

# BHA-CS: A novel cognitive composite for Alzheimer's disease and related disorders

Elena Tsoy<sup>1</sup> | Sabrina J. Erhoff<sup>1</sup> | Collette A. Goode<sup>1</sup> | Karen A. Dorsman<sup>1</sup> |  
 Suchanan Kanjanapong<sup>1</sup> | Cutter A. Lindbergh<sup>1</sup> | Renaud La Joie<sup>1</sup> | Amelia Strom<sup>1</sup> |  
 Gil D. Rabinovici<sup>1,2</sup> | Sergio C. Lanata<sup>1</sup> | Bruce L. Miller<sup>1</sup> | Sarah E. Tomaszewski  
 Farias<sup>3</sup> | Joel H. Kramer<sup>1</sup> | Katherine P. Rankin<sup>1</sup> | Katherine L. Possin<sup>1</sup>

<sup>1</sup>Department of Neurology, Memory and Aging Center, University of California San Francisco, San Francisco, California, USA

<sup>2</sup>Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, California, USA

<sup>3</sup>Department of Neurology, University of California Davis, Sacramento, California, USA

## Correspondence

Katherine L. Possin, Associate Professor in Residence, Department of Neurology, University of California San Francisco, Memory and Aging Center, Box 1207, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94158.  
 E-mail: katherine.possin@ucsf.edu

## Funding information

National Institute of Neurological Disorders and Stroke, Grant/Award Number: UG3 NS105557-01; National Institute on Aging, Grant/Award Numbers: P30AG062422, P01AG019724, P30AG010129; Alzheimer's Association, Grant/Award Number: AARF-16-443577

## Abstract

**Introduction:** Composite scores based on psychometrically rigorous cognitive assessments are well suited for early diagnosis and disease monitoring.

**Methods:** We developed and cross-validated the Brain Health Assessment-Cognitive Score (BHA-CS), based on a brief computerized battery, in 451 cognitively normal (CN) and 399 cognitively impaired (mild cognitive impairment [MCI] or dementia) older adults. We investigated its long-term reliability and reliable change indices at longitudinal follow-up (N = 340), and the association with amyloid beta ( $A\beta$ ) burden in the CN subgroup with  $A\beta$  positron emission tomography (N = 119).

**Results:** The BHA-CS was accurate at detecting cognitive impairment and exhibited excellent long-term stability. Reliable decline over one year was detected in 75% of participants with dementia, 44% with MCI, and 3% of CN. Among CN, the  $A\beta$ -positive group showed worse longitudinal performance on the BHA-CS compared to the  $A\beta$ -negative group.

**Discussion:** The BHA-CS is sensitive to cognitive decline in preclinical and prodromal neurodegenerative disease.

## KEYWORDS

Alzheimer disease, computerized cognitive assessment, dementia, early detection, neuropsychology, psychometrics

## 1 | INTRODUCTION

With the growing prevalence of dementia globally, timely detection of neurodegenerative disease is critical for optimizing treatment and care plans. Cognitive assessments play a key role in the accurate diagnosis and monitoring of Alzheimer's disease and related

disorders (ADRD).<sup>1,2</sup> Additionally, numerous studies have shown that the presence of neurodegenerative pathology, such as amyloid plaques in Alzheimer's disease (AD), precedes clinical manifestation of dementia,<sup>3,4</sup> and amyloid-positive subjects are at a greater risk of longitudinal cognitive decline.<sup>5,6</sup> Thus, cognitive measures capable of estimating longitudinal cognitive changes may be particularly sensitive to

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, Inc. on behalf of the Alzheimer's Association.

an underlying neurodegenerative etiology in preclinical stages of the disease.<sup>3,4</sup>

Brief measures of cognition that can be feasibly repeated at subsequent visits are well suited for the detection of meaningful cognitive change because they minimize time and other resource costs compared to traditional neuropsychological tests. Of the variety of brief cognitive measures available, the most commonly used, the Mini-Mental State Examination (MMSE<sup>7</sup>), has well-documented ceiling effects limiting its sensitivity to change over time, especially during preclinical or prodromal disease stages.<sup>8</sup> At the same time, the MMSE is characterized by a particular strength in that it provides a composite score capturing performance across multiple cognitive domains, which enhances its predictive validity given the variability of cognitive deficits associated with neurodegenerative disease.<sup>9</sup> Moreover, the utility of the cognitive composite measures for clinical trials has been supported by the U.S. Food and Drug Administration<sup>10</sup> due to their potential effectiveness in monitoring progression of symptoms over time and evaluating treatment effects.<sup>2,11</sup>

Several other psychometrically sound composite scores have been developed for ADRD, including the cognitive subscale of the Alzheimer's Disease Assessment Scale (ADAS-Cog<sup>12</sup>), the Alzheimer's Prevention Initiative composite cognitive test score,<sup>13</sup> and the Preclinical Alzheimer Cognitive Composite,<sup>14</sup> among others. These metrics are derived from a broader range of cognitive measures tapping into several cognitive domains, thus making them more sensitive to heterogeneity of presentations seen in ADRD.<sup>15</sup> However, most of these composites were developed based on traditional paper-and-pencil measures and require administration by highly trained clinicians or psychometricians, limiting their wide-scale application. Additionally, concerns have been raised regarding the feasibility of such measures for repeated use due to the fact that these instruments rely on rather lengthy cognitive batteries.<sup>15</sup> Computerized cognitive measures offer advantages over existing measures because they can reduce error associated with administration variability, automate scoring and interpretation, provide easy access to alternate forms, and offer efficiency of staffing and cost.<sup>16</sup> Like their traditional counterparts, these measures must be comprised of constituent tasks that evaluate the spectrum of cognitive changes seen in typical and less typical neurodegenerative syndromes for sensitive detection of decline.<sup>15</sup>

In this study, we developed and validated the Brain Health Assessment-Cognitive Score (BHA-CS), a composite score based on a brief tablet-based battery that efficiently measures associative memory, executive function and speed, language generation, and visuospatial skills, and was previously shown to exhibit high accuracy for detecting mild cognitive impairment (MCI) and dementia.<sup>17</sup> We also evaluated its long-term stability and estimated reliable change in cognitively normal (CN) older adults and individuals with MCI and dementia at 1-year follow-up. Finally, we investigated longitudinal sensitivity of the BHA-CS to early AD biomarkers based on amyloid beta ( $A\beta$ )-positron emission tomography (PET) in CN older adults over 3-year follow-up. Across all analyses, we compared the performance of the BHA-CS to the reference brief assessment, the Montreal Cognitive Assessment (MoCA<sup>18</sup>). We selected the MoCA as the reference for

## RESEARCH IN CONTEXT

1. Systematic review: We reviewed the literature on use of cognitive composite scores for early detection of neurodegenerative diseases using traditional sources (eg, PubMed). While several composite metrics have been described, there are ongoing efforts to develop more psychometrically robust and sensitive composites for use in clinical practice and drug trials.
2. Interpretation: Our findings showed that a novel cognitive composite based on a brief computerized battery accurately detected cognitive decline across the spectrum of preclinical and early stage neurodegenerative disease.
3. Future directions: Additional validation of the composite score in culturally diverse populations is needed.

comparison as it has been reported to be sensitive to cognitive changes in mild stages of cognitive decline.<sup>19,20</sup>

## 2 | METHODS

### 2.1 | Participants

The study was approved by the University of California San Francisco (UCSF) Committee on Human Research and the University of California Davis (UCD) Institutional Review Board. All participants provided written informed consent. Subjects were English- or Spanish-speaking adults aged 50 or above recruited from longitudinal observational studies at the UCSF Memory and Aging Center and the UCD Alzheimer's Disease Center. All participants including CN were diagnosed in multidisciplinary clinical consensus conferences based on neurological and neuropsychological examination, clinical interview with an informant including Clinical Dementia Rating (CDR<sup>21</sup>), and structural neuroimaging.<sup>17</sup> This clinical diagnostic work-up was completed at all baseline and follow-up visits. Exclusion criteria included severe psychiatric illness, other non-neurodegenerative neurological condition that could affect cognition, substance use disorder diagnosed in the last 20 years, significant systemic illness, and presence of subjective cognitive concerns or CDR Global Score >0 for CN. Diagnoses of MCI and dementia were made based on published criteria as previously described.<sup>17</sup>

The BHA-CS was developed and cross-validated in a sample of 451 CN, 289 MCI, and 110 mild dementia participants who completed the BHA tests and the MoCA at baseline in English or Spanish. Long-term stability and reliable change estimations were evaluated in the subset of participants who completed the BHA (265 CN, 59 MCI, 16 dementia) and the MoCA (251 CN, 50 MCI, 14 dementia) at baseline and at 1-year follow-up visits. Finally, examination of the association between

$A\beta$  burden and longitudinal changes on the BHA-CS and the MoCA was based on the subset of CN older adults who completed the BHA ( $N = 119$ ; 90  $A\beta^-$ , 29  $A\beta^+$ ) and the MoCA ( $N = 117$ ; 89  $A\beta^-$ , 28  $A\beta^+$ ) at two or more follow-up visits over the course of 3 years and also underwent  $A\beta$ -PET imaging and apolipoprotein E (APOE) genotyping.

## 2.2 | Measures and procedure

### 2.2.1 | Cognitive tests

The BHA is programmed in the TabCAT software platform (UCSF, San Francisco, CA) and consists of four subtests: Favorites (associative memory), Match (executive functioning and processing speed), Line Orientation (visuospatial skills), and Animal Fluency (language).<sup>17</sup> Detailed task descriptions and Spanish adaptation are provided in the supporting information and [memory.ucsf.edu/tabcat](http://memory.ucsf.edu/tabcat). Testing time is 10 minutes. The MoCA is a widely used cognitive screen assessing the following domains: visuoconstruction and executive functions, attention, language, abstraction, memory, and orientation, and also takes  $\approx 10$  minutes to administer.<sup>18</sup> The total score is calculated as a sum of points across domains with a maximum of 30 points. All participants completed the BHA tests and the MoCA in a private examination room independent of diagnostic assessments. For the BHA administration, participants were seated in a chair at a desk with a 9.7-inch iPad positioned horizontally in front of them, with the back of the tablet propped 1 inch up from the surface of the desk. An administrator sat next to the participant. The MoCA was administered in a standardized fashion in accordance with original instructions.<sup>18</sup>

### 2.2.2 | Neuroimaging and APOE genotyping

$A\beta$  imaging was based on PET with 18F-Florbetapir acquired at UCSF China Basin on a GE Discovery STE/VCT PET-CT scanner. Data acquisition and processing was performed in accordance with the Alzheimer's Disease Neuroimaging Initiative protocol<sup>22</sup> (for details, see supporting information). To determine  $A\beta$ -PET positivity, we applied a 1.11 standardized uptake value ratio (SUVR) threshold derived from previous publications using identical acquisition parameters and pre-processing pipelines.<sup>23,24</sup> APOE genotyping was based on DNA analysis from peripheral blood samples as described previously.<sup>25</sup> APOE  $\epsilon 4$  status was dummy coded as "1" for homozygotes and heterozygotes of  $\epsilon 4$  and "0" otherwise. There were no participants with  $\epsilon 2/\epsilon 4$  genotype in the sample.

## 2.3 | Statistical analyses

### 2.3.1 | BHA-CS construction and validation

The BHA-CS was defined as the sum of weighted demographically adjusted scaled scores of BHA component tests. The weights for each

of the component tests were logistic regression parameter estimates derived on the basis of a model discriminating cognitively intact and cognitively impaired (MCI or dementia) participants. First, the raw scores of individual subtests were transformed to z-scores using regression-based normative adjustments for age, education, sex, and language (see supporting information). Then, we built a logistic regression model in which predictors were demographically adjusted z-scores of the BHA subtests, and the outcome was a dichotomized cognitive status variable. All predictors were entered in the model simultaneously, and non-significant predictors were excluded in a stepwise fashion. The derived BHA-CS scores were then scaled to the mean and standard deviation of the CN group.

We used a repeated  $3 \times 5$ -fold cross-validation to estimate the out-of-sample accuracy of the BHA-CS model.<sup>26</sup> The key feature of repeated cross-validation procedures is that they test the effectiveness of the model derived from one part of the dataset (training set) on the reserved part of the dataset (validation set). In our analyses, the cross-validation algorithm divided the dataset into five equally sized random subsets with grossly proportional number of CN and cognitively impaired subjects in each subset and treated each of the subsets as a validation set while the other four were treated as a training set. This procedure was repeated three times based on random permutations. The results were based on the average overall accuracy and average sensitivity and specificity metrics per permutation.

### 2.3.2 | Long-term test stability and reliable change indices

To examine long-term stability, we calculated correlation coefficients (Pearson's  $r$ ) across two visits. These analyses estimate overall reliability of the scores at 1-year follow-up. Based on existing literature,<sup>27,28</sup> we applied the most conservative approaches to reliable change indices (RCI) estimates given the psychometric properties of each measure. The purpose of RCI estimation is to evaluate whether a statistically significant change in test-retest scores has occurred beyond the effects of systematic or measurement errors, such as practice effects, low reliability of the test, or regression to the mean.<sup>28</sup> For the BHA-CS, upper and lower 90% RCI were calculated using a standardized regression-based approach,<sup>29</sup> whereby the predicted follow-up performance score (derived from multiple regression using the baseline score, age, sex, education, and retest interval as predictors) is subtracted from the actual follow-up score, and the difference is standardized based on the standard error of the estimate of the regression equation. Predictors were removed from the model in a stepwise fashion with only significant predictors included in the final model. For the MoCA, 90% RCI were derived based on a formula of the difference between follow-up and baseline scores corrected for practice effects and the standardized error of difference<sup>30</sup> in light of well-documented ceiling effects of the measure and deviation from regression assumptions based on existing studies.<sup>31,32</sup> Reliable improvement or decline was defined as RCI values greater than  $\pm 1.645$  based on a two-tailed prediction with  $\alpha$  set at .05. Independent sample

**TABLE 1** Baseline demographic characteristics and group differences (N = 850)

	CN (N = 451)	MCI (N = 289)	Dementia (N = 110)	P value <sup>a</sup>
Age, years	73.3 ± 8.2	71.1 ± 8.8	70.6 ± 10.2	<.001
Education, years	17.0 ± 2.5	16.4 ± 3.1	15.7 ± 2.8	<.001
Female, n (%)	268 (59%)	128 (44%)	45 (41%)	<.001
White non-Hispanic, n (%)	346 (77%)	238 (82%)	90 (82%)	.303
CDR total	0 ± 0	0.5 ± 0.2	1.1 ± 0.4	<.001
MMSE	29.1 ± 1.1	26.5 ± 2.9	22.5 ± 4.6	<.001
MoCA	26.6 ± 2.3	22.4 ± 4.3	17.1 ± 5.2	<.001

Abbreviations: CDR, Clinical Dementia Rating; CN, cognitively normal; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment.

Values are presented as means ± standard deviations unless otherwise indicated.

<sup>a</sup>Based on differences between MCI and dementia relative to CN and estimated from t-tests for continuous variables or Pearson's  $\chi^2$  tests for categorical variables.

t-tests were performed to examine group differences on RCI on both instruments.

### 2.3.3 | Association with A $\beta$ burden in healthy older adults

Linear mixed effect models using maximum likelihood estimation with random intercepts were used to investigate the relationship between baseline A $\beta$  status and cognitive change in CN subjects. The main model included continuous A $\beta$  SUVR values, time (years since baseline), and SUVR  $\times$  time interaction to investigate changes on the BHA-CS and the MoCA. The covariates included were baseline age (centered), sex, education, APOE  $\epsilon$ 4, and the random effect of cognitive score intercepts. We also examined the APOE  $\epsilon$ 4  $\times$  time and age  $\times$  time interactions to ensure that the changes on cognitive measures over time were not explained by APOE or age variables. The supporting models were analogous to the main one but included a dichotomized A $\beta$  status (positive vs negative) instead of a continuous SUVR as a predictor.

All analyses were performed in R (v3.6.0, R Project for Statistical Computing) with two-tailed significance level set at  $P < .05$ . We used the caret package (v6.0-84)<sup>33</sup> for cross-validation analysis, lme4 package (v1.1-21)<sup>34</sup> for estimation of linear mixed effects models, and ggplot2 (v3.1.1)<sup>35</sup> package for generating figures and plots.

## 3 | RESULTS

### 3.1 | Demographics

Baseline demographic characteristics of the BHA-CS validation sample are presented in Table 1. The CN group was younger, included more women, and had higher education compared to the MCI and dementia groups. Detailed description of the cognitively impaired sample based on clinical syndromes is presented in supporting information. Baseline demographic characteristics of the Spanish-speaking subsample and the follow-up subsamples, which were subsets of the validation sample

**TABLE 2** Summary of the logistic regression model for development of the BHA-CS

	B	SE	z value	P value
Intercept	1.839	0.145	12.67	<.001
Favorites z-score	0.824	0.100	8.21	<.001
Match z-score	0.858	0.100	8.61	<.001
Animal fluency z-score	0.489	0.105	4.67	<.001

Abbreviations: B, unstandardized regression estimate; BHA-CS, Brain Health Assessment-Cognitive Score; SE, standard error.

used to evaluate long-term stability and associations with A $\beta$  burden, are provided in Tables S1-S3 in supporting information. At follow-up, there were no significant differences between the CN and cognitively impaired groups on any of the demographic characteristics, including age, education, sex, or race/ethnicity (Table S2). Similarly, A $\beta$ + and A $\beta$ -CN groups did not differ in age, sex, education, race, or prevalence of APOE  $\epsilon$ 4 status at baseline or at follow-up (Table S3).

### 3.2 | BHA-CS validation

The results of the logistic regression analysis revealed statistically significant contributions of three of the four BHA subtests: Favorites, Match, and Animal Fluency (Table 2). Therefore, the composite metric was expressed as BHA-CS =  $([a + b_{Favorites}z_{Favorites} + b_{Match}z_{Match} + b_{AnimalFluency}z_{AnimalFluency}] - M_{CN})/SD_{CN}$ , where  $a$  is the intercept of the regression equation;  $b$  is unstandardized parameter estimates for Favorites, Match, and Animal Fluency subtests, respectively;  $z$  is demographically corrected z-scores of Favorites, Match, and Animal Fluency subtests, respectively; and  $M_{CN}$  and  $SD_{CN}$  are the mean and standard deviation of the BHA-CS in CN, respectively. The final formula for the BHS-CS was: BHA-CS =  $([1.839 + 0.824*z_{Favorites} + 0.858*z_{Match} + 0.489*z_{AnimalFluency}] - 1.837)/1.516$ .

In the validation sample, 83% of cognitively impaired subjects and 15% of CN subjects had a BHA-CS of  $< -1.0$ , while 85% of CN

**TABLE 3** Rates of reliable change at 1-year follow-up visits on the BHA-CS and the MoCA across diagnostic groups

	Declined, n (%)	Stable, n (%)	Improved, n (%)
CN			
BHA-CS	8/265 (3%)	249/265 (94%)	8/265 (3%)
MoCA	16/251 (6%)	226/251 (90%)	9/251 (4%)
MCI			
BHA-CS	26/59 (44%)	33/59 (56%)	0/59 (0%)
MoCA	10/50 (20%)	38/50 (76%)	2/50 (4%)
Dementia			
BHA-CS	12/16 (75%)	4/16 (25%)	0/16 (0%)
MoCA <sup>a</sup>	3/14 (21%)	10/14 (71%)	1/14 (7%)

Abbreviations: BHA-CS, Brain Health Assessment-Cognitive Score; CN, cognitively normal; MCI, mild cognitive impairment; MoCA, Montreal Cognitive Assessment.

<sup>a</sup>Percentages do not add to 100 due to rounding.

participants and 17% of cognitively impaired subjects had a BHA-CS score of  $-1.0$  or above. Receiver operating characteristic (ROC) analyses were consistent with previously reported results<sup>17</sup> and suggested that the BHA-CS accurately discriminated cognitively intact and cognitively impaired subjects (Figure S1 in supporting information). Results of the cross-validation analysis revealed an overall accuracy of .85 (95% confidence interval [CI]: .82-.87), sensitivity of .88, and specificity of .80. When specificity was set at .85, sensitivity was .83. Analogous preliminary results in the Spanish-speaking group, while limited by a small sample size, suggested that the BHA-CS was similarly accurate at detecting cognitive impairment among Spanish speakers with an overall accuracy of .83 (95% CI: .69-.93), sensitivity of .95, and specificity of .71. When specificity was set at .85, sensitivity was .81 in the Spanish-speaking sample.

### 3.3 | Long-term stability and reliable change indices

Long-term stability at 1-year follow-up on the BHA-CS was excellent in the whole sample ( $r = .89$ , 95% CI: .85-.91) and weaker when restricted to the CN group ( $r = .70$ , 95% CI: .62-.77). MoCA demonstrated poorer stability indices both in the whole sample ( $r = .75$ , 95% CI: .70-.80) and in the CN group only ( $r = .50$ , 95% CI: .40-.59). Significant predictors of the BHA-CS follow-up score were baseline score and baseline age, such that lower baseline score and older age were associated with more decline, while sex, education, and retest interval were not significant (Table S4 in supporting information). The average practice effect on the BHA-CS was 0.2 (Table S2).

Comparing BHA-CS reliable change indices, both MCI ( $t = -9.04$ ,  $P < .001$ ) and mild dementia ( $t = -5.83$ ,  $P < .001$ ) groups showed greater rates of decline over time compared to CN. Similar results were observed on the MoCA, although the effect sizes were smaller in both MCI ( $t = -2.80$ ,  $P = .007$ ) and mild dementia ( $t = -2.03$ ,  $P = .062$ ) compared to CN. The frequency rates of reliable change estimates across diagnostic groups are presented in Table 3.

**TABLE 4** Summary of the linear mixed effect models with baseline continuous amyloid beta  $\times$  time interaction on the BHA-CS and the MoCA in CN ( $N = 119$ )

	BHA-CS			MoCA		
	B	SE	P value	B	SE	P value
Intercept	0.300	1.244	.810	25.906	1.683	<.001
Baseline age	-0.007	0.013	.601	-0.082	0.026	.002
Female sex	0.154	0.158	.332	0.257	0.317	.418
APOE $\epsilon 4$	0.127	0.195	.514	0.537	0.390	.171
A $\beta$ SUVR	0.452	0.739	.542	0.852	1.585	.592
Time	1.099	0.234	<.001	0.380	0.669	.571
A $\beta$ SUVR $\times$ time	-0.957	0.221	<.001	-0.303	0.628	.630

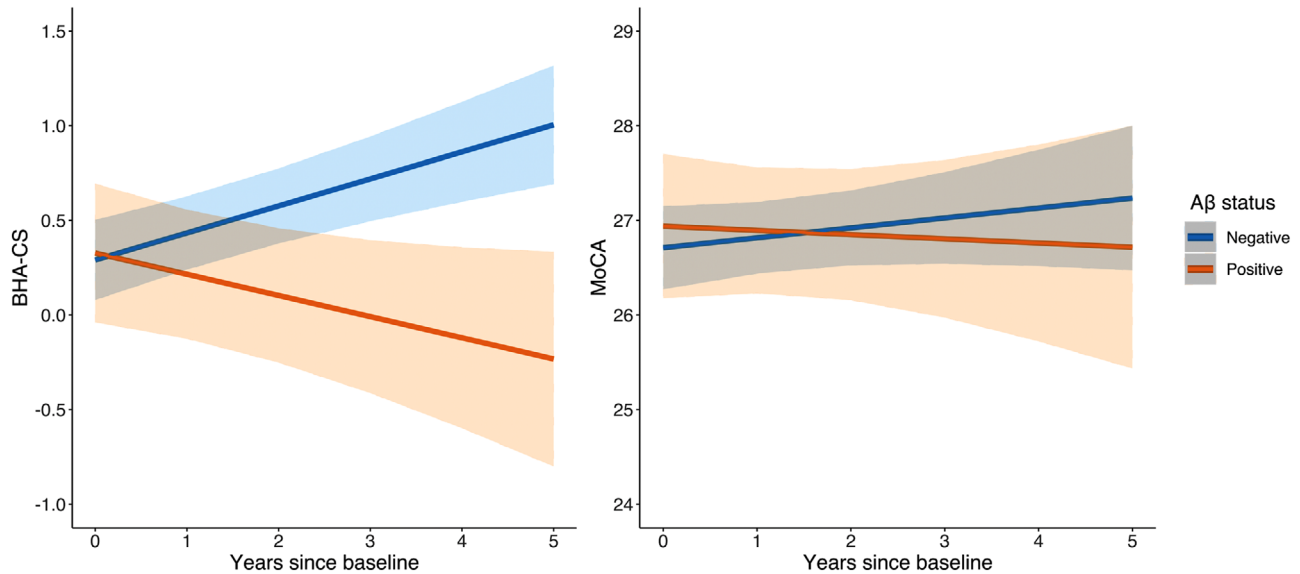
Abbreviations: APOE, apolipoprotein E; A $\beta$ , amyloid beta; B, unstandardized regression estimate; BHA-CS, Brain Health Assessment-Cognitive Score; CN, cognitively normal; MoCA, Montreal Cognitive Assessment; SE, standard error; SUVR, standardized uptake value ratio.

### 3.4 | Associations with A $\beta$ burden in CN

No significant differences between A $\beta$ + and A $\beta$ - groups were observed at baseline, second, or third follow-up visits on the MoCA scores (Table S3). On the BHA, A $\beta$ + group had significantly lower scores than A $\beta$ - group at the third visit (Table S3). Results of the linear mixed effect models showed that greater A $\beta$  SUVR at baseline was associated with poorer performance on the BHA-CS over time ( $B = -0.957$ ,  $SE = 0.221$ ,  $P < .001$ ; Table 4). The SUVR  $\times$  time interaction estimates remained significant after inclusion of APOE  $\epsilon 4 \times$  time ( $B = -0.862$ ,  $SE = 0.231$ ,  $P < .001$ ) and age  $\times$  time ( $B = -0.948$ ,  $SE = 0.218$ ,  $P < .001$ ) interaction terms in the model (Tables S5 and S6 in supporting information). Similar results were observed when a dichotomous A $\beta$  status  $\times$  time interaction was tested with the main effect of A $\beta$ + group showing poorer performance over time on the BHA-CS compared to the A $\beta$ - group ( $B = -0.250$ ,  $SE = 0.061$ ,  $P < .001$ ; Table S7 in supporting information). On the MoCA, neither the continuous ( $P = .630$ , Table 4) nor dichotomous A $\beta \times$  time interactions ( $P = .483$ , Table S7) were significant. Figure 1 depicts trajectories of change over time as a function of baseline dichotomous A $\beta$  status while holding all other fixed effects constant on the BHA-CS and the MoCA (based on Table S7).

## 4 | DISCUSSION

Psychometrically robust and sensitive measures of cognition are critical for early detection and monitoring of progression of neurodegenerative disease. Cognitive outcomes measured by multidomain composite scores have been shown to be particularly well suited to serve this purpose and to be more sensitive to preclinical changes compared to single domain metrics.<sup>36,37</sup> The BHA-CS, a cognitive composite measure optimized for the detection of cognitive impairment, is derived from brief tests of associative episodic memory, executive function and processing speed, and semantic fluency. These domains



**FIGURE 1** Predicted longitudinal performance on the Brain Health Assessment-Cognitive Score and the Montreal Cognitive Assessment in cognitively normal older adults by dichotomous amyloid beta status

have been consistently shown to be affected early in a neurodegenerative process due to AD.<sup>38</sup> The novel measures that contributed most significantly to the BHA-CS composite were Match (executive functions and speed) and Favorites (episodic memory), which may be related to their widespread neural bases in brain regions frequently affected in neurodegenerative disease. Specifically, we previously showed that performance on Match was strongly correlated with gray matter volumes in bilateral frontal, parietal, and subcortical regions, and white matter integrity in the corpus callosum.<sup>17,39</sup> Similarly, performance on Favorites localized to gray matter volumes in bilateral temporal, insular, and frontal regions with most significant associations with medial temporal lobes.<sup>17</sup> In contrast, performance on Line Orientation, which did not significantly contribute to the BHA-CS composite, narrowly localized to a cluster of gray matter in the right superior parietal lobe.<sup>17</sup> While this region represents an important area for diagnosis of neurodegenerative syndromes, it may not have predicted unique variance in cognitive impairment status due to its overlapping neural basis with Match.

Our analyses indicated greater sensitivity of the BHA-CS to change over time in the prodromal (MCI) stage of neurodegenerative disease compared to the commonly used MoCA. These results are particularly relevant to clinical settings, where robust cognitive change estimations may inform diagnostic and treatment decisions. In particular, the BHA-CS captured reliable decline in more individuals with MCI (44%) compared to the MoCA (20%). The BHA also demonstrated greater stability over 1-year period compared to the MoCA, which is consistent with past reports on suboptimal longitudinal reliability of the MoCA in older adults.<sup>31,32</sup>

Longitudinal but not cross-sectional performance on the BHA-CS was sensitive to baseline A $\beta$ -PET in CN older adults. These findings are consistent with previous reports on A $\beta$ -related cognitive decline during preclinical stages of AD.<sup>37,40</sup> Implementation of highly sensitive

cognitive measures that are able to capture cognitive changes in at-risk older adults could help better characterize subtle cognitive changes during the preclinical period. Additionally, these findings provide preliminary support for future use of the BHA-CS in secondary prevention trials with disease-modifying agents given its potential to reliably detect treatment effects over time.

Our study has a number of limitations. First, our sample was largely comprised of highly educated, white participants and may not be representative of the general populations. At the same time, inclusion of Spanish-speaking participants in our sample provided preliminary support for validity of the BHA-CS in discriminating between cognitively intact and cognitively impaired individuals in ethnically diverse groups. Second, our cross-validation procedure did not include a true external sample, which limits generalizability of these results. Data collection for culturally, linguistically, and educationally diverse external samples is ongoing. Also, our longitudinal sample of A $\beta$ + CN participants was relatively small (N = 29), and so these findings merit replication in a larger sample. Exploring associations with other neurodegenerative biomarkers (eg, tauPET, neurofilament light chain) will also be valuable in light of our findings on longitudinal sensitivity to decline in an etiologically diverse cognitively impaired group. Finally, the longitudinal analyses were based on roughly annual visits using the same versions of the BHA tests, while real-life applications may require more frequent administrations using alternate forms, particularly in clinical trials.

In summary, the BHA-CS is a novel cognitive composite measure based on a brief and easy to administer computerized battery that is optimized for the detection of cognitive impairment in older adults. Given its sensitivity to preclinical and prodromal stages of neurodegenerative disease, particularly AD, the BHA-CS represents a valuable alternative to traditional brief cognitive assessments for clinical and research applications.

## ACKNOWLEDGMENTS

We are grateful to our study participants and to Jorge Llibre-Guerra, MD, MSc; Julio Rojas-Martinez, MD, PhD; Luis Fontan, MD; Judith Salvador, PhD; Elizabeth Sequeira, MSc; and Silvia Kochen, MD, PhD for their help with Spanish adaptation of the BHA battery. This study was supported by the National Institute of Neurological Disorders and Stroke [UG3 NS105557-01], the National Institute on Aging [P30AG062422, P01AG019724, and P30AG010129], Alzheimer's Association [AARF-16-443577], the Larry L. Hillblom Foundation, Quest Diagnostics, and the Global Brain Health Institute.

## CONFLICTS OF INTEREST

No authors reported competing interests relevant to this study. Dr. La Joie has received research funding from the Alzheimer's Association. Dr. Rabinovici has received consulting fees or speaking honoraria from GE Healthcare, Axon Neurosciences, Merck, Eisai, Roche, and Genentech; grants from Avid Radiopharmaceuticals, Eli Lilly and Company, GE Healthcare, Life Molecular Imaging; and research funding from the National Institutes of Health (NIH), the Rainwater Charitable Foundation, and the Alzheimer's Association. Dr. Miller has received research funding from the NIH; serves as Medical Director for the John Douglas French Foundation; Scientific Director for the Tau Consortium; Director/Medical Advisory Board of the Larry L. Hillblom Foundation; Scientific Advisory Board Member for the National Institute for Health Research Cambridge Biomedical Research Centre and its subunit, the Biomedical Research Unit in Dementia (UK); and Board Member for the American Brain Foundation (ABF). Dr. Tomaszewski Farias has received research funding from the NIH. Dr. Kramer has received research funding from the NIH and the Larry L. Hillblom Foundation and royalties from Pearson Education, Inc. Dr. Rankin has received research funding from the NIH, Quest Diagnostics, the Rainwater Charitable Foundation, the Weill Family Foundation, and the Marcus Foundation. Dr. Possin has received research funding from the NIH, Quest Diagnostics, the Global Brain Health Institute, the Merck Foundation, and the Rainwater Charitable Foundation. No other disclosures were reported.

## REFERENCES

1. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279.
2. Snyder PJ, Kahle-Wroblewski K, Brannan S, et al. Assessing cognition and function in Alzheimer's disease clinical trials: do we have the right tools?. *Alzheimers Dement*. 2014;10(6):853-860.
3. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):280-292.
4. Mortamais M, Ash JA, Harrison J, et al. Detecting cognitive changes in preclinical Alzheimer's disease: a review of its feasibility. *Alzheimers Dement*. 2017;13(4):468-492.
5. Baker JE, Lim YY, Pietrzak RH, et al. Cognitive impairment and decline in cognitively normal older adults with high amyloid- $\beta$ : a meta-analysis. *Alzheimers Dement*. 2017;6:108-121.
6. Donohue MC, Sperling RA, Petersen R, Sun CK, Weiner MW, Aisen PS. Association between elevated brain amyloid and subsequent cognitive decline among cognitively normal persons. *JAMA*. 2017;317(22):2305-2316.
7. Folstein M, Folstein S, McHugh P. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
8. Spencer RJ, Wendell CR, Giggey PP, et al. Psychometric limitations of the mini-mental state examination among nondemented older adults: an evaluation of neurocognitive and magnetic resonance imaging correlates. *Exp Aging Res*. 2013;39(4):382-397.
9. Jacobs DM, Sano M, Dooneief G, Marder K, Bell KL, Stern Y. Neuropsychological detection and characterization of preclinical Alzheimer's disease. *Neurology*. 1995;45(5):957-962.
10. US Food and Drug Administration. *Early Alzheimer's Disease: Developing Drugs for Treatment Guidance for Industry (Draft Guidance)*; 2018. <https://www.fda.gov/media/110903/download>. Accessed November 27, 2019.
11. Rascovsky K, Salmon D, Lipton A, et al. Rate of progression differs in frontotemporal dementia and Alzheimer disease. *Neurology*. 2005;65(3):397-403.
12. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141(11):1356-1364.
13. Ayutyanont N, Langbaum JB, Hendrix SB, et al. The Alzheimer's prevention initiative composite cognitive test score: sample size estimates for the evaluation of preclinical Alzheimer's disease treatments in presenilin 1 E280A mutation carriers. *J Clin Psychiatry*. 2014;75(6):652-660.
14. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol*. 2014;71(8):961-970.
15. Weintraub S, Carrillo MC, Farias ST, et al. Measuring cognition and function in the preclinical stage of Alzheimer's disease. *Alzheimers Dement*. 2018;4:64-75.
16. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. *Alzheimers Dement*. 2008;4(6):428-437.
17. Possin KL, Moskowitz T, Erhoff SJ, et al. The Brain Health Assessment for detecting and diagnosing neurocognitive disorders. *J Am Geriatr Soc*. 2018;66(1):150-156.
18. Nasreddine Z, Phillips N, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*. 2005;53(4):695-699.
19. Gluhm S, Goldstein J, Loc K, Colt A, Liew CV, Corey-Bloom J. Cognitive performance on the Mini-Mental State Examination and the Montreal Cognitive Assessment across the healthy adult lifespan. *Cogn Behav Neurol*. 2013;26(1):1-5.
20. Lam B, Middleton LE, Masellis M, et al. Criterion and convergent validity of the Montreal Cognitive Assessment with screening and standardized neuropsychological testing. *J Am Geriatr Soc*. 2013;61(12):2181-2185.
21. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*. 1993;43(11):2412-2414.
22. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209.
23. Landau SM, Lu M, Joshi AD, et al. Comparing positron emission tomography imaging and cerebrospinal fluid measurements of  $\beta$ -amyloid. *Ann Neurol*. 2013;74(6):826-836.
24. Landau SM, Breault C, Joshi AD, et al. Amyloid-imaging with Pittsburgh Compound B and Florbetapir: comparing radiotracers and quantification methods. *J Nucl Med*. 2013;54(1):70-77.

25. Staffaroni AM, Brown JA, Casaletto KB, et al. The longitudinal trajectory of default mode network connectivity in healthy older adults varies as a function of age and is associated with changes in episodic memory and processing speed. *J Neurosci*. 2018;38(11):2809-2817.
26. Burman P. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika*. 1989;76(3):503-514.
27. Heaton RK, Taylor MJ, Marcotte TD, et al. Detecting change: a comparison of three neuropsychological methods, using normal and clinical samples. *Arch Clin Neuropsychol*. 2001;16(1):75-91.
28. Hinton-Bayre AD. Deriving reliable change statistics from test-retest normative data: comparison of models and mathematical expressions. *Arch Clin Neuropsychol*. 2010;25(3):244-256.
29. McSweeney AJ, Naugle RI, Chelune GJ, Luders H. "T scores for change": an illustration of a regression approach to depicting change in clinical neuropsychology. *Clin Neuropsychol*. 1993;7(3):300-312.
30. Chelune GJ, Naugle RI, Luders H, Sedlak J, Awad IA. Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology*. 1993;7(1):41-52.
31. Cooley SA, Heaps JM, Bolzenius JD, et al. Longitudinal change in performance on the Montreal Cognitive Assessment in older adults. *Clin Neuropsychol*. 2015;29(6):824-835.
32. Kopecek M, Bezdicek O, Sulc Z, Lukavsky J, Stepankova H. Montreal Cognitive Assessment and Mini-Mental State Examination reliable change indices in healthy older adults. *Int J Geriatr Psychiatry*. 2017;32(8):868-875.
33. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28(5):1-26.
34. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1-48.
35. Wickham H. *Ggplot2 - Elegant Graphics for Data Analysis*. 2nd ed. New York, NY: Springer-Verlag; 2016. <https://www.springer.com/gp/book/9783319242750>.
36. Langbaum JB, Hendrix SB, Ayutyanont N, et al. An empirically derived composite cognitive test score with improved power to track and evaluate treatments for preclinical Alzheimer's disease. *Alzheimers Dement*. 2014;10(6):666-674.
37. 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(9):1004-1012.
38. Wilson RS, Leurgans SE, Boyle PA, Bennett DA. Cognitive decline in prodromal Alzheimer disease and mild cognitive impairment. *Arch Neurol*. 2011;68(3):351-356.
39. Alioto AG, Mumford P, Wolf A, et al. White matter correlates of cognitive performance on the UCSF Brain Health Assessment. *J Int Neuropsychol Soc*. 2019;25(6):654-658.
40. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology*. 2013;80(14):1341-1348.

### SUPPORTING INFORMATION

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

**How to cite this article:** Tsoy E, Erloff SJ, Goode CA, et al. BHA-CS: A novel cognitive composite for Alzheimer's disease and related disorders. *Alzheimer's Dement*. 2020;12:e12042. <https://doi.org/10.1002/dad2.12042>