Baseline Characterization of the ARMADA (Assessing Reliable Measurement in

Alzheimer's Disease) Study Cohorts

Tatiana Karpouzian-Rogers,[1,2] Emily Ho,[3] Miriam Novack,[3] Miriam Chinkers,[2] Katy Bedjeti,[3] Cindy Nowinski,[3,4] Bruno Giordani,[5] Richard Gershon,[3] Sandra Weintraub[1,2]

[1]Mesulam Center for Cognitive Neurology and Alzheimer's Disease, Northwestern University Feinberg School of Medicine

[2]Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine

[3]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine

[4]Department of Neurology, Northwestern University Feinberg School of Medicine

[5]Department of Psychiatry, University of Michigan

Corresponding Author:

Tatiana Karpouzian-Rogers

300 E. Superior

Chicago, IL 60611

t-karpouzian@northwestern.edu

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ALZ.12816.

Acknowledgements: We would like to thank the study participants and their family members for their time and participation in this study. We also thank the research staff across all sites for their help in carrying out this study.

Conflicts of Interest: Authors have no conflicts of interest to disclose.

Funding Sources: 1) Mayo Clinic Florida: NIA P50 AG016574 Mayo Clinic Alzheimer's Disease Research Center; Florida Department of Health Ed & Ethel Moore Alzheimer's Disease Research Program grant 8AZ08 Evaluating The Impact of a Dementia-Caring Community Model on African Americans with Alzheimer's Disease and Their Care Partners; 2) Emory University: NIA 1P30AG066511 Goizueta Alzheimer's Disease Research Center at Emory University; 3) University of California, San Diego: NIA P30AG062429 UCSD Alzheimer's Disease Research Center; 4) NACC NIA U01 AG16976, National Alzheimer's Coordinating Center; 5) Northwestern University: NIA P30 AG013854, Northwestern Alzheimer's Disease Center; NIA R01AG045571, R56AG045571, and R01AG067781, Cognitive SuperAging studies; 6) University of Pittsburgh: NIA P50 AG005133, Alzheimer's Disease Research Center, NIA P01 AG025204, Imaging Pathophysiology in Aging and Neurodegeneration, NIA R01 AG052446, Role of Midlife Cardiovascular Disease on Alzheimer's Pathology and Cerebrovascular Reactivity in the Young-Old.; 7) Massachusetts General Hospital: NIA P30AG062421, Massachusetts ADRC; NCBI P01AG036694- Harvard Aging Brain Study; 8) University of Michigan: NIA P30 AG053760, Michigan Alzheimer's Disease Research Center; NIA R01 AG054484, Community Based Approach to Early Detection of Transitions to Mild Cognitive Impairment and Alzheimer's Disease in African Americans (ELECTRA); NIA R01 AG058724, Treating Mild Cognitive Impairment with High Definition Transcranial Direct Current Stimulation (STIM); NIH RF1 AG047866, Impact of Disclosing Amyloid Imaging Results to Cognitively Normal Individuals (REVEAL SCAN); Cure Alzheimer's Fund, Deep Phenotyping of Older African Americans at Risk of Dementia; 9) Columbia University: Washington Heights-Inwood Columbia Aging Project (WHICAP), NIA PO1AG07232, R01AG037212, RF1AG054023, National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1TR001873; 10) University of Wisconsin: NIA P50 AG033514 and NIA P30 AG062715 Wisconsin Alzheimer's Disease Research Center; 11) Oregon Health & Science University: Layton Aging and Alzheimer's Disease Center, NIA P30 AG066518.

Keywords: NIH Toolbox, cognition, aging, dementia, Alzheimer's disease, mild cognitive impairment

Abbreviations: NIHTB: National Institutes of Health Toolbox for Assessment of Neurological and Behavioral Function; AD: Alzheimer's disease; MCI: Mild Cognitive Impairment; aMCI: Amnestic Mild Cognitive Impairment; DAT: Dementia of the Alzheimer type; NC: Normal Control; UDS: Uniform Data Set; NACC: National Alzheimer's Coordinating Center; ARMADA: Assessing Reliable Measurement in Alzheimer's Disease and cognitive Aging; ADC: Alzheimer's Disease Center; CDR: Clinical Dementia Rating Scale; FAQ: Functional Activities Questionnaire; NPI-Q: Neuropsychiatric Inventory Questionnaire.

Highlights (85 characters per bullet; 5 bullets max)

- There is a need for assessments that can detect the early stages of cognitive decline in older adults.
- The ARMADA study will validate the NIHTB across the aging spectrum, including MCI and DAT.
- Here we report the characteristics of participants.
- Groups were well-matched across most demographic characteristics, and clinical characteristics differed as expected.
- ARMADA study cohorts reflect their respective clinical syndromes for validating the NIHTB.

Abstract (150 limit)

WORD COUNT: 146

Introduction: The NIH Toolbox® (NIHTB) provides computerized measures of cognition, emotion, sensation, and motor abilities across the lifespan. The ARMADA study validated the NIHTB in individuals across the cognitive aging spectrum. This paper reports the characteristics of our sample of participants ages 65-85.

Methods: Participants were recruited across nine sites and classified clinically as cognitively normal (NC), with mild cognitive impairment (MCI), or with dementia of the Alzheimer type (DAT.) They completed the NIHTB at multiple time points and many had at least one Alzheimer biomarker previously obtained.

Results: Groups differed with respect to dementia severity levels, as anticipated but were well-matched across many demographic characteristics.

Discussion: The ARMADA study demographics and baseline characteristics provide a suitable sample for validating the NIHTB across the cognitive aging spectrum. Other enriched samples (African American participants, Spanish NIHTB, 85+ years old) will be reported elsewhere.

1. Introduction:

The prevalence of cognitive decline in older adults is rising at a never-before seen rate, due to both the growing population of aging adults, and to new medical advances that are leading to longer average life expectancies. It is currently projected that by 2030, individuals over the age of 65 will comprise more than 20% of the total population in the United States, a 4% increase compared to 2020.[1] Additionally, the number of individuals with dementia worldwide is expected to increase substantially (up to three-fold) by 2050.[2] The leading cause of progressive cognitive decline in individuals over the age of 65, alone or in combination with other diseases[3] is Alzheimer's disease (AD). AD is a neurodegenerative brain disease that typically presents clinically as an amnestic cognitive syndrome progressing over years to final stages referred to as dementia of the Alzheimer type (DAT).[4] The term "dementia" (currently referred to as major neurocognitive disorder in the DSM-5[5)] connotes cognitive change that exceeds what is expected in the course of "normal aging" to the point of interfering with independent activities of daily living such as driving, managing finances, meal preparation, or grooming. Mild Cognitive Impairment (MCI) is a term that refers to the prodromal state of DAT, at a time when biomarkers for Alzheimer's are present and cognitive impairments exceed "normal" age-related decline, but are not yet significantly interfering with most activities of daily living. [6,7]

Given the increasing rates of clinical MCI and dementia, there is a great need for early detection of prognostic cognitive decline in older adults. Although there is currently no "cure" for neurodegenerative conditions, early detection allows for earlier treatment, which can slow symptom progression[8] and may help affected individuals and caregivers with future care planning. Unfortunately, at present, identifying or diagnosing a clinical dementia, especially in very early stages, can only be accomplished through lengthy neuropsychological evaluations that may not be

available to all individuals and are not always suitable for widescale use. Although brief cognitive screening tests exist, they tend not to be sensitive enough to detect mild deficits.[9] Most critically, the field lacks appropriate assessments and normative data for individuals in racially/ethnically and linguistically diverse groups.[10,11]

To address these challenges, the NIH Toolbox® for Assessment of Neurological and Behavioral Function (NIHTB)[12,13] could serve as an effective intermediate method for detecting cognitive decline in older adults, and assessing domains beyond cognition that may mark other factors associated with dementia (e.g. aspects of emotion, sensation, and motor abilities).[14] The goal of the present study, ARMADA (Assessing Reliable Measurement in Alzheimer's Disease), was to validate the four modules of the NIHTB among adults over the age of 65, and across groups with normal cognition, amnestic MCI, or early (mild) DAT. The methods for this study are described in detail in a prior publication[15] and the present report provides detailed characteristics of the study populations. Participants were drawn from established research cohorts across nine sites with existing cohorts of older adults along the cognitive aging spectrum. Research group diagnosis had been characterized prior to study enrollment by the sites (maximum of a three-month window between ARMADA study visit and corresponding diagnosis visit), and AD biomarkers were available for a portion of participants. Three additional samples were also recruited: one consisted of cognitively normal individuals over age 85; the second was an enriched sample of African Americans; and the third consisted of Spanish-speaking individuals who were administered the Spanish version of the NIHTB. These additional samples were enrolled in order to expand normative data for these populations and help address existing disparities in assessment research. The present paper describes the demographic and clinical characteristics of participants

ages 65 to 85 from our research volunteer sample. The description of special emphasis groups, including participants over 85, African American individuals, and Spanish-Speaking individuals, will be presented elsewhere.

2. Methods

2.1. Recruitment Sites

For detailed information regarding study methods, refer to Weintraub et al., 2021. Participants in the cognitively normal control, MCI, and DAT groups were recruited from existing cohorts across sites that are Alzheimer's Disease Research Centers (ADRCs), as well as affiliated sites with longitudinal studies of normal and cognitively impaired older adults that utilized assessment and diagnostic methods similar to those employed in the ADRCs. Sites included: Northwestern University (lead site), University of Michigan, University of Wisconsin-Madison, Mayo Clinic-Jacksonville, Florida, University of Pittsburgh, Emory University, University of California-San Diego, Columbia University, and Massachusetts General Hospital. Many of the sites had imaging and cerebrospinal fluid (CSF) biomarkers of AD neuropathology, available on a portion of their cohorts, as well as genotypes of AD risk. A majority of sites had utilized the Uniform Data Set (UDS)[16-18] methodology to characterize and diagnose participants prior to their recruitment for the ARMADA study. For more recruitment and enrollment details per site, refer to Table 1.

2.2 Overview of Diagnostic and Inclusion/Exclusion Criteria

Research group diagnosis was established using UDS methodology through the National Alzheimer Coordinating Center (NACC; https://www.alz.washington.edu/WEB/forms_uds.html), which maintains a large database of clinical and neuropathological data and implements guidelines for clinical diagnosis based on up-to-date research. Diagnosis was established using the Clinical Dementia Rating Scale (CDR),[19] Functional Activities Questionnaire (FAQ),[20] the UDS neuropsychological battery or comparable cognitive tests for each domain captured in the UDS, as well as other relevant clinical data. Participant characteristics and data were obtained from the participant, study partner, neuropsychological testing, neurological examination, and additional neuroimaging or laboratory data as available. A majority of sites carried out research consensus conferences in the diagnostic process. Diagnoses for participants recruited outside of the ADRCs, that is, in affiliated studies, used similar criteria.

Participants in the two groups with cognitive impairment included individuals over age 65 with a research diagnosis of DAT (CDR Global Score of approximately 1.0) or aMCI (amnestic single or multidomain; Global CDR score of approximately 0.5) using the 2011 NIA-AA criteria.[21,22] DAT is defined on the basis of progressive cognitive impairment, with prominent memory loss and additional cognitive deficits that represent a decline from previous functioning, and interfere with the ability to function in daily activities. We included individuals with mild DAT because although we have not administered the NIHTB in participants at more advanced stages of dementia, it is our suspicion that these individuals would have more difficulty completing the NIHTB. MCI is defined as impairment in one or more cognitive domains with preservation of

independence in functional abilities. Participants with amnestic MCI (single or multidomain) were included, as this clinical profile is most associated with Alzheimer neuropathology.[6] Participants with non-amnestic MCI were not included. The diagnostic procedures included prediction of primary and contributing etiologies based on currently known probabilities of association between clinical features and neuropathologic findings at postmortem brain autopsy.[23] Biomarker information was also available in some instances to further strengthen the suspected etiology. Exclusion criteria included: 1) medical conditions that may negatively affect cognitive functioning, including history of CNS disease (e.g., toxic/metabolic encephalopathy, normal pressure hydrocephalus, stroke, brain tumor), 2) history of a chronic major psychiatric disorder, or 3) alcohol or other substance abuse.

Cognitively normal individuals were recruited and enrolled using similar methods. This group included individuals over age 65, without significant complaints of cognitive decline, including memory loss, and with normal cognition as assessed by standard cognitive testing, study partner report, and other clinical data obtained with the UDS or similar research procedures. Cognitively normal participants were required to live independently, without difficulties carrying out activities of daily living. Exclusion criteria, in addition to those for the MCI and DAT groups, included serious medical conditions that may affect cognitive functioning (e.g., thyroid disorder, renal, hepatic, cardiac or pulmonary insufficiency, unstable diabetes, uncontrolled hypertension, cancer, or chronic use of neuroleptic or hypnotic medications) at the site investigator's discretion. Participants were not excluded if: 1) they had conditions such as hypertension or diabetes that were well controlled, or 2) they were taking antidepressants or anxiolytics for situational symptoms and symptoms were well controlled.

2.3 Brief Overview of NIH Toolbox, UDS Neuropsychological Battery, Clinical Data, and Biomarker Data Collection

Participants completed NIH Toolbox assessments at baseline, and were subsequently followed at 12 and 24 months. The study was significantly disrupted by the COVID pandemic which emerged just as the baseline recruitment was drawing to a close. As a result, we were not able to establish sufficient power for analysis of 24-month data but will be able to analyze 12-month data in some of the groups. With few exceptions, participants received all four modules of the NIHTB. The Cognition Battery consists of tests assessing Executive Function and Attention, Episodic Memory, Language, Processing Speed, and Working Memory. The Emotion Battery consists of self-report surveys of multiple areas of emotional functioning. The Motor Battery assesses Dexterity, Grip Strength, Standing Balance, Locomotion, and Endurance. The Sensation Battery assesses Audition, Vision, Olfaction, and Pain. For more details on each module, refer to the original publications[13,24-33] and the NIHTB website (nihtoolbox.org).

The participants recruited from each site had been studied systematically to assign research diagnoses. Those recruited from the Clinical Cores of ADRCs had completed the Uniform Data Set.[16-18] The UDS includes a neuropsychological battery, which measures attention span, processing speed, executive attention, category and letter fluency, object naming, visual constructions, and immediate and delayed memory. Those recruited from non-UDS sites contributed comparable data to that collected in the UDS including participant demographics, medical history and medications, family history, research neurological examination, dementia severity level, functional ability assessment, and behavioral and

psychiatric symptoms. Non-UDS sites administered a comparable battery of neuropsychological tests covering similar cognitive domains. All sites implemented UDS guidelines for clinical diagnosis. https://files.alz.washington.edu/documentation/uds3-ivp-guidebook.pdf

Although biomarkers were not prospectively collected as part of the ARMADA procedures, study sites contributed data on available biomarkers, including amyloid Positron Emission Tomography (PET), and/or CSF tau/amyloid levels. Apolipoprotein E (ApoE) genotype, a major genetic risk factor for Alzheimer's disease, was also collected where available. We recorded the availability of structural neuroimaging on each participant at each site for future identification of these resources by investigators interested in accessing these data site by site (see Table 5 for biomarker availability by study group).

2.4 Statistical Analysis

Differences in demographic variables and clinical characteristics were evaluated using Chi-Squared tests for categorical variables, and univariate one-way ANOVAs for continuous variables. In many cases the distributions were parametric. In the cases where they were not, robust tests were conducted (e.g., Welch's t-test rather than the conventional student's t-test). To preserve the maximum number of data points we used pairwise deletion across the variables in question. Where there were statistically significant differences (p-values < .05), post-hoc tests were

conducted using Bonferroni corrections for categorical variables and Tukey's Honestly Significant Differences for continuous variables. Differences in performance on the NIHTB will be presented separately.

3. Results

3.1 Demographic Characteristics

The final sample included 326 participants, aged 65-85 years, collected across nine sites. As noted above, this sample does not include individuals from the three special emphasis cohorts (Spanish-speaking, African American, 85+ years old) which will be described in separate publications. There were three groups that ranged from normal cognition (NC) to amnestic MCI to mild dementia presumed to be of the Alzheimer type (DAT). Groups within this sample were well matched on the majority of demographic characteristics (Table 2) including race, ethnicity, English as primary language status, handedness and education. Of note, the sample was highly educated overall. Post-hoc comparisons revealed there were slight differences between the groups with respect to age (NC were younger than MCI and DAT) and sex (NC had fewer males than MCI and DAT). Nevertheless, the demographics of our sample were well matched across groups, and roughly matched to the US census.[34]

3.2 Clinical Characteristics

Dementia severity across the three groups was characterized from a combination of Clinical Dementia Rating (CDR) scores and Functional Assessment Questionnaire ratings (Table 3). The CDR is a widely used 5-point scale that characterizes multiple domains of cognitive and functional abilities, with zero being normal and three indicating severe dementia, as rated based on clinical judgment. The CDR was designed specifically with dementia of the Alzheimer type in mind and so the global CDR rating is heavily weighted for memory loss. The FAQ is a survey completed by the informant that classifies functional abilities into four levels: i) normal, ii) independent but has difficulty, iii) requires assistance, and iv) fully dependent. As expected, the DAT group scored significantly higher (with higher scores indicating greater impairment) compared to the MCI and NC groups on the CDR and FAQ, and the MCI group scored significantly higher than the NC group for the CDR and FAQ. Specifically, the DAT group demonstrated greater cognitive changes (i.e. memory, problem solving, etc.) across multiple domains and poorer functional abilities compared to the other groups. Conversely, the MCI group was characterized by some degree of cognitive change (namely memory) and minimal changes in functional abilities. Cognitively normal controls demonstrated minimal cognitive or functional deficits. Performance on a measure of general cognitive ability, the Montreal Cognitive Assessment (MOCA) differed across groups, with the highest score in the NC group, followed by the MCI and DAT groups.

According to the Neuropsychiatric Inventory Questionnaire (NPIQ),[35] an informant report of neuropsychiatric symptoms manifested in the last month prior to testing, the DAT group was more likely to experience symptoms compared to individuals in the other two clinical groups (Table 4). Additionally, the total number of behaviors endorsed on the NPIQ, which ranged from 0-12, was significantly higher for participants in the DAT

group compared to those in the NC or MCI groups. The most commonly endorsed symptoms in the DAT group included apathy, depression, and anxiety. On the Geriatric Depression Scale (GDS- Short Form),[36] which is a self-report questionnaire of symptoms of depression, there were no differences between groups (Table 4).

3.3 Family History

A total of 223 participants reported on whether they had a family history of cognitive impairment. Across the entire sample, 147 participants indicated that they had a first-degree family member with cognitive impairment. Prevalence of family history of cognitive impairment did not differ by participant group, χ^2 (2, N = 223) = 2.78, p = .25.

3.4 Biomarkers

Although not part of data collection for ARMADA, biomarker data, including amyloid Positron Emission Tomography (PET) and Cerebrospinal fluid (CSF) tau/amyloid levels had been collected from participants at some sites as part of other protocols in which they had participated. In general, a plurality of participants across all study groups had at least one biomarker collected. Additionally, structural MRI and ApoE genotype were available for a subset of participants in all clinical groups. Table 5 includes biomarker availability across study groups.

4. Discussion

The overall goal of the ARMADA study is to validate the NIHTB in older adults with normal cognition, amnestic mild cognitive impairment, and dementia of the Alzheimer type, and examine other neurologic factors that may be predictive of cognitive status. Here, we presented demographic and clinical characteristics of one study sample including individuals 65-85 years of age at their initial visit. Briefly, the three study groups (NC, MCI, and DAT) were well matched for most demographic characteristics, with the exception of age and sex. The MCI and DAT clinical groups were slightly older than the NC group. Males made up a smaller proportion of the NC group compared to MCI and DAT cohorts. Characterization of dementia severity levels using clinical dementia ratings and functional assessment scales differed as expected, with the DAT group demonstrating higher severity compared to the MCI group, and the NC group demonstrating minimal cognitive symptoms. This pattern was also demonstrated by performance on the MoCA, which the AD group scoring the lowest, followed by MCI and NC groups. There were also differences in informant-reported neuropsychiatric symptoms: individuals in the DAT group had more symptoms than those in NC and MCI groups. There were no differences in prevalence of reported family history of dementia across groups. While family history of late onset AD may be expected to be reported at a higher rate in affected individuals, the counter force may be that cognitively healthy individuals who enroll in studies may do so because of family history. Evaluation of these demographic and clinical characteristics indicate that these study cohorts accurately represent these clinical syndromes for validating the NIHTB across the aging spectrum.

Future plans for the ARMADA study include comparing performance in all four modules of the NIHTB across cohorts of controls, individuals with aMCI, and individuals with DAT, to determine if the NIHTB will capture differences and discriminate among groups. The relationship between AD biomarkers and NIHTB measures is of great interest. Participants were followed longitudinally, over a two-year period, to track changes in

NIHTB performance over time, and we plan to investigate which measures (or combination of measures) are most predictive of future decline. Data were also collected for additional special emphasis groups. Specifically, individuals over age 85 were recruited to extend the current age-specific norms of the NIHTB. Additionally, given the need for appropriate and comprehensive norms that better represent ethnic and linguistic diversity, samples of African American participants and Spanish-Speaking participants were also followed. A potential future direction of the ARMADA study is to include various dementia subtypes, regardless of etiology.

One limitation of this study to consider is that due to the need to recruit participants from existing cohorts, there may be varying data collection methods and data availability across sites. However, the majority of sites followed a standard protocol and, where possible, the remaining sites conformed to the protocol (e.g., had the same visit windows). Given the proportion of participants with a first-degree family member with cognitive impairment, this sample may be more highly motivated to engage in research, which may somewhat limit generalizability. Another limitation is that the average education level of the three groups is higher than, and therefore not fully representative of, the general population of the US. Research centers from which the samples were recruited tend to attract individuals who are more likely to volunteer for research, which may coincide with both greater educational attainment and in general, greater accessibility. Future studies with community-based samples may also address the question of the impact of education on our findings. To address disparities in research, there have been calls for increased accessibility so that research participants may more accurately reflect the diversity of the US population, and for standardized instruments for assessing neuropsychological functioning in underserved populations. (https://www.nia.nih.gov/report-2019-2020-scientific-

advances-prevention-treatment-and-care-dementia/urgent-need-increased). In an effort to address this issue, the ARMADA study had specially recruited special emphasis groups, including cohorts of African American individuals, and Spanish-Speaking individuals.

Another limitation of the sample is that there is a relatively small number of individuals recruited with DAT. Over recent years, recruitment into the ADRCs and affiliated studies has shifted emphasis from the dementia stage of cognitive aging ever earlier, to the MCI stage, and even to what is now known as a “preclinical” stage at which individuals are cognitively healthy but have evidence of AD biomarkers. The COVID-19 pandemic, added to the difficulty of studying older adults, especially those with cognitive impairment, and, importantly, significantly interfered with the planned longitudinal follow up over two years to assess change over time in all samples. Despite these challenges, ARMADA will provide a comprehensive, cross-sectional, pan-domain study of older individuals across the cognitive aging spectrum, which will support the validation of the NIHTB as a critical assessment tool in the study of cognitive aging as well as numerous additional studies exploring the relation between the NIHTB and other measure of cognitive performance.

References

1. Alzheimer's Association. (2020). Alzheimer's disease facts and figures [published online ahead of print, 2020 Mar 10]. Alzheimers Dement. 2020.

2. *Dementia.* (2020, September 21). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/dementia

3. Schneider, J. A., Arvanitakis, Z., Bang, W., & Bennett, D. A. (2007). Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology*, *69*(24), 2197-2204.

4. Jalbert, J. J., Daiello, L. A., & Lapane, K. L. (2008). Dementia of the Alzheimer type. *Epidemiologic reviews*, *30*(1), 15-34.

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

6. Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., ... & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of neurology*, *58*(12), 1985-1992.

7. Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, *256*(3), 183-194.

8. Gillette-Guyonnet, S., Andrieu, S., Nourhashemi, F., Gardette, V., Coley, N., Cantet, C., ... & Vellas, B. (2011). Long-term progression of Alzheimer's disease in patients under antidementia drugs. *Alzheimer's & Dementia*, *7*(6), 579-592.

9. Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, *78*(8), 790-799.

10. Morlett Paredes, A., Gooding, A., Artiola i Fortuny, L., Rivera Mindt, M., Suárez, P., Scott, T. M., ... & Marquine, M. J. (2020). The state of neuropsychological test norms for Spanish-speaking adults in the United States. *The Clinical Neuropsychologist*, 1-17.

11. Rivera Mindt, M., Byrd, D., Saez, P., & Manly, J. (2010). Increasing culturally competent neuropsychological services for ethnic minority populations: A call to action. *The Clinical Neuropsychologist*, *24*(3), 429-453.

12. Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology*, *80*(11 Supplement 3), S2-S6.

13. Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., ... & Fox, N. A. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S54-S64.

14. Albers, M. W., Gilmore, G. C., Kaye, J., Murphy, C., Wingfield, A., Bennett, D. A., ... & Zhang, L. I. (2015). At the interface of sensory and motor dysfunctions and Alzheimer's disease. *Alzheimer's & Dementia*, *11*(1), 70-98.

15. Weintraub, S., Karpouzian-Rogers, T., Peipert, J. D., Nowinski, C., Slotkin, J., Wortman, K., ... & Gershon, R. (2021). ARMADA: Assessing reliable measurement in Alzheimer's disease and cognitive aging project methods. *Alzheimer's & Dementia*.

16. Morris, J. C., Weintraub, S., Chui, H. C., Cummings, J., DeCarli, C., Ferris, S., ... & Beekly, D. (2006). The Uniform Data Set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Disease & Associated Disorders*, *20*(4), 210-216.

17. Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N. R., Chui, H., ... & Peskind, E. (2009). The Alzheimer's disease centers' uniform data set (UDS): The neuropsychological test battery. Alzheimer disease and associated disorders, 23(2), 91.

18. Weintraub, S., Besser, L., Dodge, H. H., Teylan, M., Ferris, S., Goldstein, F. C., ... & Mungas, D. (2018). Version 3 of the Alzheimer Disease Centers' neuropsychological test battery in the Uniform Data Set (UDS). *Alzheimer disease and associated disorders*, *32*(1), 10.

19. Morris, J. C. (1991). The Clinical Dementia Rating (CDR): Current version and. Young, 41, 1588-1592.

20. Pfeffer, R. I., Kurosaki, T. T., Harrah Jr, C. H., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of gerontology, 37*(3), 323-329.

21. Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., ... & Snyder, P. J. (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. *Alzheimer's & dementia*, *7*(3), 270-279.

22. McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., ... & Mohs, R. C. (2011). 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. *Alzheimer's & dementia*, *7*(3), 263-269.

23. Weintraub, S. (2014). Neuropsychological assessment of dementia: A large-scale neuroanatomical network approach. In B. Dickerson & A. Atri (Eds.), *Dementia: Comprehensive principles and practice* (pp. 487–507). Oxford University Press.

24. Gershon, R. C., Cella, D., Fox, N. A., Havlik, R. J., Hendrie, H. C., & Wagster, M. V. (2010). Assessment of neurological and behavioural function: the NIH Toolbox. *The Lancet Neurology*.

25. Coldwell, S. E., Mennella, J. A., Duffy, V. B., Pelchat, M. L., Griffith, J. W., Smutzer, G., ... & Victorson, D. (2013). Gustation assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S20-S24.

26. Cook, K. F., Dunn, W., Griffith, J. W., Morrison, M. T., Tanquary, J., Sabata, D., ... & Gershon, R. C. (2013). Pain assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S49-S53.

27. Dalton, P., Doty, R. L., Murphy, C., Frank, R., Hoffman, H. J., Maute, C., ... & Slotkin, J. (2013). Olfactory assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S32-S36.

28. Dunn, W., Griffith, J. W., Morrison, M. T., Tanquary, J., Sabata, D., Victorson, D., ... & Gershon, R. C. (2013). Somatosensation assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S41-S44.

29. Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y. C., Bubela, D. J., ... & Gershon, R. C. (2013). Motor assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S65-S75.

30. Rine, R. M., Schubert, M. C., Whitney, S. L., Roberts, D., Redfern, M. S., Musolino, M. C., ... & Marchetti, G. F. (2013). Vestibular function assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S25-S31.

31. Salsman, J. M., Butt, Z., Pilkonis, P. A., Cyranowski, J. M., Zill, N., Hendrie, H. C., ... & Lai, J. S. (2013). Emotion assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S76-S86.

32. Varma, R., McKean-Cowdin, R., Vitale, S., Slotkin, J., & Hays, R. D. (2013). Vision assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S37-S40.

33. Zecker, S. G., Hoffman, H. J., Frisina, R., Dubno, J. R., Dhar, S., Wallhagen, M., ... & Newman, C. (2013). Audition assessment using the NIH Toolbox. *Neurology*, *80*(11 Supplement 3), S45-S48.

34. Ruggles, S., Flood, S., Foster, S., & Goeken, R. (2021). Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021.". *IPUMS USA: Version*, *11*.

35. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. Neurology. 1994;44(12):2308-2314.

36. Herrmann, N., Mittmann, N., Silver, I. L., Shulman, K. I., Busto, U. A., Shear, N. H., & Naranjo, C. A. (1996). A validation study of the Geriatric Depression Scale short form. *International Journal of Geriatric Psychiatry*, *11*(5), 457-460.

Table 1. Summary of recruitment and enrollment methods across sites

ARMADA SITE	Recruitment Methods	Diagnostic Methods
Columbia University	Recruited through WHICAP study, a multi-ethnic cohort of community-living participants followed longitudinally.	Core assessment includes standardized interviews, neurologic and neuropsychological exams, and biomarker studies. Diagnosis is confirmed by consensus conference with multiple research clinicians, who are blind to biomarker status and blind to prior diagnosis.
Emory University	ADRC Clinical Core, often identified during consensus conference, and Biomarker Initiative consisting of	UDS procedures and NACC guidelines; UDS demographic and diagnosis forms, notes for annual visits, and biomarkers if available. Diagnosis was reached by Consensus

	patients seen in the Cognitive Neurology Clinic.	Conference for ADRC participants, and by cognitive neurologist diagnosis for those in the Biomarker Initiative.
Massachusetts General Hospital	ADRC Clinical Core, Harvard Aging Brain Study, EARLY and LEARN studies; prioritized participants with amyloid PET imaging.	UDS procedures and NACC guidelines; standardized clinical ratings, cognitive testing, and neurologic exam. Initial diagnosis determined by clinician-related CDR, and consensus conference was consulted is discrepancy from previous year.
Mayo Clinic Jacksonville	Community events, concurrent studies within NIA-funded ADRC, referrals from Behavioral Neurology department, Mayo Older African Americans Normative Studies	UDS procedures and NACC guidelines; standardized interviews, neurologic and neuropsychological exams, and biomarker studies. Initial diagnosis rendered by neurologist and confirmed with diagnostic case consensus conference.
Northwestern University	ADRC Clinical Core, Neurobehavioral and Memory Clinic, and Center's recruitment registry.	UDS procedures and NACC guidelines; standardized interviews, neurologic and neuropsychological exams, and biomarker studies. For patients referred through the Neurobehavior and Memory Clinic, referral information and cognitive testing from clinician was provided. Diagnosis confirmed either by

		clinician or consensus conference with multiple research clinicians.
University of California-San Diego	ADRC longitudinal study cohort and recruitment sessions during memory screening events in the community	UDS procedures and NACC guidelines; self-report questionnaires, interviews, cognitive exam, and neuroimaging when available. Diagnosis reached by research diagnostic consensus via neurologists in Clinical Core.
University of Michigan	ADRC longitudinal cohort (UM Memory and Aging Project, UM-MAP), Michigan ADRC Registry, and UM clinics and local communities.	UDS procedures and NACC guidelines. Formal panel determined research consensus diagnosis.
University of Pittsburgh	Parent study staff informed participants at parent study visits on-site or via phone.	CDR & GSD, neurological exam, neuropsychological testing, and neuroimaging by the parent studies; NPI-Q, FAQ, Hachinski, UPDRS in some studies. Diagnosis confirmed via diagnostic consensus conference.
University of Wisconsin-Madison	ADRC Clinical Core study (prioritizing participants with biomarkers), ADRC recruitment and educational events, local senior centers, retirement homes,	UDS procedures and NACC guidelines; cognitive testing, CDR, and clinician reports. For participants in the Clinical Core, consensus conferences with multiple clinicians confirmed

	local memory clinics, and recruitment registry.	diagnosis. For other participants, single clinician consensus determined diagnosis.

Table 2

Demographic Characteristics

Demographic	NC (N=160)	MCI (N=97)	DAT (N=69)	p-value
*Age (years)	72.72 (5.08)	77.14 (7.34)	75.7 (7.22)	$p < .001$
*Male (%)	33.8	58.8	58	$p < 0.01$

Hispanic (%)	1.9	3.1	4.3	$p = 0.59$
White (%)	84.1	83.5	92.8	$p = 0.14$
African American (%)	14.6	2.5%	9.6%	p = 0.28
American Indian or Alaska Native	1.3%	0%	0%	p = 0.50
Asian	0%	1%	0%	p = 0.31
Primary English Lang. (%)	99.4	100	97.1	$p = 0.15$
Education				
High School or less (%)	4.3	3.6	11.9	
Some College (%)	11.7	19.3	20.9	$p < .01$
College (%)	31.9	31.3	37.3	
Graduate (%)	52.1	45.8	29.9	

Right-handed (%)	90.6	87.6	88.4	p = 0.73

[1] Although a chi-square test indicated that the three clinical groups differed in education, follow up post-hoc tests correcting for multiple comparisons did not yield significant differences.
* Indicates differences in demographic characteristic by clinical group p<.05, and confirmed by post-hoc tests using Bonferroni's correction for multiple comparisons. NC=cognitively normal older control participants; MCI= individuals with mild cognitive impairment; DAT= Dementia of the Alzheimer Type; all clinical categories based on research clinical diagnostic criteria.

Table 3

Clinical Dementia Ratings by Domain, MoCA scores, and FAQ scores

	NC	MCI	DAT	p-value
CDR Domain (mean (sd))				
*Memory	0.14 (0.26)	0.64 (0.28)	1.09 (0.45)	p < .001

*Orientation	0.02 (0.1)	0.22 (0.31)	0.86 (0.64)	$p < .001$
*Judgment and problem-solving	0.02 (0.12)	0.28 (0.34)	0.83 (0.46)	$p < .001$
*Community affairs	0 (0.04)	0.1 (0.22)	0.72 (0.52)	$p < .001$
*Home and Hobbies	0.01 (0.08)	0.12 (0.27)	0.66 (0.58)	$p < .001$
*Personal care	0 (0)	0.03 (0.17)	0.38 (0.52)	$p < .001$
*Global CDR®	0.07 (0.17)	0.48 (0.13)	0.78 (0.36)	$p < .001$
*MoCA	26.87 (2.51)	22.15 (3.05)	16.57 (5.26)	p < .001
Functional Assessment				
*FAQ (Functional Activities Questionnaire)	0.41 (1.61)	2.04 (3.35)	15.91 (7.7)	$p < .001$

Note. Summary statistics treat all variables as continuous measures. CDR scores ranged from 0-3 (0=none; 0.5=questionable; 1=mild, 2=moderate, 3=severe dementia). FAQ is a questionnaire with total scores that range from 0-30. *Indicates differences in demographic characteristic by clinical group $p<.05$, and confirmed by post-hoc tests across all pairwise comparisons across groups using Bonferroni's correction for multiple comparisons.

Table 4

Neuropsychiatric Symptoms: Percent of sample reporting each symptom

NPI-Q Symptom	NC (N=94)	MCI (N=91)	DAT (N=69)	p-value
*NPIQ Delusions (%)	0	1.1	11.6	p <.01

*NPIQ Hallucinations (%)	0	0	10.1	*p* <.01
*NPIQ Agitation or aggression (%)	2.1	13.2	34.8	*p* <.01
*NPIQ Depression or dysphoria (%)	7.5	22.2	41.2	*p* <.01
*NPIQ Anxiety (%)	7.4	20.9	40.3	*p* <.01
*NPIQ Elation or euphoria (%)	0	1.1	1.5	*p* =.53
*NPIQ Apathy or indifference (%)	4.3	17.6	43.5	*p* <.01
*NPIQ Disinhibition (%)	2.1	7.7	24.6	*p* <.01
*NPIQ Irritability or lability (%)	5.3	16.5	33.8	*p* <.01
*NPIQ Motor disturbance (%)	0	2.2	20.9	*p* <.01
*NPIQ Nighttime behaviors (%)	13.3	12.5	28.4	*p* =.02
*NPIQ Total score (max. 12)	1.05 (2.4)	1.72 (2.63)	4.59 (5.89)	*p* < .001
	NC	MCI	DAT	p-value

	(N=95)	(N=88)	(N=57)	
Geriatric Depression Scale	1.08 (1.79)	1.59 (2.07)	1.46 (1.76)	p = 0.18

Note. Numbers reported for NPIQ symptoms reflect the percent of participants within each clinical group who endorsed having experienced the symptoms within the prior month. Note that the total number of participants varies slightly for each symptom, as some individuals left answers blank. The NPIQ Total reflects the average total number of symptoms experienced by each participant (range 0-12). *Indicates differences in demographic characteristic by clinical group at p<.05, and confirmed by post-hoc tests all pairwise comparisons across groups using Bonferroni's correction for multiple comparisons.

Table 5

Biomarker availability across group

Biomarker Type	NC	MCI	DAT
Amyloid PET	92 (57%)	8 (8%)	3 (4%)
CSF	33 (21%)	35 (34%)	15 (22%)
Structural MRI	80 (50%)	74 (72%)	44 (58%)
APOE Genotype	90 (56%)	90 (87%)	66 (87%)

Note. n(%). The proportions are calculated from the total sample size for each of the respective study groups.