Early Detection of Amyloid-Related Changes in Memory among Cognitively Unimpaired Older Adults with Daily Digital Testing

Kathryn V. Papp, PhD [1,2] Roos J. Jutten, PhD,[2] Daniel Soberanes, BS,[2] Emma Weizenbaum, PhD,[3] Stephanie Hsieh, MS,[2] Cassidy Molinare, BA [2] Rachel Buckley, PhD [1,2] Rebecca A. Betensky, PhD,[4] Gad A. Marshall, MD,[1,2] Keith A. Johnson, MD,[1,2,5] Dorene M. Rentz, PsyD,[1,2] Reisa Sperling, MD,[1,2] and Rebecca E. Amariglio, PhD[1,2]

Objective: This study was undertaken to determine whether assessing learning over days reveals Alzheimer disease (AD) biomarker-related declines in memory consolidation that are otherwise undetectable with single time point assessments.

Methods: Thirty-six (21.9%) cognitively unimpaired older adults (aged 60–91 years) were classified with elevated β-amyloid (Aβ+) and 128 (78%) were Aβ− using positron emission tomography with ^{11C}Pittsburgh compound B. Participants completed the multiday Boston Remote Assessment for Neurocognitive Health (BRANCH) for 12 min/day on personal devices (ie, smartphones, laptops), which captures the trajectory of daily learning of the same content on 3 repeated tests (Digit Signs, Groceries-Prices, Face-Name). Learning is computed as a composite of accuracy across all 3 measures. Participants also completed standard in-clinic cognitive tests as part of the Preclinical Alzheimer's Cognitive Composite (PACC-5), with 123 participants undergoing PACC-5 follow-up after 1.07 (standard deviation = 0.25) years.

Results: At the cross-section, there were no statistically significant differences in performance between Aβ+/− participants on any standard in-clinic cognitive tests (eg, PACC-5) or on day 1 of multiday BRANCH. Aβ+ participants exhibited diminished 7-day learning curves on multiday BRANCH after 4 days of testing relative to Aβ− participants (Cohen $d = 0.49$, 95% confidence interval = 0.10–0.87). Diminished learning curves were associated with greater annual PACC-5 decline ($r = 0.54$, $p < 0.001$).

Interpretation: Very early Aβ-related memory declines can be revealed by assessing learning over days, suggesting that failures in memory consolidation predate other conventional amnestic deficits in AD. Repeated digital memory assessments, increasingly feasible and uniquely able to assess memory consolidation over short time periods, have the potential to be transformative for detecting the earliest cognitive changes in preclinical AD.

ANN NEUROL 2023;00:1–11

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.26833

Received Jul 18, 2023, and in revised form Sep 28, 2023. Accepted for publication Oct 23, 2023.

Address correspondence to Dr Papp, Center for Alzheimer's Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Hale Building, 60 Fenwood Rd, Boston, MA 02115, USA. E-mail: kpapp@bwh.harvard.edu

From the [1]Center for Alzheimer Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [2]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [3]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [4]Department of Biostatistics, New York University School of Global Public Health, New York, NY, USA; and [5]Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

The ability to detect and track subtle Alzheimer disease (AD)-related cognitive decrements at the preclinical stage of disease has been a significant challenge for the field. Standard cognitive assessments, administered in the clinic at a single time point, exhibit small and tenuous associations with AD biomarkers of amyloid and tau among cognitively unimpaired (CU) older adults[1,2] if they are observed at all. Despite this, CU older adults harboring AD biomarkers decline on standard assessments, when observed over longer intervals, are at greater risk for progression to dementia.[3,4] Although it is possible that subtle decrements are only observable just prior to the onset of impairment, it is more likely that standard cognitive instruments and annual administration schedules are insufficiently sensitive to capture the earliest stages of a slowly progressive neurodegenerative disease.

A few converging lines of evidence suggest that decrements in memory consolidation, a process whereby a temporary, labile memory is transformed into a more stable, lasting form, may be a very early cognitive indicator of preclinical AD. For example, diminished practice effects (lack of expected improvement on repeated testing) are related to AD pathology among CU older adults[5,6] particularly for memory tests that involve retesting using the exact same stimuli.[7] More specifically, recent work showed that decrements in learning on monthly testing of the same, but not alternate, face–name pairs was associated with elevated amyloid,[7] hinting that failures in memory consolidation could underlie the association between diminished practice effects and preclinical AD. In related work, diminished memory retention after 1 week, but not after 30 minutes, was associated with presymptomatic autosomal dominant AD[8] and *APOE* ε4[9] among CU adults. The commonality between these studies is in temporal design, meaning that memory for the same content is retested (with and without re-exposure) after the passage of time. This study design allows for the time-dependent processes involved in memory consolidation to effectively (or ineffectively) unfold over the course of hours[10] at the synaptic level and days to weeks[11] at the systems level.

Building on these promising findings, we were interested in capturing memory consolidation using daily testing over 1 week. Among CU older adults with otherwise normal performance on traditional measures, we sought to determine whether individual differences in learning curves were related to underlying AD pathology as well as predictive of future decline. By capturing learning for 7 consecutive days, we may produce a more psychometrically robust estimate of learning than afforded by a single retesting session, a critical concern in preclinical AD, where cognitive effects are subtle.

As such, we examined multiday learning curves (MDLCs) using a previously validated web-based platform: the Boston Remote Assessment for Neurocognitive Health (BRANCH).[12] CU older adults with known amyloid status were asked to complete MDLC BRANCH, which includes 3 tests with identical stimuli presented daily, for 7 days. We hypothesized that diminished learning curves, representing early deficits in memory consolidation, would be associated with elevated amyloid and prospective cognitive decline on standard measures.

Subjects and Methods

Participants

A total of 164 CU participants aged 60–91 years were recruited from 3 affiliated cohorts: 98 participants from the Harvard Aging Brain Study (2P01AG036694-11; R.S., K.A.J.), 33 from the Instrumental Activities of Daily Living Study (R01AG053184; G.A.M.), and 33 from the Subjective Cognitive Decline study (1R01AG058825-01A; R.E.A.). Study procedures were conducted in accordance with human subjects' protections, and the study protocol was approved by the Mass General Brigham institutional review board. All participants underwent informed consent. Exclusion criteria included history of alcoholism, drug abuse, head trauma, or current serious medical/psychiatric illness. Given that participants were recruited at various stages of their participation in ongoing longitudinal studies with annual assessments, they were classified as CU by either study entry criteria or via a multidisciplinary consensus,[13] depending on which source of information was most proximal to their completion of BRANCH. Study entry criteria included a Clinical Dementia Rating (CDR) global score = 0, Mini-Mental Status Examination (MMSE) >25, and Logical Memory Delayed Recall (LMDR) scores above education-adjusted cutoffs (≥9 for 16+ years of education, ≥5 for 8–15 years of education). Participants were brought to multidisciplinary consensus if they had a global CDR ≥ 0.5 and/or performance falling 1.5 standard deviations (SD) below the sample mean on any individual domain-specific cognitive composite score; if they did not fall below any of these cutoffs, they were considered CU.[13] Although a subset of participants in the current sample have a CDR = 0.5 (n = 6), they were retained in the study, because they were deemed CU via multidisciplinary consensus.

Standard Clinical and Cognitive Assessments

Participants completed in-clinic cognitive assessments including the Preclinical Alzheimer Cognitive Composite (PACC-5),[14,15] which includes LMDR, Free and Cued Selective Reminding Test (FCSRT), MMSE, Digit Symbol Substitution Test (DSST), and Category Fluency Test. Participants also completed the Buschke 6-trial Selective Reminding Test (6-SRT).[16] We examined MDLC BRANCH in relation to the PACC-5 and memory measures including LMDR, FCSRT free recall, and 6-SRT total recall. In-clinic assessments were completed within 97 days of multiday BRANCH. A total of 123 participants were followed annually with the PACC-5 for up to 2 years.

Quantification of Amyloid Burden

Participants underwent positron emission tomography (PET) with ^{11C}Pittsburgh compound B (PiB) to quantify amyloid burden.[17] PET scans were completed within 1.03 (SD = 1.36) years of multiday BRANCH. PiB images were acquired using a 60-minute dynamic acquisition on a Siemens (Erlangen, Germany) ECAT HR+ PET scanner. PET images were co-registered to corresponding T1 images using FreeSurfer (v6)-based structural regions of interest mapped into native PET space using SPM12. PiB is expressed as the distribution volume ratio, with a cerebellar gray reference region. A global cortical aggregate was calculated for each participant for the target region, comprising frontal, lateral temporal and retrosplenial regions. Participants were dichotomized into low (Aβ−) versus high (Aβ+) groups (cutoff = 1.185).[18]

MDLC Procedure. MDLCs were collected using multiday BRANCH,[12,19] which includes 3 tasks (described below) captured once per day for 7 days (Fig 1). Prior to daily testing, participants attested to completing tasks independently without recording stimuli/responses with the goal of advancing research. Participants specified the time of day they preferred to take the test and were notified at that time. Participants were also asked to rate the enjoyability of tasks each day on a Likert scale.

The Digit-Signs test is modeled on the DSST,[20] a measure of processing speed with an associative memory component, in that performance is faster if pairs are memorized. Participants shown a key of 6 street signs paired with digits must indicate "yes" or "no" regarding whether a series of digit–sign pairs are correct. The outcome is number of correct pairs completed within 120 seconds.

The Groceries Prices test is an associative memory measure.[21] Participants are asked to remember a price paired with a pictured grocery item. Following a delay, participants identify the correct price among counterbalanced incorrectly paired and partially novel price/grocery distractor pairs. The outcome is number of correct responses.

In the Face Name test (a variation of the Face Name Associative Memory Exam [FNAME])[7] participants are asked to remember a series of face–name pairs. Following a delay, the participant is shown each face and asked to select the first letter of the name paired with that face (first letter name recall). Next, they are asked to identify the correct name via multiple choice (target name, re-paired same-sex name, same-sex foil name; face–name memory). The outcome is the average number of correct responses for first letter name recall and face–name memory combined.

Computing the MDLC. To account for the different learning curve shapes (ie, learning on Digit Signs is linear, whereas learning on Face Name is logarithmic) and the potential impact of ceiling effects (ie, 23% and 8% of participants performed at ceiling levels—100% accuracy for 2 or more administrations—on the Face Name and Groceries Prices tests, respectively), we computed an area under the curve (AUC) for each task to capture both the rapidity with which an individual learns and the total accumulation of content (explained in detail elsewhere).[19] In contrast with the more common use of an AUC to be used in classification models (eg, receiver operating characteristic analyses), use of an AUC in the current context allows us to produce a summary metric for the overall proportion of information learned using a general formula from integral calculus. Additionally, to account for an individual's starting point (day 1 performance), we computed a scaled AUC that equals AUC/AUC_{max}, where AUC_{max} is the maximum value of the AUC obtained if the participant scored at the maximum value from the second through the final test

FIGURE 1: Capturing learning curves using the multiday Boston Remote Assessment for Neurocognitive Health.

administration. AUC values (herein referred to as "learning curves") were computed for an equally weighted composite across the 3 tasks as well as for each test.

Statistics

Statistical analyses were completed using R (v4.0.3). Statistical significance was set at $p < 0.05$.

Linear regression analyses were performed to determine the association between learning curves (composite, individual tests) and β-amyloid (Aβ)+/− while correcting for demographics (ie, age, sex, and education). Cohen *d* effect sizes were used to evaluate the strength of Aβ+/− group differences using the following benchmarks: small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$).[22] Comparable analyses were performed to determine whether differences between Aβ+/− groups were similarly observable on day 1 of multiday BRANCH (both the composite and individual tests) as well as on standard in-clinic measures.

We investigated whether decrements on learning curves were associated with annual change on the PACC-5 among those with longitudinal testing. Individual PACC-5 slopes were extracted using linear mixed effects models correcting for demographics, and their interaction with time. Pearson correlations were used to investigate the association between PACC-5 slopes and learning curves.

Sensitivity Analyses. To determine whether a shortened administration schedule could reduce participant burden, we examined whether Aβ+/− performance differences were retained when reducing testing to 5, 4, or 3 days. Additionally, we repeated the primary analyses excluding top performers (ie, those who reached ceiling defined as 100% accuracy on 2 or more days on Face Name and Groceries Prices).

Results

Participant Characteristics

Participants included 164 individuals (mean age = 74.3 years) of whom 36 (21.9%) were classified as Aβ+ and 128 classified

TABLE 1. Demographic Characteristics

Characteristic	Overall, N = 164	Aβ +, n = 36	Aβ−, n = 128	Difference between Aβ+/Aβ−
Age, yr	74.3 (7.31)	76.9 (5.80)	73.6 (7.55)	**$t = -2.83$, $p = 0.006$, 95% CI = −5.66 to −0.98**
Global CDR (0, 0.5)	(158, 6)	(2, 34)	(124, 4)	$\chi^2 = 0.03$, $p = 0.854$
MMSE	28.90 (1.26)	28.5 (1.7)	29.00 (1.10)	$t = 1.66$, $p = 0.105$, 95% CI = −0.11 to 1.10
LMDR	16.1 (3.87)	15.7 (4.81)	16.2 (3.57)	$t = 0.56$, $p = 0.579$, 95% CI = −1.25 to 2.22
Sex (female, male)	(105, 59)	(16, 20)	(85, 43)	$\chi^2 = 1.00$, $p = 0.316$
Race				$\chi^2 = 0.82$, $p = 0.936$
Caucasian	146	33	113	
Black	12	2	10	
Asian	4	1	3	
Native American	1	0	1	
Other	1	0	1	
Years of education	16.6 (2.45)	16.6 (2.11)	16.7 (2.55)	$t = 0.11$, $p = 0.914$, 95% CI = −0.79 to 0.88
Continuous Aβ (DVR in global composite)	1.16 (0.21)	1.47 (0.26)	1.08 (0.04)	**$t = -8.95$, $p < 0.001$, 95% CI = −0.48 to −0.30**

Note: Mean and standard deviation are reported unless otherwise noted. Bold *p* values should be considered significant if they are less than 0.05.
Abbreviations: Aβ = β-amyloid; CDR = Clinical Dementia Rating; CI = confidence interval; DVR = distribution volume ratio; LMDR = Logical Memory Delayed Recall; MMSE = Mini-Mental Status Examination.

TABLE 2. Linear Regression Models Showing Performance between Aβ+ and Aβ− Groups on Standard Single Time Point Cognitive Assessments

	PACC-5			FCSRT Free Recall			Logical Memory Delayed Recall			6-SRT Total Recall		
Predictors	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***
Intercept	0.40	−1.42 to 2.21	0.666	2.66	−0.91 to 4.40	**0.003**	−2.97	−4.80 to −1.14	**0.002**	1.02	−0.74 to 2.79	**0.254**
Age	−0.02	−0.05 to −0.00	**0.038**	−0.04	−0.06 to −0.02	**0.001**	0.02	−0.00 to 0.04	0.070	−0.02	−0.04 to −0.00	**0.050**
Sex [M]	−0.53	−0.86 to −0.20	**0.002**	−0.53	−0.85 to −0.21	**0.001**	−0.33	−0.66 to 0.01	**0.056**	−0.62	−0.94 to −0.30	**0.001**
Education	0.09	0.03 to 0.16	**0.005**	0.03	−0.03 to 0.09	0.303	0.09	0.03 to 0.16	**0.004**	0.05	−0.01 to 0.11	0.116
Aβ status [+]	−0.29	−0.67 to 0.08	0.125	0.03	−0.34 to 0.39	0.887	−0.16	−0.54 to 0.22	0.417	−0.32	−0.69 to 0.05	0.089
R^2/R^2_{adj}		0.143/0.122			0.166/0.145			0.084/0.061			0.152/0.130	

Note: N = 164. Regression coefficients are standardized. Bold *p* values should be considered significant if they are less than 0.05.
Abbreviations: 6-SRT = 6-trial Selective Reminding Test; Aβ = β-amyloid; CI = confidence interval; Est. = estimate; FCSRT = Free and Cued Selective Reminding Test; M = male; PACC-5 = Preclinical Alzheimer Cognitive Composite.

as Aβ− (Table 1). The Aβ+ group was older, but there were otherwise no group differences in terms of sex, years of education, race, MMSE, or global CDR score.

BRANCH Feasibility

Multiday BRANCH was completed over a mean of 8.1 days (maximum = 14 days) with no differences between Aβ+/− groups regarding days to completion ($t = -0.46$, 95% confidence interval [CI] = −1.96 to 1.23, $p = 0.647$). There were no Aβ+/− groups differences regarding completion time ($t = -0.58$, 95% CI = −0.54 to 0.29, $p = 0.561$), with Aβ+ individuals requiring a mean of 12.23 minutes and Aβ− participants requiring 12.11 minutes. The majority (n = 151, 92%) of participants completed all 7 days. Among the 13 participants with incomplete data, 10 were included in analyses, as they completed at least 3 days of assessments, allowing for the computation of a learning curve. There were no differences between completers/noncompleters regarding clinical severity on the MMSE (mean difference = 1.3 points, $p = .08$) or global CDR (mean difference = 0.02 points, $p = .58$); however, we did observe a slightly higher proportion of Aβ + participants among noncompleters ($\chi^2 = 6.48$, $p = 0.011$).

Participants reported increasing task enjoyability by day ($B = 0.50$, 95% CI = 0.46–0.53, $p < 0.001$). A total of 45.7% of participants completed multiday BRANCH on a smartphone, whereas the remaining participants used a laptop/desktop. Smartphone users tended to be younger, but there were otherwise no other demographic or BRANCH performance differences between groups based on device used.

Performance on Standard Cognitive Assessments

There were no statistically significant differences in performance between Aβ+/− groups on standard in-clinic cognitive testing (Table 2) including the PACC-5, FCSRT free recall, LMDR, or 6-SRT total recall, although some measures exhibited trend-level differences (ie, PACC-5, 6-SRT total recall).

Performance on Learning Curves Using Multiday BRANCH

There were no differences between Aβ+/− groups when examining composite or individual measure performance on day 1 of the learning curve (Table 3). However, Aβ+ showed diminished learning curves (composite score) relative to Aβ−, with a medium effect size of Cohen $d = 0.49$ (95% CI = 0.10–0.87; see Table 3, Fig 2). This pattern of less robust learning among Aβ+ was observed across each individual test (see Fig 2). However, the difference between Aβ+/− was not statistically significant for Face Name, whereas it was significantly different for Digit Signs and Groceries Prices (see Table 3).

Systematically reducing the number of days over which learning curves were computed, Aβ+/− group differences emerged on the 4th testing day (Table 4), indicating 4 administrations were needed to observe group differences. Overall results were comparable when top performers were removed (see Table 4).

Associations between Learning Curves and Annual Change on the PACC-5

A total of 123 participants completed at least 1 follow-up in-clinic cognitive assessment (mean age = 73.6 years, SD = 7.22; 65% female; mean years of education = 16.5,

TABLE 3. Linear Regression Models Showing Performance between Aβ+ and Aβ− Groups on Day 1 (top) and Learning Curves Using the Multiday Boston Remote Assessment for Neurocognitive Health (bottom)

	Composite Day 1			Digit Signs Day 1			Groceries Prices Test day 1			Face Name Day 1		
Predictors	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***
Intercept	2.35	0.64 to 4.07	**<0.001**	1.58	0.59 to 2.57	**0.002**	−0.46	−1.80 to 0.87	0.496	−0.60	−1.69 to 0.48	0.272
Age	−0.04	−0.06 to −0.02	**<0.001**	−0.04	−0.05 to −0.03	**<0.001**	−0.00	−0.02 to −0.01	0.566	−0.02	−0.03. −0.00	**0.007**
Sex [M]	−0.45	−0.76 to −0.13	**0.005**	0.05	−0.14 to 0.23	0.616	−0.17	−0.42 to 0.07	0.165	−0.39	−0.59 to −0.19	**<0.001**
Education	0.05	−0.01 to 0.11	0.132	0.04	0.00 to 0.07	**0.033**	0.00	−0.05 to 0.05	0.983	0.04	−0.00 to 0.07	0.062
Aβ status [+]	−0.17	−0.53 to 0.19	0.350	−0.14	−0.34 to 0.07	0.194	−0.09	−0.37 to 0.19	0.531	−0.01	−0.23 to 0.22	0.963
R^2/R^2_{adj}		0.163/0.142			0.240/0.221			0.022/0.003			0.153/0.131	
	Composite Multiday Learning Curve			**Digit Signs Test Multiday Learning Curve**			**Groceries Prices Test Multiday Learning Curve**			**Face Name Multiday Learning Curve**		
Predictors	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***
Intercept	3.18	1.61 to 4.76	**<0.001**	4.46	2.91 to 6.01	**<0.001**	2.10	0.47 to 3.74	**0.012**	1.69	0.03 to 3.36	**0.046**
Age	−0.05	−0.07 to −0.03	**<0.001**	−0.07	−0.09 to −0.05	**<0.001**	−0.03	−0.05 to −0.01	**0.001**	−0.03	−0.05 to −0.01	**0.007**
Sex [M]	−0.66	−0.95 to −0.37	**<0.001**	0.07	−0.22 to 0.35	0.649	−0.69	−0.99 to −0.39	**<0.001**	−0.88	−1.19 to −0.58	**<0.001**
Educ.	0.06	0.00 to 0.11	**0.045**	0.05	−0.00 to 0.11	0.065	0.05	−0.01 to 0.10	0.107	0.04	−0.02 to 0.10	0.153
Aβ status [+]	−0.36	−0.69 to −0.03	**0.034**	−0.37	−0.70 to −0.04	**0.028**	−0.43	−0.78 to −0.09	**0.015**	−0.13	−0.48 to 0.22	0.478
R^2/R^2_{adj}		0.323/0.305			0.310/0.293			0.255/0.236			0.248/0.229	

Note: N = 164. Regression coefficients are standardized. Bold *p* values should be considered significant if they are less than 0.05.
Abbreviations: Aβ = β-amyloid; CI = confidence interval; Est. = estimate; M = male.

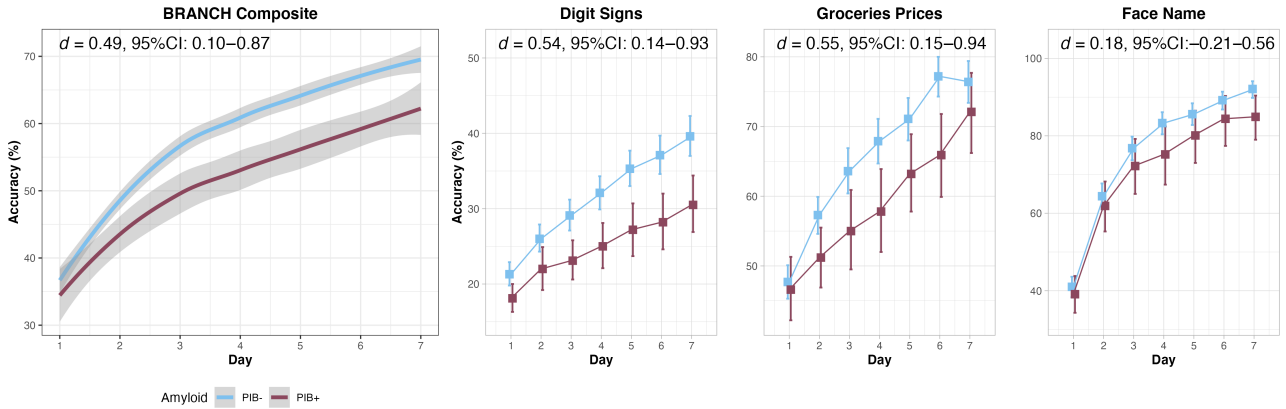

FIGURE 2: Decrements in learning curves (multiday Boston Remote Assessment for Neurocognitive Health [BRANCH]) between β-amyloid (Aβ)+ and Aβ− cognitively unimpaired older adults, N = 164. Raw (uncorrected) data are shown; y-axis refers to percent of items correct. CI = confidence interval; PIB = Pittsburgh compound B.

SD = 2.53; 20% Aβ+; mean follow-up = 1.07 years, maximum = 2 years). A diminished learning curve at baseline was associated with greater subsequent annual decline on the PACC-5 ($r = .54$, 95% = 0.40–0.65, $p < .001$; Fig 3). Weaker associations were observed between annual PACC-5 change and either day 1 BRANCH ($r = 0.36$, 95% CI = 0.19–0.50, $p < .001$) or baseline PACC-5 ($r = 0.39$, 95% CI = 0.22–0.55, $p < 0.001$).

Discussion

Our findings suggest that assessing learning over repeated evaluations can reveal disease-relevant decrements in memory during preclinical AD that are less readily observed using single time point assessments. More specifically, we found that a diminished learning curve for the same information presented for 7 days (12 min/day of unsupervised web-based testing on personal devices) was associated with

TABLE 4. Sensitivity Analysis: Reducing the Multiday BRANCH Learning Curve to 5, 4, and 3 Days (top) and Removing Top Performers (bottom)

	3-Day Multiday BRANCH Composite			4-Day Multiday BRANCH Composite			5-Day Multiday BRANCH Composite		
Predictor	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***
Intercept	0.05	−0.13 to 0.23	0.578	0.09	−0.09 to 0.27	0.311	0.09	−0.08 to 0.27	0.297
Aβ group [Aβ+]	−0.33	−0.74 to 0.08	0.110	−0.59	−0.99 to −0.20	**0.003**	−0.58	−0.97 to −0.20	**0.003**
R^2/R^2_{adj}	0.017/0.010			0.053/0.047			0.054/0.048		
Intercept	3.09	1.43 to 4.75	**<0.001**	2.92	1.30 to 4.54	**<0.001**	2.95	1.35 to 4.55	**<0.001**
Age	−0.05	−0.07 to −0.03	**<0.001**	−0.05	−0.07 to −0.03	**<0.001**	−0.05	−0.07 to −0.03	**<0.001**
Sex [M]	−0.58	−0.89 to −0.28	**<0.001**	−0.67	−0.97 to −0.38	**<0.001**	−0.68	−0.98 to −0.39	**<0.001**
Education	0.05	−0.01 to 0.11	0.118	0.05	−0.00 to 0.11	0.071	0.06	0.00 to 0.11	**0.046**
Aβ group [Aβ+]	−0.11	−0.48 to 0.25	0.539	−0.35	−0.70 to −0.00	**0.048**	−0.34	−0.67 to 0.00	0.052
R^2/R^2_{adj}	0.246/0.226			0.292/0.274			0.306/0.288		

	Composite Multiday Learning Curve			Digit Signs Test Multiday Learning Curve			Groceries Prices Test Multiday Learning Curve			Face Name Multiday Learning Curve		
Predictor	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***	**Est.**	**CI**	***p***
Intercept	2.63	.91 to 4.36	**0.003**	4.46	2.91 to 6.01	**<0.001**	2.02	.46 to 3.57	**0.012**	1.77	−.16 to 3.70	0.072
Age	−0.04	−.06 to −.02	**<0.001**	−0.07	−.09 to −.05	**<0.001**	−0.03	−.05 to −.01	**0.001**	−.02	−.05 to −.00	**0.030**
Sex [M]	−0.51	−.82 to −.20	**0.002**	0.07	−.22 to .35	0.649	−0.62	−.90 to −.33	**<0.001**	−.75	−1.09 to −.40	**<0.001**
Education	0.04	−.03 to .10	0.244	0.05	−.00 to .11	0.065	0.03	−.03 to .08	0.327	.01	−.06 to .08	0.837
Aβ status [+]	−0.40	−.76 to −.05	**0.027**	−0.37	−.70 to −.04	**0.028**	−0.36	−.69 to −.03	**0.031**	−.11	−.51 to .28	0.570
Observations	116			164			122			148		
R^2/R^2_{adj}	0.300/.274			0.310/.293			0.209/.182			0.243/.221		

Note: N = 164; regression coefficients are standardized. Bold *p* values should be considered significant if they are less than 0.05.
Abbreviation: Aβ = β-amyloid; BRANCH, Boston Remote Assessment for Neurocognitive Health; CI = confidence interval; Est. = estimate; M = male.

elevated amyloid, whereas this association was absent using single time point (day 1) or standard cognitive assessments. This same pattern was observed across 3 different memory tasks (with the strongest effects on Digit Signs and Groceries Prices) and was associated with longitudinal decline on standard assessments, suggesting that a diminished MDLC may be a harbinger of decline in the setting of otherwise normal cognitive performance. These results provide insights into the nature of the earliest detectable memory decrements in AD. That is, decrements in memory consolidation, made measurable by assessing learning over repeated daily evaluations, may predate other conventional deficits in learning and recall. Furthermore, our results provide a practical paradigm for identifying those with preclinical AD at greatest risk for short-term decline, individuals who would be ideal candidates for interventions.

Multiday Learning Curves and Memory Consolidation

Failures in initial memory encoding differentiate individuals with dementia versus mild cognitive impairment (MCI),[23] whereas failures in retention after a 30-minute delay distinguish between MCI and normal aging.[24] Our results suggest that failures in memory consolidation rather than initial encoding or retention may be the earliest observable memory changes in preclinical AD, given

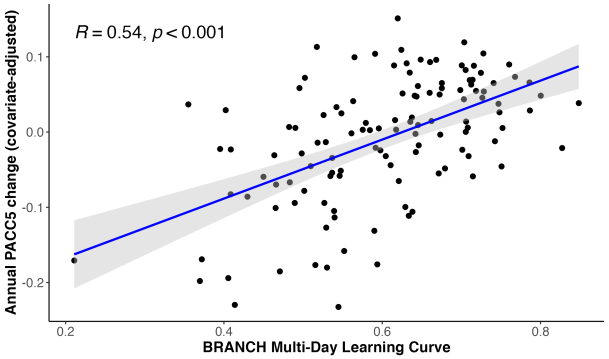

FIGURE 3: Association between a diminished learning curve (multiday Remote Assessment for Neurocognitive Health [BRANCH]) collected over 7 days and subsequent cognitive decline (Preclinical Alzheimer Cognitive Composite [PACC5]) over 1.07 years of follow-up, n = 123.

that Aβ-related decrements in memory were not observed on standard (ie, LMDR, FCSRT) or challenging (ie, 6-SRT) learning and recall tasks administered at a single time point. Instead, differentiation between Aβ+/− only became evident with repeated daily testing. Because memory consolidation is a process that requires the passage of time,[25] it is unable to be probed with a single time point assessment, regardless of the sensitivity of the measure.

There are multiple avenues by which Aβ and related AD pathology may have both direct and indirect effects on memory consolidation. Aβ is known to impair synaptic plasticity,[26] which is critical to learning. For example, in rodent models, injection of human brain-derived Aβ into rats after avoidance training impaired recall of that training at 48 hours.[27] Interestingly, the adverse effect of Aβ on recall was not significant at 24 hours and required a re-exposure to the training condition, suggesting that Aβ did not obliterate the memory, but may instead have made it more susceptible to decay or extinction. We similarly observed that significant trends toward differences in learning curves between Aβ+/− groups emerged after several days.

In addition to synaptic plasticity, it may also be that Aβ-related decrements in learning become more evident with relearning over days because of the critical role of sleep for memory consolidation. Sleep is disrupted in symptomatic AD, with recent work suggesting sleep changes may be occurring in presymptomatic stages with potentially bidirectional effects. That is, Aβ may reduce sleep quality, with reduced sleep quality in turn accelerating Aβ accumulation.[28] An important study showed that elevated Aβ among normal older adults was associated with diminished slow wave activity during non-rapid eye movement sleep (a stage of sleep critical for memory consolidation) and that the extent of this reduced activity was associated with worse overnight memory retention.[29]

Finally, tau deposition has a more anatomically direct impact on memory versus Aβ, with tau tangles impacting the integrity of medial temporal lobe functioning. With ongoing acquisition of tau PET in this sample, we will, in future studies, be able to determine whether memory consolidation deficits, measured with diminished learning curves, are the first Aβ-related sign of impaired synaptic function that portends a tau-associated amnestic syndrome typical of symptomatic AD.[30]

MDLCs: Individual Test Results

Interestingly, Aβ+/− effects were most robust on a speeded measure with an associative memory component (Digit Signs) versus on frank associative memory measures (Face Name, Groceries Prices). We hypothesize that the absence of a performance ceiling on Digit Signs, in contrast with Face Name/Groceries Prices, where participants may be able to achieve 100% accuracy, may have conferred measurement benefits. Aligned with this, removing top performers in our sensitivity analyses improved estimates. Alternatively, recent work has also shown that declines in processing speed may be a very early indicator when amyloid is accumulating, with associative memory processes impacted at later stages of tau accumulation.[31] Thus, Digit Signs may be particularly sensitive, because it is a speeded task with an associative memory component. Finally, some of our participants may have become habituated to the Face Name test, as many have encountered variations of FNAME during other parent study-related assessments, and thus may have developed their own unique strategies. Regardless, the similar patterns of findings across individual tests indicate that repeated memory testing reveals vulnerabilities in the ability to acquire specific episodic content among Aβ+ individuals.

MDLCs, Practice Effects, and Other Repeated and Digital Assessments

Our findings are also relevant to a larger body of work showing diminished practice effects signal cognitive vulnerability.[5] Whereas clinical trials actively eschew practice effects, as they can impede the ability to capture cognitive decline, the MDLC capitalizes on factors known to improve performance, including shorter retest intervals,[32] focus on memory,[33] and, most importantly, repeated stimuli.[7] Although practice effects can be related to both familiarity with test strategies and familiarity with stimuli, it seems that the latter may be most relevant for AD. For example, the present study was motivated, in part, by our previous findings showing that decrements in learning of the same face–name pairs, but not alternate face–name pairs, on monthly testing was associated with Aβ+.[7] Other experimental psychology paradigms, which leverage

repeated stimuli, have reported similar results.[34] These findings are in keeping with the notion that failures on repeated testing with the same stimuli hint at taxed memory consolidation processes.

BRANCH is one among several digital cognitive assessments whose development has been accelerated by several converging factors, including the necessity for remote assessment resulting from social distancing requirements of COVID-19, a need for scalable assessment approaches in the context of early detection and preclinical AD, and the increasingly widespread uptake of digital devices among older adults. Recent studies using web/app-based cognitive testing with older adults have shown that they are highly feasible[35-37] and that they exhibit high concordance with in-clinic cognitive testing,[12,36,38] with several showing correlations with AD biomarkers (see Öhman et al[39] for a review). These digital assessments vary in approach, with some offering more traditional measures in a digital format, others emphasizing the interrogation of specific cognitive processes, and many leveraging the digital format for repeated testing.

Notably, the MDLC differs from other types of repeated sampling of cognition such as "ecological momentary assessment" and "burst design." In these paradigms, cognition is sampled multiple times within and across days and averaged to achieve a more robust and stable estimate.[40] Reducing measurement variability may thus improve the detection of change over time, whereas the MDLC leverages improvement in performance day by day.

Associations between Diminished Learning Curves and Prospective Cognitive Decline

Generally, cross-sectional cognitive performance, even in the context of AD biomarker positivity, is not a robust predictor of risk for short-term clinical progression, particularly in preclinical stages of disease, where rates of decline are highly heterogenous. Our findings of a moderate association between MDLC and prospective PACC-5 decline are aligned with other studies showing the predictive utility of diminished practice effects for risk for cognitive decline.[41,42] Cognitive paradigms that can help identify CU subjects at greatest risk for short-term cognitive progression (and those most eligible for secondary prevention/treatment) will be critical as more accessible and cost-effective AD biomarkers (ie, blood-based biomarkers) make identification of preclinical AD more widespread.[43] However, replication in much larger and more representative samples is needed to determine whether pairing diminished MDLC with biomarker positivity among CU subjects can improve the prediction of risk for imminent clinical progression at an individual level.

Limitations and Future Directions

The racial breakdown (~10.4% from underrepresented groups) and the high education level of our sample are not aligned with the demographics of broader US population at risk for AD.[44,45] Additionally, because assessments are completed unsupervised, there are factors that cannot be experimentally controlled, such as the presence of interruptions or unauthorized use of memory aides. Although it is possible to counter this with technology (ie, recording the participant taking the test via web-video), the benefits of these strategies are outweighed by data privacy concerns. To counter this challenge, we asked participants to attest to completing the task unaided each day for the sake of advancing research. Despite this attestation, it is possible that some participants used memory aides. However, in a sensitivity analysis, removal of top performers (ie, those who hit ceiling) on FNAME/Groceries did not alter our findings. Another limitation of this study is the relatively small sample size, which may have left us underpowered to observe differences in performance between Aβ+/− groups on traditional measures. In addition to obtaining a more racially and ethnically diverse sample, it will be critical to increase the sample size to determine the extent to which this type of paradigm may be informative at the individual level for diagnosis and prognosis.

In addition to assessing MDLC in relation to tau, future work will determine whether repeating the MDLC over time with new sets of stimuli provides a more sensitive marker to track cognitive decline. For example, it may be that this paradigm is useful in detecting a change in synaptic integrity in response to a therapeutic agent over a shorter interval, facilitating brief, adaptive, and early phase clinical trials.

Conclusions

Consistent with many other cross-sectional studies in preclinical AD,[1,2] standard single time point cognitive assessments did not differentiate between Aβ+/− groups. Observing associations between biomarkers and cognition in preclinical AD is a methodological challenge given that, by restricting samples to those with a minimum level of cognitive performance, we both exclude individuals whose underlying biomarker positivity is having a demonstrable clinical impact as well as limit our ability to observe correlations given a restricted range of possible cognitive scores. However, our findings that Aβ+ participants exhibited diminished learning curves over 7 days of testing, alongside other recent work,[7,9,34] suggest that Aβ+/− differences are demonstrable with nonstandard single time point assessments. Importantly, these results suggest that failures in memory consolidation predate other conventional amnestic

deficits in AD. Implementing these promising assessments into future research may improve early detection and tracking of cognitive decline in preclinical AD.

Acknowledgments

Funding for this study was provided by NIH/NIA grants 2P01AG036694-11 (R.S., K.A.J.), R01AG053184 (G.A.M.), and 1R01AG058825-01A (R.E.A.), by the Davis Alzheimer Prevention Program, and by the Vettel Alzheimer Innovation Fund.

Author Contributions

K.V.P., R.S., and R.E.A. contributed to the conception and design of the study. R.J.J., D.S., E.W., S.H., C.M., R.B., D.M.R., G.A.M., K.A.J., and R.A.B. contributed to the acquisition and analysis of data. K.V.P., R.J.J., D.S., E.W., R.S., and R.E.A. contributed to drafting a significant portion of the manuscript or figures.

Potential Conflicts of Interest

Nothing to report.

Data Availability

To use BRANCH in academic research, please contact the corresponding author.

References

1. Han SD, Nguyen CP, Stricker NH, Nation DA. Detectable neuropsychological differences in early preclinical Alzheimer's disease: a meta-analysis. Neuropsychol Rev 2017;27:305–325.
2. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloid-cognition relations in cognitively normal older adults. Neurology 2013;80:1341–1348.
3. Ossenkoppele R, Pichet Binette A, Groot C, et al. Amyloid and tau PET-positive cognitively unimpaired individuals are at high risk for future cognitive decline. Nat Med 2022;1-7:2381–2387.
4. Donohue MC, Sperling RA, Petersen R, et al. Association between elevated brain amyloid and subsequent cognitive decline among cognitively normal persons. JAMA 2017;317:2305–2316.
5. Machulda MM, Hagen CE, Wiste HJ, et al. Practice effects and longitudinal cognitive change in clinically normal older adults differ by Alzheimer imaging biomarker status. Clin Neuropsychol 2017;31:99–117.
6. Hassenstab J, Ruvolo D, Jasielec M, et al. Absence of practice effects in preclinical Alzheimer's disease. Neuropsychology 2015;29:940–948.
7. Samaroo A, Amariglio R, Burnham S, et al. Diminished learning over repeated exposures (LORE) in preclinical Alzheimer's disease. Alzheimers Dement 2020;12:e12132.
8. Weston PS, Nicholas JM, Henley SM, et al. Accelerated long-term forgetting in presymptomatic autosomal dominant Alzheimer's disease: a cross-sectional study. Lancet Neurol 2018;17:123–132.
9. Zimmermann JF, Butler CR. Accelerated long-term forgetting in asymptomatic APOE ε4 carriers. Lancet Neurol 2018;17:394–395.
10. Bramham CR, Messaoudi E. BDNF function in adult synaptic plasticity: the synaptic consolidation hypothesis. Prog Neurobiol 2005;76:99–125.
11. Dandolo LC, Schwabe L. Time-dependent memory transformation along the hippocampal anterior–posterior axis. Nat Commun 2018;9:1205.
12. Papp KV, Samaroo A, Chou HC, et al. Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. Alzheimers Dement 2021;13:e12243.
13. Papp KV, Buckley R, Mormino E, et al. Clinical meaningfulness of subtle cognitive decline on longitudinal testing in preclinical AD. Alzheimers Dement 2019;16:552–560.
14. Papp KV, Rentz DM, Orlovsky I, et al. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: the PACC5. Alzheimers Dement 2017;3:668–677.
15. Mormino EC, Papp KV, Rentz DM, et al. Early and late change on the preclinical Alzheimer's cognitive composite in clinically normal older individuals with elevated amyloid beta. Alzheimers Dement 2017;13:1004–1012.
16. Buschke H, Fuld PA. Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology 1974;24:1019–1029.
17. Greve DNSD, Bowen S, Izquierdo-Garcia D, et al. Different partial volume correction methods lead to different conclusions: an 18F-FDG-PET study of aging. Neuroimage 2016;132:334–343.
18. Mormino EC, Betensky RA, Hedden T, et al. Synergistic effect of beta-amyloid and neurodegeneration on cognitive decline in clinically normal individuals. JAMA Neurol 2014;71:1379–1385.
19. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multi-day Boston remote assessment of neurocognitive health (BRANCH): feasibility, reliability, and validity. Neuropsychology 2023.
20. Wechsler DSCP. *Wechsler memory scale – revised*. San Antonio, TX: Psychological Corporation, 1987.
21. Castel AD. Memory for grocery prices in younger and older adults: the role of schematic support. Psychol Aging 2005;20:718–721.
22. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012;141:2–18.
23. Ally BA, Hussey EP, Ko PC, Molitor RJ. Pattern separation and pattern completion in Alzheimer's disease: evidence of rapid forgetting in amnestic mild cognitive impairment. Hippocampus 2013;23:1246–1258.
24. Greenaway MC, Lacritz LH, Binegar D, et al. Patterns of verbal memory performance in mild cognitive impairment, Alzheimer disease, and normal aging. Cogn Behav Neurol 2006;19:79–84.
25. Tonegawa S, Morrissey MD, Kitamura T. The role of engram cells in the systems consolidation of memory. Nat Rev Neurosci 2018;19:485–498.
26. Selkoe DJ. Soluble oligomers of the amyloid β-protein impair synaptic plasticity and behavior. Behav Brain Res 2008;192:106–113.
27. Borlikova GG, Trejo M, Mably AJ, et al. Alzheimer brain-derived amyloid β-protein impairs synaptic remodeling and memory consolidation. Neurobiol Aging 2013;34:1315–1327.
28. Ju Y-ES, McLeland JS, Toedebusch CD, et al. Sleep quality and preclinical Alzheimer disease. JAMA Neurol 2013;70:587–593.
29. Mander BA, Marks SM, Vogel JW, et al. β-Amyloid disrupts human NREM slow waves and related hippocampus-dependent memory consolidation. Nat Neurosci 2015;18:1051–1057.
30. Sanchez JS, Becker JA, Jacobs HI, et al. The cortical origin and initial spread of medial temporal tauopathy in Alzheimer's disease assessed with positron emission tomography. Sci Transl Med 2021;13.
31. Farrell ME, Papp KV, Buckley RF, et al. Association of emerging β-amyloid and tau pathology with early cognitive changes in clinically normal older adults. Neurology 2022;98:e1512–e1524.

32. Salthouse TA, Schroeder DH, Ferrer E. Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. Dev Psychol 2004;40: 813–822.

33. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. Clin Neuropsychol 2012;26:543–570.

34. Lim YY, Baker JE, Bruns L, et al. Association of deficits in short-term learning and Aβ and hippocampal volume in cognitively normal adults. Neurology 2020;95:e2577–e2585.

35. Berron D, Ziegler G, Vieweg P, et al. Feasibility of digital memory assessments in an unsupervised and remote study setting. Front Digit Health 2022;4:892997.

36. Nicosia J, Aschenbrenner AJ, Balota DA, et al. Unsupervised high-frequency smartphone-based cognitive assessments are reliable, valid, and feasible in older adults at risk for Alzheimer's disease. J Int Neuropsychol Soc 2023;29:459–471.

37. Thompson LI, Harrington KD, Roque N, et al. A highly feasible, reliable, and fully remote protocol for mobile app-based cognitive assessment in cognitively healthy older adults. Alzheimers Dement 2022;14:e12283.

38. Gills JL, Glenn JM, Madero EN, et al. Validation of a digitally delivered visual paired comparison task: reliability and convergent validity with established cognitive tests. Geroscience 2019;41:441–454.

39. Öhman F, Hassenstab J, Berron D, et al. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimers Dement 2021;13:e12217.

40. Sliwinski MJ. Measurement-burst designs for social health research. Soc. Personal. Psychol. Compass 2008;2:245–261.

41. Jutten RJ, Rentz DM, Fu JF, et al. Monthly At-home computerized cognitive testing to detect diminished practice effects in preclinical Alzheimer's disease. Frontiers in Aging Neuroscience 2022;13:13.

42. Duff K, Lyketsos CG, Beglinger LJ, et al. Practice effects predict cognitive outcome in amnestic mild cognitive impairment. Am J Geriatr Psychiatry 2011;19:932–939.

43. Teunissen CE, Verberk IM, Thijssen EH, et al. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. Lancet Neurol 2022;21:66–77.

44. Barnes LL, Bennett DA. Alzheimer's disease in African Americans: risk factors and challenges for the future. Health Aff 2014;33:580–586.

45. Luchsinger JA, Mayeux R. Cardiovascular risk factors and Alzheimer's disease. Curr Atheroscler Rep 2004;6:261–266.