Neurol Ther (2025) 14:865–880
https://doi.org/10.1007/s40120-025-00741-x

ORIGINAL RESEARCH

The Cognivue Amyloid Risk Measure (CARM): A Novel Method to Predict the Presence of Amyloid with Cognivue Clarity

James E. Galvin · Michael J. Kleiman · Heather M. Harris · Paul W. Estes

Received: February 12, 2025 / Accepted: March 26, 2025 / Published online: April 7, 2025
© The Author(s) 2025

ABSTRACT

***Introduction*:** At the present time, clinical detection of individuals who have amyloid in their brain is not possible without expensive biomarkers. The objective of the study was to test whether Cognivue Clarity® can differentiate True Controls, preclinical Alzheimer's disease (pAD), mild cognitive impairment (MCI) due to Alzheimer's disease (MCI-AD), AD, and MCI and dementia due to non-AD etiologies enrolled in the Bio-Hermes Study.

***Methods*:** A total of 887 individuals completed Cognivue Clarity, amyloid PET scan, and blood-based AD biomarkers. Three Cognivue Clarity subtests differentiated between True Controls and pAD, and between cognitive impairment due to AD versus non-AD processes. This finding was leveraged to develop an amyloid-specific marker, combining the three subtests with age using machine learning to create the 4-point Cognivue Amyloid Risk Measure (CARM).

***Results*:** Cognivue Clarity discriminated cognitively normal from cognitively impaired individuals ($p < 0.001$, Cohen's $d = 0.732$). The CARM differentiated between individuals with amyloid and without amyloid by PET ($p < 0.001$, Cohen's $d = 0.618$) and blood-based biomarkers (p's < 0.001). Amyloid positivity and cognitive impairment increased across four CARM thresholds ($p < 0.001$). Dichotomizing CARM thresholds into low (CARM1/CARM2) and high (CARM3/CARM4) likelihood provided excellent discrimination for amyloid PET positivity (OR: 3.67; 95% CI 2.76–4.89). CARM categories differentiated between True Controls, pAD, MCI-AD, AD, and cognitive impairment due to non-AD etiologies ($\chi^2 = 137.6$, $p < 0.001$) with the majority of True Controls and non-AD etiologies being in CARM1/CARM2, and the majority of pAD, MCI-AD, and AD being in CARM3/CARM4.

***Conclusions*:** Cognivue Clarity detects individuals with cognitive impairment, and a derivation benchmarked against amyloid PET was used to develop the CARM to predict the presence of amyloid. Combining the CARM and the Cognivue Clarity overall score could help identify individuals with and without cognitive impairment due to AD or non-AD etiologies, help screen for treatment protocols with anti-amyloid

Prior Presentation: This work was previously presented in abstract form at the 2024 Clinical Trials on Alzheimer's Disease (CTAD) meeting on October 31, 2024 in Madrid, Spain.

J. E. Galvin (✉) · M. J. Kleiman
Department of Neurology, Comprehensive Center for Brain Health, University of Miami Miller School of Medicine, 7700 W Camino Real, Suite 200, Boca Raton, FL 33433, USA
e-mail: jeg200@miami.edu

J. E. Galvin · H. M. Harris · P. W. Estes
Cognivue, Inc, 7911 Rae Blvd, Victor, NY 14564, USA

therapies, enrich clinical trial recruitment, and help to identify pAD for prevention studies.
Trial Registration: ClinicalTrials. gov identifier, NCT04733989.

Keywords: Alzheimer's disease; Amyloid PET; Bio-Hermes; Cognivue; Cognitive assessment; Digital biomarker; Mild Cognitive Impairment; Preclinical Alzheimer's disease

Key Summary Points

Why carry out this study?

At the present time, clinical detection of individuals who have amyloid in their brain is not possible without expensive biomarkers. This could delay diagnosis and miss opportunities for treatment and clinical trial enrollment.

Cognivue Clarity, the first FDA-cleared computerized cognitive test, can detect cognitive impairment, distinguish biomarker-confirmed clinical diagnoses, and identify people with preclinical Alzheimer's disease (AD).

The Cognivue Amyloid Risk Measure (CARM) was developed using machine learning to provide a risk of amyloid presence and, when combined with Cognivue Clarity, global scores can characterize True Controls, preclinical AD (pAD), cognitive impairment due to AD, and cognitive impairment due to a non-AD process.

What was learned from the study?

CARM discriminated individuals with amyloid and without amyloid by PET ($p < 0.001$, Cohen's $d = 0.618$) and blood-based biomarkers ($ps < 0.001$).

While not a diagnostic test, the CARM is "added value" provided along with Cognivue Clarity scores suggesting that amyloid could be present and providing additional information which could be useful to researchers and clinicians in deciding next steps.

Cognivue Clarity can help screen for treatment protocols with anti-amyloid therapies, enrich clinical trial recruitment before expensive biomarkers, and identify individuals likely to have pAD for prevention studies.

INTRODUCTION

Alzheimer's disease (AD) affects nearly 7 million people in the US and more than 55 million people worldwide [1]. Detection of mild AD in the community is a clinical challenge, with many people coming to medical attention and diagnosis at the moderate stage, particularly in minoritized populations [2, 3]. This is likely too late for treatment with amyloid-lowering therapies or to enroll in most AD clinical trials. Mild cognitive impairment (MCI) often represents a prodromal state of AD (MCI-AD), and recent reports suggest that nearly 80% of people living with MCI are not diagnosed until progression to dementia [4, 5]. This is a missed opportunity for earlier intervention before loss of functional independence with amyloid-lowering therapies or to enroll in a clinical trial. Just as disease-modifying therapy trials are recruiting individuals with MCI and mild AD, there are prevention trials targeting individuals with preclinical AD (pAD) [6, 7]. Individuals with pAD are cognitively normal individuals who demonstrate evidence of AD pathology [8]. Individuals with pAD are not typically recognized in clinical practice as they are cognitively normal and would therefore not undergo biomarker testing. Thus, there is an unmet need for a brief but sensitive way to discriminate individuals who are likely to have amyloid from those who are less likely to have amyloid. This strategy, when combined with brief cognitive testing, could enrich recruitment into clinical trials for pAD, MCI-AD, and AD; lower costs (e.g., time, effort, financial) related to expensive amyloid PET scans by prescreening individuals more likely to have an abnormal scan; and allow clinicians to intervene with amyloid-lowering therapies at the earliest possible stages.

Cognivue Clarity (Cognivue, USA) is the first FDA-cleared digital biomarker that uses computerized cognitive assessment and adaptive psychophysics to detect the risk of cognitive impairment. It evaluates baseline visual and motor skills to reliably assess information processing and memory and adapt subsequent testing to these baseline abilities, thus minimizing many of the biases found in common cognitive testing mechanisms and brief screening tests [9]. We previously demonstrated the ability of Cognivue Clarity to (a) detect cognitive impairment across different age ranges, sociodemographic groups, and clinical diagnoses [9, 10], (b) characterize the presence of amyloid by PET scan [10, 11], (c) distinguish between biomarker-confirmed groups (AD vs non-AD process) [10], and (d) identify individuals with pAD [11]. We leveraged the Bio-Hermes Study [12], funded by the Global Alzheimer Platform Foundation, to explore the ability of Cognivue Clarity to detect the presence of amyloid in individuals with and without cognitive impairment.

We identified three subtests contained within Cognivue Clarity (adaptive motor, visual salience, shape discrimination, Fig. 1) that are sensitive to the presence of amyloid and assist in the differentiation between True Controls (cognitively normal/biomarker negative) and pAD (cognitively normal/biomarker positive) [11]. Although both groups scored within the normal range, significant differences in the three individual tests and the 3-test mean were found between True Controls and pAD. Further, performance on the three individual tests and the 3-test mean was similar between pAD and MCI-AD ,supporting that these three tests could be further explored as a screen for the presence of amyloid. Here, we describe the development and validation of the Cognivue amyloid risk measure (CARM) as a brief, sensitive, and valid marker of amyloid that can be used to screen for individuals who have amyloid present in their brains, regardless of cognitive status. When the CARM is combined with the overall score from the full 10-min Cognivue Clarity battery, individuals can not only be screened for the likelihood of cognitive impairment but it may also be possible to screen them for the likelihood of amyloid and characterize them as True Controls, pAD, cognitive impairment due to AD, and cognitive impairment due to a non-AD process.

METHODS

Study Design and Participants

The Bio-Hermes Study was designed to explore the relationship between blood-based and digital biomarkers and amyloid status as determined by PET scan. The complete study design has been previously described [12]. Between April 2021 and November 2022, 1001 participants were enrolled, of which 887 completed both Cognivue Clarity and amyloid PET scans. Bio-Hermes inclusion criteria included age 60–85 years, fluency in English or Spanish, and a Mini-Mental State Examination (MMSE) [13] score between 20 and 30 inclusive. Exclusion criteria included history of depression, strokes or seizures in the past year, or cancer within the past 5 years, or a negative amyloid PET scan in the past year. Written informed consent was obtained from all study participants. This study was reviewed and approved by Advarra, a central institutional review board (Reference Number Pro00046018), and was conducted in accordance with the Helsinki Declaration of 1964. The study was registered on ClinicalTrials.gov (NCT04733989).

Clinical Diagnoses

The Bio-Hermes study stratified individuals into three clinical cohorts (Cognitively Normal, MCI, probable AD) using National Institute on Aging-Alzheimer Association (NIA-AA) consensus clinical criteria for MCI-AD [14] and AD [15]. Diagnostic characterization was based on the participants' performance on the MMSE, the Rey Auditory Verbal Learning Test (RAVLT) [16] and the Functional Activities Questionnaire (FAQ) [17]. The Cognitively Normal cohort had MMSE scores 26–30 inclusive, age- and race-adjusted normative RAVLT scores, and no functional decline on the FAQ. The MCI-AD cohort met NIA-AA criteria [14], scored 24–30 (inclusive) on the MMSE, scored

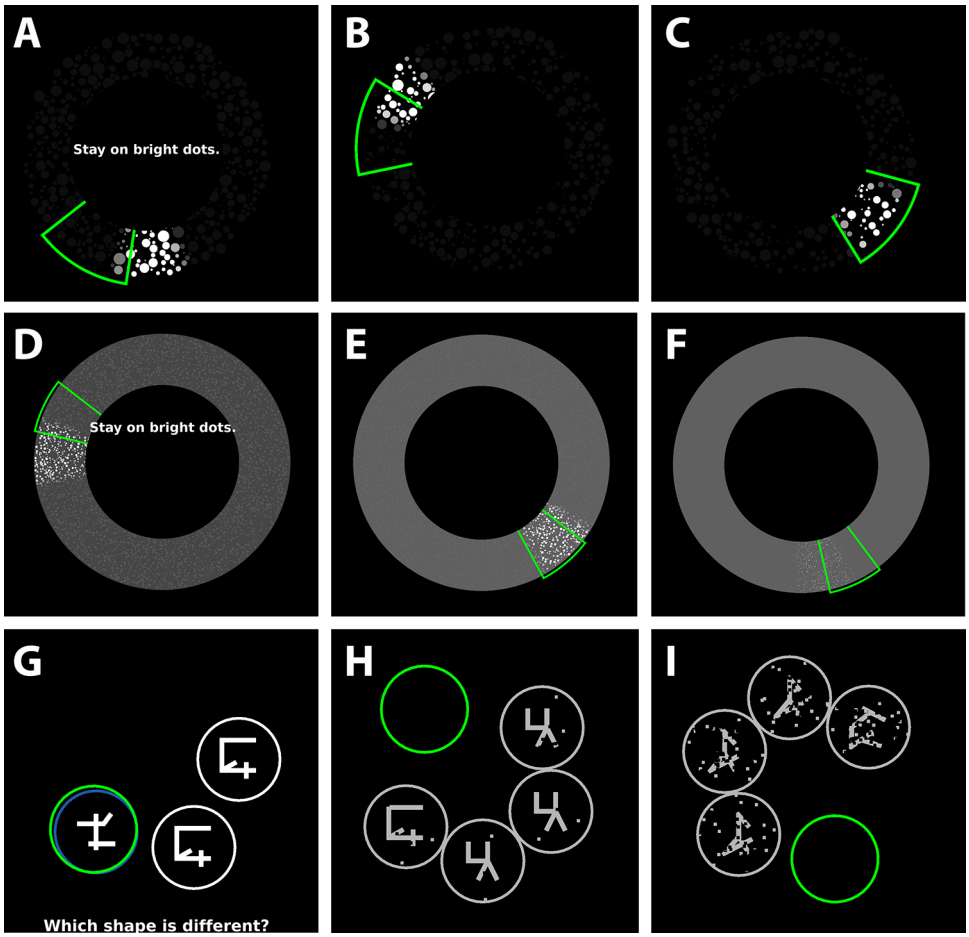

Fig. 1 Cognivue Clarity tests that are amyloid sensitive. Three cognivue clarity subtests (adaptive motor control, visual salience, shape discrimination) were determined to be amyloid-sensitive and able to differentiate True Controls (cognitively normal/amyloid negative) individuals from preclinical AD (cognitively normal/amyloid positive). The adaptive motor control subtest (**A**) evaluates the individual's ability to control their rotatory movement response of the cognivue wheel in response to a rotating visual cue (**B**) with the speed of response measured once the respondent can keep up with the moving stimuli (**C**). The visual salience subtest (**D**) measures the threshold of the respondent to determine visual contrast of the target stimuli presented locating a wedge of bright dots on the screen and follow them as they begin to fade (**E**). The stimuli will continue to fade until the respondent can no longer detect the stimuli (**F**). The shape discrimination subtest (**G**) quantifies the individual's ability to discriminate one regular shape from others in the annular display (**H**). This test will continue with the stimuli slowly degrading until the respondent can no longer correctly identify the stimuli (**I**)

at least 1 standard deviation below age- and race-adjusted norms on the RAVLT, and had no functional decline on the FAQ. Mild AD individuals met NIA-AA criteria [15] and scored 20–24 (inclusive) on the MMSE, scored ≥ 1 standard deviation below age- and race-adjusted norms on the RAVLT, and had functional decline on the FAQ.

Alzheimer's Disease Biomarkers

Participants underwent amyloid PET using 18F-Florbetapir (Eli Lilly and Company) and the standardized uptake value ratio (SUVR) was calculated [18, 19]. Amyloid status (elevated vs. not elevated) was established using a SUVR of 1.1, which is equivalent to a Centiloid Level of

24.1 consistent with thresholds used in recent AD clinical trials on amyloid-lowering monoclonal antibodies [20, 21]. Participants also had blood-based biomarkers collected and analyzed by C_2N Diagnostics laboratories (PrecivityAD) [22] for Aβ40, Aβ42, Aβ42/Aβ40, ApoE ε4, and the Amyloid Probability Score (APS), Quanterix laboratories for pTau181 [23], and Eli Lilly and Company for pTau217 [24]. The amyloid status for the cohort included 567 amyloid-negative and 353 amyloid-positive individuals.

Clinical-Pathologic Diagnosis

The clinical diagnosis was combined with amyloid status in 887 individuals who completed both Cognivue Clarity and amyloid PET to develop clinicopathologic groups consisting of True Controls (cognitively normal + the absence of amyloid, n = 297), pAD (cognitively normal + the presence of amyloid, n = 91), MCI-AD (cognitive impairment + the presence of amyloid, n = 111), MCI due to a process other than AD (MCI-non-AD, cognitive impairment + the absence of amyloid, n = 171), dementia due to AD (dementia + the presence of amyloid, n = 130) and dementia due to a process other than AD (non-AD, dementia + the absence of amyloid, n = 87).

Development of the CARM

The CARM was developed as an indicator of the likelihood of a patient exhibiting amyloid positivity as defined by SUVR threshold. Primary strategy was to explore the utility of five methodologies: (1) a standardized mean strategy, (2) linear modeling of centiloids using regression models, (3) classification modeling of amyloid positivity, (4) hyperplane separation using support vector machine (SVM) classification coefficients, and (5) ensemble modeling of the combined above models that perform best, with and without weighting of each model.

Prior to any examination, one-third of the data (stratified based on both SUVR threshold positivity and impairment status) was split into a separate test file and only used in final evaluation. The remaining two-thirds were used in all model development. Distribution and evaluation of statistical differences using pairwise t tests was performed. No significant differences in age, sex, race, or years of education were found between the two sets after applying Bonferroni correction for multiple comparisons. After visually exploring the distribution of the data of each measure with respect to amyloid positivity, the following strategies were used.

In the standardized mean strategy, each standard deviation away from the mean is assigned a distinct value (i.e., 1 point for one SD, 2 points for 2 SDs, etc.) with weighting assigned to one or more measures identified via factor analysis and/or K-means clustering. The best threshold for classifying amyloid positivity was then determined from the derived score using Youden's J. The linear modeling of centiloids strategy utilized machine learning regression models, including RF regression, linear regression, gradient boosted machine (GBM) regression, and SVM regression, to predict centiloid score using only the identified Cognivue measures. Centiloids were derived from the Lilly SUVR using the following formula [25], with S signifying SUVr and c signifying Florbetapir centiloid score:

$$c = 183S - 177.$$

The centiloid score was then further refined by applying a lower threshold of 0 (converting all negative numbers to 0) and an upper threshold of 100 (lowering all numbers greater than 100), resulting in a 101-point scale. This final scale was used as the dependent variable in regressions. Regressions were evaluated using error metrics (mean absolute error, root mean squared error), as well as classification metrics using the known centiloid threshold of 24.1.

The classification modeling strategy utilized logistic regression, RF classification, SVM classification, and GBM classification, with each model used to predict belongingness to either negative or positive amyloid groups defined by a SUVR-derived centiloid threshold of 24.1. Metrics for evaluating model performance included sensitivity, specificity, positive predictive value, negative predictive value, F1

score, and diagnostic odds ratio. Scikit-learn [26] was used for all model development, with the exception of GBM where Microsoft's LightGBM [27] package was utilized. Combining methods from both classification and regression, the hyperplane separation strategy utilized coefficients derived from a trained SVM classification model [28] to apply weights to each of the identified Cognivue measures. The final score was then thresholded to identify the best performing value for indication of amyloid positivity.

Finally, the ensemble modeling strategy aimed to develop a scale based on the outputs of the best performing models outlined above [29]. Classification output probabilities were assigned 3 points if above 0.8 (80% confidence) and 1 point if above 0.5 (50% confidence), with 1 point removed if confidence was less than 20%. Regressions were assigned 1 point if above the defined SUVR threshold of amyloid positivity. Hyperplane separation score was assigned 1 point if above the identified best threshold. A standardized mean was assigned 1 point if above the identified best threshold. Hyperparameter optimization was performed for all models on two-thirds of the available training set, with the remaining one-third of the training set used as an intermediary validation set. Final testing utilized hyperparameters established by the training set.

Post-development Modifications

Following the development of the original CARM and selection of the final model strategy, further refinements were undertaken to enhance the accuracy of the model's sensitivity, specificity, overall accuracy, and precision in determining amyloid positivity. These refinements included examining the effects of rescaling or unscaling centroid levels, adding additional variables including demographics information to the model, and recalibrating the scale. The final model resulted in partially unscaling centroids to allow negative values but not values over 100, including age as an additional variable, and recalibrating the CARM to be a 4-point instead of a 3-point scale, with a score of 1 indicating a low specificity (below 80% sensitivity), a score of 2 indicating good specificity but low sensitivity (above 80% sensitivity but below Youden's J), a score of 3 indicating below 80% specificity but above Youden's J, and a score of 4 indicating high specificity (above 80%).

RESULTS

Sample Characteristics

A total of 887 individuals enrolled in BioHermes completed both Cognivue Clarity testing and amyloid PET scan representing 89% of the total sample. The 388 individuals rated as cognitively normal included 297 True Controls and 91 pAD. The 282 individuals diagnosed with MCI included 111 MCI-AD and 171 MCI-non-AD cases. The 217 individuals diagnosed with dementia cases included 130 AD and 87 non-AD cases. The mean age of the sample was 71.8 ± 6.7 years (range 59–85) with 15.5 ± 2.7 years of education (range 4–24). The sample was 56.4% female, 37.3% were ApoE e4 carriers, and the ethnoracial self-identification included 78.8% non-Hispanic white, 9.6% Black, 9.8% Hispanic, and 1.8% Asian. The mean MMSE score was 26.9 ± 2.7 (range 17–30), the mean FAQ score was 3.6 ± 5.2 (range 0–30), and the mean total score for the RAVLT was 40.6 ± 14.1 (range 0–67). The mean Cognivue Clarity score was 63.8 ± 16.7 (range 0–93). Table 1 shows the sample characteristics (demographics, cognitive performance, biomarkers) across the clinical-pathologic diagnostic groups.

Defining the CARM

The first version of the CARM was used to identify the best performing model, after which additional versions refined model performance. For this first version's ensemble model strategy, regressions and hyperplane separation were assigned 1 point above the best identified thresholds; a score of 26 for linear regressions, 24.1 (standard SUVR threshold) for RF and GBM regressions, and 63 for hyperplane separation. Three points were assigned if the score was two

Table 1 Sample characteristics

	True control	pAD	MCI-AD	AD	MCI non-AD	Non-AD	*p* value[B]
Age, years	69.5 (6.2)	72.7 (6.6)	74.2 (6.1)	75.4 (5.7)	70.7 (6.9)	72.8 (6.6)	< 0.001[a,d,e,f]
Education, years	15.8 (2.4)	15.8 (2.4)	15.9 (3.1)	15.1 (3.3)	15.3 (2.5)	14.9 (2.5)	0.027
Sex, % female	60.9	58.2	47.7	51.5	56.7	56.3	0.203
Race, % white	86.9	86.8	91.9	89.2	84.8	81.6	0.593
Ethnicity, % hispanic	8.1	2.2	7.2	15.4	12.3	12.6	0.013
ApoE4 carriers, %	25.2	57.1	64.9	62.8	20.8	15.7	< 0.001
MMSE	28.4 (1.5)	28.5 (1.4)	27.0 (1.9)	23.0 (2.5)	27.5 (1.9)	24.3 (2.3)	< 0.001[a,b,d,f]
FAQ	0.6 (1.4)	1.0 (1.9)	4.9 (5.2)	9.5 (6.1)	2.7 (3.4)	7.5 (6.9)	< 0.001[b,c,d,e,f]
RAVLT, total score	47.9 (13.4)	46.9 (13.6)	38.3 (10.4)	30.6 (10.4)	36.9 (12.1)	33.9 (13.6)	< 0.001[b,c,d]
Cognivue, total score	71.9 (12.5)	68.0 (12.2)	60.1 (16.2)	49.4 (17.4)	65.6 (14.9)	54.6 (17.1)	< 0.001[a,b,c,e,f]
Adaptive motor control[A]	42.6 (15.4)	37.3 (14.8)	35.0 (16.6)	27.3 (15.6)	39.2 (16.7)	29.7 (15.9)	< 0.001[a,b,c,e]
Visual salience[A]	69.9 (13.7)	65.3 (13.9)	60.4 (16.5)	51.6 (16.3)	65.4 (16.6)	55.9 (15.3)	< 0.001[a,b,c,e,f]
Shape discrimination[A]	74.6 (19.7)	67.2 (20.8)	57.4 (24.4)	43.7 (26.5)	68.2 922.3)	48.4 (28.3)	< 0.001[a,b,c,e,f]
Amyloid PET SUVR	0.96 (0.06)	1.31 (0.17)	1.38 (0.21)	1.46 (0.21)	0.96 (0.70)	0.98 (0.08)	< 0.001[a,b,c,d,e,f]
Aβ42/40 ratio	0.100 (0.009)	0.091 (0.008)	0.089 (0.006)	0.092 (0.008)	0.101 (0.009)	0.101 (0.008)	< 0.001[a,d,e,f]
APS	16.4 (19.4)	55.8 (26.9)	59.7 (26.5)	56.4 (29.4)	17.9 (20.3)	18.9 (23.9)	< 0.001[a,d,e,f]
pTau217	0.17 (0.06)	0.26 (0.10)	0.36 (0.23)	0.47 (0.25)	0.18 (0.11)	0.23 (0.17)	< 0.001[a,b,c,d,e,f]
pTau181	16.2 (7.5)	22.9 914.6)	24.2 (11.1)	30.7 (19.2)	17.9 (10.0)	19.3 (12.2)	< 0.001[a,c,e,f]

Mean (SD) or %

AD Alzheimer's disease, *APS* Amyloid Probability Score, *CARM* cognivue amyloid risk measure, *FAQ* functional activities questionnaire, *MMSE* Mini Mental State Exam, *MCI-AD* mild cognitive impairment due to Alzheimer's disease, *MCI non-AD* mild cognitive impairment due to a non-Alzheimer process,, *Non-AD* dementia due to a non-Alzheimer process, *pAD* preclinical Alzheimer's disease *RAVLT* rey auditory verbal learning task, *SUVR* standardized uptake value ratio

[A]Congivue Clarity subtests included in the CARM

[B]Corrected for multiple comparisons using Bonferroni correction

Post-hoc comparisons

[a]True Controls different than pAD

[b]pAD different from MCI-AD

[c]MCI-AD different from AD

[d]pAD different from MCI non-AD

[e]MCI-AD different from MCI non-AD

[f]AD different from non-AD

standard deviations above that threshold: 85 for hyperplane separation, 52 for logistic and GBM regressions, and 54 for RF regression. The cutoffs for risk were then determined to be 3 points for moderate risk and 6 points for high risk.

The validation set (1/3 of training set used for validating models) indicated that the ensemble model was the best performing model (see Table 2) with sensitivity of 72.6%, specificity of 63.4%, and a diagnostic odds ratio (DOR) of 4.59. However, running all models on the holdout test set revealed that the ensemble model was likely overfit, resulting in the lowest available DOR and F1 score. Instead, the best performing model on the final test set was the GBM Regression model with a DOR of 3.28 (Table 2), which also achieved the second highest score in the validation set. While specificity was consistently lower than the ensemble model, sensitivity was higher. Thus, the GBM Regression model was chosen to calculate the CARM. For comparison, a model utilizing the GBM regression strategy was developed using only demographics data: age, sex, years of education, and race; this model was the worst-performing model on the validation set and was only comparable to the overfit ensemble model on the holdout set (Table 2).

Refining the CARM

The refined CARM introduced age as a predictor variable and unscaled the lower levels of the dependent variable. New thresholds were developed based on specificity, sensitivity, and Youden's J metrics: CARM 1 used a raw score cutoff of 5 or below, determined to be the threshold for 80% sensitivity; CARM 2 used a raw score cutoff of 11.7 or below, determined to be the Youden's J; CARM 3 used a raw score cutoff of 22.1 or below, determined to be the threshold for 80% specificity; CARM 4 was assigned to those with a raw score above 22.1. Raw scores for the CARM are not directly interpretable, instead, the thresholds were utilized to establish the likelihood of amyloid presence.

Table 3 describes the association between the CARM thresholds, amyloid PET, and clinical diagnoses. Amyloid positivity increased across the 4 CARM thresholds (CARM1: 19%, CARM2: 12%, CARM3: 26%, CARM4: 43%, $p < 0.001$). Cognitive impairment also increased across

Table 2 Modeling the cognivue amyloid risk measure (CARM)

	Sensitivity	Specificity	PPV	NPV	F1 score	DOR
Validation set						
Linear regression	0.74	0.60	0.52	0.80	0.61	4.29
Random forest regression	0.75	0.50	0.47	0.78	0.58	3.11
GBM regression	0.81	0.52	0.50	0.82	0.62	4.57
Hyperplane separation	0.71	0.63	0.54	0.79	0.61	4.29
Ensemble model	0.73	0.63	0.54	0.80	0.62	4.59
Demographics only (GBM)	0.61	0.61	0.66	0.66	0.61	2.39
Holdout set						
Linear regression	0.64	0.59	0.48	0.73	0.55	2.51
Random forest regression	0.68	0.54	0.47	0.74	0.55	2.47
GBM regression	0.73	0.55	0.49	0.77	0.59	3.28
Hyperplane separation	0.60	0.61	0.47	0.72	0.53	2.30
Ensemble model	0.60	0.60	0.47	0.72	0.53	2.25
Demographics only (GBM)	0.60	0.60	0.67	0.67	0.60	2.25

DOR diagnostic odds ratio, *GBM* gradient boosted machine, *NPV* negative predictive value, *PPV* positive predictive value

Table 3 Cognivue amyloid risk measure (CARM) association with biomarkers and diagnoses

Condition	CARM 1	CARM 2	CARM 3	CARM 4	CARM 3+
% amyloid-positive	19.9%	11.8%	25.7%	42.6%	68.3%
% amyloid-negative	44.0%	17.0%	19.1%	19.9%	39.0%
% preclinical AD	35.2%	15.4%	24.2%	25.3%	49.5%
% MCI due to AD	14.2%	10.4%	26.3%	49.2%	75.4%
% non-AD impairment	34.6%	14.8%	23.3%	27.2%	50.6%
% healthy control	52.2%	18.9%	15.5%	13.5%	29.0%

AD Alzheimer's Disease, *CARM 3+* combining CARM 3 and CARM 4, *CI* cognitive impairment, *MCI* mild cognitive impairment

CARM thresholds (CARM1: 40%, CARM2: 47%, CARM3: 64%, CARM4: 75%, $p < 0.001$). Dichotomizing the CARM into low likelihood of amyloid (CARM1, CARM2) and high likelihood of amyloid (CARM3, CARM4) provided excellent discrimination for amyloid positivity by PET (OR 3.67, 95% CI 2.76–4.89) with 68.3% of amyloid-positive individuals correctly classified, including half of individuals with pAD, discriminating individuals with amyloid from individuals without amyloid ($p < 0.001$, Cohen's $d = 0.618$).

Independent samples ANOVAs with Bonferroni-corrected pairwise post hoc tests revealed significant differences in clinical measures and biomarkers across CARM thresholds in all comparisons except sex (Table 4). Of note, many regional cortical PET SUVR values were found to significantly differ between CARM 1 and CARM 3, likely a result of the placement of MCI and other cognitively impaired patients into this group. CARM 1 and 2 were the most similar to each other, potentially due to the predicted raw CARM values being significantly lower than the centiloid cutoff for amyloid positivity. CARM thresholds differentiated True Controls, pAD, MCI-AD AD, and cognitive impairment due to non-AD etiologies ($\chi^2 = 137.6$, $p < 0.001$) with the majority of True Controls, MCI-non-AD, and non-AD dementia being in CARM1 and CARM2, and the majority of pAD, MCI-AD, and AD being in CARM3 and CARM4. CARM thresholds also differentiated by increasing levels of plasma biomarkers: APS, Ab42/40 ratio, ptau181, and ptau217 (all ANOVA $p < 0.001$).

Comparison with Combined Cognivue components

Combining the three Cognivue components of the CARM (adaptive motor, visual salience, shape discrimination) into a single score allowed us to also examine the efficacy of the CARM's centiloid prediction model compared to producing a simple mean. When predicting amyloid positivity, the mean score performed similarly to the raw CARM's centiloid model (Eta^2 0.826 vs. 0.820), however the thresholded CARM performed significantly better than the thresholded mean score (using the same sensitivity, Youden's J, specificity strategy) at determining amyloid positivity; a Chi-squared test showed that the CARM produced a Cramer's V (a measure of power) of 0.301 ($\chi^2 = 79.98$, $p < 0.001$) compared to the thresholded mean score's V of 0.234 ($\chi^2 = 48.32$, $p < 0.001$). Additionally, the prediction of four-way diagnosis (healthy control, preclinical AD, AD MCI, non-AD MCI) was also more powerful when using the CARM with a binary cutoff of 3 ($V = 0.360$, $\chi^2 = 114.7$, $p < 0.001$) compared to a binary thresholded mean score ($V = 0.172$, $\chi^2 = 26.4$, $p < 0.001$). Combining the CARM and the Cognivue Clarity overall score could help identify individuals with and without cognitive

Table 4 Cognivue amyloid risk measure (CARM) group differences by clinical, cognitive and biomarker measures

	CARM 1	CARM 2	CARM 3	CARM 4	*p* value
% cognitively impaired	0.4 (0.5)	0.5 (0.5)	0.6 (0.5)	0.7 (0.50)	< 0.001[b]
Age	66.2 (4.9)	70.8 (4.9)	74.8 (4.9)	77.0 (4.9)	< 0.001[a]
% Female	0.6 (0.5)	0.6 (0.5)	0.5 (0.5)	0.5 (0.5)	0.120
Education	15.9 (2.5)	15.8 (2.7)	15.3 (2.7)	15.2 (2.7)	0.007[c]
MMSE	27.9 (2.1)	27.4 (2.3)	26.5 (2.3)	25.4 (2.3)	< 0.001[d]
FAQ	1.9 (3.6)	2.1 (3.8)	4.4 (3.8)	5.9 (3.8)	< 0.001[a]
Cognivue average score	75.0 (10.3)	69.1 (12.5)	61.4 (12.5)	49.4 (12.5)	< 0.001[a]
SUVR	1.0 (0.2)	1.1 (0.25)	1.1 (0.2)	1.2 (0.2)	< 0.001[b,c,e,f]
Centiloid level	0.3 (30.8)	7.6 (38.9)	16.1 (38.9)	29.6 (38.9)	< 0.001[b,c,e]
Ab42/40 ratio	0.099 (0.01)	0.098 (0.01)	0.096 (0.01)	0.095 (0.01)	< 0.001[b,c,e]
APS	22.0 (27.0)	27.9 (28.7)	37.9 (28.7)	43.1 (28.7)	< 0.001[d]
pTau 217	0.22 (0.16)	0.22 (0.11)	0.27 (0.11)	0.35 (0.11)	< 0.001[d]
pTau181	17.9 (10.3)	19.7 (13.4)	19.8 (9.00	25.6 (16.8)	< 0.001[g]

Mean (SD)

APS Amyloid Probability Score, *CARM* cognivue amyloid risk measure, *FAQ* Functional Activities Questionnaire, *MMSE* Mini Mental State Exam, *SUVR* standardized uptake value ratio

Post hoc comparisons:

[a]Significantly different between all groups

[b]Significantly different between 1 and 3

[c]Significantly different between 1 and 4

[d]Significantly different between all except 1 and 2

[e]Significantly different between 2 and 4

[f]Significantly different between 3 and 4

[g]Significantly different between 4 and all other groups

impairment due to AD or non-AD etiologies (Fig. 2).

DISCUSSION

Cognivue Clarity, a 10-min computerized cognitive battery, detects individuals with cognitive impairment [9–11] and a derivation of Cognivue scores was used to develop the CARM to predict the risk of amyloid. The generated CARM performed well at screening for amyloid positivity and moderately well at discriminating between pAD and True Controls. Further refinement of the GBM regression model, including redetermination of risk thresholds, was examined using additional data, and age was added to the model to provide a more robust classification. Classification thresholds were optimized at three points: 80% specificity, Youden's Index *J* (best mix of sensitivity and specificity), and 80% sensitivity, enabling standardized interpretation of the CARM across clinical and research applications. The CARM 3 cutoff has a slightly reduced sensitivity for detecting abnormal amyloid and pAD (68.3% vs. 72.7% and 49.5% vs. 56.7%), but a higher specificity for ruling out

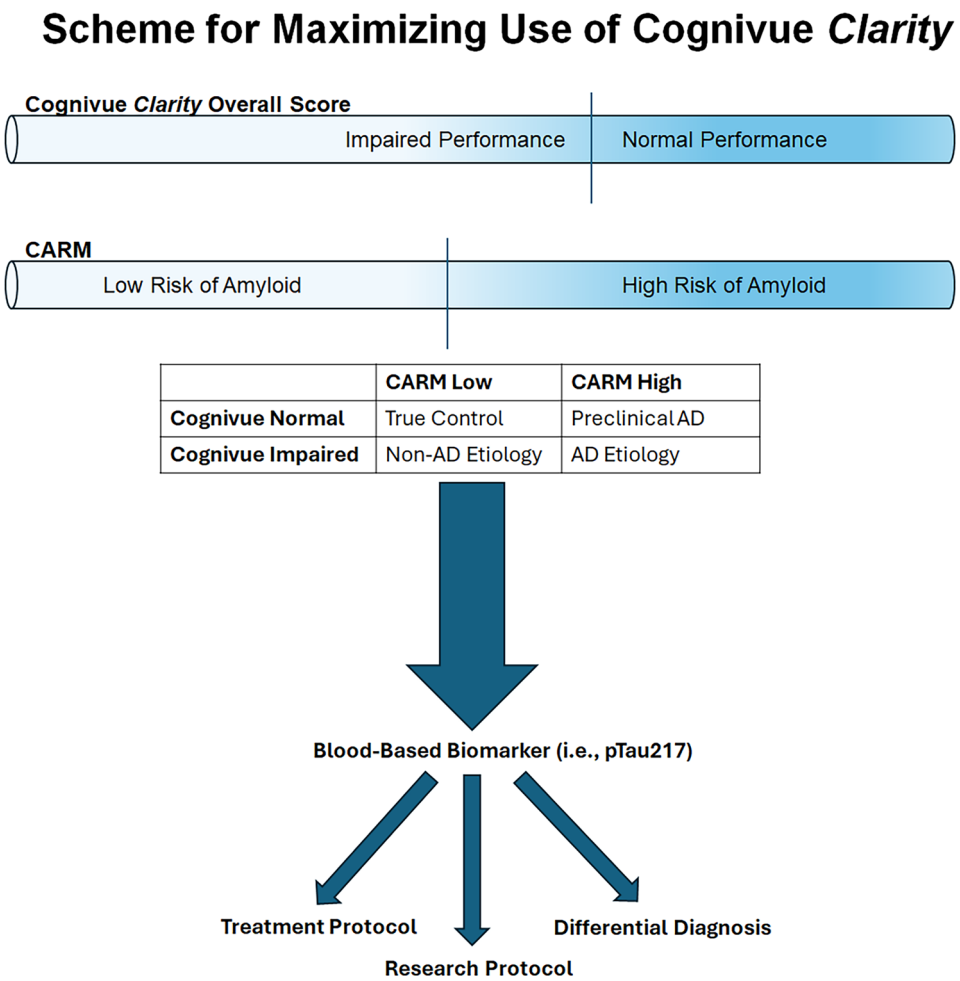

Fig. 2 Proposed scheme for maximizing use of cognivue clarity and CARM. Combining the CARM and the cognivue clarity overall score could help identify individuals with and without cognitive impairment due to AD or non-AD etiologies. Individuals with low cognivue clarity overall scores represent increased likelihood of cognitive impairment. Individuals with high CARM thresholds present increased likelihood of amyloid. Combining the CARM and the cognivue clarity overall score in a 2 × 2 paradigm could help identify true controls (cognitively normal + the absence of amyloid), pAD (cognitively normal + the presence of amyloid), MCI-AD/AD (cognitive impairment + the presence of amyloid), or MCI/dementia due to a process other than AD (cognitive impairment + the absence of amyloid). This can be followed by a blood-based biomarker such as pTau217. Based on the results, patients can be directed to a treatment protocol, a clinical trial or other research protocol, or a further search for etiological cause of cognitive impairment. *AD* Alzheimer's Disease

amyloid-negative individuals. Additionally, the use of algorithmically fine-tuned threshold levels allows for a better dialing in of risk tolerance, while a cutoff of 3 is recommended for the best combination of sensitivity and specificity, higher or lower scores could be chosen as cutoff values for recruitment depending on whether further amyloid testing is preferred (lower values more useful) or whether further testing would not be possible (higher values more useful).

Screening for MCI and AD is not universally agreed to and will depend on the population at-risk being evaluated, and the inherent properties on the screening instruments used [30–33]. If the screening process elicited a large group of false positive individuals, unnecessary health

care dollars might be expended. If the screening process elicited a large group of false negative individuals, opportunities to be treated with emerging disease-modifying therapies or enrollment into clinical trials might be missed. The emergence of biomarkers has helped define the likely cause of cognitive impairment (i.e., AD vs. non-AD); however, it is important to note that no imaging, CSF or blood-based biomarker can by itself determine whether cognitive impairment is present or stage severity. Traditional brief pencil-and-paper cognitive tests, especially those not fit-for-purpose, may misclassify individuals as impaired, such as those with older age, lower education, illiteracy, visual impairment, and/or being from a minoritized group [31]. These same tests could fail to detect cognitive impairment in highly educated individuals. Misclassification could affect eligibility for treatment protocols or enrollment into clinical trials [34]. Gold Standard Neuropsychological test batteries are generally considered the Gold Standard, but are lengthy and require specialized skills, usually making them impractical during the office visit [33].

Cognivue Clarity is a brief, validated global assessment of cognitive functioning providing a global score, along with 4 domain and 10 subtest scores for more detailed characterization of patients [9–11]. The CARM provides a risk estimate of whether amyloid is present, irrespective of cognitive status. Combining the CARM and the Cognivue Clarity overall score could help identify individuals with and without cognitive impairment due to AD or non-AD etiologies (See Fig. 1). Cognivue Clarity could be used to help screen older adults for treatment protocols with anti-amyloid therapies, enrich clinical trial recruitment before obtaining expensive biomarkers, and identify individuals likely to have pAD for prevention studies in a valid, brief, and cost-effective fashion.

Study Limitations and Strengths

The CARM was developed as risk measure of the presence of amyloid using defined thresholds—raw scores are not interpretable. The CARM is not a diagnostic test and should not be used independently of other clinical, cognitive, or biomarker data. Instead, the CARM is "added value" provided along with Cognivue Clarity global, domain, and subtest scores providing a fuller picture of the individual's cognitive performance and risk for impairment. This additional information could be useful to researchers and clinicians in deciding next steps. As a cross-sectional study, longitudinal changes in Cognivue Clarity or other biomarkers could not be determined. The goal of Bio-Hermes was not to conduct comprehensive clinical–cognitive–behavioral evaluations. Instead, the Bio-Hermes study was designed to explore relationships between blood-based and digital biomarkers and amyloid PET scans. The clinical diagnoses were made by experienced AD clinical trialists and the inclusion criteria were similar to inclusion criteria used to screen participants for many AD clinical trials. This limitation, however, does not preclude studying the relationship between Cognivue Clarity and the CARM with amyloid PET, which was designated a priori as the Gold Standard. The Bio-Hermes clinical classifications were based on limited testing (e.g., MMSE, RAVLT, FAQ). The absence of a neuropsychological examination, Clinical Dementia Rating, or MRI limits the full characterization of participants, and may miss features such as non-memory impairments, small vessel disease, or extrapyramidal symptoms. This could be important, since a substantial proportion of clinically diagnosed MCI and probable AD cases did not have amyloid and thus likely had a non-AD etiology for their cognitive impairment.

Bio-Hermes was offered in English and Spanish and ~ 20% of the sample was from an underrepresented group (mostly African American or Hispanic). There were few individuals who self-identified as Asian, American Indian/Alaska Native, or Native Hawaiian/Pacific, so no conclusions could be made regarding performance of the CARM in these ethnoracial groups. Ongoing research including these other underrepresented groups should help to better generalize the findings from the current report.

The strengths of this study include a large sample of older adults similar to cohorts recruited for AD prevention and treatment trials, with ~ 20% representation of African

American and Hispanic individuals and extensive biomarker characterization. The biomarker characterization used in Bio-Hermes is consistent with the NIA-AA research framework [35, 36] for biological definitions of AD and non-AD causes of cognitive impairment.

CONCLUSION

With a large number of AD prevention and treatment trials on-going, screening for individuals with cognitive impairment and amyloid is an expensive and labor-intensive proposition, since neuropsychological testing alone cannot detect amyloid and PET scans, nor can lumbar puncture characterize cognitive performance. Cognivue Clarity, a 10-min computerized battery, can detect individuals with cognitive impairment, identify individuals likely to have amyloid positivity, and could thereby capture individuals with pAD [9–11]. To further increase the efficiency and cost-effectiveness of cognitive screening, a staged screening approach likely makes the most sense [37]. Cognivue Clarity could be used to establish whether there is (1) cognitive impairment with the Cognivue Clarity global score, and (2) a high likelihood of amyloid presence with the CARM. This could be followed by measuring a readily accessible AD biomarker such as plasma pTau217 [24, 38, 39]. Such a strategy would increase the likelihood of identifying early AD for treatment or trial enrollment, avoiding the cost of expensive PET scans in a brief, reliable, valid, and time- and cost-effective fashion.

ACKNOWLEDGEMENTS

The authors would like to thank all the participants and Bio-Hermes sites that contributed data to this project, and Ernest Fung, PsyD for helpful comments in preparing the manuscript.

Author Contributions. Study conception and design was performed by James E. Galvin, MD, MPH, Heather M. Harris and Paul W. Estes. Material preparation, data collection and analysis were performed by James E. Galvin, MD, MPH and Michael J. Kleiman, PhD. The first draft of the manuscript was written by James E. Galvin, MD, MPH and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding. The Bio-Hermes study was sponsored by the Global Alzheimer Platform Foundation. Eli Lilly and Company provided quantitative readings of 18F-Florbetapir PET scans and pTau217 assays. Entities that provided direct funding to Bio-Hermes were AbbVie, Alzheimer's Drug Discovery Foundation (ADDF), Aural Analytics, Biogen, Cognivue, C_2N, Gates Ventures, Linus Health, Merck & Co, Quanterix, Retispec, and Roche. Entities that did not provide direct funding but did provide in-kind laboratory testing and/or use of devices and testing support were Alzheimer's Disease Data Initiative, Inc. (ADDI), AITIA, EMTherapro, Lilly, Somalogic, and University of Gothenburg. Cognivue, Inc is the source of funding for the Journal's Rapid Service Fees.

Data Availability Statement. A de-identified dataset for this project is available to all interested parties. Please contact JEG at jeg200@miami.edu.

Declarations

Conflict of Interest. James E. Galvin, MD, MPH is Chief Scientific Officer for Cognivue, Inc and receives consulting fees. Michael J. Kleiman, PhD received consulting fees from Cognivue, Inc to conduct statistical analyses. Heather M. Harris and Paul W. Estes are full-time employees of Cognivue, Inc. The authors take full responsibility for the data and have the right to publish all data.

Ethical Approval. This study was reviewed and approved by Advarra, a central institutional review board (Reference Number Pro00046018), and was conducted in accordance with the Helsinki Declaration of 1964. The study was registered on ClinicalTrials.gov (NCT04733989).

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. Alzheimer's Association. Alzheimer's disease facts and figures. Alzheimers Dementia. 2024;2024(20):3708–821.

2. Silva-Rudberg JA, Carrión CI, Pérez-Palmer N, Li J, Mehta SK, Diab NS, et al. Assessment of disparities in timely diagnosis and comprehensive workup of cognitive impairment between English and Spanish speakers. Am J Geriatr Psychiatry. 2024;32:773–86.

3. Giebel C, Silva-Ribeiro W, Watson J, Volkmer A, Chirico I, Diaz A, et al. A systematic review on the evidence of misdiagnosis in dementia and its impact on accessing dementia care. Int J Geriatr Psychiatry. 2024;39: e6158.

4. Liu Y, Jun H, Becker A, Wallick C, Mattke S. Detection of rates of mild cognitive impairment in primary care for the United States Medicare population. J Prev Alz Dis. 2024;11:7–12.

5. Casagrande M, Marselli G, Agostini F, Forte G, Favieri F, Guarino A. The complex burden of determining prevalence rates of mild cognitive impairment: a systematic review. Front Psychiatry. 2022;13: 960648.

6. Rafii MS, Sperling RA, Donohue MC, Zhou J, Roberts C, Irizarry MC, et al. The AHEAD 3–45 study: design of a prevention trial for Alzheimer's disease. Alzheimers Dement. 2023;19:1227–33.

7. Sperling RA, Donohue MC, Raman R, Rafii MS, Johnson K, Masters CL, et al. Trial of solanezumab in preclinical Alzheimer's disease. N Engl J Med. 2023;389:1096–107.

8. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, 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;7:280–92.

9. Galvin JE, Chang LC, Estes P, Harris HM, Fung E. Cognitive assessment with cognivue clarity: psychometric properties and normative ranges in a diverse population. J Alzheimers Dis. 2024;100:509–23.

10. Galvin JE, Kleiman MJ, Estes PW, Harris HM, Fung E. Cognivue clarity characterizes mild cognitive impairment and Alzheimer's disease in biomarker confirmed cohorts in the Bio-Hermes study. Sci Rep. 2024;14:24519.

11. Galvin JE, Kleiman MJ, Estes PW, Harris HM, Fung E. Cognivue clarity characterizes amyloid status and preclinical Alzheimer's disease in biomarker confirmed cohorts in the Bio-Hermes study. J Alzheimer Dis. 2025;104:83–94.

12. Mohs RC, Beauregard D, Dwyer J, Gaudioso J, Bork J, MaGee-Rogers T, et al. The bio-hermes study: biomarker database developed to investigate blood-based and digital biomarkers in community-based, diverse populations clinically screened for Alzheimer's disease. Alzheimers Dement. 2024;20:2752–65.

13. 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;12:189–98.

14. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, 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;7:270–9.

15. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7:263–9.

16. Rosenberg SJ, Ryan JJ, Prifitera A. Rey auditory-verbal learning test performance of patients with

and without memory impairment. J Clin Psychol. 1984;40:785–7.

17. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol. 1982;37:323–9.

18. Pemberton HG, Collij LE, Heeman F, Bollack A, Shekari M, Salvado G, et al. Quantification of amyloid PET for future clinical use: a state-of-the-art review. Eur J Nucl Med Mol Imaging. 2022;49:3508–28.

19. Su Y, Flores S, Hornbeck RC, Speidel B, Vlassenko AG, Gordon BA, et al. Utilizing the Centiloid scale in cross-sectional and longitudinal PiB PET studies. NeuroImage Clin. 2018;19:406–16.

20. van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen C, Gee M, et al. Lecanemab in early Alzheimer's disease. N Engl J Med. 2023;388:9–21.

21. Sims JR, Zimmer JA, Evans CD, Lu M, Ardayfio P, Sparks JD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. JAMA. 2023;330:512–27.

22. Kirmess KM, Meyer MR, Holubasch MS, Knapik SS, Hu Y, Jackson EN, et al. The PrecivityAD™ test: accurate and reliable LC-MS/MS assays for quantifying plasma amyloid beta 40 and 42 and apolipoprotein E proteotype for the assessment of brain amyloidosis. Clin Chim Acta. 2021;519:267–75.

23. Ding X, Zhang S, Jiang L, Wang L, Li T, Lei P. Ultrasensitive assays for detection of plasma tau and phosphorylated tau 181 in Alzheimer's disease: a systematic review and meta-analysis. Transl Neurodegener. 2021;10:1–14.

24. Janelidze S, Berron D, Smith R, Strandberg O, Proctor NK, Dage JL, et al. Associations of plasma phospho-tau217 levels with tau positron emission tomography in early Alzheimer disease. JAMA Neurol. 2021;78:149–56.

25. Navitsky M, Joshi AD, Kennedy I, Klunk WE, Rowe CC, Wong DF, et al. Standardization of amyloid quantitation with florbetapir standardized uptake value ratios to the Centiloid scale. Alzheimers Dement. 2018;14:1565–71.

26. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–30.

27. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. "LightGBM: a highly efficient gradient boosting decision tree." In: Proceedings of the 31st international conference on neural information processing systems. NIPS'17. Red Hook: Curran Associates Inc.; 2017. p. 3149–57.

28. Awad M, Khanna R. Support vector machines for classification. In: Awad M, Khanna R, editors. Efficient learning machines: theories, concepts, and applications for engineers and system designers. New York: Apress Media; 2015. p. 39–66.

29. Dietterich TG. Ensemble methods in machine learning. In: Multiple classifier systems: proceedings of the first international workshop on multiple classifier systems. 2000. p. 1–15.

30. Chithiramohan T, Santhosh S, Threlfall G, Hull L, Mukaetova-Ladinska EB, Subramaniam H, et al. Culture-fair cognitive screening tools for assessment of cognitive impairment: a systematic review. J Alzheimers Dis Rep. 2024;8:289–306.

31. Ranson JM, Kuźma E, Hamilton W, Muniz-Terrera G, Langa KM, Llewellyn DJ. Predictors of dementia misclassification when using brief cognitive assessments. Neurol Clin Pract. 2019;9:109–17.

32. Cordell CB, Borson S, Boustani M, Chodosh J, Reuben D, Verghese J, et al. Medicare Detection of Cognitive Impairment Workgroup. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. Alzheimers Dement. 2013;9:141–50.

33. Galvin JE. Using informant and performance screening methods to detect mild cognitive impairment and dementia. Curr Geriatr Rep. 2018;7:19–25.

34. Jutten RJ, Papp KV, Hendrix S, Ellison N, Langbaum JB, Donohue MC, et al. Why a clinical trial is as good as its outcome measure: a framework for the selection and use of cognitive outcome measures for clinical trials of Alzheimer's disease. Alzheimers Dement. 2023;19:708–20.

35. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14:535–62.

36. Therriault J, Janelidze S, Benedet AL, Ashton NJ, Arranz Martínez J, Gonzalez-Escalante A, et al. Diagnosis of Alzheimer's disease using plasma biomarkers adjusted to clinical probability. Nat Aging. 2024;4:1529–37.

37. Kleiman MJ, Ariko T, Galvin JE. Alzheimer's Disease Neuroimaging Initiative. Hierarchical two-stage cost-sensitive clinical decision support system for screening prodromal Alzheimer's

disease and related dementias. J Alzheimers Dis. 2023;91:895–909.

38. Jonaitis EM, Janelidze S, Cody KA, Langhough R, Du L, Chin NA, Mattsson-Carlgren N, et al. Plasma phosphorylated tau 217 in preclinical Alzheimer's disease. Brain Commun. 2023;5:fcad057.

39. Ashton NJ, Brum WS, Di Molfetta G, Benedet AL, Arslan B, Jonaitis E, et al. Diagnostic accuracy of a plasma phosphorylated tau 217 immunoassay for Alzheimer disease pathology. JAMA Neurol. 2024;81:255–63.