



Article

<https://doi.org/10.1038/s41591-025-03965-4>

Primary care detection of Alzheimer's disease using a self-administered digital cognitive test and blood biomarkers

Received: 13 March 2025

Accepted: 13 August 2025

Published online: 15 September 2025

Pontus Tideman^{1,2,6}✉, Linda Karlsson^{1,6}, Olof Strandberg¹,
Susanna Calling^{3,4}, Ruben Smith^{1,2}, Patrik Midlöv¹, Philip B. Verghese^{1,5},
Joel B. Braunstein⁵, Niklas Mattsson-Carlsgren^{1,2}, Erik Stomrud^{1,2},
Sebastian Palmqvist^{1,2} & Oskar Hansson¹✉

Check for updates

After the clinical implementation of amyloid- β -targeting therapies for people with cognitive impairment due to Alzheimer's disease (AD), there is an urgent need to efficiently identify this patient population in primary care. Therefore, we created a brief and self-administered digital cognitive test battery (BioCog). Based on its sub-scores, a logistic regression model was developed in a secondary care cohort ($n = 223$) and then evaluated in an independent primary care cohort comprising 19 primary care centers ($n = 403$). In primary care, BioCog had an accuracy of 85% when using a single cutoff to define cognitive impairment, which was significantly better than the assessment of primary care physicians (accuracy 73%). The accuracy increased to 90% when using a two-cutoff approach. BioCog had significantly higher accuracy than standard paper-and-pencil tests (that is, Mini-Mental State Examination, Montreal Cognitive Assessment, Mini-Cog) and another digital cognitive test. Furthermore, BioCog combined with a blood test could detect clinical, biomarker-verified AD with an accuracy of 90% (one cutoff), significantly better than standard-of-care (accuracy 70%) or when using the blood test alone (accuracy 80%). In conclusion, this proof-of-concept study shows that a brief, self-administered digital cognitive test battery can detect cognitive impairment and, when combined with a blood test, accurately identify clinical AD in primary care.

AD is the most common cause of dementia, characterized by deposits of amyloid- β (A β), aggregated tau and progressive neurodegeneration¹. The clinical syndrome of AD typically starts with subjective cognitive decline (SCD), in which individuals report memory problems and/or other cognitive difficulties but perform normally on cognitive tests. This phase is followed by mild cognitive impairment (MCI), marked by objective cognitive symptoms, and later by dementia, characterized by clear functional deficits in activities of daily living².

Diagnosing AD is challenging, especially in the early stages of the disease and in primary care. The prevalence of both underdiagnosis and misdiagnosis is high with 20–30% of people being misdiagnosed in specialized care and approximately 40% in primary care when the AD diagnosis is not supported by biomarkers^{3–6}. The implementation of biomarker testing such as positron emission tomography (PET) with A β aggregate-binding tracer or by lumbar puncture measuring concentrations of A β and tau in the cerebrospinal fluid (CSF) has improved

¹Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Lund, Sweden. ²Memory Clinic, Skåne University Hospital, Malmö, Sweden. ³Center for Primary Health Care Research, Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden. ⁴Office for Primary Care, Skåne University Hospital, Lund, Sweden. ⁵C2N Diagnostics LLC, St Louis, MO, USA. ⁶These authors contributed equally: Pontus Tideman, Linda Karlsson. ✉e-mail: pontus.tideman@med.lu.se; oskar.hansson@med.lu.se

Table 1 | Characteristics of the primary and secondary care cohorts

	Care cohort		Standardized between-group difference	
	Primary	Secondary	Median (95% CI) ^a	Percent (95% CI) ^b
Number of participants	403	223		
Age median (IQR), years	77 (72 to 82)	73 (66 to 78)	0.52 (0.26 to 0.81)	
Sex, n (%)				
Female	209 (52)	95 (43)		2.2 (0.33 to 4.1)
Male	194 (48)	128 (57)		-2.2 (-4.1 to -0.35)
Length of education median (IQR), years	11 (9 to 13)	13 (11 to 16)	-0.60 (-0.92 to -0.30)	
Carrier of APOE ε4, n; total (%)	169; 402 (42)	130; 220 (59)		-4.1 (-6.1 to -2.2)
Mini-Mental State Examination				
Number of participants	402	223		
Score, median (IQR)	27 (25 to 29)	27 (24 to 29)	-0.14 (-0.32 to 0)	
Cognitive impairment (RBANS < 78), n (%) ^c				
Negative	174 (43)	104 (47)		-0.83 (-2.7 to 1.1)
Positive without dementia	124 (31)	66 (30)		0.31 (-1.6 to 2.4)
Positive with dementia	105 (26)	53 (24)		0.64 (-1.4 to 2.6)
Aβ status ^d , n; total (%)				
Negative	196; 403 (49)	84; 222 (38)		2.6 (0.8 to 4.8)
Positive	207; 403 (51)	138; 222 (62)		-2.6 (-4.8 to -0.7)
Medical history, n (%)				
Cardiovascular disease	303 (75)	120 (54)		5.4 (3.4 to 7.5)
Hyperlipidemia	249 (62)	110 (49)		3.0 (0.87 to 5.0)
Chronic kidney disease	108 (27)	46 (21)		1.8 (-0.20 to 3.8)
Diabetes	78 (19)	34 (15)		1.3 (-0.59 to 3.4)

^aStandard median difference (difference between medians divided by pooled s.d.). ^bStandardized difference in proportions. ^cProxy variable based on similar paper-and-pencil tests for secondary care cohort (see details in Supplementary Methods). ^dBiomarker confirmed, based on either CSF Aβ42/Aβ40 or Aβ-PET.

diagnostic precision in specialized clinics^{7–9}. Nevertheless, these methods are either costly or invasive, and the availability is very limited globally. Recently, blood tests based on tau phosphorylated at threonine 217 (phosphorylated-tau217) have been shown to detect AD pathology with accuracies of around 90% in both primary and secondary care⁶, demonstrating performances comparable to clinically used CSF tests¹⁰. Therefore, blood tests for AD have the potential to revolutionize the diagnostic work-up of AD in people with cognitive symptoms, especially in primary care centers^{9,11}.

The prevalence of AD brain pathology in the population varies by age and apolipoprotein E (APOE) status, but even more so by the severity of cognitive symptoms¹¹. SCD can arise from a variety of causes, resulting in approximately only 20–25% of individuals with SCD exhibiting AD pathology. By contrast, about 40–60% of individuals with MCI and 70–80% with dementia have AD pathology¹². As a result, the pretest probabilities vary by cognitive stage, resulting in lower positive predictive values (PPV) of AD biomarkers in people with SCD compared to those with dementia^{11,12}. Consequently, a positive AD blood test in an individual with SCD carries a higher risk of being a false-positive than a positive result for an individual with cognitive impairment (MCI or dementia)⁶. The implications of such a false-positive result are substantial, because a person may be incorrectly informed of having AD associated with progressive cognitive decline in the coming years. Furthermore, only individuals with AD with cognitive impairment (but not SCD) are eligible for treatment with the newly implemented Aβ-targeting immunotherapies^{1,13–15}. Biomarker testing in clinical practice is currently only recommended in patients with cognitive impairment^{1,13,15,16}. This underscores the importance of identifying objective impairment (MCI and dementia) in individuals with cognitive symptoms, which in

international guidelines is recommended to always precede AD biomarker testing^{5,11,17,18}.

The administration and interpretation of cognitive tests vary substantially among primary care providers, both in and between countries. In addition, primary care providers often lack the time, expertise or resources to conduct assisted, in-clinic, cognitive tests¹⁹. Brief yet accurate digital cognitive testing, performed either remotely or unsupervised at the primary care center, may offer a viable solution to overcome this gap. We therefore developed a digital, self-administered, cognitive test battery (BioCog) and examined its ability to accurately identify cognitive impairment among people presenting with cognitive symptoms in primary care in a proof-of-concept study. Furthermore, we explored its utility when combined with a blood test to identify individuals with cognitive impairment due to AD.

Results

Study participants

We included 223 participants with cognitive symptoms from a secondary care cohort: the ongoing BioFINDER-2 study ([NCT03174938](#)) (Table 1)²⁰. The mean (s.d.) age was 73 (8.8) years, 128 (57%) were male and 119 (53%) had objectively verified cognitive impairment. In addition, 403 participants were included from a primary care cohort, BioFINDER-Primary Care study ([NCT06120361](#)) (Table 1 and Supplementary Fig. 1), which recruits people seeking help for cognitive symptoms at 19 primary care units⁶. Only people for whom the primary care physician (PCP) thought neurodegenerative disease was a reasonably possible cause of the symptomatology were invited to participate in the study. The mean (s.d.) age of the primary care cohort was 77 (8.0) years, 194 (48%) were male and 229 (57%) had objectively verified cognitive impairment.

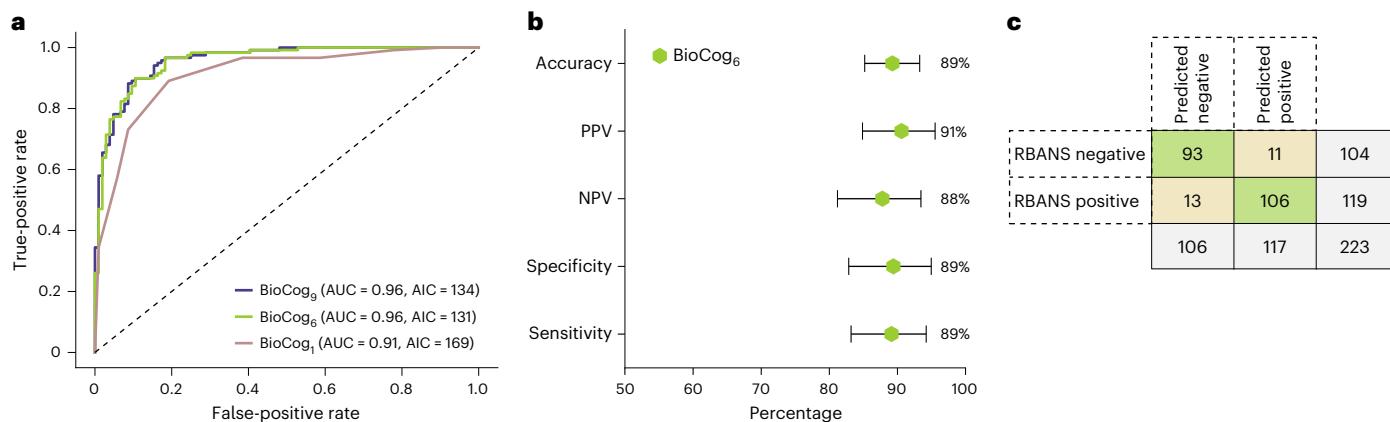


Fig. 1 | Establishing and evaluating models in the secondary care cohort.

Model performances in the secondary care cohort (which was also used to fit the models and to establish the cutoff), $n = 223$. **a**, ROC curves, AUCs and AICs for the three BioCog models using one (BioCog₁), six (BioCog₆) or nine (BioCog₉) input

variables. **b**, Evaluation of BioCog₆ selecting an optimal cutoff point based on the maximum value of Youden's index (probability = 0.575). Error bars indicate 95% CI, with the center point corresponding to the mean value. **c**, Cross-tabulation analysis using BioCog₆.

Compared with secondary care, participants in the primary care cohort were older, had fewer years of education, a higher prevalence of cardiovascular disease and hyperlipidemia, and a larger proportion were female. A slightly smaller proportion of participants in primary care were A β -positive and *APOE* $\epsilon 4$ carriers. Similar proportions were cognitively impaired (Table 1).

All participants performed the self-administered digital test battery, BioCog, comprising: (1) a word list test with immediate recall, delayed recall and recognition of ten words; (2) a cognitive processing speed task; and (3) questions about orientation to time (Supplementary Fig. 2). The study's primary outcome was objectively verified cognitive impairment determined on the basis of a standard battery of cognitive tests, the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), administered by neuropsychologists at a memory clinic. The secondary care cohort was used for training to establish cognitive test models and cutoffs. The primary care cohort was used for validation where the models and cutoffs were applied.

Psychometric properties of BioCog

The subtests of BioCog demonstrated moderate negative Spearman's correlations with age (-0.31 to -0.58), and weak positive correlations with education (0.13 to 0.33). Performance across subtests did not differ by sex in either cohort, except for immediate recall and recognition in the primary care cohort, where female participants performed slightly better (Supplementary Tables 1 and 2). Strong correlations (0.57 to 0.83) were observed between BioCog subtests and their corresponding paper-and-pencil analogs (for example, BioCog processing speed test and Symbol Digit Modalities Test), supporting convergent validity. Correlations between individual BioCog subtests and a short-term memory task were weak (0.11 to 0.29). Similarly, only weak correlations were observed between BioCog subtests and a measure of visuoperceptual or spatial functioning (0.13 to 0.27), indicative of divergent validity (Supplementary Tables 3 and 4). Participants with cognitive impairment scored significantly lower on all BioCog measures compared to cognitively unimpaired individuals (Supplementary Tables 5 and 6). Internal consistency of the subtests, estimated by McDonald's omega (ω), a reliability coefficient, ranged from acceptable to excellent (0.70 to 0.90) (Supplementary Table 7). Self-reported difficulty in understanding test instructions was very low (2%), and most participants reported that the tasks were engaging or were neutral (word list test 80%; processing speed test 91%) (Supplementary Fig. 3). Average completion time for BioCog was 11.2 min (s.d. 1.02 min), similarly across both cognitively unimpaired and cognitively

impaired participants (Supplementary Table 8). Overall, these results suggest that BioCog is a valid, reliable and feasible tool for assessing cognitive function.

Establishing a BioCog model in the secondary care cohort

In the secondary care cohort, multiple logistic regression models that could predict the presence or absence of cognitive impairment were established using basic demographics (three variables: age, sex and education level) and BioCog (six variables from the subtests), in total nine variables (Supplementary Table 9). When creating models, recursive variable selection was performed to decrease the risk of overfitting. The best model based on the Akaike information criterion (AIC) was selected. The model with lowest AIC included six variables (BioCog₆): delayed ten-word recall (number of correct answers), cognitive processing speed task (number of correct answers), three repetitions of immediate ten-word recall (number of correct answers), age, delayed ten-word recognition (number of correct answers) and three repetitions of immediate ten-word recall (total time); listed in the order of feature importance.

For comparison, receiver operating characteristic (ROC) curves from BioCog₆ (area under the curve (AUC) = 0.96) together with models containing all variables (BioCog₉, AUC = 0.96, in addition including orientation to time (correct answers), education level and sex), or only the best performing single subtest (BioCog₁, AUC = 0.91, only delayed ten-word recall), are provided in Fig. 1a. In the secondary care cohort, BioCog₆ demonstrated an accuracy of 89% (95% confidence interval (CI) 84–93%), PPV of 91% (95% CI 85–95%), negative predictive value (NPV) of 88% (95% CI 81–93%), specificity of 89% (95% CI 83–95%) and sensitivity of 89% (95% CI 83–94%) for detection of cognitive impairment (Fig. 1b). Comparisons of true versus predicted classes are shown in a cross-tabulation analysis (Fig. 1c).

To enhance diagnostic performance, a two-cutoff approach (as previously proposed^{3,6,9,21–23}) was applied using BioCog₆. This way, participants with predicted probabilities close to the one-cutoff decision boundary formed an intermediate group and were not considered when evaluating performance (representing individuals who would need further evaluation to determine their cognitive status). The two-cutoff probabilities of 0.332 and 0.769 were chosen to achieve 95% sensitivity and 95% specificity, respectively. Using these cutoffs, BioCog₆ had an accuracy of 96% (95% CI 92–98%), PPV of 96% (95% CI 91–99%), NPV of 96% (95% CI 90–98%), specificity of 96% (95% CI 90–98%) and sensitivity of 96% (95% CI 91–99%) (Extended Data Fig. 1). The intermediate group comprised 18% of the cohort (95% CI 12–22%).

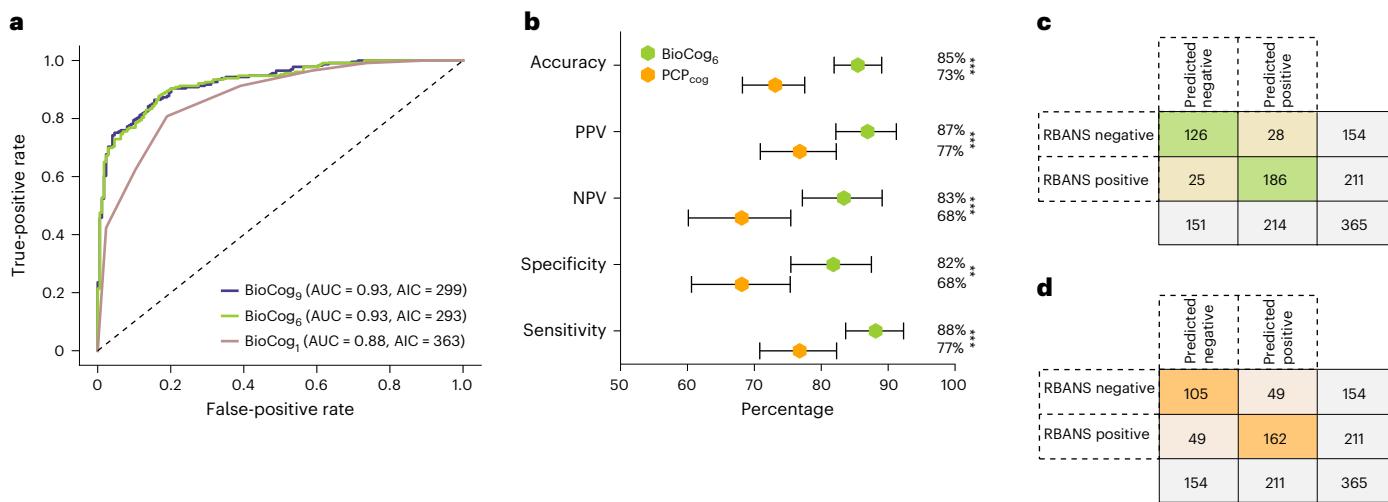


Fig. 2 | Evaluating models in the primary care cohort. Model performances in the primary care cohort (which was not used to fit the models or to establish the cutoff). **a**, ROC curves, AUCs and AICs for the three BioCog models using one (BioCog₁), six (BioCog₆) or nine (BioCog₉) input variables, $n = 403$. **b**, Evaluation of BioCog₆ based on the optimal cutoff point established in the secondary care cohort (probability = 0.575), $n = 365$. Compared to a PCP diagnosis of SCD versus MCI or dementia (PCP_{cog}). Error bars indicate 95% CI, with the center point corresponding to the mean value. CI and two-sided P values were computed using bootstrapping ($n = 5,000$ resamples with replacement) and

adjusted for multiple comparisons (Benjamini–Hochberg). **c**, Cross-tabulation analysis using BioCog₆. **d**, Cross-tabulation analysis using PCP_{cog} diagnoses. Comparisons in **b**–**d** were made on a subset of individuals with existing PCP_{cog} evaluation data ($n = 365$). ** $P < 0.01$, *** $P < 0.001$ (assessed with bootstrapping and false discovery rate (FDR) corrected). Exact P values in **b**: $P_{\text{Accuracy}} = 0.0007$, $P_{\text{PPV}} = 0.001$, $P_{\text{NPV}} = 0.0007$, $P_{\text{Specificity}} = 0.005$, $P_{\text{Sensitivity}} = 0.0007$. Nominal significance <0.0001 was set to 0.0001 during FDR correction, yielding a lowest adjusted P value of 0.0007.

Validating the BioCog model in the primary care cohort

Next, we evaluated the predictive ability of the previously established BioCog models and cutoffs in the independent primary care cohort. BioCog₆ achieved higher or similar AUC (0.93) compared to BioCog₁ (AUC = 0.88) and BioCog₉ (AUC = 0.93) when identifying cognitive impairment (Fig. 2a). In the primary care cohort, BioCog₆ had an accuracy of 85% (95% CI 81–89%), PPV of 87% (95% CI 82–91%), NPV of 83% (95% CI 77–89%), specificity of 82% (95% CI 75–87%) and sensitivity of 88% (95% CI 83–92%) (Fig. 2b). Of the 96 cognitively impaired participants without AD, 82 were correctly identified as impaired by the BioCog₆ model (85% accuracy), suggesting similar performance in this subpopulation. The BioCog₆ model was compared to standard assessment of PCPs (PCP_{cog}), which included brief cognitive testing with Mini-Mental State Examination (MMSE) and The Montreal Cognitive Assessment (MoCA), computed tomography scan of the brain and a clinical assessment of the participant. BioCog₆ had significantly higher performance across all evaluation metrics (Fig. 2b) (accuracy BioCog₆ = 85% versus PCP = 73%; $P < 0.001$). Comparisons of true versus predicted classes for BioCog₆ and PCP_{cog} assessments can be seen in cross-tabulation analyses (Fig. 2c,d).

Using two cutoffs (applying the cutoffs previously established in the secondary care cohort), BioCog₆ showed an accuracy of 90% (95% CI 86–93%), PPV of 91% (95% CI 87–95%), NPV of 89% (95% CI 82–93%), specificity of 87% (95% CI 81–92%) and sensitivity of 92% (95% CI 88–95%) (Fig. 3). The intermediate group comprised 18% of the participants (95% CI 14–22%).

Next, a head-to-head comparison was performed between the BioCog₆ model and several widely used standard paper-and-pencil tests (MMSE, MoCA, Mini-Cog) and another digital test (Cambridge Neuropsychological Test Automated Battery (CANTAB)). Using one cutoff, the BioCog₆ model significantly outperformed all the other cognitive tests on accuracy (Table 2). This result remained consistent when cognitive tests were treated as continuous variables and adjusted for demographics (Supplementary Fig. 4). It also remained when removing age as an input in the BioCog₆ model (Supplementary Table 10). For the two-cutoff approach, BioCog₆ demonstrated significantly higher

accuracy than MoCA, the only other test adapted for a two-cutoff approach (90% versus 76%, $P = 0.001$), and with a smaller, but not significantly different intermediate group size (18% versus 23%, $P = 0.07$).

Digital test and blood biomarkers compared to standard care

Finally, we simulated a two-step diagnostic workflow in primary care that can be performed in addition to the standard clinical evaluation made by a physician to identify biomarker-verified clinical (symptomatic) AD: step 1, detection of objective cognitive impairment using BioCog (BioCog₆); and step 2, blood biomarker testing (PrecivityAD2; Amyloid Probability Score-2 (APS2)) in those who are cognitively impaired according to BioCog (Fig. 4a). We compared this workflow in the primary care cohort against current standard clinical evaluation by PCPs, where the physician first assessed whether the patient had cognitive impairment and second, whether the impairment was caused by AD (PCP_{AD}).

When cognitive impairment (based on neuropsychological assessment) and a clinical AD diagnosis (based on a consensus diagnosis done by dementia experts including CSF analysis or A β -PET assessments) was used as outcome, the two-step diagnostic workflow of BioCog (one cutoff) and blood biomarker achieved an overall clinical AD diagnostic accuracy of 90% (95% CI 86–92%), PPV of 83% (95% CI 75–89%), NPV of 93% (95% CI 89–95%), specificity of 92% (95% CI 88–95%) and sensitivity of 84% (95% CI 77–90%) (Fig. 4b). By contrast, PCPs had an overall accuracy of 70% (95% CI 65–74%), PPV of 52% (95% CI 43–61%), NPV of 79% (95% CI 73–84%), specificity of 76% (95% CI 71–81%) and sensitivity of 56% (95% CI 47–66%), all significantly lower than the new digital test and blood biomarker workflow (Fig. 4b).

In comparison, relying solely on the blood biomarker to identify clinical AD without using BioCog maintained a high sensitivity of 92% (95% CI 86–96%) and NPV of 95% (95% CI 92–98%), but showed significantly lower accuracy of 80% (95% CI 76–84%), PPV of 63% (95% CI 55–69%) and specificity of 75% (95% CI 69–80%) compared to the combined BioCog and blood biomarker workflow (Fig. 4b and Supplementary Table 11). Notably, a high proportion (17%) of individuals were classified as false positives when only using the blood test (Fig. 4e).

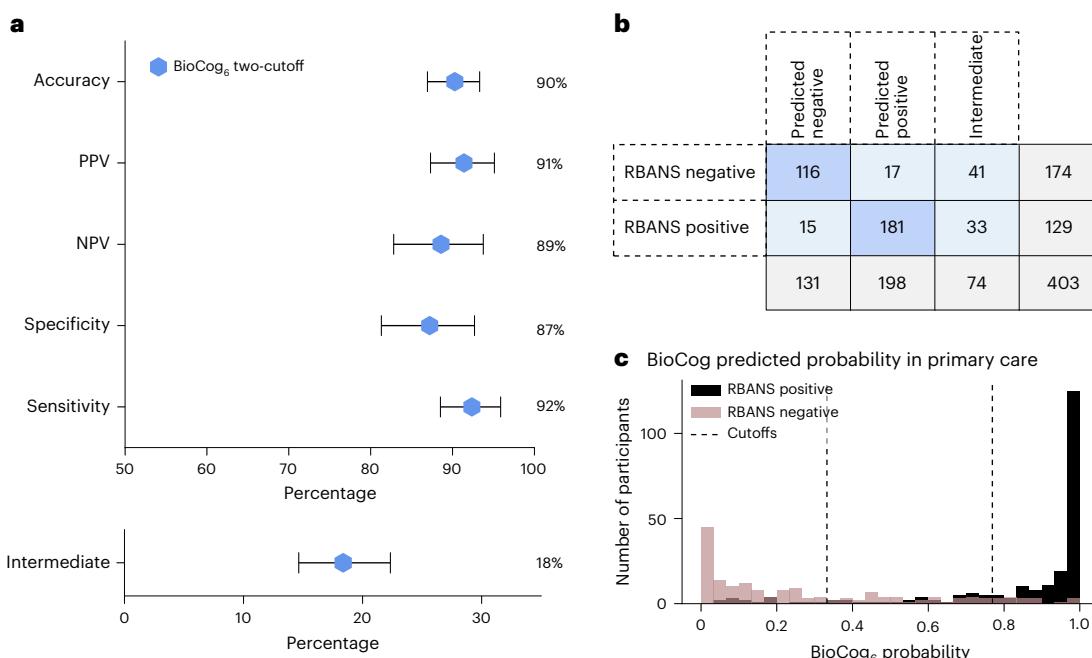


Fig. 3 | Evaluating a two-cutoff approach for the BioCog₆ model in the primary care cohort. BioCog₆ two-cutoff model performance in the primary care cohort (which was not used to fit the model nor to establish the cutoffs), $n = 403$. **a**, Evaluation of the BioCog₆ two-cutoff model ($n = 329$, without the participants in the intermediate group). Error bars indicate 95% CI,

with the center point corresponding to the mean value. **b**, Cross-tabulation analysis using the BioCog₆ two-cutoff model. **c**, Histogram of the predicted probabilities from the BioCog₆ model (x axis) colored by the true binary RBANS composite score (outcome variable).

Using the BioCog₆ and APS2 workflow resulted in significantly higher accuracy than combining APS2 with the traditional cognitive assessments MMSE and MoCA (Supplementary Table 12).

If using a two-cutoff approach for both BioCog and the blood biomarker, the accuracy increased to 95% (95% CI 91–97%), with 30% (95% CI 25–34%) of individuals in the intermediate group (Supplementary Fig. 5). All results were similar when using cognitive impairment and an abnormal Aβ CSF or PET biomarker as outcome instead of a clinical AD diagnosis (Supplementary Fig. 5), and when applying the clinically used APS2 cutoff (Supplementary Fig. 6).

Sensitivity analyses

A Clinical Dementia Rating (CDR) global score of ≥ 0.5 was evaluated as the reference standard for cognitive impairment instead of RBANS. When directly applying the BioCog₆ model in the primary care cohort using this outcome, the performance remained high (Supplementary Fig. 7 and Supplementary Table 13). The AUC decreased from 0.93 to 0.90, accuracy from 85% to 82%, and PPV from 87% to 86%. These results highlight the robustness and generalizability of the BioCog₆ model in predicting cognitive impairment, also when using a different reference standard.

Additional evaluation metrics for the main analyses in the primary care cohort (including balanced accuracy, harmonic mean of precision and recall (F1 score), and positive and negative likelihood ratios) (Supplementary Tables 14 and 15) supported our main findings.

Discussion

This study demonstrated that a brief digital cognitive assessment battery detects cognitive impairment with an accuracy of 85% when validated in an elderly primary care population presenting with cognitive symptoms. The accuracy of the self-administered BioCog test was significantly higher than that of standard-of-care methods, including paper-and-pencil tests MMSE, MoCA and Mini-Cog, which require the presence of healthcare personnel when administered. The digital test

on its own also outperformed clinical evaluation by the PCP, which included patient history and standard cognitive tests. Notably, higher performance was consistent whether cognitive impairment was defined using RBANS or CDR, underscoring BioCog's robustness and generalizability in detecting objective cognitive impairment. Furthermore, in a two-step workflow combining BioCog with a blood biomarker test, clinical AD (that is, cognitive impairment due to AD) could be identified with significantly higher accuracy than the standard PCP assessment (90% versus 70%), or using a blood biomarker test without previous digital cognitive assessment (90% versus 80%). Beyond high accuracy, digital tests like BioCog offer distinct advantages over paper-and-pencil assessments. BioCog enables time-efficient cognitive evaluations and eliminates variability introduced by differences in scoring or interpretation, which are critical features in primary care¹⁹. Furthermore, it facilitates the simultaneous collection of different variable categories, such as time, alongside correct answers, allowing for more comprehensive data. Digital platforms also make it easier to utilize advanced statistical models, moving beyond simple threshold-based scoring of sums (used in MMSE, MoCA and Mini-Cog) to leverage more nuanced, data-driven insights without adding significant processing time or requiring advanced interpretation skills.

Compared to another digital test that has been CE marked and US Food and Drug Administration approved and is currently used in clinical practice for dementia evaluations (CANTAB), BioCog demonstrated significantly higher accuracy in identifying cognitive impairment (85% versus 76%) (Table 2). Our study did not evaluate all available digital tablet solutions, such as Cognivue²⁴, NIH Toolbox²⁵, cCOG²⁶, Cogstate²⁷ or TabCAT²⁸, considering the load this would impose on participants, and to avoid learning biases. To our knowledge, none of these digital tests have been evaluated in primary care. Future research comparing various brief and self-administered digital tools in primary care will be essential to ensure that clinicians have access to the most effective options for detecting cognitive impairment. Importantly, a central conclusion of this study is that combining digital cognitive testing and

Table 2 | Head-to-head comparison between BioCog₆ and other cognitive tests when predicting the binary RBANS composite in the primary care cohort

Cutoff(s) for positivity	Accuracy (95% CI); FDR-corrected P value compared to BioCog ₆	PPV (95% CI); FDR-corrected P value compared to BioCog ₆	NPV (95% CI); FDR-corrected P value compared to BioCog ₆	Specificity (95% CI); FDR-corrected P value compared to BioCog ₆	Sensitivity (95% CI); FDR-corrected P value compared to BioCog ₆	Intermediate (95% CI); FDR-corrected P value compared to BioCog ₆	
One-cutoff approaches							
BioCog ₆	>0.575	84% (81–88%)	84% (79–89%)	84% (79–90%)	79% (73–85%)	88% (84–92%)	–
MMSE	<27	71% (67–76%); 0.0007	81% (75–87%); 0.30	64% (58–70%); 0.0007	80% (75–86%); 0.70	64% (58–71%); 0.0007	–
MoCA	<26	67% (62–71%); 0.0007	63% (58–68%); 0.0007	93% (86–100%); 0.06	27% (20–34%); 0.0007	98% (96–100%); 0.0007	–
Mini-Cog	<4	75% (71–79%); 0.0007	78% (72–83%); 0.02	71% (64–78%); 0.0007	72% (65–79%); 0.08	77% (71–83%); 0.0007	–
CANTAB	>41	76% (71–80%); 0.0007	78% (73–83%); 0.03	73% (66–79%); 0.001	72% (65–79%); 0.08	78% (73–84%); 0.001	–
Two-cutoff approaches							
BioCog ₆	>0.769, positive <0.332, negative	90% (86–93%)	90% (86–94%)	89% (83–94%)	86% (80–92%)	92% (88–96%)	18% (14–22%)
MoCA	<24, positive >26, negative	76% (71–81%, 0.001)	74% (68–79%, 0.001)	90% (80–100%, 0.8)	31% (22–40%, 0.001)	98% (96–100%, 0.008)	24% (20–28%, 0.07)

Diagnostic performance (accuracy, PPV, NPV, sensitivity and specificity) was calculated using the BioCog₆ model (established in the secondary care cohort) or prespecified cutoffs from the literature. Comparisons were made on a subset of individuals with existing data for all cognitive tests ($n=381$). Confidence intervals and two-sided P values were computed using bootstrapping ($n=5,000$ resamples with replacement) and adjusted for multiple comparisons (Benjamini–Hochberg). Nominal significance <0.0001 was set to 0.0001 during FDR correction, yielding a lowest adjusted P value of 0.0007.

a blood biomarker can enhance not only the efficiency of the diagnostic process in primary care, but also the diagnostic certainty of AD.

A diagnostic workflow (Fig. 4) combining a positive BioCog assessment with an accurate blood biomarker test offers an efficient and reliable way to diagnose clinical AD with high certainty (accuracy of 90% compared to 70% for current standard-of-care without biomarkers and 80% for blood biomarkers only). These results suggest that digital cognitive tests, when combined with blood biomarker tests, have potential to significantly enhance the clinical work-up of AD and can become highly valuable to efficiently align clinical AD diagnostics with global recommendations. The World Health Organization's preferred product profile for AD blood tests emphasizes that such tests should be used exclusively in individuals with cognitive impairment¹⁷, where the pretest probability is high enough to achieve a PPV of $>90\%$ without the need for confirmatory CSF or PET biomarkers. This aligns with other international recommendations^{11,16}. In addition, only individuals with cognitive impairment are eligible for treatment with acetylcholinesterase inhibitors or recently developed A β -targeting immunotherapies^{1,13–15}. This also underscores the importance of identifying cognitive impairment in primary care with efficient and accurate tools like BioCog, to enable referral of as many eligible patients as possible to specialist clinics for treatment with A β immunotherapies without substantial burden on the primary care system. Further evaluation (for example magnetic resonance imaging scans, *APOE*, contraindicated medications) of potential treatment candidates is required at specialty clinics before initiating treatment²⁹.

For digital assessments and blood tests to be effective in diagnosing clinical AD, it is essential that they demonstrate very high PPVs, particularly before initiating expensive treatments. In addition, it is of high importance that their predictive properties are robust, with strong generalization to new patients. Blood biomarker tests have shown high performance across many secondary care cohorts^{9,30}, as well as in primary care¹¹, but their PPV varies by age and cognitive status because of differences in pretest probabilities³¹, so their use should be adapted to the clinical context. In this study, using BioCog to assess cognitive impairment unexpectedly showed a slight decrease in performance between the secondary care cohort (used to fit the model, AUC = 0.96 and PPV = 91%) and the more clinically diverse primary care cohort (only used for model evaluation, AUC = 0.93 and PPV = 87%). To develop a robust BioCog model, we applied recursive feature selection with AIC penalization and focused on external model validation in the independent primary care cohort.

We also investigated the ability of BioCog₆ to identify cognitive impairment with CDR as outcome instead of RBANS. Whereas RBANS offers a more objective measure of cognitive impairment, CDR captures functional decline^{32,33}. Both are widely used in clinical trials and routine clinical practice^{14,15,34,35}. When using a CDR global score ≥ 0.5 as outcome, performance remained high, with only a slight decrease (for example, accuracy from 85% to 82%, and PPV from 87% to 86%). Because the BioCog₆ model was optimized to detect cognitive impairment (RBANS) and not functional impairment (CDR), a drop in performance is expected. Nevertheless, this decrease was minimal, and BioCog₆ still outperformed both the clinical judgment of PCPs and other cognitive tests. To further increase diagnostic performance of the BioCog₆ model, we applied a two-cutoff approach, which resulted in a PPV that remained $>90\%$ even in primary care and with 18% of individuals in the intermediate group. These results are encouraging, with BioCog demonstrating a high level of robustness and adaptability, even when applied to a cohort with differing demographic characteristics and a higher prevalence of comorbidities (Table 1), and when defining cognitive impairment using a different reference standard. For the two-cutoff approach, people with intermediate results would need more comprehensive cognitive evaluations. Alternatively, repeated cognitive testing incorporating changes in test scores may help identify cognitive impairment in those individuals³⁶.

The findings of our study highlight the key role digital cognitive tools can play in optimizing diagnostic workflows, particularly when combined with blood biomarker tests for a more precise evaluation of clinical AD. Beyond individual test performance, the broader impact of digital assessments lies in their ability to enhance efficiency, standardization and scalability in cognitive assessments³⁷. Importantly, their greatest value emerges when integrated into a broader clinical evaluation, in which results can be interpreted alongside other clinical information to support well-informed diagnostic decisions and strengthen the healthcare professional's confidence in making a correct diagnosis³⁸. Even though this study focuses on highlighting the clinical value of BioCog by showing its independent contribution in identifying objective cognitive impairment and clinical AD, our proposal is not to use BioCog or other digital tests as stand-alone tools. BioCog is aimed to complement, not replace, clinical judgment, and further validation in additional research studies and integration into structured care pathways is essential before deployment. In addition, future implementation-based research in different healthcare systems is needed to evaluate the impact on clinical management and to

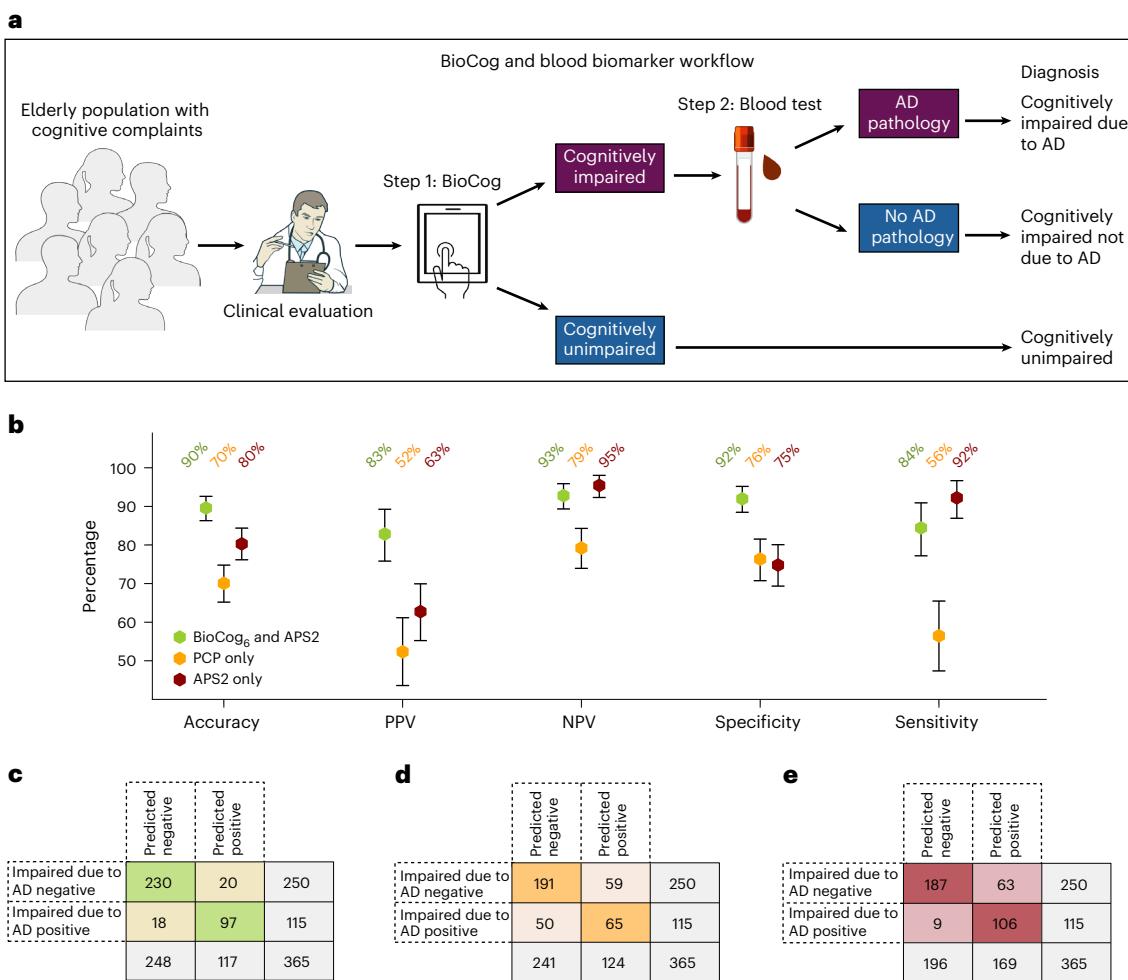


Fig. 4 | Comparing a digital testing and blood biomarker-based diagnostic workflow to the current standard clinical evaluation in the primary care cohort. Comparisons were made on a subset of individuals with all existing data available ($n = 365$; see Supplementary Table 17 for detailed population characteristics). **a**, Our proposed primary care two-step workflow consisting of step 1, detection of cognitive impairment using the BioCog, followed by step 2 in cognitively impaired individuals, a blood biomarker assessment to evaluate whether AD pathology is present. **b**, Evaluation of the workflow using our BioCog₆ model for step 1 and the plasma biomarker APS2 (green). The workflow was compared against a standard clinical evaluation by PCPs in

which the physician assesses both whether the patient had cognitive impairment (MCI or dementia) and whether the impairment was caused by AD (without any biomarkers) (PCP_{AD}, orange), and against a workflow using only the plasma biomarker APS2 without any cognitive assessment (red). Error bars indicate 95% CI, with the center point corresponding to the mean value. Significance comparisons between all workflows are given in Supplementary Table 11. **c–e**, Corresponding cross-tabulation analyses for BioCog and blood marker workflow (**c**), PCP workflow (**d**) and blood biomarker only workflow (**e**). Credit: Images in **a** adapted from NIAID NIH BIOART (<https://bioart.niaid.nih.gov>).

establish clinical practice guidelines before digital assessments combined with blood biomarkers can be used as part of everyday primary care in real-world settings.

There are limitations to this study. First, we established the BioCog model in a secondary care cohort, which differed on demographic factors like age and education level, and on comorbidity prevalence compared to the primary care cohort, representative of the intended use population. However, the robust results across cohorts despite these differences also highlight the generalizability of the model and cutoffs. Second, although BioCog performed well in identifying cognitive impairment, further research is needed to evaluate its utility in longitudinal settings and its ability to monitor disease progression. Third, BioCog was developed and evaluated in Swedish cohorts. It has been translated into several languages, including English, Finnish, Dutch and Spanish, and validation in these languages and in more diverse populations is important to further assess the generalizability. Such validation will be crucial before clinical implementation. Notably, paper-and-pencil versions of similar types of tests have performed very well across different languages and cultures. Fourth, this study

did not directly assess how the use of a digital tool influences the clinical work-up in primary care, or whether physicians find it beneficial in practice. These questions will be addressed in future work. Fifth, BioCog was not evaluated as a screening tool for the general population. Instead, it was assessed in individuals who presented with cognitive complaints in primary care, where a physician considered a neurodegenerative disease to be a possible cause. People whose symptoms were clearly attributable to other causes (for example, depression or sleep disorders) were treated or referred accordingly and were not included in the study. This limits our ability to assess BioCog's utility for screening of the general population. However, such screening currently lacks evidence for improving decision-making or other important outcomes³⁹. We therefore argue that the current study, limited to participants with cognitive symptoms possibly caused by an underlying neurodegenerative disease, is clinically more relevant and reflects a realistic scenario for how digital cognitive tools may be integrated to support physicians in primary care.

In summary, this proof-of-concept study demonstrates that a self-administered digital test is effective for identifying objective

cognitive impairment in a population with cognitive symptoms for whom the PCP thought a neurodegenerative disease was a possible cause of the symptomatology. Furthermore, when combined with a blood biomarker test, the digital cognitive test could reliably diagnose clinical AD with high accuracy, outperforming the current standard-of-care in primary care. These findings have the potential to transform clinical diagnostics of AD in primary care, enabling detection and treatment in a resource efficient manner.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-025-03965-4>.

References

1. Jack, C. R. Jr et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement.* **20**, 5143–5169 (2024).
2. Petersen, R. C. Mild cognitive impairment as a diagnostic entity. *J. Intern. Med.* **256**, 183–194 (2004).
3. Hansson, O. et al. The Alzheimer's Association appropriate use recommendations for blood biomarkers in Alzheimer's disease. *Alzheimers Dement.* **18**, 2669–2686 (2022).
4. Mattke, S. et al. Expected and diagnosed rates of mild cognitive impairment and dementia in the U.S. Medicare population: observational analysis. *Alzheimers Res. Ther.* **15**, 128 (2023).
5. Mielke, M. M. et al. Considerations for widespread implementation of blood-based biomarkers of Alzheimer's disease. *Alzheimers Dement.* **20**, 8209–8215 (2024).
6. Palmqvist, S. et al. Blood biomarkers to detect Alzheimer disease in primary care and secondary care. *JAMA* **332**, 1245–1257 (2024).
7. Rabinovici, G. D. et al. Association of amyloid positron emission tomography with subsequent change in clinical management among Medicare beneficiaries with mild cognitive impairment or dementia. *JAMA* **321**, 1161–1169 (2019).
8. Hansson, O. Biomarkers for neurodegenerative diseases. *Nat. Med.* **27**, 954–963 (2021).
9. Hansson, O., Blennow, K., Zetterberg, H. & Dage, J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. *Nat. Aging* **3**, 506–519 (2023).
10. Barthélémy, N. R. et al. Highly accurate blood test for Alzheimer's disease is similar or superior to clinical cerebrospinal fluid tests. *Nat. Med.* **30**, 1085–1095 (2024).
11. Schindler, S. E. et al. Acceptable performance of blood biomarker tests of amyloid pathology — recommendations from the Global CEO Initiative on Alzheimer's Disease. *Nat. Rev. Neurol.* **20**, 426–439 (2024).
12. Jansen, W. J. et al. Prevalence estimates of amyloid abnormality across the Alzheimer disease clinical spectrum. *JAMA Neurol.* **79**, 228–243 (2022).
13. Dubois, B. et al. Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol.* **20**, 484–496 (2021).
14. Sims, J. R. et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA* **330**, 512–527 (2023).
15. van Dyck, C. H. et al. Lecanemab in early Alzheimer's disease. *N. Engl. J. Med.* **388**, 9–21 (2023).
16. Hansson, O. & Jack, C. R. Jr. A clinical perspective on the revised criteria for diagnosis and staging of Alzheimer's disease. *Nat. Aging* **4**, 1029–1031 (2024).
17. World Health Organization. *Preferred Product Characteristics of Blood-Based Biomarker Diagnostics for Alzheimer Disease* (2024).
18. Palmqvist, S. et al. Alzheimer's Association Clinical Practice Guideline on the use of blood-based biomarkers in the diagnostic workup of suspected Alzheimer's disease within specialized care settings. *Alzheimers Dement.* **21**, e70535 (2025).
19. Mattke, S. et al. Expanding the use of brief cognitive assessments to detect suspected early-stage cognitive impairment in primary care. *Alzheimers Dement.* **19**, 4252–4259 (2023).
20. Palmqvist, S. et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *JAMA* **324**, 772–781 (2020).
21. Thomann, A. E., Berres, M., Goettel, N., Steiner, L. A. & Monsch, A. U. Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the montreal cognitive assessment. *Alzheimers Res. Ther.* **12**, 39 (2020).
22. Mattsson-Carlgren, N. et al. Plasma biomarker strategy for selecting patients with Alzheimer disease for antiamyloid immunotherapies. *JAMA Neurol.* **81**, 69–78 (2024).
23. Brum, W. S. et al. A two-step workflow based on plasma p-tau217 to screen for amyloid beta positivity with further confirmatory testing only in uncertain cases. *Nat. Aging* **3**, 1079–1090 (2023).
24. Cahn-Hidalgo, D., Estes, P. W. & Benabou, R. Validity, reliability, and psychometric properties of a computerized, cognitive assessment test (Cognivue®). *World J. Psychiatry* **10**, 1 (2020).
25. Weintraub, S. et al. Cognition assessment using the NIH Toolbox. *Neurology* **80**, S54–S64 (2013).
26. Rhodius-Meester, H. F. M. et al. cCOG: a web-based cognitive test tool for detecting neurodegenerative disorders. *Alzheimers Dement. (Amst.)* **12**, e12083 (2020).
27. Maruff, P. et al. Validity of the CogState brief battery: relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and AIDS dementia complex. *Arch. Clin. Neuropsychol.* **24**, 165–178 (2009).
28. Possin, K. L. et al. The brain health assessment for detecting and diagnosing neurocognitive disorders. *J. Am. Geriatr. Soc.* **66**, 150–156 (2018).
29. Cummings, J. et al. Lecanemab: appropriate use recommendations. *J. Prev. Alzheimers Dis.* **10**, 362–377 (2023).
30. Hampel, H. et al. Blood-based biomarkers for Alzheimer's disease: current state and future use in a transformed global healthcare landscape. *Neuron* **111**, 2781–2799 (2023).
31. Therriault, J. et al. Diagnosis of Alzheimer's disease using plasma biomarkers adjusted to clinical probability. *Nat. Aging* **4**, 1529–1537 (2024).
32. Hughes, C. P., Berg, L., Danziger, W., Coben, L. A. & Martin, R. L. A new clinical scale for the staging of dementia. *Br. J. Psychiatry* **140**, 566–572 (1982).
33. Morris, J. C. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int. Psychogeriatr.* **9**, 173–176 (1997).
34. Rabin, L. A., Paolillo, E. & Barr, W. B. Stability in test-use practices of clinical neuropsychologists in the United States and Canada over a 10-year period: a follow-up survey of INS and NAN members. *Arch. Clin. Neuropsychol.* **31**, 206–230 (2016).
35. Wessels, A. M. et al. Cognitive outcomes in trials of two BACE inhibitors in Alzheimer's disease. *Alzheimers Dement.* **16**, 1483–1492 (2020).
36. Borland, E. et al. Individualized, cross-validated prediction of future dementia using cognitive assessments in people with mild cognitive symptoms. *Alzheimers Dement.* **20**, 8625–8638 (2024).
37. Tsay, E. et al. Scalable plasma and digital cognitive markers for diagnosis and prognosis of Alzheimer's disease and related dementias. *Alzheimers Dement.* **20**, 2089–2101 (2024).
38. Ozawa, T. et al. Effects of the Davos Alzheimer's Collaborative early detection of cognitive impairment program on clinician attitudes, engagement, and confidence. *J. Prev. Alzheimers Dis.* **12**, 100038 (2025).

39. Lin, J. S., O'Connor, E., Rossom, R. C., Perdue, L. A. & Eckstrom, E. Screening for cognitive impairment in older adults: a systematic review for the US preventive services task force. *Ann. Intern. Med.* **159**, 601–612 (2013).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits 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/4.0/>.

© The Author(s) 2025

Methods

Participants

All participants provided written informed consent before enrollment in the study, and ethical approval for the study was obtained from the Swedish Ethical Review Authority. Participants from two cohorts were included, one from secondary care and one from primary care. Data for this study were collected between February 2022 and December 2024. The secondary care cohort was used as the training cohort to establish cognitive test models and cutoffs. The primary care cohort was used as the validation cohort where the models and cutoffs were applied.

Secondary care

Secondary care participants were recruited from the ongoing BioFINDER-2 study ([NCT03174938](#)), described in detail elsewhere²⁰. Briefly, the study consecutively enrolls participants at the secondary care memory clinic of Skåne University Hospital and the memory clinic of Ängelholm Hospital, and includes a diverse population of study participants. For the current study, patients with either SCD, MCI or dementia were included. Further details are given in the Supplementary Methods.

Primary care

The BioFINDER-Primary Care study ([NCT06120361](#)), described previously⁶, is an ongoing study that recruits patients from primary care centers in southern Sweden. The study consecutively includes patients seeking medical help due to cognitive symptoms either self-reported or reported by a close relative or spouse. In addition, inclusion may also be initiated based on the PCP's suspicion of a neurocognitive disorder. To ensure a realistic and ethical study design aligned with standard care, the primary care cohort included only individuals for whom the treating PCPs, after the first assessment, considered a potential neurodegenerative disease. People whose symptoms were assessed to be clearly attributable to other causes (for example, depression, sleep disorders) were not enrolled and instead treated at the primary center or referred for appropriate care elsewhere. Patients already diagnosed with dementia or suffering from an unstable severe systemic illness as well as those with ongoing significant alcohol or substance abuse were excluded. At their respective primary care center, patients underwent standard clinical evaluation ('standard-of-care'), including medical examination by the PCP, cognitive testing, standard blood assessments to rule out other causes and structural brain imaging. Based on the assessments of PCPs, only 29.1% (106 of 364) of the participants would have been referred to secondary care if not included in this study, underscoring that the primary and secondary care cohorts represented different patient populations. For this study patients from 19 primary care centers were included. Further details of the cohort, cognitive testing and inclusion and exclusion criteria are provided in Supplementary Methods.

The digital cognitive battery (BioCog)

The digital cognitive battery, BioCog, was developed using Python and Kivy to run on an Android tablet. The battery was designed to include a primary part and an optional secondary part that contained more demanding tasks (and was therefore not used in this study). During development of the battery, it was tested by cognitively unimpaired volunteers to improve the user experience. Following this pilot phase, participants in the BioFINDER-Primary Care study completed the battery and provided feedback on how they perceived the tasks and instruction comprehensibility. After additional adjustments, digital testing was introduced in the study design and data collection started. None of the participants in this study had previous experience with the BioCog test battery. The first part of the self-administered test battery, which was used in the current study, had an average completion time of 11.2 min, and is comprised of a word list test with immediate and delayed recall as well as a delayed recognition part, a cognitive

processing speed task and questions about orientation to the weekday, date, month and year.

Comparison between the digital test battery and other cognitive tests

To evaluate the performance of the digital test to detect cognitive impairment, it was compared to paper-and-pencil tests frequently used in primary care settings (which require healthcare personal to perform the tests) and to a well-known, CE marked and US Food and Drug Administration approved digital test. The brief paper-and-pencil tests used, were: (1) the MMSE⁴⁰, the world's most used cognitive test, which consists of several shorter tasks and has a maximum score of 30 points; (2) MoCA⁴¹, a widely used screening tool with a maximum of 30 points, which evaluates a broader range of cognitive domains than the MMSE; (3), Mini-Cog⁴², which was developed as a very brief screening test that includes a three-word delayed recall task and clock drawing and has a maximum of 5 points; and (4) the digital, tablet-based paired associates learning task (PAL)⁴³ from the CANTAB, which measures visual associative memory. For the main analyses, externally established and validated cutoffs were used for all measures: MMSE <27 points⁴⁴, MoCA <26 points⁴¹, Mini-Cog <4 points⁴⁵ and for PAL >41 adjusted errors in total⁴⁶. For MoCA a two-cutoff approach was also included using <24 points as the lower cutoff and 26 as the upper cutoff²¹. For the other tests no validated two-cutoff approaches were available. As sensitivity analyses, the cognitive tests were treated as continuous variables and adjusted for demographics.

PCP estimation of cognitive stage and etiology

The standard clinical evaluation at the primary care units included brief cognitive testing (MMSE and MoCA), computed tomography scan of the brain, and a clinical assessment of the patient. After completing this evaluation, PCPs were asked to report the cognitive stage (SCD, MCI or dementia) and if assessed as MCI or dementia, the most likely cause of the cognitive impairment. The latter was binarized as AD or non-AD. Mixed etiologies including AD (for example, AD with vascular co-pathology) were coded as AD.

Plasma sampling and analysis

Blood was collected in EDTA tubes and centrifuged within 1 h. Plasma was then pipetted to 1-ml LoBind tubes and shipped the same day to be frozen at -80 °C. Plasma samples were then shipped to and analyzed by C2N Diagnostics in the United States continuously during the study period (47 occasions). Mass spectrometry assays were used to analyze Aβ42, Aβ40, phosphorylated-tau217 and nonphosphorylated-tau217, as previously described⁶. These biomarkers were combined in a logistic regression model (PrecivityAD2 test) to yield a score from 0 to 100 called the APS2 (ref. [47](#)). An APS2 cutoff value of 36 was used, previously defined in a clinical practice research study established at 90% specificity for AD pathology⁶. In addition, a two-cutoff approach with values corresponding to 95% sensitivity and 95% specificity (31 lower, 62 upper) was used, also previously defined⁶.

Outcomes

The primary outcome was objectively verified cognitive impairment. In the primary care cohort, this was established using the RBANS⁴⁸, administered by neuropsychologists at the memory clinic of Skåne University Hospital. The RBANS is a widely used cognitive test battery that has shown high accuracy in detecting and characterizing cognitive deficits associated with AD⁴⁹ and MCI in general⁵⁰. The test battery comprises 12 subtests assessing 5 cognitive domains: immediate and delayed memory, visuospatial and visuoconstructional ability, verbal ability and attention, and requires about 30 min to perform. Using normative data from the Swedish manual generated an age-corrected standard ($M = 100$, s.d. = 15) total scale score for each participant. Based on the widely accepted neuropsychological criteria for defining

cognitive impairment^{51,52}, participants with an RBANS total scale score of 1.5 s.d. below the normative mean or less were considered cognitively impaired. In the secondary care cohort, RBANS was not available. Instead, an RBANS proxy variable was created, based on similar paper-and-pencil neuropsychological tests (Supplementary Methods and Supplementary Table 18). The RBANS proxy was used as outcome in the secondary care cohort.

The secondary outcome 'clinical AD' was defined as having cognitive impairment (as described above) where AD was determined as the primary etiology. The latter was determined in consensus discussions including dementia specialists and neuropsychologists. The consensus discussions included access to medical examination by a dementia specialist at the memory clinic, a computed tomography or magnetic resonance imaging scan of the brain, CSF AD biomarkers or A β -PET results as well as neuropsychological assessment (Supplementary Methods). The diagnosis of clinical AD was defined in accordance with criteria from the International Working Group¹³, which includes a typical clinical presentation of AD in patients with objective cognitive impairment and confirmation with AD biomarkers. Participants fulfilling the criteria for clinical AD but also having signs of other concomitant diseases (for example, AD mixed with vascular disease) were labeled as having clinical AD in the analyses. Individuals that were negative according to this outcome were either cognitively unimpaired or cognitively impaired not due to AD (see Supplementary Table 16 for details of comorbidities and diagnoses).

The CDR scale is a semiobjective, qualitative staging instrument for the assessment dementia severity³². It assesses the cognitive and functional performance of everyday activities in six categories: memory, orientation, judgment and problem-solving, community affairs, home and hobbies, as well as personal affairs. Each item in the respective subscales is rated and a global CDR score is derived. A score of 0 is considered normal, with a score of 0.5, 1, 2 or 3 indicating questionable, mild, moderate and severe dementia³³. Participants with a CDR global score of 0.5 and above were considered cognitively impaired and this was used as an outcome in the sensitivity analysis. In this study, CDR was performed by a dementia specialist at the memory clinic blinded to BioCog.

Statistical analysis

The analyses were performed in Python v.3.9. Multiple logistic regression models were created, and cutoffs were established in the secondary care cohort. Model performance was then evaluated in the same cohort, and more importantly, in the external and completely independent primary care cohort. To decrease the risk of overfitting when creating multiple logistic regression models, the number of input variables was reduced with recursive feature elimination based on the standardized logistic regression model coefficients. The best model was selected using the AIC, requiring a decrease of 2 or more in AIC to indicate a better model fit and justifying the addition of another input variable^{53,54}. Model performance was evaluated from the ROC AUC. Cutoffs were established for the probabilities from the logistic regression models. An optimal cutoff point was selected based on the maximum value of Youden's index and the corresponding model was evaluated on accuracy, PPV, NPV, specificity and sensitivity. For the two-cutoff approach, cutoffs were selected based on 95% specificity and 95% sensitivity in the independent secondary care cohort. Participants in the intermediate group were not considered when evaluating model performance. The 95% CI and two-sided P values were computed using bootstrapping ($n = 5,000$ resamples with replacement), with a P value <0.05 indicating statistical significance. P values were adjusted for multiple comparisons by the Benjamini–Hochberg method. Differences between cognitively unimpaired and cognitively impaired on the various subtests of BioCog were assessed using the Mann–Whitney U-test. Spearman's rank correlation coefficients were used to evaluate associations between BioCog subtests and their paper-and-pencil

counterparts, as well as with age and education. Internal consistency was assessed using McDonald's ω .

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Anonymized data can be shared by request from qualified academic investigators for the purpose of replicating procedures and results presented in the article. Data transfer is required to be in agreement with EU legislation on the general data protection regulation and decisions by the Ethical Review Board of Sweden and Region Skåne, which should be regulated in a material transfer agreement.

Code availability

No custom code or mathematical algorithm that was central to the conclusions was used in this study.

References

40. Folstein, M. F., Folstein, S. E. & McHugh, P. R. 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
41. Nasreddine, Z. S. et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).
42. Borson, S., Scanlan, J., Brush, M., Vitaliano, P. & Dokmak, A. The Mini-Cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. *Int. J. Geriatr. Psychiatry* **15**, 1021–1027 (2000).
43. Sahakian, B. J. et al. A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain* **111**, 695–718 (1988).
44. O'Bryant, S. E. et al. Detecting dementia with the Mini-Mental State Examination in highly educated individuals. *Arch. Neurol.* **65**, 963–967 (2008).
45. McCarten, J. R. et al. Finding dementia in primary care: the results of a clinical demonstration project. *J. Am. Geriatr. Soc.* **60**, 210–217 (2012).
46. Junkkila, J., Oja, S., Laine, M. & Karrasch, M. Applicability of the CANTAB-PAL computerized memory test in identifying amnestic mild cognitive impairment and Alzheimer's disease. *Dement. Geriatr. Cogn. Disord.* **34**, 83–89 (2012).
47. Meyer, M. R. et al. Clinical validation of the PrelevityAD2™ blood test: a mass spectrometry-based test with algorithm combining %p-tau217 and Ab42/40 ratio to identify presence of brain amyloid. *Alzheimers Dement.* **20**, 3179–3192 (2024).
48. Randolph, C. *Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)* (Psychological Corporation, 1998).
49. Duff, K. et al. Utility of the RBANS in detecting cognitive impairment associated with Alzheimer's disease: sensitivity, specificity, and positive and negative predictive powers. *Arch. Clin. Neuropsychol.* **23**, 603–612 (2008).
50. Karantzoulis, S., Novitski, J., Gold, M. & Randolph, C. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): utility in detection and characterization of mild cognitive impairment due to Alzheimer's disease. *Arch. Clin. Neuropsychol.* **28**, 837–844 (2013).
51. Petersen, R. C. et al. Mild cognitive impairment: clinical characterization and outcome. *Arch. Neurol.* **56**, 303–308 (1999).
52. Jak, A. J. et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am. J. Geriatr. Psychiatry* **17**, 368–375 (2009).
53. Olofsen, E. & Dahan, A. Using Akaike's information theoretic criterion in mixed-effects modeling of pharmacokinetic data: a simulation study. *F1000Research* **2**, 71 (2014).

54. Burnham, K. P. & Anderson, D. R. Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304 (2004).

Acknowledgements

We thank the research volunteers who participated in the studies and their families. We also express our appreciation to H. Nordström, neuropsychologist at the Memory Clinic in Malmö, for her work with the collection of the data used in the study. Work at the authors' research center was supported by the National Institute of Aging (grant no. R01AG083740), European Research Council (grant no. ADG-101096455), Alzheimer's Association (grant nos. ZEN24-1069572 and SG-23-1061717), GHR Foundation, Michael J. Fox Foundation (grant no. MJFF-025507), Lilly Research Award Program, WASP and DDLS Joint call for research projects (grant no. WASP/DDLS22-066), Swedish Research Council (grant nos. 2021-02219, 2022-00775 and 2018-02052), ERA PerMed (grant no. ERAPERMED2021-184), Knut and Alice Wallenberg Foundation (grant no. 2022-0231), Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, Swedish Alzheimer Foundation (grant nos. AF-1011949, AF-994229 and AF-980907), Swedish Brain Foundation (grant nos. FO2023-0163, FO2024-0284 and FO2021-0293), Parkinson foundation of Sweden (grant no. 1412/22), Cure Alzheimer's fund, Rönström Family Foundation (grant nos. FRS-0003, FRS-0004, FRS-0011 and FRS-0013), Berg Family Foundation, Ingvar Kamprad Foundation (grant no. 20243058), Avid Pharmaceuticals, Bundy Academy, Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, Skåne University Hospital Foundation (grant no. 2020-O000028), Regionalt Forskningsstöd (grant nos. 2022-1259 and 2022-1346) and Swedish federal government under the ALF agreement (2022-Projekt0080, 2022-Projekt0107). The funding sources had no role in the design and conduct of the study; in the collection, analysis and interpretation of the data; or in the preparation, review or approval of the paper.

Author contributions

P.T., S.P. and O.H. conceived of and designed the study. P.T. and L.K. contributed equally as first authors. O.H. supervised and was

principal investigator for the study. L.K. performed statistical analyses and prepared the figures. P.T. provided statistical support. O.S. programmed BioCog and E.S. was responsible for logistics. All authors contributed substantially to the subsequent version of the paper and approved the final version for submission.

Funding

Open access funding provided by Lund University.

Competing interests

S.P. has acquired research support (for the institution) from Avid and ki elements through ADDF. In the past 2 years, he has received consultancy and/or speaker fees from Bioartic, Biogen, Esai, Eli Lilly, Novo Nordisk and Roche. R.S. has received consultancy and/or speaker fees from Lilly, Novo Nordisk, Roche and Triolab. N.M.-C. received consultancy and/or speaker fees from Biogen, Merck and Owkin. O.H. is an employee of Eli Lilly and Lund University. P.B.V. and J.B.B. are salaried employees for C2N Diagnostics and receiving compensation from the company in the form of a salary or equity. The other authors declare no competing interests.

Additional information

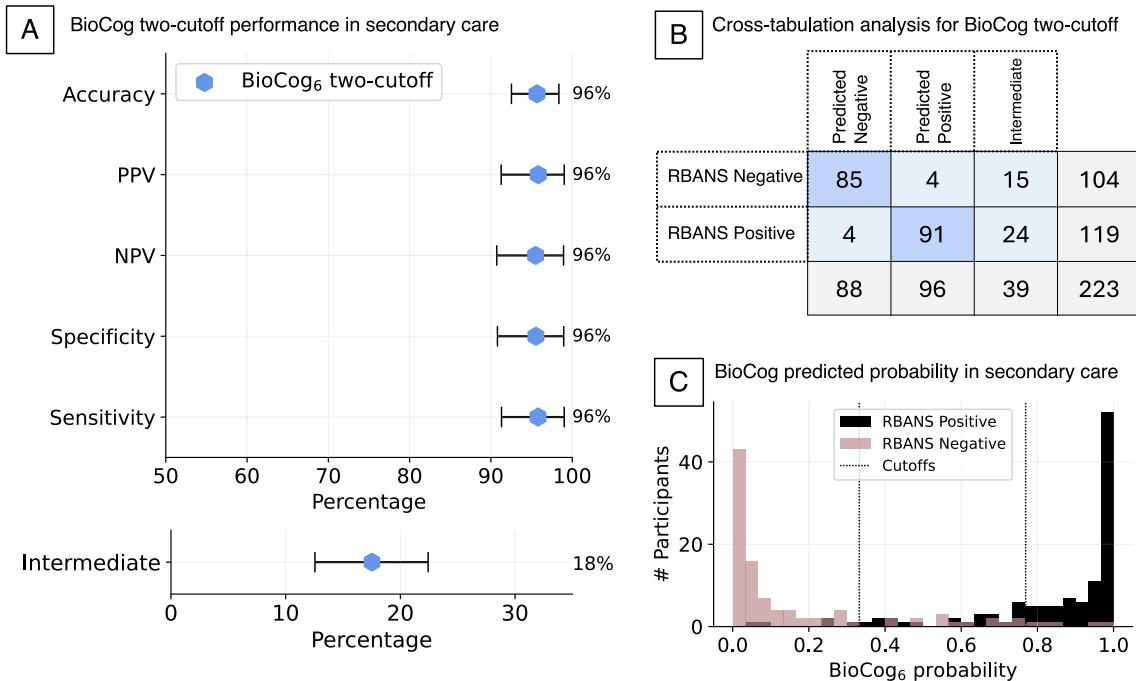
Extended data is available for this paper at <https://doi.org/10.1038/s41591-025-03965-4>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41591-025-03965-4>.

Correspondence and requests for materials should be addressed to Pontus Tideman or Oskar Hansson.

Peer review information *Nature Medicine* thanks Vijaya Kolachalam and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Jerome Staal in collaboration with the *Nature Medicine* team.

Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | Evaluating a two-cutoff approach for the BioCog₆ model in the secondary care cohort. BioCog₆ two-cutoff model performance in the secondary care cohort (which was also used to fit the model and to establish the cutoffs). The two cutoffs were set at probabilities 0.332 and 0.769, corresponding to 95% sensitivity and 95% specificity respectively. **a)** Evaluation

of the BioCog₆ two-cutoff model. Error bars indicate 95% CIs, with the centre point corresponding to the mean value. **b)** Cross-tabulation analysis the BioCog₆ two-cutoff model. **c)** Histogram of the predicted probabilities from the BioCog₆ model (x-axis) colored by the true binary RBANS proxy composite score (outcome variable).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection The BioCog Application for Andriod was used to collect data.

Data analysis Python version 3.9 (packages: sklearn 0.0, pandas 1.4.4, numpy 1.23.3, matplotlib 3.5.3, statsmodels 0.13.2) and SPSS version 27.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Anonymized data can be shared by request from qualified academic investigators for the purpose of replicating procedures and results presented in the article. Data transfer is required to be in agreement with EU legislation on the general data protection regulation and decisions by the Ethical Review Board of Sweden and Region Skåne.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

We used the term "sex" throughout the manuscript. Sex was determined based on self-reporting. Statistical analyses included sex as predictor or covariate; the study included two independent cohorts altogether comprised of balanced numbers of males (n=194 in the primary care cohort and 128 in the secondary care cohort) and females (n=209 in the primary care cohort and 95 in the secondary care cohort); therefore, we believe the findings apply to both sexes.

Reporting on race, ethnicity, or other socially relevant groupings

This study did not include categorization of race, ethnicity and/or other socially relevant groupings.

Population characteristics

We included 223 participants with cognitive symptoms from a secondary care cohort: the ongoing BioFINDER-2 study (NCT03174938). The mean (SD) age was 73 (8.8) years, 128 (57%) were male, and 119 (53%) had objectively verified cognitive impairment. Additionally, 403 participants were included from a primary care cohort, BioFINDER-Primary Care study (NCT06120361, Supplementary Fig. 1), which recruits patients seeking help for cognitive symptoms at 19 primary care units. Only patients for whom the primary care physician thought a neurodegenerative disease was reasonably possible to cause the symptomatology were offered to participate in the study. The mean (SD) age of the primary care cohort was 77 (8.0) years, 194 (48%) were male, and 229 (57%) had objectively verified cognitive impairment. Detailed information is given in Table 1.

Recruitment

Participants from two cohorts were included, one from secondary care and one from primary care. Data for this study were collected between February 2022 and December 2024.

Secondary care participants were recruited from the ongoing BioFINDER-2 study (NCT03174938). Briefly, the study consecutively enrolls participants at the secondary care Memory Clinic of Skåne University Hospital and the Memory Clinic of Ängelholm Hospital and includes a diverse population of study participants.

The BioFINDER-Primary Care study (NCT06120361), is an ongoing study that recruits patients from primary care centers in southern Sweden. The study consecutively includes patients seeking medical help due to cognitive symptoms either self-reported or reported by a close relative or spouse. Additionally, inclusion may also be initiated based on the primary care physician's suspicion of a neurocognitive disorder.

Ethics oversight

All participants provided written informed consent prior to enrollment in the study, and ethical approval for the study was obtained from the Swedish Ethical Review Authority.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

We included all available participants with complete data from the secondary care cohort (n=223) and the primary care cohort (n=403). Following the commonly cited rule of thumb for logistic regression analyses of a minimum of 10 events per predictor variable (Peduzzi, P. et al. 1996), our sample satisfied this criterion in all models.

Data exclusions

Data were limited to participants that had performed BioCog and had complete data for each analysis. See flowchart Supplementary Figure 1 for detailed exclusion after enrollment.

Replication

To develop a robust BioCog model and reduce the risk of overfitting, we applied recursive feature selection with AIC penalization and focused on external model validation in the independent primary care cohort. Performance metrics included bootstrapped confidence intervals to quantify uncertainty and support the reliability of the findings.

Randomization

In these two prospective cohort studies (observational studies) no allocation into experimental groups were performed, therefore randomization is not relevant to this study. Statistical analyses were controlled for potential confounding effects of age, sex and education.

Blinding

CSF, plasma and PET analyses were performed by individuals blinded to all clinical data. RBANS was performed by neuropsychologists and CDR was performed by dementia specialists, all at the memory clinic blinded to BioCog.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input checked="" type="checkbox"/>	Animals and other organisms
<input type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern
<input checked="" type="checkbox"/>	Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration BioFINDER-2 study: NCT03174938. BioFINDER-Primary Care study: NCT06120361.

Study protocol BioFINDER-2: <https://clinicaltrials.gov/ct2/show/NCT03174938>
BioFINDER-Primary Care study: <https://clinicaltrials.gov/ct2/show/NCT06120361>

Data collection Data for this study were collected between February 2022 and December 2024.

Secondary care participants were recruited from the ongoing BioFINDER-2 study (NCT03174938). Briefly, the study consecutively enrolls participants at the secondary care Memory Clinic of Skåne University Hospital and the Memory Clinic of Ängelholm Hospital and includes a diverse population of study participants.

The BioFINDER-Primary Care study (NCT06120361), is an ongoing study that recruits patients from primary care centers in southern Sweden. The study consecutively includes patients seeking medical help due to cognitive symptoms either self-reported or reported by a close relative or spouse. Additionally, inclusion may also be initiated based on the primary care physician's suspicion of a neurocognitive disorder

Outcomes The primary outcome was objectively verified cognitive impairment, established using RBANS (primary care) or an RBANS proxy variable (secondary care).
The secondary outcome "clinical AD" was defined as having cognitive impairment where AD was determined as the primary etiology.
In a sensitivity analysis, CDR was used to define objectively verified cognitive impairment.

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.