Temporal order of Alzheimer's disease-related cognitive marker changes in BLSA and WRAP longitudinal studies

Murat Bilgel, Ph.D.[a,*], **Rebecca L. Koscik, Ph.D.**[b], **Yang An, M.S.**[a], **Jerry L. Prince, Ph.D.**[c], **Susan M. Resnick, Ph.D.**[a], **Sterling C. Johnson, Ph.D.**[b,d,e,f], and **Bruno M. Jedynak, Ph.D.**[g]

[a]Laboratory of Behavioral Neuroscience, Intramural Research Program, National Institute on Aging, Baltimore, MD

[b]Wisconsin Alzheimer's Institute, University of Wisconsin School of Medicine and Public Health, USA, Madison, WI 53705

[c]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD

[d]Geriatric Research Education and Clinical Center, Wm. S. Middleton Veterans Hospital, USA, Madison WI 53705

[e]Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health, USA, Madison, WI 53705

[f]Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin-Madison, USA, Madison, WI 53705

[g]Department of Mathematics and Statistics, Portland State University, Portland, OR, USA

Abstract

Investigation of the temporal trajectories of currently used neuropsychological tests is critical to identifying earliest changing measures on the path to dementia due to Alzheimer's disease (AD). We used the Progression Score (PS) method to characterize the temporal trajectories of measures of verbal memory, executive function, attention, processing speed, language, and mental status using data spanning normal cognition, mild cognitive impairment (MCI), and AD from 1661 participants with a total of 7839 visits (age at last visit 77.6 SD 9.2) in the Baltimore Longitudinal Study of Aging and 1510 participants with a total of 3473 visits (age at last visit 59.5 SD 7.4) in the Wisconsin Registry for Alzheimer's Prevention. This method aligns individuals in time based on the similarity of their longitudinal measurements to reveal temporal trajectories. As a validation of our methodology, we explored the associations between the individualized cognitive progression scores (Cog-PS) computed by our method and clinical diagnosis. Digit span tests were the first to show declines in both data sets, and were detected mainly among cognitively normal individuals. These were followed by tests of verbal memory, which were in turn followed by Trail Making Tests, Boston Naming Test, and Mini-Mental State Examination. Differences in Cog-PS across the clinical diagnosis and *APOE* ε4 groups were statistically significant, highlighting the potential use of Cog-PS as individualized indicators of disease progression. Identifying cognitive

*Corresponding author: Murat Bilgel, murat.bilgel@nih.gov, National Institute on Aging, Laboratory of Behavioral Neuroscience, 251 Bayview Blvd. Suite 100, Rm 04B329, Baltimore, MD 21224, USA, Phone: 410-558-8151, Fax: 410-558-8674.

measures that are changing in preclinical AD can lead to the development of novel cognitive tests that are finely tuned to detecting earliest changes.

Keywords

Alzheimer disease; preclinical; cognition; neuropsychological tests; longitudinal studies

INTRODUCTION

Cognitive changes in Alzheimer's disease (AD), in particular declines in episodic memory, are detectable on neuropsychological testing up to fifteen years prior to clinical diagnosis [1–4]. However, cognition has only limited cross-sectional association with cerebral amyloid burden [5–11], which marks the beginning of preclinical AD according to the National Institute on Aging–Alzheimer's Association (NIA–AA) criteria [12]. Longitudinal studies have consistently shown that amyloid levels are associated with greater rates of decline on tests of episodic memory [13–21], suggesting that amyloid changes precede episodic memory declines. These findings indicate that despite the fifteen-year period prior to diagnosis for detecting cognitive change, currently used neuropsychological tests fall short of detecting changes in the earliest disease stages where therapeutic intervention is hypothesized to be most promising. Therefore, there is a need to develop cognitive tests and batteries that are more sensitive to changes in early preclinical AD and that correlate better with AD-related neuroimaging measures. Cognitive measures that are dynamic in preclinical disease can facilitate clinical trials aimed at this early stage, as they can serve as outcome measures that are non-invasive and cheaper to administer than neuroimaging evaluations.

To develop such cognitive tests or batteries, it is necessary to study the neuroimaging correlates and temporal trajectories of currently used tests to identify earliest changing measures. Various studies have investigated correlations between cognitive measures and early AD-related neuroimaging markers [22–30]. In this work, we focus on characterizing the trajectories of a collection of cognitive markers widely used in studies of aging and AD. The ideal study of preclinical AD markers would follow cognitively normal individuals until they are diagnosed with dementia due to AD, and retrospectively analyze the time courses of the markers of those who developed AD. However, currently available sample sizes do not allow for cross-validation nor yield adequate statistical power to conduct such a study given that preclinical stages span a decade or longer [19]. To overcome this limitation, several statistical analysis approaches have been developed [31–36]. An important concept in a subset of these approaches is the time-alignment of individuals. Here, we used the Progression Score Model [36,37], a multivariate nonlinear mixed effects model, to construct cognitive marker trajectories spanning normal cognition, mild cognitive impairment (MCI), and AD stages by aligning individuals in time based on the similarity of their marker profiles. The method incorporates longitudinal information in performing this alignment and accounts for inter-individual differences in baseline measurement levels and rate of progression. The main premise of the method is that rather than using age as a proxy for disease progression, we can obtain better disease progression indicators as well as temporal trajectories for the biomarkers by aligning individuals in time based on the similarity of their

biomarker profiles. Time alignment of individuals allows us to study the long-term trajectories of cognitive measures despite the availability of only a small number of individuals who have traversed large extents of the cognitive trajectories. We conducted separate analyses on two well-characterized longitudinal studies, the Baltimore Longitudinal Study of Aging and the Wisconsin Registry for Alzheimer's Prevention, to delineate the trajectories of measures of verbal memory, executive function, attention, processing speed, language, and mental status among individuals who are cognitively normal, have MCI, or have dementia due to AD.

The main aims of the analyses were to: 1) develop a standardized template of cognitive changes against which individuals can be quantified, 2) present evidence of the validity of the individualized scores obtained using the standardized cognitive template by exploring their associations with clinical diagnosis, *APOE* ε4 status, and time to dementia diagnosis, 3) identify the order in which detectable changes begin to appear across cognitive tests, and 4) validate this identified ordering using an independent approach at the individual level.

METHODS

Statistical methods used in the following analyses were applied separately to two ongoing longitudinal cohort studies of human aging: the Baltimore Longitudinal Study of Aging (BLSA) [38], initiated in 1958 and conducted by the Intramural Research Program of the National Institute on Aging, and the Wisconsin Registry for Alzheimer's Prevention (WRAP) [39], initiated in 2001 and conducted by the University of Wisconsin Alzheimer's Institute.

Participants

BLSA participants—BLSA analyses were based on data from 1661 participants (Table 1). We included visits where participants were cognitively normal or had a clinical diagnosis of MCI or dementia due to AD. Baseline and last visit were defined for each participant as the first and last BLSA visit between 1/1992 and 11/2015 where they were aged ≥ 60 and met the inclusion criteria based on clinical diagnosis and number of available cognitive scores (see Section 2.2). A total of 7839 visits were selected for analysis. The Institutional Review Board associated with the NIA Intramural Research Program approved the research protocol for this study, and informed consent was obtained at each visit from all participants.

WRAP participants—WRAP analyses were based on data from 1510 participants (Table 1), 72.6% of whom had a family history of AD. We included all visits where participants were cognitively normal or had a clinical diagnosis of MCI (including an "early MCI" diagnosis, defined in Section 2.3) or dementia due to AD. Any visit that had not been assigned a consensus diagnosis at the time of the analyses was excluded. Baseline and last visit were defined for each participant as the first and last WRAP visit up until Wave 5 (inclusive) where they met the inclusion criteria based on clinical diagnosis and number of available cognitive scores (see Section 2.2). A total of 3473 visits were selected for analysis. All activities for this study were approved by the Institutional Review Board and completed in accordance with the Helsinki Declaration.

Cognitive measures

Eight cognitive measures were selected for our analyses: Trail Making Tests (Trails) A and B [40] to assess processing speed and executive function, Wechsler Adult Intelligence Scale (WAIS) digit span forward and backward (revised edition [41] in BLSA, 3rd edition [42] in WRAP) to assess attention and executive function, California Verbal Learning Test (CVLT) [43] in BLSA and Rey Auditory Verbal Learning Test (AVLT) [44] in WRAP measuring immediate recall (sum of total recall across five learning trials) and 20-minute delayed recall to assess verbal memory, Boston Naming Test [45] to assess language, and Mini-Mental State Examination (MMSE) [46] to assess mental status. Trails A and B were truncated at 300 seconds and log-transformed. We included visits where at least four of these eight measures were available. We did not include WAIS digit symbol substitution [41,42] or animal fluency [47] in our analyses even though they were common to both data sets, since they were introduced in later stages in BLSA and/or WRAP.

CVLT was added to the BLSA battery in 3/1993 and digit span was added in 1/1992, while the remainder of the measures were added between 5/1984 and 1/1990. BLSA participants under age 60 did not receive the Trail Making Test until 2005. MMSE was added to the WRAP battery in Wave 2, while the remaining eight measures were available starting with Wave 1 in 2001.

Clinical Diagnoses

BLSA methods—All BLSA participants are reviewed for cognitive impairment at a consensus case conference if they have a score 4 on the Blessed Information–Memory–Concentration Test [48], if their Clinical Dementia Rating (CDR) score [49] is 0.5 on the subject or informant report or if they screen abnormal on the Dementia Questionnaire. Consensus diagnoses are determined by a research team that includes neurologists, psychiatrists, neuropsychologists, research nurses, and research assistants. Diagnoses of dementia and AD are based on Diagnostic and Statistical Manual of Mental Disorders 3rd edition–revised [50] and National Institute of Neurological and Communication Disorders and Stroke–AD and Related Disorders Association (NINCDS-ADRDA) [51] criteria, respectively. Mild cognitive impairment (MCI) is diagnosed based on the Petersen criteria [52] when cognitive impairment is evident for a single domain or multiple domains without significant functional loss in activities of daily living. Out of the eight cognitive measures described in Section 2.2, only Trails A and B, Boston Naming, and MMSE are used in clinical diagnosis determination.

WRAP methods—WRAP participant visits are reviewed at a consensus case conference if they meet one or more of the following criteria: 1) cognitive abnormalities relative to WRAP peers (i.e., at least 1.5 standard deviations below expected relative to robust internal norms adjusting for age, sex, and literacy-level on the most recent assessment for factor scores or individual measures of memory, executive function, language, working memory, or attention [53], [54]; 2) cognitive performance on one or more tests below values used in other studies as cutpoints for mild cognitive impairment (MCI) diagnoses (e.g., WMS-R Logical Memory II [55] story A score <9, Alzheimer's Disease Neuroimaging Initiative [56]); or 3) an abnormal informant report indicating subjective cognitive or functional decline. Consensus

diagnoses are determined by a research team that includes physicians, clinical neuropsychologists, and clinical nurse practitioners. Cognitive status includes cognitively normal, early MCI, clinical MCI, other cognitive impairment (e.g., due to other medical conditions), or dementia. The diagnosis of clinical MCI is based on NIA-AA criteria [57] and includes a) concern regarding change in cognition, b) impairment in one or more cognitive domains, c) preservation of functional abilities, and d) does not meet criteria for dementia. The status of early MCI was developed to identify individuals in the cohort who exhibit lower than expected objective performance in one or more cognitive domains (relative to internal robust norms), but may not yet report subjective cognitive complaints. This experimental construct is thought to represent a phenotype of early cognitive decline expected to precede a clinical diagnosis of MCI [54,58]. Dementia diagnosis is based on NINCDS-ADRDA criteria [51]. Both demented WRAP individuals included in these analyses had dementia due to AD. The consensus review process was initiated in late 2012.

Statistical analyses

Removal of cognitive effects of aging and other covariates in the absence of neurodegenerative disease—To estimate the cognitive effects of aging in the absence of neurodegenerative disease, we selected a subset of visits corresponding to individuals we refer to as the "robust normal group". For inclusion in this group, individuals needed to have at least two years of cognitive follow-up in which they remained cognitively normal. Individuals whose clinical diagnosis reverted from any type of cognitive impairment to cognitively normal were excluded. All visits prior to impairment, excluding last two years of cognitively normal follow-up, were included in the subset.

Linear mixed effects models were fitted separately for each cognitive measure using this subset. Fixed effects included baseline age, sex, race, years of education, time from baseline in years, and interactions of time with baseline age and sex. These variables were selected based on previous reports demonstrating associations with cognition [59]. Random effects included intercept per individual. Estimated fixed effects were used to make predictions for each individual and visit in the larger sample, and these predictions were subtracted from observed cognitive measures to yield residuals.

Progression score model—We used an improved version [37] of the Progression Score (PS) method [36] to compute cognitive progression scores (Cog-PS) for individuals in BLSA and WRAP in separate models using residuals for the eight cognitive tests described in Section 2.2 (see Supplementary Material for model details). The Progression Score method is based on a multivariate model that enables the computation of a score for each visit using a collection of longitudinal biomarker measures to reflect the state of the visit relative to the rest of the sample. Cog-PS is an affine transformation of age, and these transformations are allowed to vary across individuals via subject-specific variables that model inter-subject differences in rate and baseline levels of progression. The scale of the Cog-PS at the end of the model fitting procedure is arbitrary (i.e., a Cog-PS of 0 does not convey any meaning on its own, but compared to a Cog-PS of 1, it indicates better overall performance on the cognitive measures included in the model). To give meaning to the Cog-PS scale, we shifted all Cog-PS values such that a Cog-PS of 0 corresponds to the mean

value across cognitively normal individuals, and then scaled all Cog-PS values such that the standard deviation of Cog-PS among cognitively normal individuals is 1. The Cog-PS calibration, when combined with the appropriate transformations of the model parameters, does not affect the likelihood of the model (i.e., the calibrated and uncalibrated models are identical in terms of their fit to the observed data). This calibration is intended to make Cog-PS values computed on separate data sets comparable, with the underlying assumption that cognitively normal groups are comparable. At the end of model fitting, we obtained Cog-PS for each visit as well as estimates of cognitive measure trajectories as a function of Cog-PS. The overall procedure is summarized in Figure 1.

Temporal ordering of cognitive measures—To compare the estimated trajectories for the eight cognitive measures, we linearly scaled them such that their estimated values at the minimum and maximum Cog-PS observed in the sample were 0 and 1, respectively. In this standardized space, we refer to a scaled cognitive value of 0 as "normal" and a scaled cognitive value of 1 as "abnormal". We visually evaluated the estimated trajectories to make determinations about temporal ordering of cognitive changes. To complement this qualitative evaluation, we developed a quantitative procedure based on bootstrapping. The number of bootstrap experiments was limited to 20 due to the time-consuming model fitting procedure. For each bootstrap, we computed the PS value at which the estimated trajectory attained the midway point of 0.5 in the standardized normal-to-abnormal scale for each cognitive measure. We compared the range of these PS values across biomarkers to make quantitative assessments about ordering.

Validation of Cog-PS results—To show evidence for the validity of the individualized Cog-PS, we first investigated differences in Cog-PS cross-sectionally (at last visit) across diagnosis and *APOE* ε4 (carrier of at least one ε4 allele vs. non-carrier) groups using two-sided Wilcoxon ranksum tests. *APOE* ε4 is the most influential known genetic risk factor for AD, with each ε4 allele leading to an earlier onset of the disease by about a decade [60,61]. Since this variable was not used in the PS model, and since it is not dependent on the cognitive measures that went into our analyses, it served as an unbiased validator of Cog-PS as a disease stage indicator. Finally, using data for 87 BLSA individuals who had an AD diagnosis at their last visit, we investigated the relationship between Cog-PS at baseline and years to dementia diagnosis using partial Pearson correlation, adjusting for age.

To provide further evidence for the temporal ordering determined by the Cog-PS model, we performed an analysis using linear mixed effects models that was independent from the Cog-PS approach. These models allowed us to assess the ordering of a pair of cognitive measures. For a pair of measures both of which decline over time, we expected to find a negative association between baseline values of the later changing cognitive measure and the rate of change of the earlier changing measure. We also expected to find a positive association between baseline values of the earlier changing measure and the rate of change of the later changing measure. These concepts are illustrated in Figure S1. Longitudinal measures on one of the two cognitive tests being compared constituted the outcome in the linear mixed effects model. Covariates were included for baseline age, time (measured from baseline), sex, years of education, interaction of baseline age with time, and interaction of

sex with time. Baseline measure on the other cognitive test, as well as its interaction with time, were included as predictors. Random effects included intercept per individual. We used the estimates and p-values for the time×baseline cognitive measure interaction term to assess our hypothesis about temporal ordering.

RESULTS

Temporal ordering of cognitive measures

Estimated cognitive marker trajectories as function of Cog-PS are presented in Figure 2. While digit span tests have a marked downward slope in the left-hand side of the figure at a Cog-PS of –1, other cognitive measures do not begin exhibiting marked declines until later along the Cog-PS scale. CVLT/AVLT measures are next to exhibit declines. These are followed by Trails, Boston Naming, and MMSE in BLSA, and MMSE, Trails, and Boston Naming in WRAP.

The standardized cognitive template, obtained by scaling the cognitive measures based on fitted values corresponding to the minimum and maximum Cog-PS values in the sample, as well as the bootstrap results, are presented in Figure 3. In both BLSA and WRAP, digit span forward and backward were the first measures to attain the midway point of the biomarker abnormality scale, followed by CVLT/AVLT immediate and delayed recall.

Validation of Cog-PS results

Cog-PS at last visit was associated with concurrent diagnosis, with cognitively normal individuals having lower Cog-PS compared to impaired individuals (Figure 4). All pairwise group comparisons (normal vs. CI/MCI, MCI vs. AD, and normal vs. AD) were significant within each data set (two-sided Wilcoxon rank-sum test, all $p<.0004$). Cog-PS was higher in the *APOE* ε4+ group compared to non-carriers in both BLSA (ε4- mean 0.320 SD 1.089 vs. ε4- mean 0.161 SD 1.077, two-sided Wilcoxon rank-sum test $p=0.015$) and WRAP (ε4+ mean 0.235 SD 0.995 vs. ε4- mean 0.116 SD 1.020, $p=0.02$). Years to dementia diagnosis was negatively correlated with Cog-PS at baseline in BLSA (partial Pearson correlation given age=−0.61, $p=<.00001$).

We used linear mixed effects models to further assess the finding that digit span measures are the earliest measures to change among of the eight considered in this study. To this end, we investigated their temporal relationship with the CVLT/AVLT measures. Model results are presented in Table S1. As we expected, in BLSA, we found that higher baseline scores on digit span forward were associated with attenuated declines on CVLT immediate (time×baseline digit span forward estimate=0.027, SE=0.011, $p=0.012$) as well as delayed recall (estimate=0.0066, SE=0.0030, $p=0.026$), and that higher baseline CVLT delayed recall scores were associated with steeper declines on digit span forward (time×baseline CVLT delayed recall estimate=−0.0043, SE=0.0021, $p=0.043$). We did not find any associations with digit span backward or with AVLT in WRAP.

DISCUSSION

We applied the progression score model to align individuals in time based on the similarity of their cognitive marker profiles to learn long-term cognitive trajectories from shorter-term longitudinal measurements. From this model, we obtained individualized scores, termed Cog-PS, which we validated by demonstrating their association with clinical diagnosis, *APOE* ε4 status, and time to dementia diagnosis. Our analyses were conducted separately on data from two well-characterized longitudinal studies of aging, BLSA and WRAP, and revealed similar results.

Cognitive measures reflect not only disease-related changes but also changes due to aging. While we used residualized measures in our analyses to distill effects due to disease, changes due to aging may not have been completely removed. Therefore, our results regarding the ordering of cognitive measure changes may be confounded by changes due to aging rather than due to disease.

Using the estimated trajectories and the approximate confidence intervals obtained via bootstrapping, we showed that digit span tests exhibit their most extensive dynamic range while individuals are cognitively normal, and that these changes precede those observed on verbal memory tests, which in turn precede changes on Trails and Boston naming. MMSE was the last to change among the eight measures in BLSA, but in WRAP, changes in MMSE followed those in AVLT. We must consider the limited range of MMSE values observed in WRAP data set when interpreting this discrepancy. As evident in the MMSE panel in Figure 2b, the MMSE residual attained at highest values of Cog-PS in WRAP is about −1, whereas in BLSA, MMSE exhibits a much wider range, with highest values of Cog-PS corresponding to an MMSE residual of about −5. This means when we standardize the cognitive measures to visualize their trajectories on a common *y*-axis as presented in Figure 3, we base the standardization of the MMSE trajectory on very different values in the two data sets. Therefore, the standardized MMSE trajectory reflects the path to an MMSE residual of −1 in WRAP, and to −5 in BLSA. This explains the difference between WRAP and BLSA in the temporal order of MMSE relative to the other cognitive measures.

Our finding that digit span changes occur prior to verbal memory changes was further supported by our linear mixed effects modeling approach. Neither the progression score model nor the linear mixed effects model is a perfect representation of the disease process. There are different sets of underlying assumptions in each approach. However, results of these different techniques, taken collectively, provide complementary perspectives through which we can build confidence regarding our results about temporal ordering.

The progression score approach enabled the computation of an individualized indicator of cognitive performance based on a collection of longitudinal measurements. We demonstrated that there are statistically significant differences in progression score across cognitively normal, CI/MCI, and AD groups. Since the progression score model is agnostic to clinical diagnosis, this analysis served as a validation of the computed scores. However, it should be noted that this is a partial validation since there is a degree of circularity as four out of the eight cognitive measures were used in diagnosis determination in BLSA, and eight

were used in WRAP. Our demonstration of associations between Cog-PS and time to dementia diagnosis as well as *APOE* ε4 status avoided this circularity and provided further evidence that Cog-PS is indicative of the disease process.

While we used a calibration approach to render the progression scores computed on BLSA and WRAP comparable, it is important to note that due to differences in the composition of cognitively normal groups in the two data sets, Cog-PS values do not convey the same meaning across the studies. Despite this limitation of our methodology, we found similar temporal cognitive patterns in both studies. Cognitively normal as well as the CI/MCI groups in BLSA had lower cognitive performance on all tests compared to WRAP. As mentioned previously in the discussion of the MMSE trajectories, our scaling procedure for establishing correspondences across cognitive markers is sensitive to the range of cognitive measurements present in the study. Since this scaling procedure is intended to identify normal and abnormal cognitive test values, it is important that the data set contain measurements spanning the entire normal-to-abnormal range for each cognitive test. This condition is not fully satisfied in either data set we considered in this work, particularly in WRAP since WRAP participants are about 18 years younger than BLSA participants on average. Therefore, the scaling procedure may be less accurate in WRAP than in BLSA, especially for later-changing measures such as Boston naming and MMSE.

Several previous studies have reported that changes in digit span tests are detectable prior to a clinical diagnosis of cognitive impairment. For example, a study that evaluated non-demented individuals with subjective memory complaints found that those with normal digit span scores (defined using age- and education-adjusted neuropsychological test scores) at baseline did not exhibit significant declines on verbal memory, visual memory, or executive function after a mean follow-up of 6.6 years, but had significant declines on the sum score of digit span forward and backward [62]. On the other hand, age-, sex-, and education-matched individuals with impaired digit span scores at baseline had significant declines on tests of verbal learning and animal fluency but not on any other cognitive test. These findings agree with our estimated ordering of cognitive trajectories on the path to cognitive impairment, with digit span measures declining first, followed by measures of verbal memory. Another study that investigated changes in cognition prior to autopsy in a sample of individuals who remained cognitively normal found that while longitudinal changes were not significant when assessed separately among individuals with and without AD neuropathology, longitudinal decline in the attention/working memory domain (assessed using digit span forward and backward tests) was greater among those with neuropathology compared to those without [63]. Other domains, including episodic memory, language, and executive function, did not show statistically significant longitudinal differences between the two groups [63].

A limitation of our method is that it does not indicate whether the detected cognitive changes are due to AD-related mechanisms, and further studies are needed to delineate normal aging versus disease. Therefore, our results cannot be interpreted as evidence for using digit span to predict individualized diagnoses. Digit span is not a good predictor of concurrent diagnosis; computing cognitive measure thresholds based on our Cog-PS model and using these thresholds to classify individuals as normal versus CI/MCI/AD would yield

many false positives. This lack of specificity of digit span for dementia has been documented previously [64,65]. What our results demonstrate is that changes in digit span are most evident early on in cognitively normal stages. Detectability of these changes and their associations with future outcomes at the individual-level, including AD diagnosis in following years or decades, remain to be fully elucidated; however, prior studies have reported that sensitivity to change on digit span is small [66]. Another limitation of our method is that it treats discrete cognitive measures such as digit span as continuous variables, and does not account for practice effects seen on serial administrations of neuropsychological tests. These factors may introduce bias into the characterization of the longitudinal cognitive trajectories. Despite these limitations, understanding the association of digit span tests with brain changes can be informative for designing novel cognitive tests or batteries that are more sensitive to changes in preclinical AD and that correlate with functional and structural brain changes over the course of disease.

Several studies of a small number of healthy individuals found that higher performance on digit span backward is associated with activation of the right dorsolateral prefrontal cortex (DLPFC) [67–69], bilateral inferior parietal lobule, anterior cingulate, left DLPFC, and Broca's area, with a subset of these regions also implicated in relation to digit span forward [68,69]. DLPFC is one of the amyloid accumulating regions, and therefore it may be possible to demonstrate associations between digit span performance and DLPFC amyloid levels among cognitively normal individuals in studies with large sample sizes. Analyses conducted in parallel in multiple data sets or in combined samples will accelerate efforts to further elucidate the relationships among cognitive measures implicated in preclinical AD. Identifying cognitive measures that are dynamic in preclinical AD can lead to the development of novel cognitive tests that are finely tuned to detecting earliest changes. Such measures will facilitate clinical trials aimed at this early stage by serving as outcome measures that are non-invasive and cheaper to administer than neuroimaging.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Bäckman L, Jones S, Berger A-K, Laukka EJ, Small BJ. Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. Neuropsychology. 2005; 19:520–31. [PubMed: 16060827]

2. Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. J. Int. Neuropsychol. Soc. 2008; 14:266–78. [PubMed: 18282324]

3. Kawas CH, Corrada MM, Brookmeyer R, Morrison a, Resnick SM, Zonderman aB, Arenberg D. Visual memory predicts Alzheimer's disease more than a decade before diagnosis. Neurology. 2003; 60:1089–1093. [PubMed: 12682311]

4. Elias MF, Beiser A, Wolf PA, Au R, White RF, D'Agostino RB. The preclinical phase of Alzheimer disease: A 22-year prospective study of the Framingham Cohort. Arch. Neurol. 2000; 57:808–813. [PubMed: 10867777]

5. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloidcognition relations in cognitively normal older adults. Neurology. 2013; 80:1341–1348. [PubMed: 23547267]

6. Johnson SC, Christian BT, Okonkwo OC, Oh JM, Harding S, Xu G, Hillmer AT, Wooten DW, Murali D, Barnhart TE, Hall LT, Racine AM, Klunk WE, Mathis CA, Bendlin BB, Gallagher CL, Carlsson CM, Rowley HA, Hermann BP, Dowling NM, Asthana S, Sager MA. Amyloid burden and neural function in people at risk for Alzheimer's Disease. Neurobiol. Aging. 2014; 35:576–584. [PubMed: 24269021]

7. Perrotin A, Mormino EC, Madison CM, Hayenga AO, Jagust WJ. Subjective cognition and amyloid deposition imaging: A Pittsburgh compound B positron emission tomography study in normal elderly individuals. Arch. Neurol. 2012; 69:223–229. [PubMed: 22332189]

8. Song Z, Insel PS, Buckley S, Yohannes S, Mezher A, Simonson A, Wilkins S, Tosun D, Mueller S, Kramer JH, Miller BL, Weiner MW. Brain Amyloid-β Burden Is Associated with Disruption of Intrinsic Functional Connectivity within the Medial Temporal Lobe in Cognitively Normal Elderly. J. Neurosci. 2015; 35:3240–3247. [PubMed: 25698758]

9. Amariglio RE, Becker JA, Carmasin J, Wadsworth LP, Lorius N, Sullivan C, Maye JE, Gidicsin C, Pepin LC, Sperling RA, Johnson KA, Rentz DM. Subjective cognitive complaints and amyloid burden in cognitively normal older individuals. Neuropsychologia. 2012; 50:2880–2886. [PubMed: 22940426]

10. Doraiswamy PM, Sperling RA, Coleman RE, Johnson KA, Reiman EM, Davis MD, Grundman M, Sabbagh MN, Sadowsky CH, Fleisher AS, Carpenter A, Clark CM, Joshi AD, Mintun MA, Skovronsky DM, Pontecorvo MJ. Amyloid-β assessed by florbetapir F 18 PET and 18-month cognitive decline: A multicenter study. Neurology. 2012; 79:1636–1644. [PubMed: 22786606]

11. Sperling RA, Johnson KA, Doraiswamy PM, Reiman EM, Fleisher AS, Sabbagh MN, Sadowsky CH, Carpenter A, Davis MD, Lu M, Flitter M, Joshi AD, Clark CM, Grundman M, Mintun MA, Skovronsky DM, Pontecorvo MJ. Amyloid deposition detected with florbetapir F 18 (18F-AV-45) is related to lower episodic memory performance in clinically normal older individuals. Neurobiol. Aging. 2013; 34:822–831. [PubMed: 22878163]

12. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. 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. Alzheimer's Dement. 2011; 7:280–292. [PubMed: 21514248]

13. Lim YY, Maruff P, Pietrzak RH, Ames D, Ellis Ka, Harrington K, Lautenschlager NT, Szoeke C, Martins RN, Masters CL, Villemagne VL, Rowe CC. Effect of amyloid on memory and non-memory decline from preclinical to clinical Alzheimer's disease. Brain. 2014; 137:221–231. [PubMed: 24176981]

14. Lim YY, Villemagne VL, Pietrzak RH, Ames D, Ellis KA, Harrington K, Snyder PJ, Martins RN, Masters CL, Rowe CC, Maruff P. APOE ɛ4 moderates amyloid-related memory decline in preclinical Alzheimer's disease. Neurobiol. Aging. 2015; 36:1239–1244. [PubMed: 25559335]

15. Mormino EC, Betensky RA, Hedden T, Schultz AP, Ward A, Huijbers W, Rentz DM, Johnson KA, Sperling RA. Amyloid and APOE ɛ4 interact to influence short-term decline in preclinical Alzheimer disease. Neurology. 2014; 82:1760–1767. [PubMed: 24748674]

16. Pietrzak RH, Lim YY, Neumeister A, Ames D, Ellis KA, Harrington K, Lautenschlager NT, Restrepo C, Martins RN, Masters CL, Villemagne VL, Rowe CC, Maruff P. Amyloid-β, anxiety,

and cognitive decline in preclinical Alzheimer disease: a multicenter, prospective cohort study. JAMA Psychiatry. 2015; 72:284–291. [PubMed: 25629787]

17. Roe C, Fagan A, Grant E, Hassenstab J, Moulder K, Dreyfus DM, Sutphen C, Benzinger T, Mintun M, Holtzman D, Morris J. Amyloid imaging and CSF biomarkers in predicting cognitive impairment up to 7.5 years later. Neurology. 2013; 80:1784–1791. [PubMed: 23576620]

18. Stonnington CM, Chen K, Lee W, Locke DEC, Dueck AC, Liu X, Roontiva A, Fleisher AS, Caselli RJ, Reiman EM. Fibrillar amyloid correlates of preclinical cognitive decline. Alzheimer's Dement. 2014; 10:e1–e8. [PubMed: 23583233]

19. Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, Szoeke C, Macaulay SL, Martins R, Maruff P, Ames D, Rowe CC, Masters CL. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol. 2013; 12:357–367. [PubMed: 23477989]

20. Yotter RA, Doshi J, Clark V, Sojkova J, Zhou Y, Wong DF, Ferrucci L, Resnick SM, Davatzikos C. Memory decline shows stronger associations with estimated spatial patterns of amyloid deposition progression than total amyloid burden. Neurobiol. Aging. 2013; 34:2835–2842. [PubMed: 23859610]

21. Resnick SM, Sojkova J, Zhou Y, An Y, Ye W, Holt DP, Dannals RF, Mathis CA, Klunk WE, Ferrucci L, Kraut MA, Wong DF. Longitudinal cognitive decline is associated with fibrillar amyloid-beta measured by [11C]PiB. Neurology. 2010; 74:807–815. [PubMed: 20147655]

22. Besson FL, La Joie R, Doeuvre L, Gaubert M, Mezenge F, Egret S, Landeau B, Barre L, Abbas A, Ibazizene M, de La Sayette V, Desgranges B, Eustache F, Chetelat G. Cognitive and brain profiles associated with current neuroimaging biomarkers of preclinical Alzheimer's disease. J. Neurosci. 2015; 35:10402–10411. [PubMed: 26203136]

23. Toledo JB, Bjerke M, Chen K, Rozycki M, Jack CR, Weiner MW, Arnold SE, Reiman EM, Davatzikos C, Shaw LM, Trojanowski JQ. Memory, executive, and multidomain subtle cognitive impairment: Clinical and biomarker findings. Neurology. 2015; 85:144–153. [PubMed: 26085606]

24. Wirth M, Villeneuve S, Haase CM, Madison CM, Oh H, Landau SM, Rabinovici GD, Jagust WJ. Associations between Alzheimer disease biomarkers, neurodegeneration, and cognition in cognitively normal older people. JAMA Neurol. 2013; 70:1512–1519. [PubMed: 24166579]

25. Doré V, Villemagne VL, Bourgeat P, Fripp J, Acosta O, Chetélat G, Zhou L, Martins R, Ellis Ka, Masters CL, Ames D, Salvado O, Rowe CC. Cross-sectional and longitudinal analysis of the relationship between Aβ deposition, cortical thickness, and memory in cognitively unimpaired individuals and in Alzheimer disease. JAMA Neurol. 2013; 70:903–911. [PubMed: 23712469]

26. Buckner RL. Molecular, Structural, and Functional Characterization of Alzheimer's Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory. J. Neurosci. 2005; 25:7709–7717. [PubMed: 16120771]

27. Fleisher AS, Sherzai A, Taylor C, Langbaum JBS, Chen K, Buxton RB. Resting-state BOLD networks versus task-associated functional MRI for distinguishing Alzheimer's disease risk groups. Neuroimage. 2009; 47:1678–1690. [PubMed: 19539034]

28. Kennedy KM, Rodrigue KM, Devous MD, Hebrank AC, Bischof GN, Park DC. Effects of beta-amyloid accumulation on neural function during encoding across the adult lifespan. Neuroimage. 2012; 62:1–8. [PubMed: 22569063]

29. Oh H, Madison C, Villeneuve S, Markley C, Jagust WJ. Association of gray matter atrophy with age, β-amyloid, and cognition in aging. Cereb. Cortex. 2014; 24:1609–1618. [PubMed: 23389995]

30. Ahn H-J, Seo SW, Chin J, Suh MK, Lee BH, Kim ST, Im K, Lee J-M, Lee JH, Heilman KM, Na DL. The cortical neuroanatomy of neuropsychological deficits in mild cognitive impairment and Alzheimer's disease: a surface-based morphometric analysis. Neuropsychologia. 2011; 49:3931–3945. [PubMed: 22019776]

31. Schiratti, J-B., Allassonniere, S., Colliot, O., Durrleman, S. Learning spatiotemporal trajectories from manifold-valued longitudinal data. In: Cortes, C.Lawrence, N.Lee, D.Sugiyama, M., Garnett, R., editors. Advances in Neural Information Processing Systems. Vol. 28. 2015. p. 2395-2403.

32. Schulam P, Wigley F, Saria S. Clustering longitudinal clinical marker trajectories from electronic health data: Applications to phenotyping and endotype discovery. AAAI Conference on Artificial Intelligence. 2015:2956–2964.

33. Donohue MC, Jacqmin-Gadda H, Le Goff M, Thomas RG, Raman R, Gamst AC, Beckett La, Jack CR, Weiner MW, Dartigues JF, Aisen PS. Estimating long-term multivariate progression from short-term data. Alzheimer's Dement. 2014; 10:S400–S410. [PubMed: 24656849]

34. Schmidt-Richberg A, Guerrero R, Ledig C, Molina-Abril H, Frangi AF, Rueckert D. Multi-stage biomarker models for progression estimation in Alzheimer's disease. Lect. Notes Comput. Sci. (9123), Inf. Process. Med. Imaging. 2015; 9123:387–398.

35. Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, Schott JM, Alexander DC. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain. 2014; 137:2564–2577. [PubMed: 25012224]

36. Jedynak BM, Lang A, Liu B, Katz E, Zhang Y, Wyman BT, Raunig D, Jedynak CP, Caffo B, Prince JL. A computational neurodegenerative disease progression score: Method and results with the Alzheimer's Disease Neuroimaging Initiative cohort. Neuroimage. 2012; 63:1478–1486. [PubMed: 22885136]

37. Bilgel M, Prince JL, Wong DF, Resnick SM, Jedynak BM. A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. Neuroimage. 2016; 134:658–670. [PubMed: 27095307]

38. Shock, NW., Greulich, RC., Andres, R., Arenberg, D., Costa, PT., Jr, Lakatta, EG., Tobin, JD. Normal human aging: The Baltimore Longitudinal Study of Aging. U.S. Government Printing Office; Washington, DC: 1984.

39. Sager MA, Hermann B, La Rue A. Middle-aged children of persons with Alzheimer's disease: APOE genotypes and cognitive function in the Wisconsin Registry for Alzheimer's Prevention. J. Geriatr. Psychiatry Neurol. 2005; 18:245–249. [PubMed: 16306248]

40. Reitan, RM. Trail making test: Manual for administration and scoring. Reitan Neuropsychological Laboratory; Tucson, AZ: 1992.

41. Wechsler, D. Wechsler adult intelligence scale-revised. Psychological Corporation; San Antonio, TX: 1981.

42. Wechsler, D. Wechsler adult intelligence scale. The Psychological Corporation; San Antonio, TX: 1997.

43. Delis, DC., Kramer, JH., Kaplan, E., Ober, BA. The California Verbal Learning Test. Psychological Corporation; San Antonio, TX: 1987.

44. Lezak, M., Howieson, D., Loring, DW. Neuropsychological assessment. Oxford University Press; London: 2004.

45. Kaplan, E., Goodglass, H., Weintraub, S. Boston naming test. Lea & Febiger; Philadelphia, PA: 1983.

46. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 1975; 12:189–198. [PubMed: 1202204]

47. Newcombe, F. Missile wounds of the brain: A study of psychological deficits. Oxford University Press; Oxford, UK: 1969.

48. Fuld, PA. Psychological testing in the differential diagnosis of the dementias. In: Katzman, R.Terry, RD., Bick, KL., editors. Alzheimer's disease: Senile dementia and related disorders. Raven Press; New York, NY: 1978. p. 185-193.

49. Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. Neurology. 1993; 43:2412–2414.

50. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-III-R. American Psychiatric Association; Washington, DC: 1987.

51. 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; 34:939–944. [PubMed: 6610841]

52. Petersen RC. Mild cognitive impairment as a diagnostic entity. J. Intern. Med. 2004; 256:183–194. [PubMed: 15324362]

53. Clark LR, Koscik RL, Nicholas CR, Okonkwo OC, Engelman CD, Bratzke LC, Hogan KJ, Mueller KD, Bendlin BB, Carlsson CM, Asthana S, Sager MA, Hermann BP, Johnson SC. Mild Cognitive

Impairment in late middle age in the Wisconsin Registry for Alzheimer's Prevention (WRAP) study: Prevalence and characteristics using robust and standard neuropsychological normative data. Arch. Clin. Neuropsychol. 2016; 31:675–688.

54. Koscik RL, La Rue A, Jonaitis EM, Okonkwo OC, Johnson SC, Bendlin BB, Hermann BP, Sager MA. Emergence of mild cognitive impairment in late middle-aged adults in the Wisconsin Registry for Alzheimer's Prevention. Dement. Geriatr. Cogn. Disord. 2014; 38:16–30. [PubMed: 24556849]

55. Wechsler, D. Wechsler memory scale-revised. Psychological Corporation; San Antonio, TX: 1987.

56. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, Jack CR, Jagust WJ, Shaw LM, Toga AW. Alzheimer's Disease Neuroimaging Initiative (ADNI) clinical characterization. Neurology. 2010; 74:201–209. [PubMed: 20042704]

57. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH. 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 Dement. 2011; 7:270–279. [PubMed: 21514249]

58. Clark LR, Racine AM, Koscik RL, Okonkwo OC, Engelman CD, Carlsson CM, Asthana S, Bendlin BB, Chappell R, Nicholas CR, Rowley HA, Oh JM, Hermann BP, Sager MA, Christian BT, Johnson SC. Beta-amyloid and cognitive decline in late middle age: Findings from the Wisconsin Registry for Alzheimer's Prevention study. Alzheimer's Dement. 2016; 12:805–814. [PubMed: 26806386]

59. McCarrey AC, An Y, Kitner-Triolo MH, Ferrucci L, Resnick SM. Sex differences in cognitive trajectories in clinically normal older adults. Psychol. Aging. 2016; 31:166–175. [PubMed: 26796792]

60. Corder EH, Lannfelt L, Bogdanovic N, Fratiglioni L, Mori H. The role of APOE polymorphisms in late-onset dementias. Cell. Mol. Life Sci. 1998; 54:928–934. [PubMed: 9791536]

61. Liu C-C, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat. Rev. Neurol. 2013; 9:106–118. [PubMed: 23296339]

62. Kurt P, Yener G, Oguz M. Impaired digit span can predict further cognitive decline in older people with subjective memory complaint: a preliminary result. Aging Ment. Health. 2011; 15:364–369. [PubMed: 21491221]

63. Monsell SE, Mock C, Hassenstab J, Roe CM, Cairns NJ, Morris JC, Kukull W. Neuropsychological changes in asymptomatic persons with Alzheimer disease neuropathology. Neurology. 2014; 83:434–440. [PubMed: 24951474]

64. Kiewel NA, Wisdom NM, Bradshaw MR, Pastorek NJ, Strutt AM. A retrospective review of digit span-related effort indicators in probable Alzheimer's disease patients. Clin. Neuropsychol. 2012; 26:965–974. [PubMed: 22703555]

65. Loring DW, Goldstein FC, Chen C, Drane DL, Lah JJ, Zhao L, Larrabee GJ. False-Positive Error Rates for Reliable Digit Span and Auditory Verbal Learning Test Performance Validity Measures in Amnestic Mild Cognitive Impairment and Early Alzheimer Disease. Arch. Clin. Neuropsychol. 2016; 31:313–331. [PubMed: 27084732]

66. Bossers WJR, van der Woude LHV, Boersma F, Scherder EJA, van Heuvelen MJG. Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: a systematic review. Dement. Geriatr. Cogn. Dis. Extra. 2012; 2:589–609. [PubMed: 23341825]

67. Hoshi Y, Oda I, Wada Y, Ito Y, Yutaka Yamashita, Oda M, Ohta K, Yamada Y, Mamoru Tamura. Visuospatial imagery is a fruitful strategy for the digit span backward task: A study with near-infrared optical tomography. Cogn. Brain Res. 2000; 9:339–342.

68. Gerton BK, Brown TT, Meyer-Lindenberg A, Kohn P, Holt JL, Olsen RK, Berman KF. Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. Neuropsychologia. 2004; 42:1781–1787. [PubMed: 15351627]

69. Aleman A, van't Wout M. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex disrupts digit span task performance. Neuropsychobiology. 2008; 57:44–8. [PubMed: 18451637]

ABBREVIATIONS

|	
AD	Alzheimer's disease
AVLT	Rey Auditory Verbal Learning Test
BLSA	Baltimore Longitudinal Study of Aging
CI	Cognitive impairment
Cog-PS	Cognitive progression score
CVLT	California Verbal Learning Test
MCI	Mild cognitive impairment
MMSE	Mini-Mental State Examination
WRAP	Wisconsin Registry for Alzheimer's Prevention
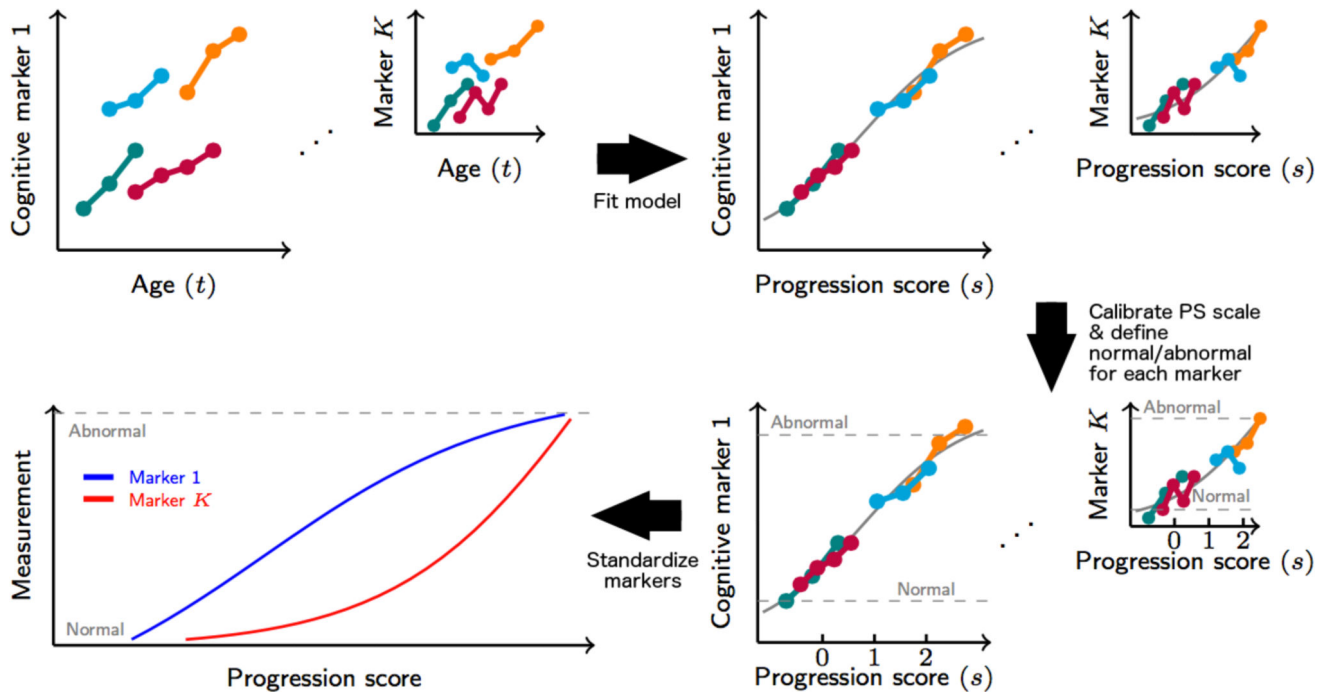
Figure 1.
Diagram illustrating progression score (PS) model fitting, PS calibration, and cognitive measurement scaling to obtain standardized space of cognitive markers. Lower values for the illustrated cognitive markers indicate lower cognitive performance. PS values are calibrated such that lower progression scores indicate better overall cognitive performance.
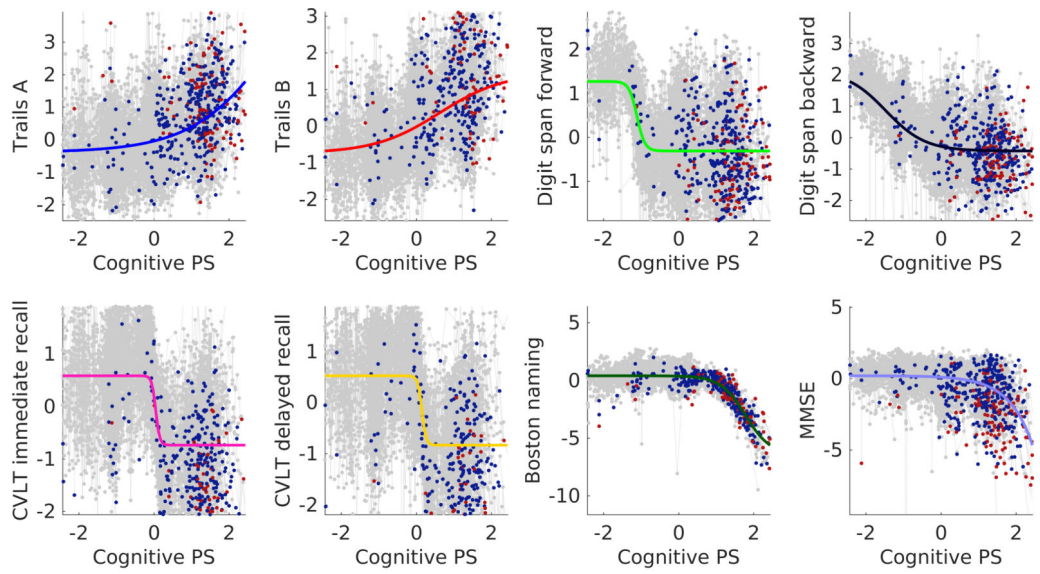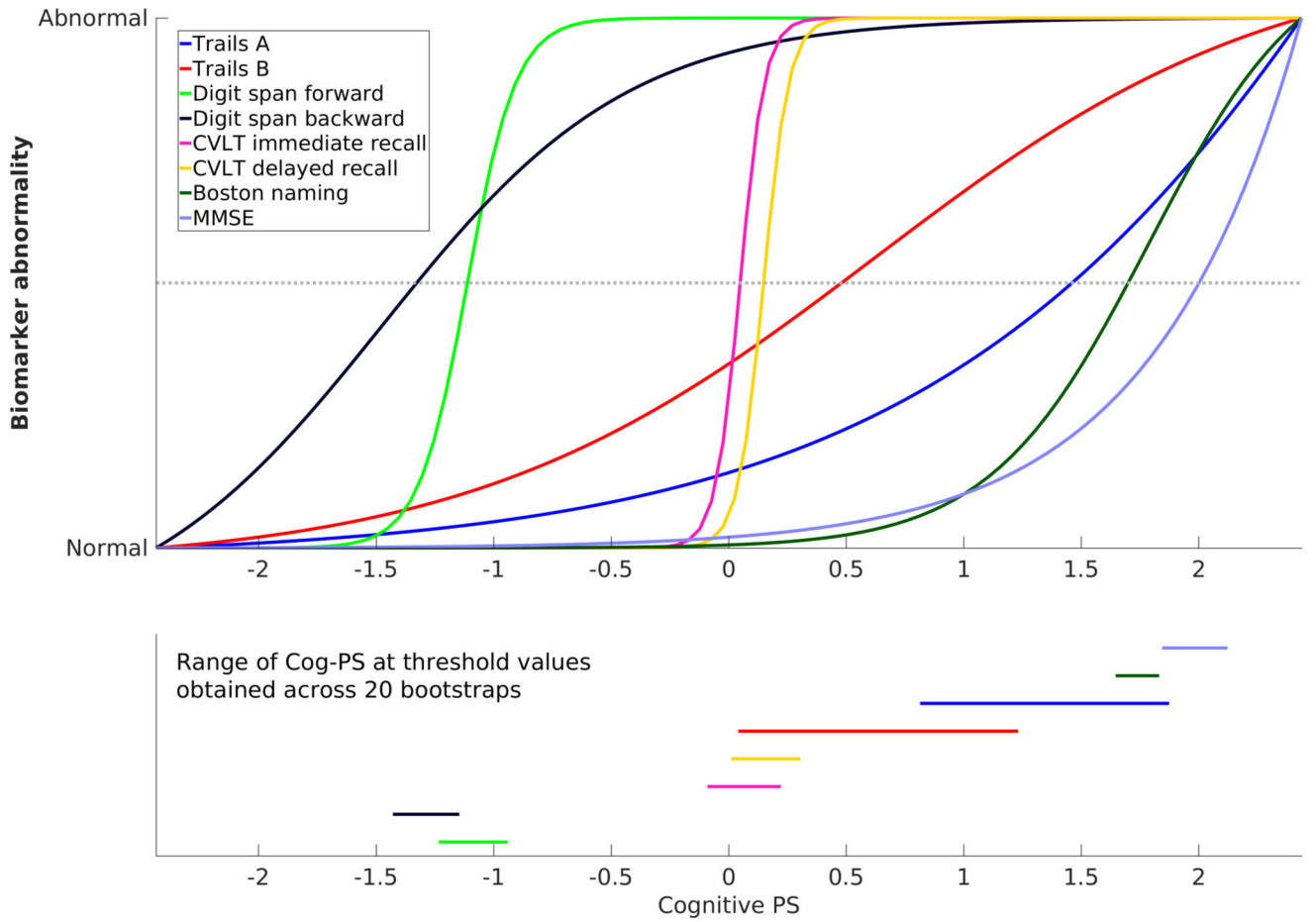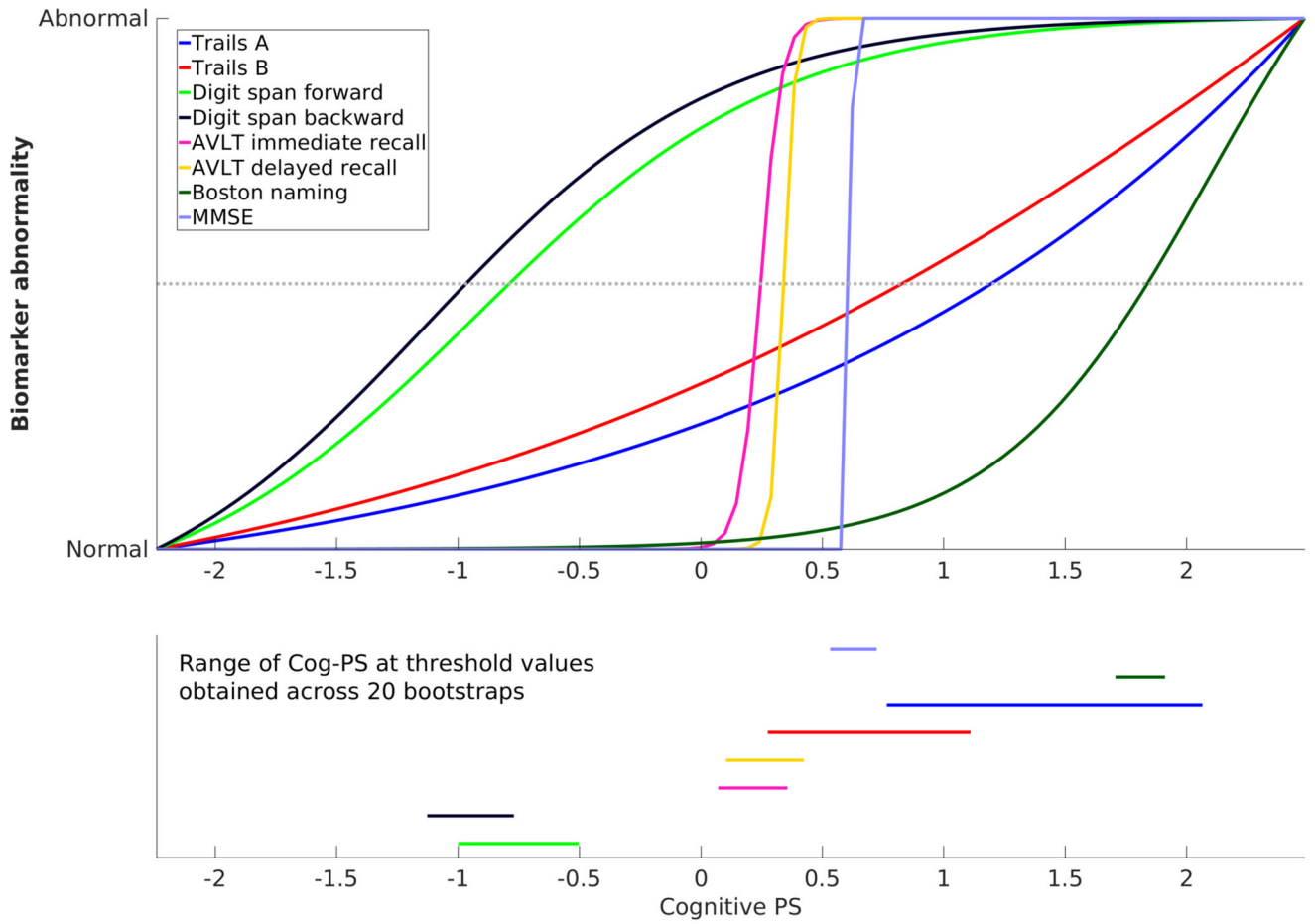
a.
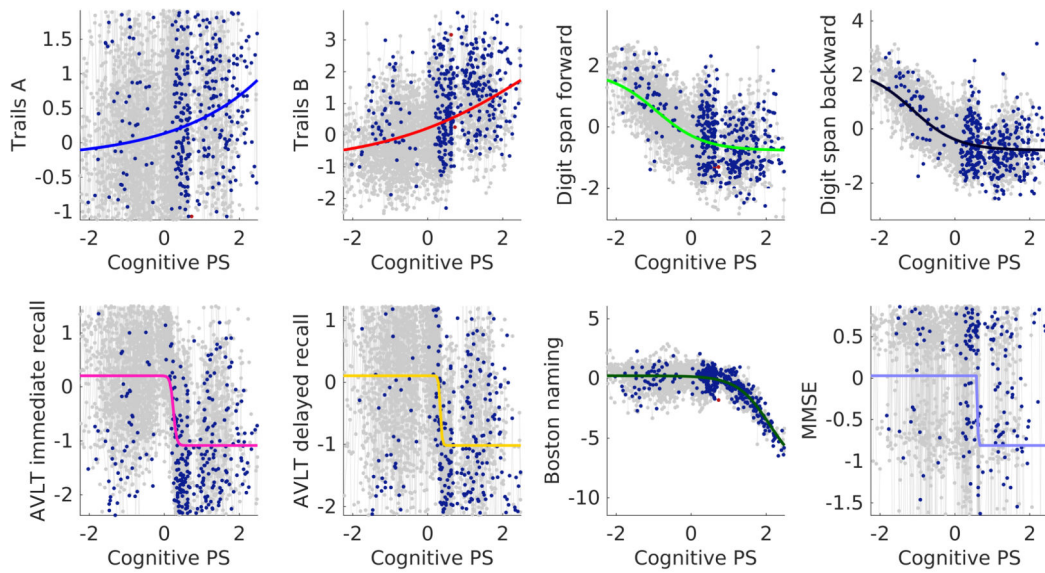
b.

Figure 2.
Cognitive measure trajectories in (a) BLSA and (b) WRAP. Gray, blue, and red dots indicate cognitively normal, MCI or CI, and AD visits, respectively.
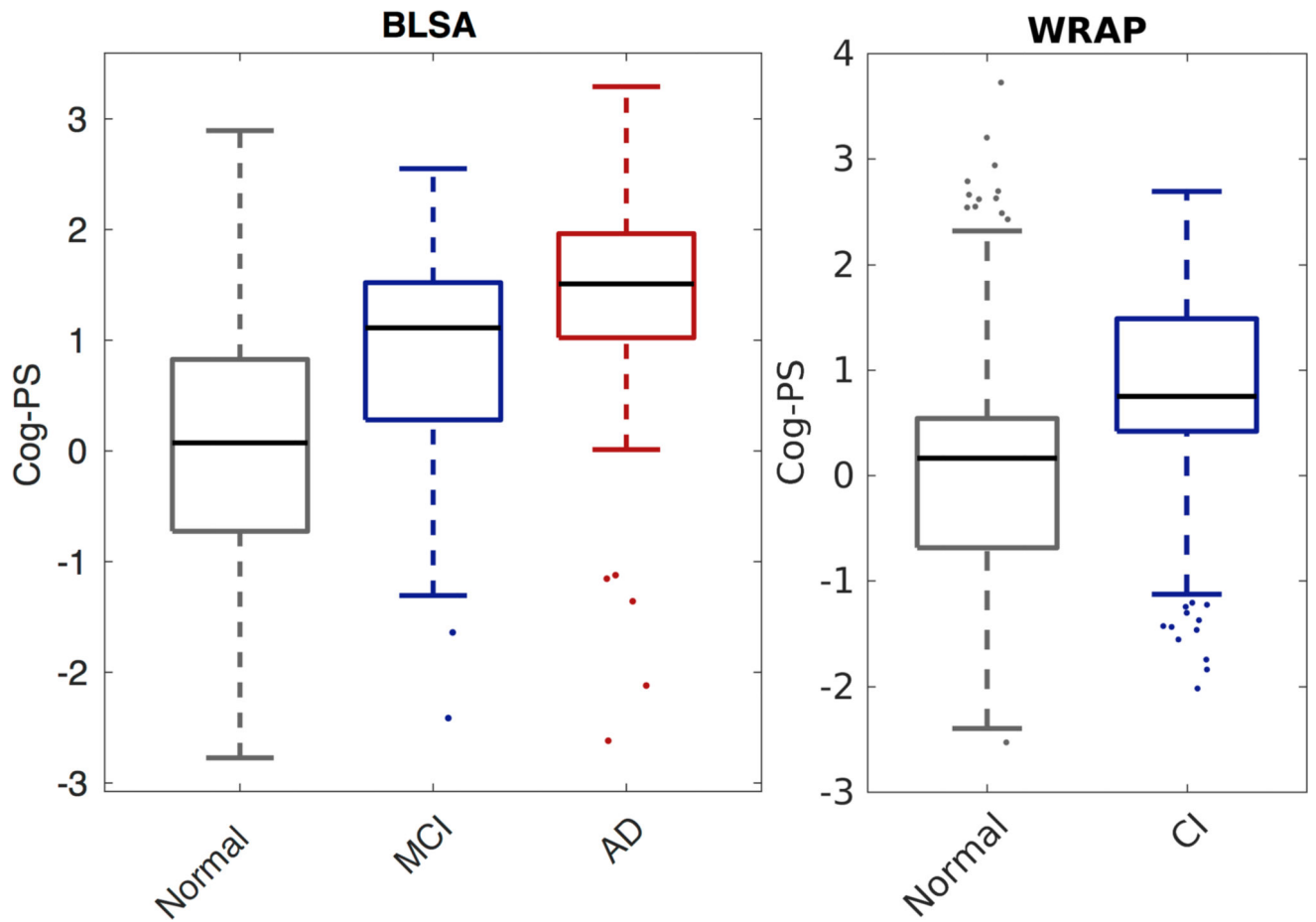
a.

Figure 3.
Scaled cognitive trajectories obtained using (a) BLSA and (b) WRAP data. Dotted line in the top panel corresponds to the midway point between Normal and Abnormal. Line segments in the bottom panel indicate the range of the Cog-PS values that attain the midway point along the standardized normal-to-abnormal *y*-axis across 20 bootstrap experiments.

Figure 4.
Box plot of Cog-PS by clinical diagnosis at last visit in BLSA (left) and WRAP (right). Central mark is the median, edges of the box correspond to the interquartile range, and whiskers extend to the range of non-outlier Cog-PS values. Outliers are plotted individually. All pairwise group comparisons were significant within each data set (two-sided Wilcoxon rank-sum test, $p < 0.0004$).

Table 1

Participants.

Final cognitive status	BLSA				WRAP		
	Normal (N=1428)	MCI (N=146)	AD (N=87)	Whole sample (N=1661)	Normal (N=1275)	CI* (N=233)	Whole sample (N=1510**)
Age at last visit, mean (SD), y	76.4 (9.0)	84.3 (6.4)	85.8 (6.7)	77.6 (9.2)	59.2 (7.4)	61.2 (7.3)	59.5 (7.4)
Range	60.0–101.0	68.0–99.3	66.4–98.6	60.0–101.0	38–78	40–76	38–78
Male, no. (%)	687 (48%)	58 (40%)	36 (41%)	781 (47%)	346 (27%)	100 (43%)	447 (30%)
White, no. (%)	1082 (76%)	128 (88%)	76 (87%)	1286 (77%)	1159 (91%)	194 (83%)	1355 (90%)
Education, mean (SD), y	16.7 (2.7)	16.2 (2.9)	17.2 (2.4)	16.7 (2.7)	16.1 (2.8)	15.9 (3.1)	16.1 (2.8)
Range	8–21	8–20	9–21	8–21	9–29	10–26	9–29
Number of visits, mean (SD)	4.5 (3.8)	5.5 (4.4)	6.3 (4.5)	4.7 (3.9)	2.3 (1.0)	2.1 (1.0)	2.3 (1.0)
Range	1–22	1–23	1–22	1–23	1–5	1–5	1–5
Follow-up duration, mean (SD), y	7.1 (6.5)	7.6 (6.2)	9.1 (6.5)	7.3 (6.5)	5.7 (3.8)	5.9 (4.1)	5.7 (3.8)
Range	0.0–23.4	0.0–23.0	0–22.9	0–23.4	0–13	0–12	0–13
Trails A at last visit, mean (SD) ***	38.8 (19.6)	54.5 (20.8)	87.1 (56.7)	42.3 (25.0)	26.0 (8.4)	32.3 (10.4)	26.9 (9.0)
Range	11–300	21–145	29–300	11–300	9–88	13–69	9–88
Trails B	98.4 (50.2)	168.6 (70.2)	198.8 (73.9)	106.2 (57.8)	61.8 (24.6)	93.6 (49.7)	66.8 (32.0)
	26–300	51–300	78–300	26–300	24–300	29–300	24–300
Digit span forward	7.7 (2.5)	7.2 (2.3)	6.2 (2.6)	7.6 (2.5)	10.5 (2.2)	9.9 (2.3)	10.4 (2.2)
	0–14	0–13	0–12	0–14	4–16	5–16	4–16
Digit span backward	6.7 (2.4)	5.8 (2.1)	4.7 (2.2)	6.5 (2.4)	7.2 (2.2)	6.3 (2.3)	7.1 (2.2)
	0–14	0–13	0–11	0–14	2–14	2–13	2–14
CVLT/AVLT# immediate recall	49.5 (12.4)	33.1 (9.3)	21.4 (10.6)	46.8 (14.1)	52.2 (7.5)	41.1 (8.1)	50.5 (8.6)
	0–80	15–61	0–47	0–80	24–72	21–62	21–72

Final cognitive status	BLSA				WRAP		
	Normal (N=1428)	MCI (N=146)	AD (N=87)	Whole sample (N=1661)	Normal (N=1275)	CI* (N=233)	Whole sample (N=1510**)
CVLT/AVLT delayed recall	10.2 (3.6)	5.2 (3.3)	1.7 (2.1)	9.5 (4.1)	11.0 (2.6)	7.2 (3.1)	10.4 (3.1)
	0–16	0–13	0–8	0–16	1–15	0–15	0–15
Boston naming test	53.6 (6.4)	48.2 (8.0)	42.3 (12.2)	52.6 (7.4)	57.2 (3.4)	55.3 (4.6)	56.9 (3.7)
	6–60	18–60	5–58	5–60	34–60	40–60	34–60
MMSE	28.3 (1.8)	26.4 (2.7)	22.3 (5.9)	27.8 (2.6)	29.4 (0.9)	29.0 (1.2)	29.4 (1.0)
	13–30	13–30	0–30	0–30	24–30	24–30	23–30

*
Cognitively impaired (CI) group in WRAP includes both clinical MCI (N=22) as well as individuals who were determined to have early MCI based on internal norms for cognitive measures (N=211).

**
Two participants were diagnosed with dementia due to Alzheimer's disease at last visit.

***
All cognitive measure statistics are computed using available values at last visit and reported as mean (SD); Range.

#
Verbal memory is measured using CVLT in BLSA and AVLT in WRAP.

BLSA = Baltimore Longitudinal Study of Aging; WRAP = Wisconsin Registry for Alzheimer's Prevention; CI = Cognitive impairment; MCI = Mild cognitive impairment; AD = Alzheimer's disease; SD = standard deviation; CVLT = California verbal learning test; AVLT = Auditory verbal learning test; MMSE = Mini mental state exam