 Kaplan–Meier survival curves showing progression to a consensus diagnosis of MCI or dementia in the cluster-derived groups within NACC's "normal cognition" sample ($n = 16{,}005$ at baseline). An event was defined as the visit at which a participant first received a diagnosis of either MCI or dementia. Participants who did not progress during their follow-up period were censored at their last visit. All groups differed significantly from one another ($Ps < 0.001$) except for similar progression rates in the Low-WM and Low-Memory groups ($P = 0.57$). MCI, mild cognitive impairment; NACC, National Alzheimer's Coordinating Center; WM, working memory.

higher rates of progression. Similar to the full sample, most non-White participants who progressed were presumed to have AD as the primary etiology (82%) and secondary etiologies were present in many cases. A previous study[22] using the NACC cohort also found that, relative to non-Hispanic White participants, non-Hispanic Black participants were more likely to meet criteria for a data-driven diagnosis of MCI based on neuropsychological testing, despite being classified as CN or impaired-not-MCI by consensus diagnosis. Other work comparing the utility of diagnostic methods found that consensus MCI diagnoses best predicted incident dementia in a mixed-raced sample, but data-driven diagnoses based on neuropsychological testing were more sensitive to early signs of decline and better predicted functional changes, particularly among Black older adults.[23]

The overrepresentation of racially/ethnically diverse participants in our MCI groups may reflect increased risk of cognitive decline and dementia in minoritized populations secondary to disparities in quality of education, access to health care, quality of health care, socioeconomic opportunities, chronic stress due to racism, and other social determinants of health.[24–27] Additionally, there may be implications of using a "robust" CN sample to determine normative performance in racial/ethnic minority participants. This method excludes individuals with prodromal neurodegenerative disease from normal aging comparison samples.[28,29] Robust norms may create a higher standard for "normal" and thus identify more individuals as falling into an MCI group. Indeed, previous work has shown that use of a robust normative sample results in higher sensitivity for detecting preclinical dementia.[28]

While the robust norms provided an adjustment for demographic variables at an individual level, it is noted that raw scores and *z* scores were highly correlated, both within the full sample ($r = 0.92$ to 0.98 for normally distributed variables, $\rho = 0.85$ to 0.88 for non-normally distributed variables) and the non-White subsample ($r = 0.90$ to 0.98, $\rho = 0.92$ to 0.97). Thus, our normative method may not fully account for the higher rates of diverse participants in our MCI groups, and future studies should explore the benefit of stratifying normative approaches by race/ethnicity and/or including additional variables that could be adjusted for during the normative process (e.g., education quality, socioeconomic status).

Comparison of our cluster-derived groups to the NACC consensus diagnoses showed that our method classified a greater number of MCI cases. Consensus diagnoses closely corresponded to global CDR scores, which likely explains the considerable discrepancy between diagnostic methods. Specifically, the consensus panel may have been less likely to diagnose a participant with MCI, despite low test scores, if they or their study partner did not report subjective complaints (a requirement of conventional diagnostic methods[1,2]) on the CDR. While this approach may have applicability in clinical settings, it may have less utility in research studies aimed at detecting early objective cognitive changes in at-risk older adults.

All three of our cluster-derived MCI subgroups demonstrated objective memory impairment, with two memory scores that were ≥ 1 SD below the demographic mean. This is generally consistent with published neuropsychologically-defined criteria for MCI, which define objective impairment as having two test scores within a particular cognitive domain > 1 SD below the normative mean.[7,22,30] However, conventional criteria for MCI[1,2] typically define objective memory impairment as having at least one score > 1.5 SD below the mean. This difference in the definition of what constitutes memory impairment may have contributed to more participants being classified with MCI using our data-driven approach.

The current finding of a discrepancy between the cluster-derived classifications and consensus diagnoses differs from a previous study[23] that found 90% concordance between statistically-determined MCI diagnoses from a latent profile analysis and consensus diagnoses. While both studies leveraged demographically-adjusted neuropsychological *z* scores, only the current study derived *z* scores from the performance of robust cognitively normal participants. Additionally, the previous study was conducted in a sample of participants from one ADRC while we used the larger NACC UDS sample. Despite these differences, the subgroups identified by the previous study were similar to our cluster-derived subgroups, as that study identified two normal groups (Low-Normal and High-Normal) and three MCI groups (Memory-Only, Memory/Language, Memory/Executive).[23]

Our data-driven method of classifying participants into five neuropsychological subgroups provided more precise information about progression risk than did the consensus method of dividing the sample into three groups (normal cognition, impaired-not-MCI, MCI). Additionally, a higher percentage of participants who progressed to dementia had been classified as MCI at baseline by the cluster analysis (80%) than by the consensus approach (68%), suggesting the cluster method is more sensitive for detecting subtle cognitive changes early in the course of the disease.

We further explored subtle cognitive weaknesses by restricting the sample to those classified as "normal cognition" by NACC's consensus diagnostic approach. Cluster analysis within this subsample revealed five neuropsychological subgroups, including above average performance (High-All), and groups with low scores in particular cognitive domains (Low-Attention/Working Memory and Low-Memory) that potentially reflect preclinical AD.[31,32] Additionally, there were participants whose performance was categorized as amnestic or non-amnestic MCI who were missed by the consensus approach. The focus on memory on the CDR may explain why some participants with non-amnestic impairments were not captured by consensus diagnosis. Across the "normal cognition" cluster groups, the extent and pattern of cognitive weaknesses was predictive of progression to MCI/dementia. Importantly, progression rates did not differ between the subgroups with weaknesses in attention/working memory versus weaknesses in memory, emphasizing that there are different initial presentations, and paths to a diagnosis of MCI/dementia may vary.

Other studies have demonstrated multiple cognitive subgroups within subsets of NACC's MCI[33–35] or CN sample[36] using statistical methods such as latent profile analysis. These studies have shown that subgroups have unique cognitive features, and differing longitudinal clinical outcomes and neuropathological findings,[33–36] highlighting the significant heterogeneity within the NACC sample. Our study extends this work by including all NACC UDS participants without dementia, thereby increasing the sample size substantially, and applying a data-driven approach to define CN versus MCI.

The current study used the very large, well-characterized NACC sample with longitudinal clinical outcome data spanning up to 17 years. A limitation of the study, and of the NACC more generally, is that participants tend to be highly educated and are largely White and non-Hispanic, which limits generalizability of results. Further, differences in enrollment factors (e.g., referral source) by race have been shown to impact observed associations between MCI status and rate of progression to dementia in NACC.[37] Continued research, particularly studies incorporating AD biomarkers in more representative cohorts, will be critical for further determining the utility of data-driven diagnoses in diverse samples.

Results of the present study have implications for future research by demonstrating a method to identify empirically-derived subtypes of subtle cognitive decline and MCI that optimizes prediction of future risk for MCI/dementia. This is critically important given the costs and stakes of past and ongoing clinical trials. Our previous work[38] with the Alzheimer's Disease Cooperative Study (ADCS) Vitamin E/donepezil trial dataset revealed a stronger effect of donepezil in patients with MCI diagnosed post hoc using the more rigorous cluster analytic methods that were applied in the current study. This work showed that comprehensive neuropsychological testing and a data-driven diagnostic approach can result in more efficient clinical trials and improved ability to detect treatment effects.[38] Application of our data-driven methods can provide improved diagnostic classification and a more precise assessment of relationships among cognitive decline and underlying AD biomarkers, clinical trial outcomes, and various risk and resilience factors associated with AD.

ACKNOWLEDGMENTS

The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADRCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI John Morris, MD), P30 AG066518 (PI Jeffrey Kaye, MD), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI David Bennett, MD), P30 AG072978 (PI Neil Kowall, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Eric Reiman, MD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Todd Golde, MD, PhD), P30 AG066508 (PI

Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Justin Miller, PhD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD). This work was supported by the Arizona Alzheimer's Consortium and Arizona DHS (CTR057001 to E.C.E.) and the NIH (P30 AG062429 to D.P.S., M.W.B.; R03 AG070435 to K.R.T.; RF1 AG082726 to K.R.T., E.C.E., M.W.B.). Dr. Thomas receives salary support from the U.S. Department of Veterans Affairs Clinical Sciences Research and Development Service (Career Development Award-2 1IK2CX001865 to K.R.T.).

CONFLICT OF INTEREST STATEMENT

Dr. Salmon is a consultant for Aptinyx and Biogen. Dr. Bondi receives royalties from Oxford University Press. The other authors have no disclosures. Author disclosures are available in the supporting information.

CONSENT STATEMENT

Written informed consent to participate in the study was obtained from all participants or their caregivers at each individual ADRC, as approved by individual institutional review boards (IRBs) at each site; the current study was approved by the Banner Health IRB.

REFERENCES

1. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256(3):183-194. doi:10.1111/j.1365-2796.2004.01388.x
2. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment–beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Intern Med*. 2004;256(3):240-246. doi:10.1111/j.1365-2796.2004.01380.x
3. Morris JC, Weintraub S, Chui HC, et al. The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis Assoc Disord*. 2006;20(4):210-216. doi:10.1097/01.wad.0000213865.09806.92
4. Beekly DL, Ramos EM, Lee WW, et al. The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Dis Assoc Disord*. 2007;21(3):249-258. doi:10.1097/WAD.0b013e318142774e
5. Besser L, Kukull W, Knopman DS, et al. Version 3 of the National Alzheimer's Coordinating Center's Uniform Data Set. *Alzheimer Dis Assoc Disord*. 2018;32(4):351-358. doi:10.1097/WAD.0000000000000279
6. Clark LR, Delano-Wood L, Libon DJ, et al. Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? *J Int Neuropsychol Soc*. 2013;19(6):635-645. doi:10.1017/S1355617713000313
7. Bondi MW, Edmonds EC, Jak AJ, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimers Dis*. 2014;42(1):275-289. doi:10.3233/JAD-140276
8. Edmonds EC, Delano-Wood L, Clark LR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimers Dement*. 2015;11(4):415-424. doi:10.1016/j.jalz.2014.03.005
9. Bangen KJ, Clark AL, Werhane M, et al. Cortical amyloid burden differences across empirically-derived mild cognitive impairment subtypes and interaction with APOE ε4 genotype. *J Alzheimers Dis*. 2016;52(3):849-861. doi:10.3233/JAD-150900
10. Edmonds EC, Eppig J, Bondi MW, et al. Heterogeneous cortical atrophy patterns in MCI not captured by conventional diagnostic criteria. *Neurology*. 2016;87(20):2108-2116. doi:10.1212/WNL.0000000000003326
11. Eppig JS, Edmonds EC, Campbell L, et al. Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: a latent profile analysis. *J Int Neuropsychol Soc*. 2017;23(7):564-576. doi:10.1017/S135561771700039X
12. Machulda MM, Lundt ES, Albertson SM, et al. Neuropsychological subtypes of incident mild cognitive impairment in the Mayo Clinic Study of Aging. *Alzheimers Dement*. 2019;15(7):878-887. doi:10.1016/j.jalz.2019.03.014
13. Edmonds EC, Weigand AJ, Hatton SN, et al. Patterns of longitudinal cortical atrophy over 3 years in empirically derived MCI subtypes. *Neurology*. 2020;94(24):e2532-e2544. doi:10.1212/WNL.0000000000009462
14. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat Med*. 2021;27(6):1034-1042. doi:10.1038/s41591-021-01348-z
15. Edmonds EC, Smirnov DS, Thomas KR, et al. Data-driven vs consensus diagnosis of MCI: enhanced sensitivity for detection of clinical, biomarker, and neuropathologic Outcomes. *Neurology*. 2021;97(13):e1288-e1299. doi:10.1212/WNL.0000000000012600
16. Thomas KR, Bangen KJ, Weigand AJ, et al. Cognitive heterogeneity and risk of progression in data-driven subtle cognitive decline phenotypes. *J Alzheimers Dis*. 2022;90(1):323-331. doi:10.3233/JAD-220684
17. Weintraub S, Salmon D, Mercaldo N, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord*. 2009;23(2):91-101. doi:10.1097/WAD.0b013e318191c7dd
18. Weintraub S, Besser L, Dodge HH, et al. Version 3 of the Alzheimer Disease Centers' neuropsychological test battery in the Uniform Data Set (UDS). *Alzheimer Dis Assoc Disord*. 2018;32(1):10-17. doi:10.1097/WAD.0000000000000223
19. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984;34(7):939-944. doi:10.1212/wnl.34.7.939
20. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia 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*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005
21. Monsell SE, Dodge HH, Zhou XH, et al. Results from the NACC Uniform Data Set Neuropsychological Battery Crosswalk Study. *Alzheimer Dis Assoc Disord*. 2016;30(2):134-139. doi:10.1097/WAD.0000000000000111
22. Graves LV, Edmonds EC, Thomas KR, Weigand AJ, Cooper S, Bondi MW. Evidence for the utility of actuarial neuropsychological criteria across the continuum of normal aging, mild cognitive impairment, and dementia. *J Alzheimers Dis*. 2020;78(1):371-386. doi:10.3233/JAD-200778
23. Devlin KN, Brennan L, Saad L, et al. Diagnosing mild cognitive impairment among racially diverse older adults: comparison of consensus, actuarial, and statistical methods. *J Alzheimers Dis*. 2022;85(2):627-644. doi:10.3233/JAD-210455
24. Sisco S, Gross AL, Shih RA, et al. The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *J Gerontol B Psychol Sci Soc Sci*. 2015;70(4):557-567. doi:10.1093/geronb/gbt133

25. Chin AL, Negash S, Hamilton R. Diversity and disparity in dementia: the impact of ethnoracial differences in Alzheimer disease. *Alzheimer Dis Assoc Disord*. 2011;25(3):187-195. doi:10.1097/WAD.0b013e318211c6c9
26. Galvin JE, Chrisphonte S, Chang LC. Medical and social determinants of brain health and dementia in a multicultural community cohort of older adults. *J Alzheimers Dis*. 2021;84(4):1563-1576. doi:10.3233/JAD-215020
27. Babulal GM, Quiroz YT, Albensi BC, et al. Perspectives on ethnic and racial disparities in Alzheimer's disease and related dementias: update and areas of immediate need. *Alzheimers Dement*. 2019;15(2):292-312. doi:10.1016/j.jalz.2018.09.009
28. Sliwinski M, Lipton RB, Buschke H, Stewart W. The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *J Gerontol B Psychol Sci Soc Sci*. 1996;51(4):P217-P225. doi:10.1093/geronb/51b.4.p217
29. Grober E, Mowrey W, Katz M, Derby C, Lipton RB. Conventional and robust norming in identifying preclinical dementia. *J Clin Exp Neuropsychol*. 2015;37(10):1098-1106. doi:10.1080/13803395.2015.1078779
30. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry*. 2009;17(5):368-375. doi:10.1097/JGP.0b013e31819431d5
31. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):280-292. doi:10.1016/j.jalz.2011.03.003
32. Jack CR Jr, Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562. doi:10.1016/j.jalz.2018.02.018
33. Blanken AE, Jang JY, Ho JK, et al. Distilling heterogeneity of mild cognitive impairment in the National Alzheimer Coordinating Center database using latent profile analysis. *JAMA Netw Open*. 2020;3(3):e200413. doi:10.1001/jamanetworkopen.2020.0413
34. Hanfelt JJ, Wuu J, Sollinger AB, et al. An exploration of subgroups of mild cognitive impairment based on cognitive, neuropsychiatric and functional features: analysis of data from the National Alzheimer's Coordinating Center. *Am J Geriatr Psychiatry*. 2011;19(11):940-950. doi:10.1097/JGP.0b013e31820ee9d2
35. Hanfelt JJ, Peng L, Goldstein FC, Lah JJ. Latent classes of mild cognitive impairment are associated with clinical outcomes and neuropathology: analysis of data from the National Alzheimer's Coordinating Center. *Neurobiol Dis*. 2018;117:62-71. doi:10.1016/j.nbd.2018.05.025
36. Hayden KM, Kuchibhatla M, Romero HR, et al. Pre-clinical cognitive phenotypes for Alzheimer disease: a latent profile approach. *Am J Geriatr Psychiatry*. 2014;22(11):1364-1374. doi:10.1016/j.jagp.2013.07.008
37. Gleason CE, Norton D, Zuelsdorff M, et al. Association between enrollment factors and incident cognitive impairment in Blacks and Whites: data from the Alzheimer's Disease Center. *Alzheimers Dement*. 2019;15(12):1533-1545. doi:10.1016/j.jalz.2019.07.015
38. Edmonds EC, Ard MC, Edland SD, Galasko DR, Salmon DP, Bondi MW. Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: a secondary analysis of the ADCS vitamin E and donepezil in MCI study. *Alzheimers Dement (N Y)*. 2018;4:11-18. doi:10.1016/j.trci.2017.11.001

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Edmonds EC, Thomas KR, Rapcsak SZ, et al. Data-driven classification of cognitively normal and mild cognitive impairment subtypes predicts progression in the NACC dataset. *Alzheimer's Dement*. 2024;20:3442–3454. https://doi.org/10.1002/alz.13793