Title
Cognitive test completion time as an indicator of early brain atrophy: Findings from the Cognitive Assessment for Dementia, iPad version 2 (CADi2) in healthy adults

Author(s)
IWASA Kenichi / Takamura Masahiro / Ishida Manabu / Kuroda Yoko / Abe Satoshi / NAGAI Atsushi

Journal
Journal of Alzheimer's Disease 109 (2) , 799 - 809

Published
2025-11-28

URL (The Version of Record)
https://doi.org/10.1177/13872877251400746

Cognitive Test Completion Time as an Indicator of Early Brain Atrophy: Findings from the CADi2 in Healthy Adults

Kenichi Iwasa*, Masahiro Takamura, Manabu Ishida, Yoko Kuroda, Satoshi Abe, Atsushi Nagai

Department of Neurology, Faculty of Medicine, Shimane University, Izumo, Japan

*Corresponding author:
Kenichi Iwasa
Department of Neurology, Faculty of Medicine, Shimane University
89-1 Enya-cho, Izumo, Shimane 693-8501, Japan
Email: ken1wasa@med.shimane-u.ac.jp
ORCID: 0000-0001-8569-2410

Abstract

Background: While several cognitive screening tools are available, early-stage detection remains challenging because individuals in prodromal or mild cognitive impairment stages often score within normal ranges. Assessment of response time, rather than test score alone, may offer greater sensitivity to subtle decline.

Objective: This study aimed to examine whether completion time on the Cognitive Assessment for Dementia, iPad version 2 (CADi2), is associated with early brain changes suggestive of preclinical dementia.

Methods: In this cross-sectional study of 511 cognitively unimpaired adults (290 men, 221 women; mean age 61 years) at the Health Science Center, Shimane, Japan, participants completed the 10-item CADi2 and underwent structural magnetic resonance imaging. Based on the ceiling effects in total scores, we analyzed associations between total test duration and regional brain atrophy.

Results: CADi2 time was significantly associated with hippocampal atrophy, while total scores were not. To assess predictive value, we applied a deep survival analysis model using Alzheimer's Disease Neuroimaging Initiative data to estimate 5-year dementia risk. Using a threshold of ≥0.25, receiver operating characteristic analysis showed that CADi2 time (area under the curve, 0.632) outperformed CADi2 (0.554) and Mini-Mental State Examination (0.582) scores. Only CADi2 time showed a significant odds ratio [OR] for high dementia risk (OR = 1.94; 95% confidence interval: 1.12–3.40; p = .018) after adjusting for covariates.

Conclusions: These findings suggest that completion time from a cognitive test without ceiling effects is more suitable than score-based measures for evaluating the pre-dementia stage, and that combining time and score metrics may further improve stratification in cognitively normal individuals.

Keywords: Alzheimer's disease; brain atrophy; Cognitive Assessment for Dementia, iPad version 2; cognitive screening; dementia; hippocampus; magnetic resonance imaging

Introduction

Dementia, particularly Alzheimer's disease (AD), is a global health challenge due to its growing prevalence and lack of accessible early diagnostic tools. Early risk assessment for dementia has become important because of increasing attention to early intervention. However, only a limited number of facilities can perform established tests, such as amyloid positron emission tomography and spinal fluid tau testing, and simpler screening assessments are preferred.

Although screening tests for dementia, such as the Mini-Mental State Examination (MMSE)[1,2] and Montreal Cognitive Assessment,[3] are well-known, they cannot be used to assess the predementia stage.[4] We developed and implemented a dementia mass-screening test called the Cognitive Assessment for Dementia, iPad version 2 (CADi2).[5] The CADi2 comprised 10 tablet-based questions, yielding two metrics—total score and completion time—both of which differentiated patients with AD from healthy controls. Notably, a CADi2 score of ≤6 or a completion time of ≥200 seconds generally corresponded to an MMSE score of ≤21 based on our institutional study.[6] However, their discriminative performance in conditions that are more prodromal than clinical AD remains unknown.

One important limitation of traditional cognitive assessment tools is the ceiling effect, which reduces their sensitivity to subtle cognitive impairment in cognitively normal individuals. Notably, CADi2 offers a novel dual-output format by simultaneously capturing both total scores and item-level response times through automatic logging. This feature enables more granular cognitive profiling, enhancing the potential to detect preclinical neurodegenerative changes.

The natural history of AD suggests that brain atrophy begins earlier than a clear decline in cognitive function.[7–9] In addition, dementia-related atrophy can be detected using magnetic

resonance imaging (MRI), whereas cognitive tests primarily assess functional performance. These modalities therefore provide complementary yet distinct information. Recent studies have shown that MRI volumetric analysis can reliably distinguish progressive from stable mild cognitive impairment (MCI), underscoring its value for structural rather than functional assessment.[11]

Recent studies have explored the relationship between psychological testing and brain atrophy in patients with dementia and healthy individuals. In AD, brain atrophy measured using MRI volumetric analysis correlates with cognitive dysfunction assessed using the AD Assessment Scale–Cognitive Subscale test.[12] In older Japanese adults without dementia, significant correlations have been observed between hippocampal atrophy and neuropsychological test scores, including those for the Raven's Coloured Progressive Matrices test, Trail-Making Test (TMT) part B, and surfactant protein-A test.[13] Specific neuropsychological deficits are associated with regional atrophy in patients with AD. Right amygdala–hippocampal complex volumes correlate with non-verbal visual, declarative memory, while left temporoparietal regions are linked to semantic memory.[14]

The MMSE, a commonly used cognitive screening test, correlates with hippocampal volume when a cutoff score of 24 points is applied, indicating a link between performance and structural change.[15] However, this score is already within the range for dementia and is not useful as an early predictor of atrophy.

A novel cognitive test without ceiling effects, designed to detect subtle functional changes associated with early hippocampal atrophy in healthy older adults, has recently been reported.[16] Assessing the time taken to complete cognitive tests may help avoid ceiling effects and detect signs of cognitive decline or brain atrophy in healthy individuals. The existence of brain atrophy subtypes in dementia has garnered considerable attention. The TMT is a test that

assesses the time required to respond to questions.[17] Although negative correlations have been reported between frontal and parietal lobe volumes and TMT in seconds in healthy and dementia groups,[18] to our knowledge, no reports have focused only on their assessment in healthy participants. Furthermore, unlike conventional paper-based screening tools, the CADi2 offers a unique digital interface that includes standardized voice instructions and automatic time logging for each item. These features potentially enable broader applicability in remote or unsupervised settings, as well as more granular analysis of processing speed beyond total scores.

Recent work by Weiner et al. highlighted that digital cognitive assessments can serve as low-burden markers of progression to mild cognitive impairment, showing that time-based and performance metrics obtained in remote or unsupervised settings can detect subtle cognitive changes earlier than conventional assessments.[19] In parallel, novel approaches that integrate digital cognitive markers with blood-based biomarkers have demonstrated promise in predicting brain amyloid load.[20] Together, these findings underscore the growing importance of scalable and accessible tools, providing a rationale for our investigation of CADi2 completion time. In this study, we examined the relationship between CADi2 performance (both accuracy and completion time) and brain atrophy, focusing on early structural changes as potential markers of preclinical dementia. We hypothesized that CADi2 metrics could predict such early alterations.

Methods

Dataset

In Japan, brain scans consist of head MRIs and other tests for primary prevention or early detection of brain diseases in healthy participants.

Participants were volunteers who attended community-based brain health screening (brain

dock) at the Health Science Center, Shimane, as part of primary prevention for cerebrovascular disease. They were not recruited on the basis of cognitive complaints, and no financial compensation was provided.

Data from 511 volunteers Participants were community-dwelling adults who voluntarily attended the Health Science Center's brain health screening program (Brain Dock) for preventive purposes. They were not recruited based on memory complaints, and no financial compensation was provided. All procedures were conducted as part of routine health screening. who visited a brain clinic at the Health Science Center, Shimane, between March 2019 and March 2021, for whom MRI and CADi2 data were available, were included in the analysis. Although we initially planned to exclude individuals with a documented history of dementia, no such cases were identified in the cohort. This study was approved by the Shimane University Medical Research Ethics Committee, and informed consent was obtained from all participants. This study was conducted in accordance with the ethical standards set out in the 1964 Declaration of Helsinki and its subsequent amendments (Approval No. 2225; Date of Approval: May 12, 2016).

Psychological tests

In addition to standard psychological tests— including the MMSE, Frontal Assessment Battery (FAB), Kohs Cube test, and Self-Rating Depression Scale (SDS)—participants who underwent MRI scans also took the CADi2 test. The details of the 10 questions in CADi2 are shown in Fig. 1 and Table 1. After a brief practice session with the application, the participants completed the test using a tablet. No other intervention by examiners was performed during the test. Participants were allowed to use their habitual visual or hearing aids during the administration of the CADi2. The CADi2 score (out of 10), CADi2 time for each question, and total score

were automatically calculated.

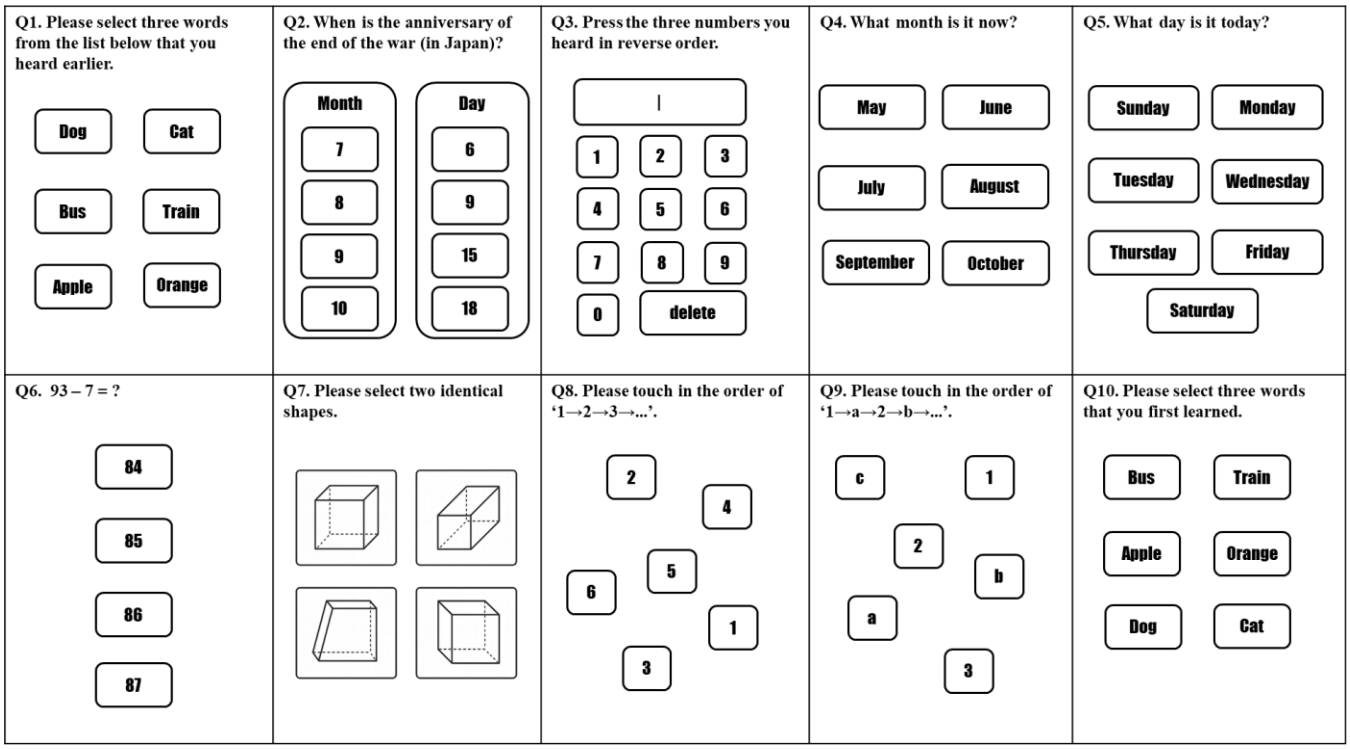

Fig. 1 CADi2 test

The figure displays an English-translated reproduction of the original Japanese CADi2 tablet interface. Details of the individual items are provided in Table 1. Depending on the specific item, voice instructions may be delivered, and responses are recorded via touch input.

Table 1. Ten questions of CADi2

Description
Q1. Immediate recognition
Three words (cat, bus, and orange) were presented slowly, audibly, and individually. The list was presented twice. Participants were asked to select the three words they had

learned from six options (cat, dog, bus, train, apple, and orange). They were then asked to recall the three words.
Q2. Long-term memory
Participants were asked about the date of the end of the hostility during World War II. Participants chose the correct answer from a list of months (July, August, September, and October) and days (July 6th, 9th, 15th, and 18th). Note: In Japan, August 15 is widely recognized as the day commemorating the end of World War II.
Q3. Reverse-ordered digits
Three digits (5, 1, and 8) were presented slowly and audibly, one at a time. The participants were then asked to key the digits in reverse order.
Q4. Orientation (month)
Participants were asked to select the current month from six options.
Q5. Orientation (day of the week)
Participants were asked to choose the current day of the week from seven options.
Q6. Calculation
Participants were asked to choose the answer to the question '93 minus 7' from four options (84, 85, 86, and 87).
Q7. Cube rotation
Four three-dimensional figures (two cubes, a rectangle, and a trapezoidal corpus drawn from different perspectives) were presented. Participants were asked to choose one matching pair.
Q8. Sequence-making A

Six-digit numbers (1, 2, 3, 4, 5, and 6) were presented at random positions on the screen. Participants were asked to touch numbers 1–6 sequentially on the screen.
Q9. Sequence-making B
Three-digit numbers (1, 2, 3) and three hiragana (a, i, u) were presented at random positions on the screen. The participants were asked to alternately touch the numbers and hiragana on the screen in sequence (1, a, 2, i, 3, and u).
Q10. Delayed recognition
Participants were again asked to select the three words presented in Question 1 from the six options.

MRI acquisition

MRI was performed using a 3.0 T machine (Philips Ingenia 3TCX Philips Healthcare, Best, Netherlands). As a routine brain health scanning, longitudinal relaxation (T1)-weighted, transverse relaxation (T2)-weighted, T2*-weighted, and fluid-attenuated inversion recovery images were acquired in axial sections with a 7-mm slice thickness. For T1-weighted images, coronal and sagittal images were also acquired. Asymptomatic cerebral infarction, deep white matter lesions, and microbleeds were classified according to the Guidelines for Brain Docs 2019, published by the Japan Brain Doc Society.[21] For voxel-based morphometry analysis, a three-dimensional T1-weighted structural image was acquired using an MP-RAGE scan with the following parameters: repetition time, 6.8 ms; echo time, 3.1 ms; flip angle, 9°; 170 sagittal slices; in-plane resolution, 256 × 256; field of view, 256 × 256 mm; and voxel size, 1.0 × 1.0 × 1.2 mm.

MRI analysis

Preprocessing and statistical analyses of 3D T1-weighted images were performed using the Computational Anatomy Toolbox (CAT12) and Statistical Parametric Mapping (SPM12), both running on MATLAB R2021b (MathWorks, Natick, MA, USA). The standard preprocessing pipeline of CAT12 was conducted. T1-weighted images were denoised, resampled, normalized to standard brain template space (MNI152NLin2009cAsym), and segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). These process are based on SPM's unified segmentation[22] for initial segmentation, an Adaptive Maximum A Posterior (AMAP) technique[23] for detailed segmentation, and Geodesic Shooting (Ashburner & Friston, 2011) for spatial normalization. The resampled and normalized GM images[24] were then spatially smoothed with a Gaussian kernel (FWHM = 6 mm) prior to subsequent analysis.

Voxel-based morphometric analysis using a generalized linear model with age, sex, and intracranial volume input as covariates was employed to examine the association between CADi2 performance, execution time, and regional brain volume for the whole brain. The threshold-free cluster enhancement method was used to determine significant brain regions with a family-wise error ($p < 0.05$). Two-tailed p-values were reported for all hypothesis tests. In addition to the main analyses, we also examined voxel-wise associations between completion times of individual CADi2 items and regional gray matter volume. Significant clusters for each item are summarized in Supplementary Table S1.

Deep survival analysis (DSA)-based risk prediction and cutoff validation

To further evaluate the clinical significance of CADi2 performance metrics in predicting future cognitive decline, we performed an additional analysis using a DSA framework based on brain MRI volumetric data. Specifically, we applied a published DSA model developed by Nakagawa

et al. using the Alzheimer's Disease Neuroimaging Initiative dataset, which estimates 5-year conversion risk to AD based on regional brain atrophy. In this analysis, a DSA-estimated risk of ≥0.25 was defined as high risk for future dementia onset.[25]

Receiver operating characteristic (ROC) curves were generated to assess the discriminative performance of CADi2 completion time, CADi2 total score, and Mini-Mental State Examination (MMSE) score in identifying individuals classified as high-risk by DSA. Optimal cutoff thresholds were determined using the point with the highest Youden index. These thresholds were subsequently used to compute odds ratios (ORs) for the presence of high dementia risk.

Adjusted ORs were estimated using multivariate logistic regression models that included age, sex, asymptomatic cerebral infarction, cerebral microbleeds, and deep white matter hyperintensities (classified according to Fazekas grades for subcortical white matter hyperintensities [DSWMH] and periventricular hyperintensities [PVH]), as well as subjective memory complaints and 14 dementia risk factors proposed by Livingston, Gill, and colleagues in 2024.[26] These factors comprised history of hypertension, diabetes, alcohol consumption, smoking history, physical activity, cohabitation status, years of education, Self-Rating Depression Scale score obtained during brain dock, visual acuity, hearing at 2000 Hz and 4000 Hz, body mass index, and LDL cholesterol. Missing covariate values were imputed using the linear trend at point method, a single imputation approach based on linear interpolation. Although history of head trauma was initially planned as a covariate, none of the 511 participants reported such history; therefore, it was omitted. Data on residential air pollution exposure were not collected and thus could not be assessed. All statistical analyses were

215 performed using SPSS (IBM Corp., Armonk, NY, USA).

216

217 **Results**

218 *Performance characteristics of CADi2 and its association with brain atrophy*

219 Table 2 summarizes the results of the CADi2 test. The mean CADi2 score was 9.47 (standard
220 deviation [SD], 0.79), with a minimum of 7. Most participants scored within the range of 8 to
221 10, which is generally considered indicative of normal cognitive function.

222 The mean total time required to complete the test was 89.5 seconds (SD, 25.92), significantly
223 below the 200-second threshold often associated with cognitive decline. The larger SD in
224 completion time compared with the score suggests a greater variability in processing speed
225 among participants.

226

227 Table 2. CADi2 test results

n = 511	Mean	SD	Min	Max
CADi2 total score	9.47	0.79	7	10
Total test duration (s)	89.5	25.92	41.81	259.99
Item 1 response time (s)	9.31	3.89	3.63	60.93
Item 2 response time (s)	10.89	6.86	3.67	77.57
Item 3 response time (s)	9.89	6.19	2.97	72.76
Item 4 response time (s)	5.78	1.34	2.52	15.43
Item 5 response time (s)	6.43	3.78	2.63	57.36
Item 6 response time (s)	9.4	4.19	3.29	34.04
Item 7 response time (s)	8.28	2.92	3.75	44.71

Item 8 response time (s)	7.65	3.56	4.27	53.49
Item 9 response time (s)	10.18	5.74	4.71	62.68
Item 10 response time (s)	11.68	6.79	4.23	75.54

228 SD: standard deviation

Significant clusters were negatively correlated with CADi2 time (Table 3). The left parahippocampal region showed the strongest correlation and was the largest cluster in terms of volume ($p = 0.008$). Significant clusters were also found in the right inferior occipital and right inferior temporal gyri (Fig. 2). No significant clusters were observed in the CADi2 scores.

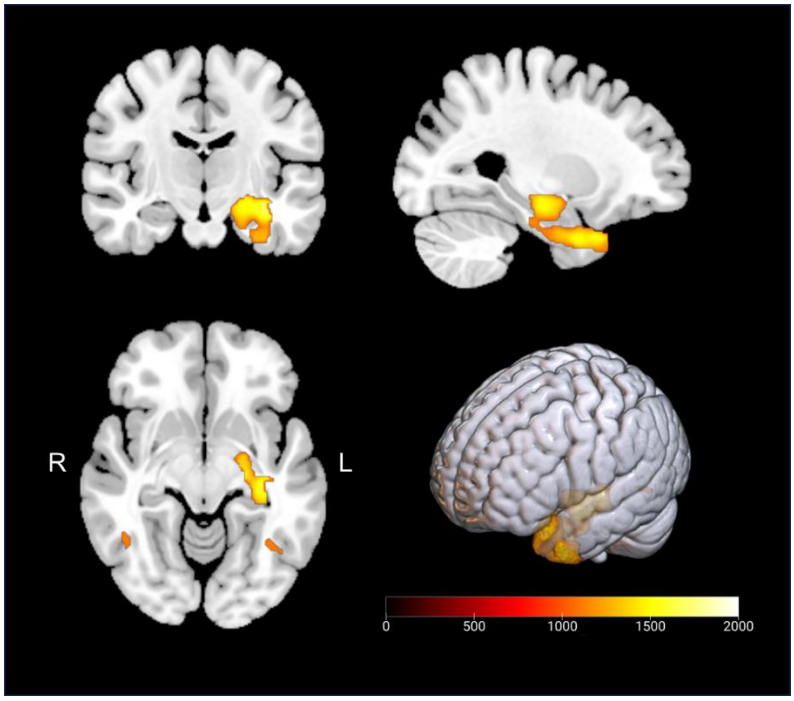

Fig. 2 Correlations between the total time required for completion of the CADi2 (total time) and brain volume

Voxels with significant correlations (all negative correlations) are shown in color. Age, sex, and intracranial volume were used as covariates, and the threshold-free cluster enhancement method was used to determine significant brain regions with a family-wise error of $p < 0.05$. R, right; L, left. The color bar reflects the threshold-free cluster enhancement values, with warmer colors indicating stronger negative correlations between CADi2 completion time and voxel-based gray matter volume. Only clusters that survived family-wise error correction ($p < 0.05$) are displayed.

Significant clusters associated with completion times of individual CADi2 items are presented in Supplementary Table S1 (items without significant clusters are not shown).

Table 3. Brain atrophy areas that correlated with CADi2 time in the VBM analysis

CADi2 Parameter	Cluster Size	P(FWE-corr)	TFCE	MNI coordinate X	y	z	Anatomical Label	
CADi2 time	6914	0.008	1680.54	-38	-18	-15	Left Hippocampus	Left Temporal Pole
	81	0.046	1112.24	48	-54	-14	Right Inferior Temporal Gyrus	
	54	0.043	1078.76	-38	-57	-9	Left Fusiform Gyrus	
	26	0.045	1065.72	38	-80	-4	Right Inferior Occipital Gyrus	
	6	0.049	1044.92	50	-45	-15	Right Inferior Temporal Gyrus	
CADi2 score	No significant clusters							

VBM: voxel-based morphomety

ROC-based cutoff identification and odds ratio estimation for dementia risk

Figure 3 illustrates ROC curves for CADi2 completion time, CADi2 score, and MMSE score in predicting high-risk dementia status as defined by the DSA model (risk ≥0.25). The area under the curve (AUC) values were 0.632, 0.554, and 0.582 for CADi2 time, CADi2 score, and MMSE score, respectively. The optimal cutoff points, determined by Youden's index, were approximately 80–90 seconds for CADi2 time, ≤8–9 points (out of 10) for CADi2 score, and ≤28–29 points (out of 30) for MMSE score. As a supplementary analysis, logistic regression was conducted using CADi2 completion time and score as predictors, with a cutoff of DSA-predicted probability ≥0.25 at 5 years. The resulting AUC was 0.634 (95% CI, 0.582–0.686; $p < 0.0001$), comparable to—or only marginally better than—the performance of CADi2 completion time alone. Among the 511 participants, 278 individuals (54.4%) had CADi2 completion times ≥80 seconds.

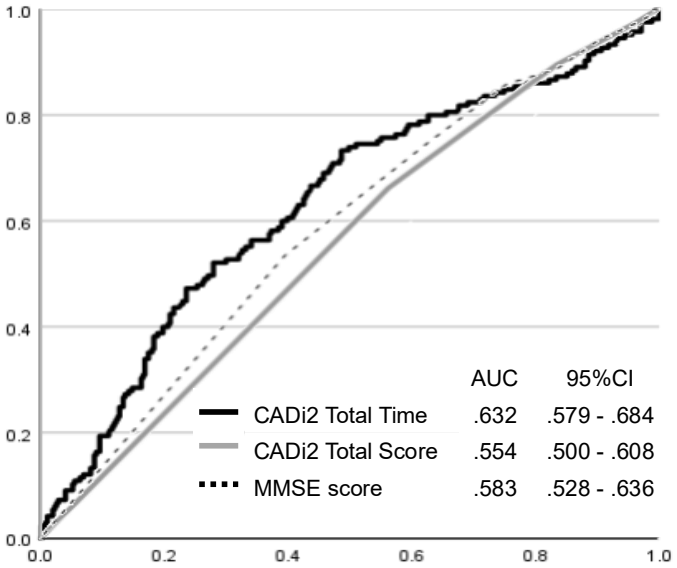

Fig. 3 Receiver operating characteristic (ROC) curves for predicting 5-year dementia risk ≥0.25 based on deep survival analysis

The ROC curves illustrate the predictive performance of three cognitive measures: CADi2 Total Time (black solid line), CADi2 Total Score (gray solid line), and MMSE score (black dotted line). The area under the curve (AUC) was highest for CADi2 Total Time (AUC = 0.632, 95% confidence imterval [CI]: 0.579–0.684), followed by MMSE score (AUC = 0.583, 95% CI: 0.528–0.636) and CADi2 Total Score (AUC = 0.554, 95% CI: 0.500–0.608). The diagonal line represents the reference (random) classifier.

Based on these thresholds, individuals exceeding the cutoff were considered to be at a higher clinical risk. The demographic and baseline data stratified by DSA categories are presented in Tables 4a and 4b.

Table 4a. Baseline Characteristics by 5-Year Risk Categories (DSA ≥0.25 vs <0.25)

Variable	DSA ≥0.25 at 5 years (n=164)	DSA <0.25 at 5 years (n=347)	p value
Sex: male	99	188	0.27
Hypertension	61	94	0.030*
Diabetes	24	23	0.008**
SBI	23	21	0.005**
MBs	26	36	0.09
PVH Fazekas grade			<0.001***
0	63	234	
1	60	86	
2	35	17	
3	6	3	
DSWMH Fazekas grade			<0.001***
0	27	125	
1	70	148	
2	58	62	

3	9	4	
Mental disorder	3	5	0.74
Subjective cognitive decline	62	112	0.22
Sleep_disorder	53	29	0.53
decreased_concentration	34	19	0.54
malignant_tumor	21	20	0.02
Ever smoker	75	130	0.1
Alcohol consumption			0.35
Non-drinker	74	164	
<21 g ethanol/day	32	68	
21–<63 g ethanol/day	54	96	
≥63 g ethanol/day	0	5	
Exercise habit			0.36
Rarely / almost never	53	120	
Monthly (≈ once/month)	8	32	
Weekly (≈ once/week)	27	46	
2–3 times/week	40	84	
Daily	32	56	
Living Alone	17	27	0.4

Hearing impairment, Right 1000Hz	31	32	0.002**
Hearing impairment, Left 1000Hz	27	31	0.015*
Hearing impairment, Right 4000Hz	37	43	0.003**
Hearing impairment, Left 4000Hz	41	44	<0.001***
Job Category			
MedicalStaff	51	7	<0.001***
PublicServantTeacher	75	48	0.06
AgriFisheryWorker	5	6	0.11
DoctorLawyerPolitician	9	1	0.13
TechnicalEmployee	17	11	0.40
Manager	10	7	0.41
EngineerArchitect	5	4	0.42
Freelance	1	0	0.49
SmallBusiness	7	2	0.52
Clergy	1	1	0.59
ConstructionWorker	7	4	0.76
ClerkService	95	46	0.87
None_or_Housewife	5	4	0.42
Other Worker	41	18	0.78

Values are expressed as number of participants. p values were calculated using χ^2 tests. *p<0.05, **p<0.01, ***p<0.001. Abbreviations: DSA,

deep survival analysis; SBI, silent brain infarction; MBs, cerebral microbleeds; PVH, periventricular hyperintensity; DSWMH, deep subcortical white matter hyperintensity.

Table 4b. Baseline Clinical, Cognitive, and Laboratory Variables by 5-Year Risk Categories (DSA ≥0.25 vs <0.25)

Variable	DSA ≥0.25 at 5 years (n=164) Mean (SD)	DSA <0.25 at 5 years (n=347) Mean (SD)	p value
Age, years	65.6 (12.6)	58.9 (12.4)	<0.001***
CADi2 score	9.36 (0.84)	9.53 (0.75)	0.025*
CADi2 completion time, s	96.4 (28.7)	86.1 (23.9)	<0.001***
MMSE	28.4 (1.88)	29.0 (1.59)	0.002**
SDS	33.2 (7.45)	33.8 (7.43)	0.38
Kohs Cube test	99.1 (20.2)	107 (18.3)	<0.001***
FAB total	16.2 (1.54)	16.7 (1.6)	<0.001***
LDL-cholesterol, mg/dL	113 (28.4)	121.9 (31.5)	0.003**
HbA1c, %	6.01 (0.79)	5.81 (0.56)	0.001**
SBP, mmHg	127 (18.6)	123 (16.9)	0.013*
DBP, mmHg	72.9 (12.0)	73.0 (11.6)	0.899
BMI, kg/m²	22.4 (3.25)	22.8 (3.0)	0.156
Years of education	13.5 (2.37)	14.1 (2.5)	0.014*
Visual acuity, right	0.74 (0.37)	0.84 (0.33)	0.012*
Visual acuity, left	0.76 (0.37)	0.86 (0.35)	0.006**

Between-group differences were tested using independent samples t-tests. p values <0.05 were considered statistically significant. *p<0.05, **p<0.01, ***p<0.001. Abbreviations: DSA = Deep Survival Analysis; CADi2 = Cognitive Assessment for Dementia, iPad version 2; MMSE

= Mini-Mental State Examination; SDS = Self-rating Depression Scale; FAB = Frontal Assessment Battery; LDL = low-density lipoprotein; HbA1c = glycated hemoglobin A1c; SBP = systolic blood pressure; DBP = diastolic blood pressure; BMI = body mass index.

Adjusted ORs for predicting DSA-based high-risk status were calculated. CADi2 time ≥80 seconds was significantly associated with elevated dementia risk, with an adjusted OR of 1.94 (95% confidence interval [CI]: 1.12–3.40, p = 0.018). In contrast, the associations for a CADi2 score of ≤8 (p = 0.632) and a MMSE score of ≤28 (p = 0.572) were not statistically significant (Fig. 4). Similar non-significant results were observed when alternative cutoff values of ≤9 for the CADi2 score and ≤29 for the MMSE score were applied. In a sensitivity analysis additionally adjusting for all occupational categories summarized in the table 4b, only CADi2 time ≥80 seconds remained significantly associated with the outcome (aOR, 2.04; 95% CI, 1.16–3.60; p = 0.013), while no other variables reached statistical significance. Analyses using the DSA-predicted 3-year risk did not yield statistically significant results, although similar trends were observed. Detailed results for each cutoff are provided in the Supplementary Figure S2 and Supplementary Table S2.

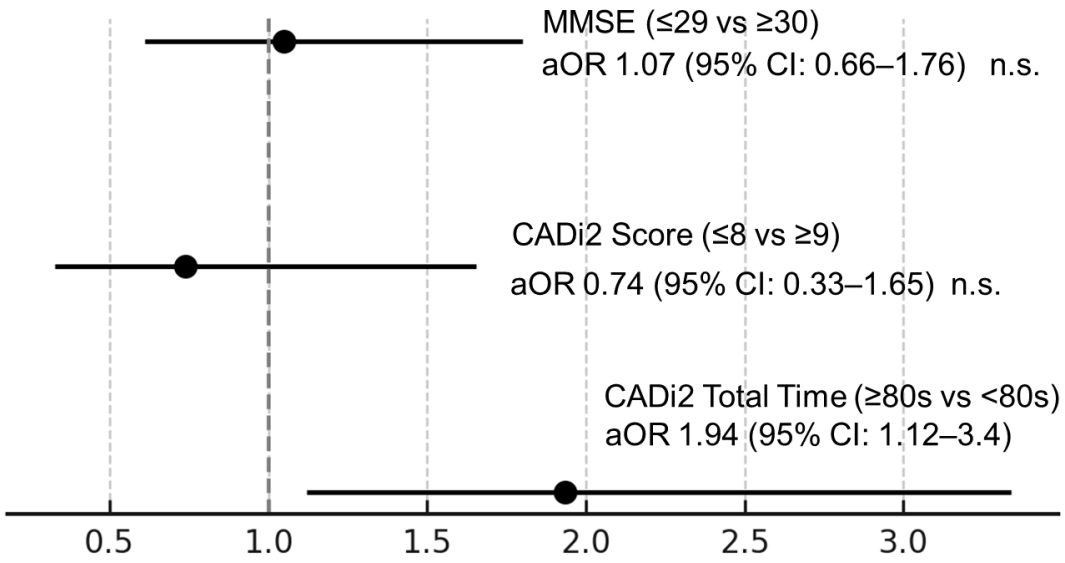

Fig. 4 Adjusted odds ratio (aORs) of cognitive indicators predicting 5-year dementia risk ≥0.25 based on deep survival analysis (DSA)

aORs for predicting 5-year risk of Alzheimer's disease ≥0.25, as defined by the DSA model, were calculated for each cognitive indicator after adjustment for covariates. These are illustrated in the forest plot.

n.s.: not significant, **: $p < 0.01$.

Discussion

Summary of key findings

In this study, we examined the association between CADi2 performance and brain atrophy in cognitively healthy adults. Completion time, but not total score, correlated significantly with localized atrophy, particularly in the hippocampus and medial temporal regions. These findings suggest that latency-based measures may be more sensitive to early structural changes than accuracy-based assessments. CADi2 completion time may therefore serve as an indicator in the preclinical stage of AD, whereas the CADi2 score may be more informative in the clinical stage (Fig. 5). Time and accuracy may also trade off, as some individuals respond quickly but inaccurately, while others are slower but accurate. In line with prior work, latency measures appear to complement accuracy-based outcomes. Libon and colleagues have shown that latency indices from the digital Clock Drawing Test predict impairment beyond conventional scoring, supporting our conclusion that CADi2 completion time captures subtle processing inefficiencies not evident from test scores alone.[27-28]

Fig. 5. Conceptual model of CADi2 measures across AD stages

Completion time may reflect subtle atrophy in preclinical AD, whereas scores may become more informative in clinical stages. A single CADi2 administration could thus contribute to disease staging by capturing different aspects of inefficiency across the continuum.

Sensitivity of completion time

Traditional screening tools, such as MMSE or CADi2 scores, are prone to ceiling effects, whereas completion time appears less affected and may detect early brain changes even among cognitively normal individuals. The hippocampus, highly vulnerable in AD, showed the strongest

association with completion time. Importantly, completion time incorporates processing speed, attention, and working memory, though it may also be influenced by factors such as mood, medication, or digital literacy, which were not fully assessed. In item-level analyses, Trail Making A/B and Cube Rotation tasks were disproportionately associated with regional atrophy, while orientation contributed minimally. This supports the potential of item-level latency measures to complement overall completion time, though validation in larger samples is needed.

Clinical validation using DSA

The analysis using DSA highlighted the clinical value of completion time. While score-based measures such as the MMSE and CADi2 scores were not associated with elevated dementia risk, CADi2 completion time showed a significant OR in identifying high-risk individuals. However, at this stage of our findings, thresholds such as 80 seconds should not be interpreted as diagnostic cutoffs, because even individuals classified as low-risk by DSA had mean completion times exceeding 80 seconds. Moreover, more than half of the entire sample required ≥80 seconds to complete the test, which is substantially higher than the prevalence of dementia in the general population. These results should therefore be interpreted only as indicating that time-based measures may detect early risk more effectively than accuracy-based tests, with the 80-second threshold serving as a reference value rather than a diagnostic criterion.

Study limitations

This study has several limitations. First, CADi2 was administered in Japanese to a Japanese population, limiting generalizability to other languages and cultures. Second, the sample consisted mainly of cognitively normal adults, restricting heterogeneity. Although we adjusted for age, sex,

education, vascular risk factors, and MRI findings (asymptomatic infarcts, microbleeds, DSWMH, PVH), other potential confounders such as psychiatric history, medication use, digital literacy, and occupational complexity were not fully assessed. Third, structural MRI alone has limited diagnostic utility in early AD. Knight et al. reported that MRI can detect early molecular and cellular changes, but clinical use in prodromal stages is constrained.30 A Cochrane review similarly concluded that MRI shows some promise for MCI but insufficient accuracy for standalone diagnosis.[31] Recent work integrating deep learning has reported >80% sensitivity and specificity.[32] Fourth, although stroke and Parkinson's disease were screened out, undetected cases cannot be excluded. Visual acuity, another known risk factor, was not assessed. Fifth, the relatively modest sample size limits generalization. Finally, DSA reflects risk of AD conversion rather than general dementia, so observed associations should be interpreted as correlational rather than diagnostic.

Clinical implications and future directions

Taken together, our findings suggest that combining completion time with scores could improve staging and risk prediction for dementia. Implementing multivariate models that integrate both metrics into the CADi2 platform may allow real-time risk estimation and feedback. Item-level analyses and integration of accuracy with latency may further enhance performance. Compared with emerging biomarkers such as plasma p-tau or Aβ42/40 (e.g., PrecivityAD2™), CADi2 is inevitably less sensitive and specific. Its major strength, however, lies in its practicality: a free, tablet-based tool enabling home self-assessment, which could support early intervention among at-risk individuals in the community.

Conclusion

CADi2 completion time was significantly associated with subtle atrophy in cognitively normal adults and may serve as a sensitive indicator of early dementia risk. While traditional cognitive scores are prone to ceiling effects, latency-based measures provide complementary value, particularly in preclinical stages. Future screening should consider incorporating response time alongside accuracy metrics to enhance risk stratification and staging.

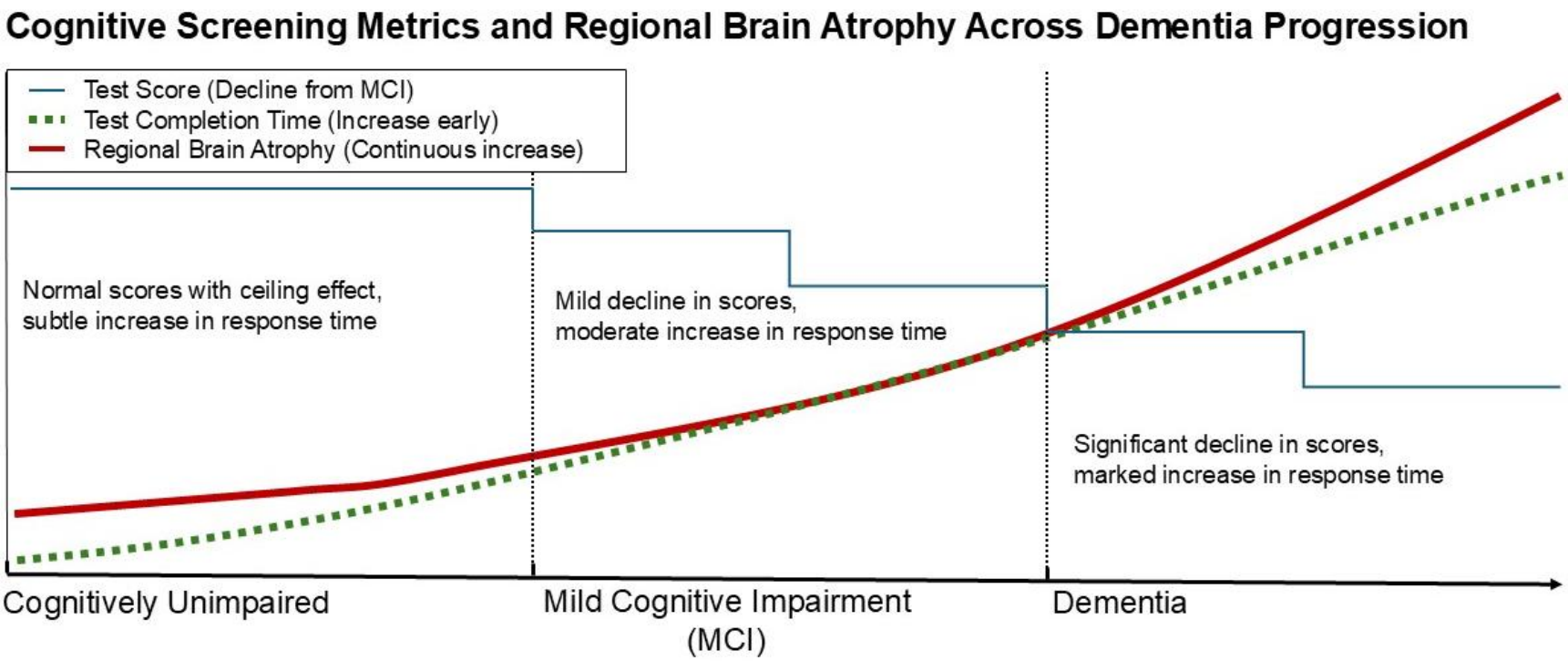

Fig. 5 Conceptual model of CADi2 measures across Alzheimer's disease stages

Completion time in the CADi2 test may be more sensitive in the preclinical stage of Alzheimer's disease, reflecting subtle brain atrophy, whereas CADi2 scores may become more informative in the clinical stage. This conceptual framework suggests that a single administration of CADi2 could contribute to disease staging by capturing different aspects of cognitive inefficiency across the disease continuum.

Acknowledgments

None

Author Contributions

Conceptualization, K. I. and M. T.; methodology, K. I., M. I., and M. T.; investigation, M. I. and Y. K.; writing—original draft preparation, K. I. and M. T.; writing—review and editing, M.I., Y.K., S. A. and A. N. All the authors have read and agreed to the published version of the manuscript.

Ethical Considerations

The study was approved by the Medical Ethics Committee of Shimane University School of Medicine and conducted in accordance with the Declaration of Helsinki (approval number: 2225, approval date: May 12, 2016).

Consent to Participate

Written informed consent was obtained from all participants.

Consent for Publication

Not applicable.

Declaration of Conflicting Interests

The authors have no conflicts of interest to declare.

Funding

This study was supported by a 2021 grant from the Taiju Life Social Welfare Foundation. The funder played no role in the design, data collection, data analysis, or reporting of this study.

Data Availability Statement

Raw data were generated at Shimane University. The data supporting the findings of this study are available from the corresponding author (K.I.) upon reasonable request.

References

1. Folstein MF, Folstein SE and McHugh PR. “Mini-mental state.” A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189-198.
2. Tsoi KKF, Chan JYC, Hirai HW, et al. Cognitive tests to detect dementia: a systematic review and meta-analysis. *JAMA Intern Med* 2015; 175: 1450-1458.
3. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005; 53: 695-699.
4. Davis DHJ, Creavin ST, Yip JLY, et al. Montreal Cognitive Assessment for the diagnosis of Alzheimer’s disease and other dementias. *Cochrane Database Syst Rev* 2015; 2015: CD010775.
5. Onoda K, Hamano T, Nabika Y, et al. Validation of a new mass screening tool for cognitive impairment: Cognitive Assessment for Dementia, iPad version. *Clin Interv Aging* 2013; 8: 353-360.
6. Onoda K and Yamaguchi S. Revision of the cognitive assessment for dementia, iPad version (CADi2). *PLoS One* 2014; 9: e109931.

7. Fox NC, Warrington EK, Freeborough PA, et al. Presymptomatic hippocampal atrophy in Alzheimer's disease. A longitudinal MRI study. *Brain* 1996; 119: 2001-2007.
8. Tondelli M, Wilcock GK, Nichelli P, et al. Structural MRI changes are detectable up to ten years before clinical Alzheimer's disease. *Neurobiol Aging* 2012; 33: 825.e25-825.e36.
9. Scheltens P, Fox N, Barkhof F, et al. Structural magnetic resonance imaging in the practical assessment of dementia: beyond exclusion. *Lancet Neurol* 2002; 1: 13-21.
10. Jack CR, Shiung MM, Gunter JL, et al. Comparison of different MRI brain atrophy rate measures with clinical disease progression in AD. *Neurology* 2004; 62: 591-600.
11. Moradi E, Pepe A, Gaser C, et al. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. *Neuroimage* 2015; 104: 398-412.
12. Edison P, Femminella GD, Nowell J, et al. Assessing the relationship between cognitive dysfunction and brain atrophy in Alzheimer's disease. *Alzheimers Dement* 2020; 16: e046004.
13. Sue K, Hirabayashi H and Osawa M. Relationship between neuropsychological test scores and hippocampal atrophy in non-demented Japanese older adults. *Interdiscip Neurosurg* 2022; 30: 101605.

14. Pantel J, Schönknecht P and Essig M. Distribution of cerebral atrophy assessed by magnetic resonance imaging reflects patterns of neuropsychological deficits in Alzheimer's dementia. *Neurosci Lett* 2004; 361: 17-20.

15. Sakurai R, Kim Y, Inagaki H, et al. MMSE cutoff discriminates hippocampal atrophy: neural evidence for the cutoff of 24 points. *J Am Geriatr Soc* 2021; 69: 839-841.

16. Kasai W, Goto T and Aoyama Y. Association between cognitive quotient test score and hippocampal volume: a novel, rapid application-based screening tool. *Sci Rep* 2020; 10: 16728.

17. Reitan RM. The relation of the Trail Making Test to organic brain damage. *J Consult Psychol* 1955; 19: 393-394.

18. Arlt S, Buchert R, Spies L, et al. Association between fully automated MRI-based volumetry of different brain regions and neuropsychological test performance in patients with amnestic mild cognitive impairment and Alzheimer's disease. *Eur Arch Psychiatry Clin Neurosci* 2013; 263: 335-344.

19. Weiner MW, Petersen RC, Aisen PS, et al. Digital cognitive assessments as low-burden markers for predicting progression to mild cognitive impairment. *Alzheimer's Dement* 2023; 19: 1234–1245.

20. Chen et al., 2025, Predicting brain amyloid load with digital and blood-based biomarkers, https://pubmed.ncbi.nlm.nih.gov/40618175/

21. The Japan Society of Brain Dock. Guidelines for Brain Dock 2019 [in Japanese]. Tokyo: Igaku-Shoin; 2019.

22. Ashburner J and Friston KJ. Unified segmentation. *NeuroImage* 2005; 26, 839-851.

23. Ashburner J and Friston KJ. Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation. *NeuroImage* 2011; 55, 954-967.

24. Rajapakse JC, Giedd JN, and Rapoport JL. Statistical approach to segmentation of single-channel cerebral MR images. *IEEE Transactions on Medical Imaging* 1997; 16, 176-186.

25. Nakagawa T, Ishida M, Naito J, et al. Prediction of conversion to Alzheimer's disease using deep survival analysis of MRI images. Brain Commun 2020; 2: fcaa057 .

26. Livingston G, Frankish H, Sommerlad A, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care: 2024 update of the Lancet Commission. Lancet 2024; 403: 1853-1912.Commission. *Lancet* 2024; 403: 1853-1912

27. .Emrani S, Libon DJ, Lamar M, Price CC, Giovannetti T, Swenson R, et al. Clock drawing latency predicts cognitive impairment. *Alzheimer's Dement (Amst).* 2022;14:e12345.

28. Libon DJ, Emrani S, Price CC, Lamar M, Swenson R, Delano-Wood L, et al. Latency indices in the digital Clock Drawing Test: novel markers of cognitive impairment. *J Alzheimers Dis.* 2023;95:567–579.

29. Ashford JW, Kolm P, Colliver JA, Bekian C, Hsu LN. Alzheimer patient evaluation and the mini-mental state: item characteristic curve analysis. J Gerontol. 1989 Sep;44(5):P139-46. doi: 10.1093/geronj/44.5.p139. PMID: 2768772.

30. Knight MJ, McCann B and Kauppinen RA. Magnetic resonance imaging to detect early molecular and cellular changes in Alzheimer's disease. *Front Aging Neurosci* 2016; 8: 139.

31. Lombardi G, Crescioli G, Cavedo E, et al. Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer's disease in people with mild cognitive impairment. *Cochrane Database Syst Rev* 2020; 3: CD009628.

32. Wang LX, Wang YZ, Han CG, et al. Revolutionizing early Alzheimer's disease and mild cognitive impairment diagnosis: a deep learning MRI meta-analysis. *Arq Neuropsiquiatr* 2024; 82: 1–10.