



OPEN Stepwise approach to alzheimer's disease diagnosis in primary care using cognitive screening, risk factors, neuroimaging and plasma biomarkers

Miren Altuna^{1,2,3,4}✉, Maite García-Sebastián¹, Raffaella Cipriani⁵, Estibaliz Capetillo-Zarate^{5,6,7}, Elena Alberdi^{5,6}, Ainara Estanga¹, Mirian Ecay-Torres¹, Ane Iriondo¹, Jon Saldias¹, Marta Cañada¹, Carolina López¹, Maria Arriba¹, Mikel Tainta⁸ & Pablo Martínez-Lage^{1,9}✉

Early identification of Alzheimer's disease (AD) pathology is essential for timely intervention, particularly in primary care. We evaluated the diagnostic performance of a scalable, multimodal framework in a real-world, population-based cohort. A total of 277 community-dwelling individuals aged ≥ 60 years from the STOP-ALZHEIMER DEBA study (Basque Country, Spain) underwent brief cognitive screening (MMSE, M@T, Fototest, AD8) with optimized cut-offs, along with clinical risk assessment. Among them, 181 participants also completed structural MRI, plasma biomarker profiling (p-tau181, A β 42/40, GFAP, NfL), and cerebrospinal fluid (CSF) analysis. We assessed performance for detecting cognitive impairment, CSF amyloid positivity (A+), and combined amyloid-tau positivity (A + T+). Optimized cognitive tests showed moderate accuracy (AUC 0.66–0.77), with the Fototest performing best. For biological outcomes, GFAP and p-tau181 had the highest predictive value (AUCs: 0.813 and 0.755 for A+; 0.852 and 0.710 for A + T+), and their combination further improved accuracy (AUC = 0.842). Fully adjusted models incorporating optimized cognitive scores, plasma biomarkers, APOE genotype, MRI, and demographics achieved high diagnostic performance (AUC = 0.886 for A+; 0.893 for A + T+). Results were consistent across sex and age strata. These findings support a stepwise diagnostic strategy combining brief, minimally invasive tools to enhance early AD detection in community settings.

Keywords Alzheimer's disease, Plasma biomarkers, Cognitive screening, Primary care, ATN framework, Mild cognitive impairment

Early detection of cognitive impairment, particularly Alzheimer's disease (AD), is a cornerstone of public health strategy in aging societies^{1,2}. In 2022, over 909,000 people in Spain were living with dementia, with projections exceeding 1.7 million by 2050³. Globally, dementia-related costs have surpassed \$1 trillion annually, expected to double by 2030³. AD accounts for 60–70% of dementia cases and remains the leading cause of progressive

¹Center for Research and Memory Clinic, CITA-Alzheimer Foundation, Donostia-San Sebastián 20009, Spain.

²Debabarrena Integrated Health Organization, Osakidetza Basque Health Service, Mendaro 20690, Spain.

³Department of Neurology, Bioaraba Health Research Institute, Araba University Hospital-Txagorritxu, Vitoria-Gasteiz 01009, Spain. ⁴Department of Medicine, Faculty of Health Sciences, University of Deusto, Bilbo 48007, Spain. ⁵Achucarro Basque Center for Neuroscience Foundation, Leioa, Bizkaia 48940, Spain. ⁶Department of Neuroscience, University of the Basque Country UPV/EHU, Leioa, Bizkaia 48940, Spain. ⁷Ikerbasque Foundation, Bilbao, Bizkaia 48009, Spain. ⁸Neurology Department, Osakidetza Basque Health Service, Donostialdea Integrated Health Organization, Donostia-San Sebastián 20014, Spain. ⁹Network Center for Biomedical Research in Neurodegenerative Diseases, CIBERNED, National Institute of Health Carlos III, Madrid 28220, Spain. ✉email: maltuna@cita-alzheimer.org; pmlage@cita-alzheimer.org

cognitive decline⁴. Nevertheless, vascular, frontotemporal, and Lewy body dementias are also prevalent and often coexist, complicating diagnosis.

Primary care represents the first—and often only—point of contact for individuals with cognitive concerns, yet current diagnostic tools are constrained by accessibility, diagnostic sensitivity, and demographic biases. Widely used brief cognitive screening tools such as the Mini-Mental State Examination (MMSE)⁵, Fototest^{6,7}, Ascertain Dementia 8-item Questionnaire (AD8 questionnaire)⁸, and Memory Alteration Test (M@T)⁹ are practical and low-cost, but often fail to detect early or atypical impairment. Their performance is also affected by sociodemographic variables such as age, sex, education, and cardiovascular risk profile. The Fototest, in particular, has demonstrated utility in populations with low educational attainment and in bilingual contexts that include a minority language—such as Basque—where validated cognitive screening tools are scarce.

Recent advances in blood-based biomarkers (BBMs), including plasma phosphorylated tau (p-tau181), the Aβ42/40 ratio, glial fibrillary acidic protein (GFAP), and neurofilament light (NfL), measured via Single Molecule Array (SIMOA) platforms, offer a minimally invasive and scalable alternative to cerebrospinal fluid (CSF) or positron emission tomography (PET). Furthermore, newer markers such as p-tau217 have shown exceptional diagnostic accuracy in recent studies^{10,11}. However, their current implementation is limited by availability, cost, and standardization in population-level settings. Our study instead focuses on how clinically validated, mature biomarkers—that is, biomarkers such as CSF Aβ42, total tau, and p-tau181 that are well-established, widely used in research and clinical trials, and supported by large-scale validation—can be combined with other accessible tools to develop scalable diagnostic strategies.

Mild cognitive impairment (MCI) is a particularly relevant target for early detection, as it often represents the first syndromic expression of neurodegenerative disease. Defined by measurable cognitive decline without loss of functional independence, MCI bridges normal aging and dementia, and is increasingly recognized as a critical window for therapeutic intervention. Distinguishing MCI from dementia—defined by progressive loss of autonomy due to cognitive decline—is essential for prognosis, care planning, and eligibility for emerging disease-modifying treatments.

A sequential diagnostic strategy—beginning with brief cognitive screening, incorporating clinical risk scores like CAIDE¹², and escalating to neuroimaging and plasma biomarkers when needed—may help close the gap between clinical syndromes and biological definition of AD. Optimized cut-offs for cognitive tests, as recently proposed by Tainta et al.¹³, may improve accuracy and mitigate sociodemographic bias. Magnetic Resonance Imaging (MRI) markers of medial temporal atrophy and vascular burden can refine diagnostic accuracy, particularly in mixed pathology. The integration of APOE ε4 genotype, sex, education, and comorbidities supports a precision medicine approach aligned with population health realities.

In this study, we aimed to evaluate the diagnostic performance of a stepwise model combining plasma biomarkers (p-tau181, Aβ42/40, GFAP, NfL), brief cognitive screening tools (using both traditional and optimized cut-offs), and MRI variables, both individually and in combination. We further examined performance across stratified subgroups (by sex, education, CAIDE risk, APOE ε4 status, and vascular comorbidity), to propose a scalable and personalized framework for the early detection of cognitive impairment and biologically defined Alzheimer's disease in primary care settings.

Results

Characteristics of the study cohort

Of the 1,568 residents aged ≥60 years in Deba (Basque Country, Spain), 678 individuals participated in the community-based cognitive screening phase of the STOP-ALZHEIMER study. Of these, 277 underwent full clinical and biomarker evaluation, including cognitive testing; lumbar puncture, blood analysis, and brain MRI were performed in 181 participants (65.3%) (Fig. 1).

CSF = cerebrospinal fluid. MRI = magnetic resonance imaging.

The mean age of participants was 70.73 years (SD = 8.3), 50.9% were women, and the mean years of education was 9.32 (SD = 4.1). Regarding clinical syndromic diagnosis, 67.1% were cognitively normal (CN), 6.9% presented with subjective cognitive decline (SCD), 17% had mild cognitive impairment (MCI), and 9% had dementia. To assess potential selection bias and ensure the representativeness of the subsample with CSF data, we compared demographic, clinical, and cognitive characteristics between participants who underwent lumbar puncture and those who did not, given that willingness to undergo cerebrospinal fluid (CSF) collection may differ across clinical profiles in population-based studies. Those with CSF were slightly older and more educated and had lower systolic and diastolic blood pressure, cholesterol, levels. In turn, slightly better performances were observed in cognitive scores (MMSE, M@T, Fototest, AD8) in those with available CSF samples, but no differences were detected in relation to CAIDE dementia risk score. Only 10.9% of those with CSF were APOE E4 carrier, no available information about those without CSF. Notably, the distribution of syndromic diagnoses also differed between groups, with a higher proportion of cognitively unimpaired individuals among those with CSF data (70.7% vs. 60.4%) and a markedly higher proportion of dementia cases among those without CSF data (19.8% vs. 3.3%) (Table 1).

Diagnostic accuracy of brief cognitive tests for detecting cognitive impairment

We evaluated the discriminative accuracy of brief cognitive screening instruments using previously validated optimized cut-offs. The Fototest demonstrated the highest discriminative ability for detecting cognitive impairment (AUC = 0.767), followed by the AD8 (AUC = 0.735), M@T (AUC = 0.711), and MMSE (AUC = 0.668). The composite binary variable (BCTo), defined as positive if any individual test exceeded its threshold, yielded an AUC of 0.694 (Fig. 2).

These findings suggest that combining test results in the BCTo index does not markedly improve overall discriminative performance compared to the best-performing single tests such as the Fototest and AD8.

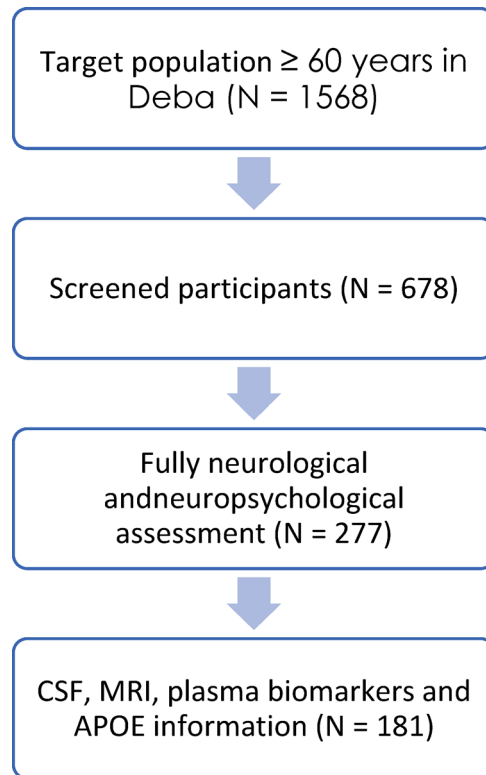


Fig. 1. Flowchart of participant inclusion and evaluation steps in the STOP-ALZHEIMER project.

	With CSF samples (N = 181)	Without CSF samples (N = 96)	
Age (mean ± SD)	74.49 ± 9.19 years	68.73 ± 7.06 years	$p < 0.001^*$
Sex (female, N, %)	92 (50.8%)	49 (51%)	$p = 0.973$
Education years	10.01 ± 3.90	8.04 ± 4.13	$p = 0.432$
CAIDE score	7.54 ± 1.99	8.01 ± 2.08	$p = 0.947$
Syndromic diagnoses (Number of subjects, % inside category)	CU 128 (70.7%) SCD 16 (8.8%) MCI 31 (17.1%) DEM 6 (3.3%)	CU 58 (60.4%) SCD 3 (3.1%) MCI 16 (16.7%) DEM 19 (19.8%)	$p < 0.001^*$
SBP (mean ± SD)	147.11 ± 20.21 mmHg	150.28 ± 22.26 mmHg	$p = 0.512$
DBP (mean ± SD)	82.36 ± 8.77 mmHg	82.64 ± 12.24 mmHg	$p = 0.095$
MBP (mean ± SD)	108.08 ± 37.09 mmHg	106.77 ± 14.26 mmHg	$p = 0.616$
Cholesterol total (mean ± SD)	207.75 ± 36.96 mg/dL	211.98 ± 38.39	$p = 0.762$
Glucose (mean ± SD)	106.75 ± 29.34 mg/dL	106.33 ± 19.28 mg/dL	$p = 0.228$
BMI (mean ± SD)	27.49 ± 4.68 kg/m ²	26.91 ± 3.98	$p = 0.954$
AD8 (mean ± SD)	0.82 ± 1.50	2.09 ± 2.69	$p < 0.001^*$
Fototest (mean ± SD)	37.33 ± 5.61	33.45 ± 7.82	$p = 0.006^*$
MMSE (mean ± SD)	27.13 ± 2.82	25.99 ± 4.19	$p < 0.001^*$
M@T (mean ± SD)	43 ± 6.69	38.98 ± 10.05	$p < 0.001^*$
Pfeiffer (mean ± SD)	1.03 ± 4.04	1.67 ± 5.51	$p = 0.024^*$
HADS anxiety score (mean ± SD)	5.15 ± 3.07	4.41 ± 3.48	$p = 0.013^*$
HADS depression score (mean ± SD)	3.03 ± 2.83	3.21 ± 2.95	$p = 0.465$
NPI total score (mean ± SD)	2.74 ± 7.39	3.52 ± 9.62	$p = 0.162$

Table 1. Comparison of demographic, cognitive, and clinical variables between participants with and without CSF biomarker data. CU = Cognitively Unimpaired. SCD = Subjective cognitive decline. MCI = Mild Cognitive Impairment. DEM = Dementia. SBP = systolic blood pressure. DBP = diastolic blood pressure. MBP = medium blood pressure. BMI = body mass index. MMSE = Minimental Test. M@T = Memory alteration test. HADS = Hospital Anxiety and Depression Scale. NPI = Cummings neuropsychiatric inventory. * indicates significant difference ($p < 0.05$).

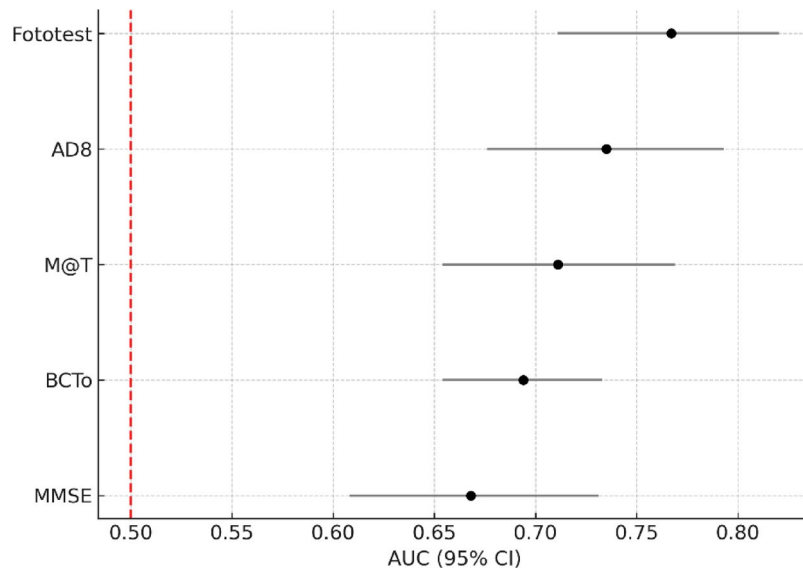


Fig. 2. Forest plot showing the area under the curve (AUC) and 95% confidence intervals for each optimized cognitive screening test in detecting cognitive impairment.

MMSE = Minimental Test. M@T = Memory alteration test. BCTo = brief cognitive tests with optimized cut-offs. Red dashed line indicates the no-discrimination threshold (AUC = 0.5).

Subgroup analyses revealed differential test performance across sex, age, and education levels. AUCs were generally higher among women than men across most cognitive tests, although no statistically significant differences were observed (Figure S1). A modest decline in test performance was observed with increasing age, with individuals aged 60–69 achieving the highest AUCs, though differences across age groups were less pronounced (Figure S2). Educational attainment strongly influenced diagnostic accuracy, with substantially lower AUCs for participants with ≤ 6 years of education, especially for the MMSE and M@T, where differences reached statistical significance (Figure S3).

Detection of alzheimer's continuum with cognitive screening

We assessed the ability of brief cognitive tests to detect biological AD, operationalized as CSF amyloid positivity (A+) and combined amyloid and tau positivity (A + T+). Among the individual instruments, the Fototest yielded the highest AUC for detecting A+ (0.710) and A + T+ (0.709), followed closely by the AD8 (AUC = 0.684 and 0.695, respectively). M@T and MMSE achieved moderate discriminative accuracy, while the composite index BCTo showed intermediate values (AUC = 0.683 for A+, and 0.652 for A + T+) (Fig. 3).

A+: Amyloid positivity assessed by CSF (ptau/AB42 ratio). T: Tau positivity assessed by CSF (ptau). MMSE: Mini Mental Test. M@T: Memory Alteration Test. BCTo = Brief cognitive test optimized cut-offs.

These results suggest that single tools such as the Fototest and AD8 may be at least as effective—or even superior—to composite screening for detecting early stages of AD pathology.

In the subgroup with SCD, 25.0% were A+, and 5.3% were A + T+, compared to 14.8% and 3.8% among cognitively normal individuals, respectively suggesting a possible trend toward increased pathological burden in individuals with subjective complaints, although this should be interpreted cautiously due to the small sample size (Figure S4).

Plasma biomarkers for predicting CSF-defined amyloid and Tau positivity

Plasma biomarkers p-tau181, GFAP, and NfL showed a direct relationship with CSF-defined amyloid and tau positivity (A + T+): higher concentrations were associated with increased probability of pathology. In contrast, the A β 42/40 ratio demonstrated an inverse relationship, with lower values corresponding to greater likelihood of A + T+; hence, its inverse was used in all analyses. GFAP showed the highest individual discriminative capacity for A + T+ (AUC = 0.852). While GFAP and p-tau181 combined yielded high discriminative performance (AUC = 0.842), the addition of the A β 42/40 ratio did not improve the model (AUC = 0.830), suggesting limited added value in this context (Fig. 4).

Red dashed line indicates the no-discrimination threshold (AUC = 0.5).

Optimal cut-off points for plasma biomarkers

Optimal cut-off points for predicting CSF-defined amyloid and tau positivity (A + T+) were derived using the Youden index. For GFAP, a threshold of 115.98 pg/mL provided the best diagnostic trade-off (sensitivity: 0.92, specificity: 0.70). p-tau181 performed more conservatively, with a cut-off of 24.76 pg/mL yielding a balanced profile (sensitivity: 0.62, specificity: 0.81). The A β 42/40 ratio (used in inverse form) had lower specificity (0.56) despite acceptable sensitivity (0.77). NfL showed limited utility (cut-off: 20.37 pg/mL, sensitivity: 0.54, specificity: 0.76) (Table 2).

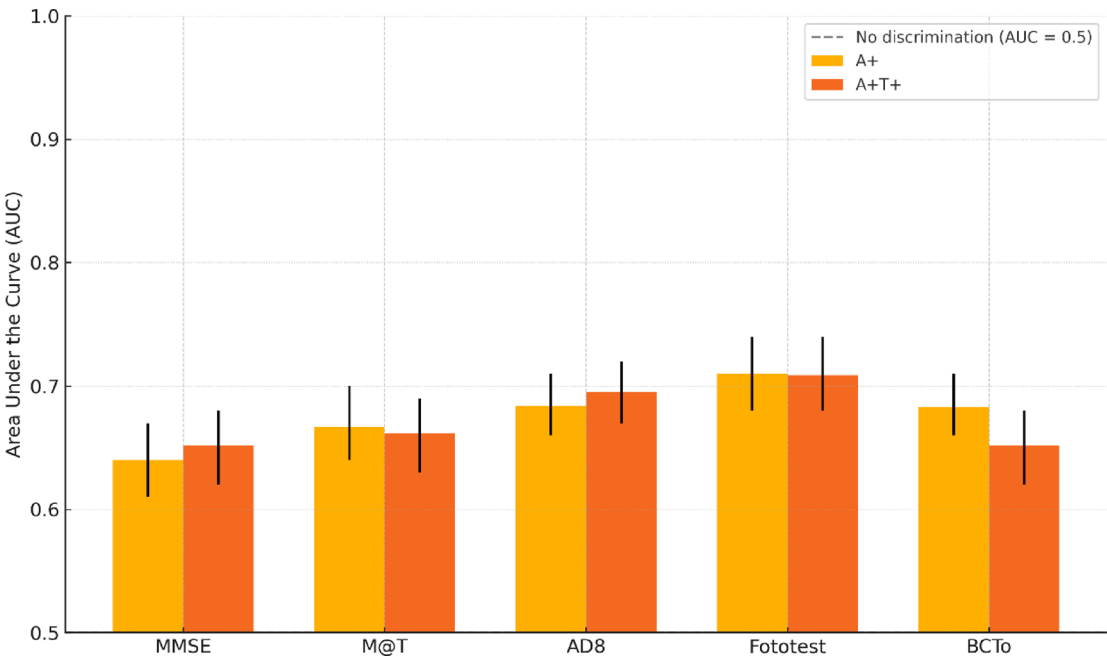


Fig. 3. AUC values for brief cognitive screening tests to detect CSF-defined amyloid positivity (A+) and combined amyloid and tau positivity (A + T+).

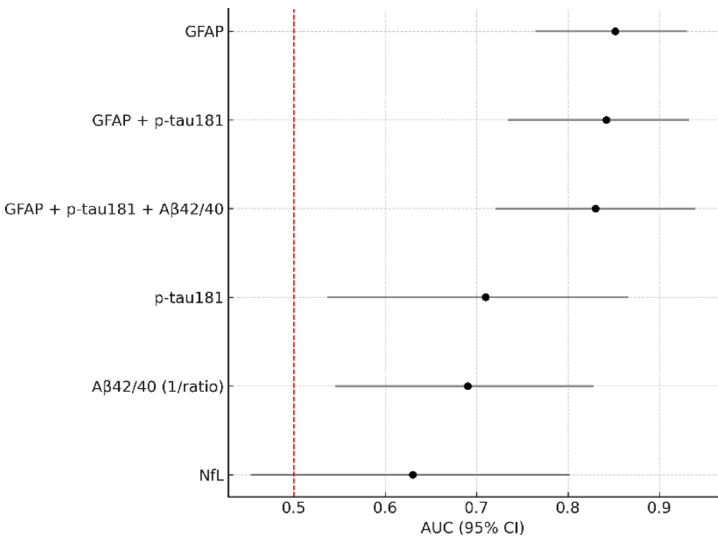


Fig. 4. Forest plot showing area under the curve (AUC) and 95% confidence intervals for individual and combined plasma biomarker models in predicting CSF-defined amyloid and tau positivity (A + T+).

	CSF A + T+				CSF A+			
	Cut-off	Sensitivity	Specificity	AUC	Cut-off	Sensitivity	Specificity	AUC
p-tau181 (pg/mL)	≥ 24.76	0.62	0.81	0.710	≥ 22.78	0.66	0.80	0.755
GFAP (pg/mL)	≥ 115.98	0.92	0.70	0.852	≥ 115.98	0.80	0.77	0.813
NfL (pg/mL)	≥ 20.37	0.54	0.76	0.630	≥ 21.76	0.46	0.86	0.663
Aβ42/40 ratio	≤ 0.0544	0.77	0.56	0.690	≤ 0.0558	0.80	0.57	0.737

Table 2. Youden-optimal cut-offs for plasma biomarkers with sensitivity and specificity.

Outcome	AUC
MCI + A+	0.934
CSF A+	0.886
MCI	0.826
CSF A + T+	0.893

Table 3. Area under the curve (AUC) for fully adjusted multimodal models across clinical, biological, and combined outcomes.

Assessment Level	AUC	N
Interview level (Level 1)	0.779	239
Interview + MRI (Level 2)	0.868	239
Interview + MRI + Biomarkers (Level 3)	0.938	163

Table 4. Area under the curve (AUC) and number of participants included (N) for each assessment level in the prediction of mild cognitive impairment with CSF amyloid positivity (MCI + A+).

When focusing exclusively on CSF amyloid positivity (A+), GFAP and Aβ42/40 ratio both achieved high sensitivity (0.80), though GFAP showed superior specificity (0.77 vs. 0.57). p-tau181 offered a more balanced profile (sensitivity: 0.66, specificity: 0.80), while NfL had limited utility (sensitivity: 0.46, specificity: 0.86) (Table 2). These values support GFAP as the most broadly effective individual biomarker for amyloid detection.

Multimodal prediction of MCI and alzheimer’s disease pathology

We evaluated the performance of fully adjusted multimodal models integrating brief cognitive screening (BCTo), plasma biomarkers (p-tau181, GFAP, NfL, and Aβ42/40 ratio inverted), APOE ε4 status, demographic variables (age, sex), and structural MRI features (medial temporal atrophy ≥ 2 and vascular burden). The highest diagnostic accuracy was observed for the combined clinical and biological model predicting MCI with underlying amyloid positivity (MCI + A+), which reached an AUC of 0.934. Models predicting isolated biological endpoints such as CSF-defined A + T + or A + also achieved excellent performance (AUC = 0.893 and 0.886, respectively), while the model for syndromic MCI alone showed good accuracy (AUC = 0.826) (Table 3).

Diagnostic yield of assessment levels

We evaluated the diagnostic performance of sequential assessment levels to predict MCI with underlying amyloid pathology (MCI + A+). The first level, based on clinical interview and brief cognitive screening (BCTo), CAIDE score, and neuropsychiatric symptoms (total Neuropsychiatric Inventory (NPI) score, calculated as the sum of frequency × severity across all domains), achieved an AUC of 0.779. Adding MRI-derived features (medial temporal atrophy (MTA, dichotomized as ≥ 2) and white matter hyperintensities using the Fazekas scale (dichotomized as ≥ 2)) increased the AUC to 0.868. Incorporating plasma biomarkers (p-tau181, GFAP, NfL, Aβ42/40) further improved diagnostic accuracy to an AUC of 0.938. Biomarkers were included as binary variables using exploratory cut-offs derived from Youden’s index. APOE ε4 status was entered as a binary variable, and all models were adjusted for age and sex. These results support the value of multimodal assessment in identifying individuals with early biological AD (Table 4; Fig. 5).

Level 1 includes BCTo (brief cognitive test with optimized cut-offs), CAIDE, and neuropsychiatric symptoms; Level 2 adds structural MRI (MTA -medial temporal atrophy- and vascular burden); and Level 3 incorporates plasma biomarkers, APOE ε4 status, and demographic variables. Diagnostic accuracy improves with each level.

To statistically validate these stepwise improvements, DeLong tests for correlated ROC curves were performed using the subset of participants with complete data across all three levels (N = 163). The increases in AUC from Level 1 to Level 2 (p = 0.01) and from Level 2 to Level 3 (p = 0.03) were both statistically significant. These findings support the additive diagnostic value of MRI and plasma biomarkers beyond interview-based tools alone.

Stratified analyses by subgroup

To explore diagnostic consistency across demographic and clinical strata, we applied the full multimodal model for predicting MCI + A + to subgroups defined by age, sex, and CAIDE dementia risk. Diagnostic accuracy was high in women and in the medium CAIDE group. In contrast, estimates in men, individuals aged < 75 years, and those in the low or high CAIDE categories were limited by very small numbers of MCI + A + cases and should therefore be interpreted with caution, highlighting the need for replication in larger samples. Detailed subgroup performance is presented in Supplementary Table S1.

Discussion

In this population-based study, we evaluated a stepwise diagnostic framework integrating brief cognitive screening, clinical variables, MRI features, and blood-based biomarkers for detecting cognitive impairment and AD pathology in a real-world primary care setting. The study cohort—drawn from the STOP-ALZHEIMER

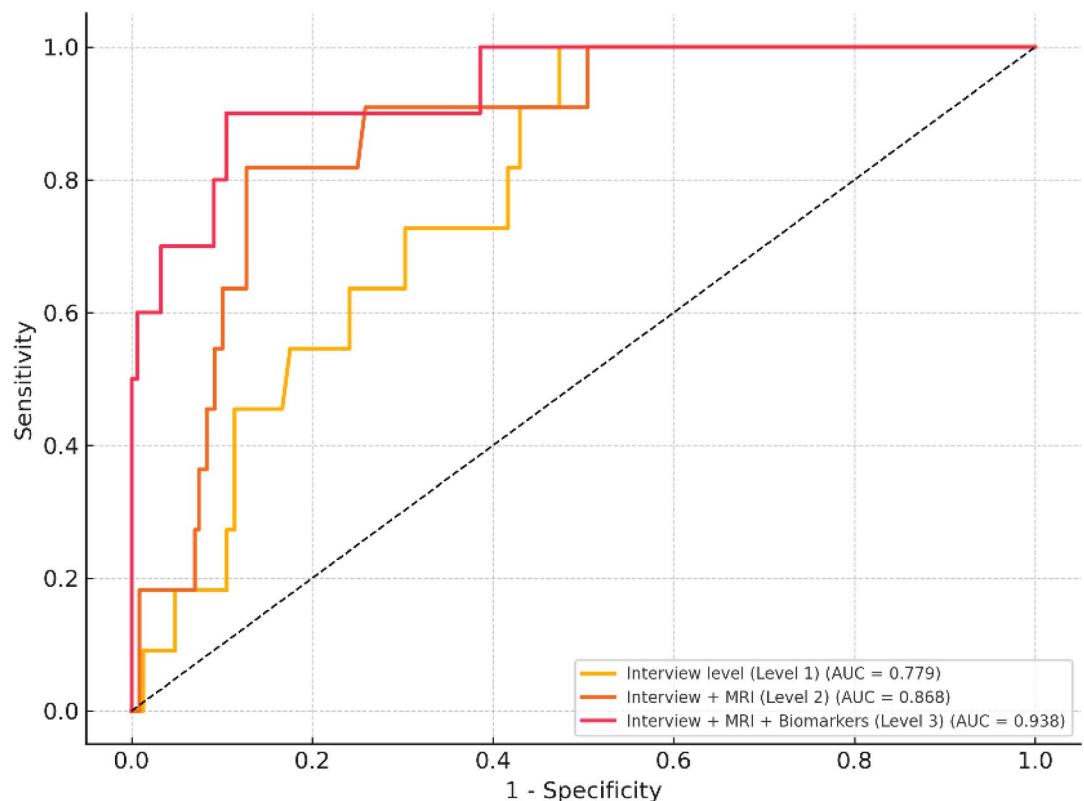


Fig. 5. ROC curves for sequential assessment levels predicting MCI + A+.

DEBA initiative—was representative of the general population aged ≥ 60 years, avoiding referral bias typical of memory clinics and ensuring the applicability of results to routine care.

Our findings support the use of optimized brief cognitive tests such as the Fototest and AD8, which demonstrated good discriminative performance for cognitive impairment (AUCs 0.73–0.77). As emphasized by Tainta et al.¹³, updated thresholds adapted to contemporary educational and demographic characteristics improve sensitivity and specificity in screening. The Fototest, in particular, proved especially valuable in our bilingual cohort, which includes a minority language (Basque) with very limited availability of validated cognitive screening tools. While combining multiple cognitive tests did not yield substantial improvement over the best individual tools, using optimized cut-offs tailored to educational level and sex added value in stratified performance.

Notably, the educational profile of our cohort—composed of individuals born primarily in the 1940s, 1950 and 1960s—reflects the structure of the Spanish education system at the time, when compulsory schooling typically ended at age 14. This translates to a predominant educational attainment of 6–8 years, consistent with “primary education” as defined in national statistics. Given the rural setting of the sample and known regional disparities in access to education, the overall low-to-moderate schooling level observed in our participants is representative of older adults in Spain, particularly outside large urban centers. Crucially, this contrasts with many previously published cohorts in the AD literature, which often originate from specialized urban memory clinics or research centers. These referral-based cohorts typically recruit individuals with higher educational attainment and distinct socioeconomic profiles, potentially limiting the generalizability of their findings to the broader population. Our cohort, drawn from a real-world primary care setting in a rural context, provides a more representative sample of the general older adult population. This demographic characteristic underscores the utility of the Fototest, a screening tool specifically validated for its reduced educational bias and applicability in low-literacy populations, thereby mitigating potential misclassification often observed with traditional cognitive assessments in such cohorts^{6,7}.

Among plasma biomarkers, GFAP consistently outperformed other markers in predicting both CSF amyloid (A+) and combined amyloid and tau (A + T+) positivity^{14,15}. The inverse of the A β 42/40 ratio was also predictive, but less specific^{16,17}. Multimodal models combining GFAP and p-tau181 reached an AUC of 0.842 for A + T+ detection, with no further improvement when A β 42/40 or NfL were added¹⁸. These results align with previous reports but emphasize the practical utility of available plasma markers in primary care settings where p-tau217 is not yet accessible. While recent studies have demonstrated superior performance of plasma p-tau217^{10,11,19}, our findings highlight that other combinations remain effective in pragmatic environments. It is worth noting that the recent FDA clearance of the Lumipulse G pTau217/A β 42 plasma ratio (in 2025)²⁰ marks a significant step toward the clinical implementation of blood-based biomarkers, underscoring the rapidly evolving nature of this diagnostic landscape.

Regarding their performance across dementia severity, literature suggests that GFAP, an astrogliosis marker, often shows a robust association with amyloid pathology (A+) and general disease progression from preclinical stages²¹. Its levels may reflect the ongoing neuroinflammatory response throughout the disease course. Conversely, p-tau181, a marker of tau pathology and neuronal injury, is highly specific for AD and tends to increase as cognitive impairment progresses from MCI to dementia, more closely reflecting the accumulation of neurofibrillary tangles²². While both are highly accurate, their dynamic range and diagnostic utility can vary across the AD continuum²³, with GFAP potentially offering earlier detection of amyloid burden and p-tau181 more strongly correlating with later-stage neuronal degeneration and clinical decline^{24,25}. The synergistic effect observed in our combined models suggests these markers capture complementary facets of AD pathophysiology, enhancing diagnostic precision.

Although several studies have proposed plasma biomarker thresholds—particularly for p-tau181, p-tau217, and the A β 42/40 ratio—there are currently no universally accepted or clinically approved diagnostic cut-offs for SIMOA-based assays. Cut-offs reported in the literature often vary across cohorts, assay versions, and population characteristics (e.g., age, comorbidities, ethnicity), limiting their generalizability. Our decision to derive cut-offs internally using Youden's index reflects this lack of consensus and the absence of reference standards for real-world, population-based cohorts. These thresholds should therefore be considered exploratory and context-specific, and external validation will be required before broader clinical adoption. This limitation is consistent with broader calls for harmonization and standardization in biomarker interpretation²⁶.

MRI-derived markers such as medial temporal atrophy and vascular burden also improved classification, particularly in hybrid models combining syndromic and biological criteria. This underlines the relevance of imaging in detecting mixed pathologies, a common scenario in aging populations^{27,28}.

Fully adjusted multimodal models integrating clinical, cognitive, MRI, and plasma variables achieved excellent performance in detecting both biological and syndromic outcomes. The model predicting CSF A + T+ reached an AUC of 0.893, while the model for MCI + A+ achieved 0.934. These results indicate that plasma biomarkers and MRI features provide clear incremental diagnostic value beyond interview-level tools, and can significantly enhance early detection strategies. This diagnostic focus was intentionally placed on individuals with MCI and biological evidence of AD (A+ or A+T+), as they represent the most actionable stage for early intervention. Dementia cases were less prevalent in our CSF-validated subsample, limiting statistical power for robust modeling at later stages of disease.

From a methodological standpoint, our stratified modeling approach revealed that neuropsychiatric symptoms and vascular comorbidities (e.g., hypertension, diabetes) were more strongly associated with clinical diagnoses than with biological positivity²⁹, reinforcing the need for multi-layered assessment strategies that differentiate symptomatic burden from underlying pathology. The CAIDE score proved valuable for stratifying long-term vascular risk but did not significantly enhance models predicting CSF amyloid or tau positivity once molecular and imaging data were integrated. These findings align with previous reports highlighting CAIDE's prognostic rather than diagnostic utility³⁰.

We also simulated sequential diagnostic scenarios, finding that accuracy improved with each level of integration: from 0.779 for interview-based data alone to 0.938 with full inclusion of plasma biomarkers and MRI. This supports a scalable, tiered diagnostic approach, adaptable to resource availability and clinical suspicion.

Stratified analyses showed stable performance across sex and education strata. However, some models—particularly those involving the oldest participants (≥ 75 years) or high CAIDE risk—suffered from limited sample size and must be interpreted cautiously. In these subgroups, apparent high AUCs likely reflect overfitting rather than true generalizability.

While digital biomarkers and AI-based tools are under development³¹, their clinical deployment is currently constrained by digital literacy gaps and lack of standardized thresholds. Thus, enhancing the diagnostic value of currently validated brief cognitive tests—particularly when adapted for the language and educational background of the population—remains an essential priority. In this regard, the STOP-ALZHEIMER DEBA study made a specific effort to adapt and translate the screening questionnaires into Basque, the co-official language of the region alongside Spanish. This allowed participants to choose their preferred language—Basque or Spanish—for completing the interviews and cognitive screening tests.

Looking forward, digital cognitive assessments and AI-supported diagnostics offer promise, but current barriers—including digital exclusion and lack of standardization—necessitate ongoing investment in optimizing existing tools³². Validated cognitive tests tailored to linguistic and cultural contexts remain foundational.

However, this study has several limitations. First, the cross-sectional design prevents conclusions about longitudinal disease progression or conversion to dementia. Second, although population-based, the sample was exclusively Caucasian and relatively homogeneous. Third, the number of CSF-positive cases—particularly for A+T+ and MCI+A+—was limited, leading to small cell sizes in subgroup analyses and inflated AUCs in some models. Fourth, while our panel included clinically available markers (e.g., p-tau181, GFAP), we did not include newer candidates such as p-tau217, which may yield superior performance. Finally, external validation is required to confirm the stability and reproducibility of these models.

Despite these limitations, our findings advocate for a sequential, stratified diagnostic model tailored to primary care that incorporates brief cognitive screening, clinical comorbidity data, structural imaging, and blood-based biomarkers. This layered approach aligns with current movements toward early detection and disease-modifying therapies in AD and supports the role of personalized medicine based on individual risk profiles (sex, education, comorbidities, APOE $\epsilon 4$ status). In turn, minimizes unnecessary referrals, and empowers healthcare systems to identify patients eligible for advanced diagnostic and therapeutic interventions. Given the growing availability of disease-modifying therapies, including FDA and EMA-approved anti-amyloid drugs like Lecanemab³³ and Donanemab has been recently approved also by EMA³⁴, compels health systems to optimize diagnostic

pathways. APOE genotyping is poised to become a critical component in treatment selection algorithms, and its integration into multimodal models—as shown in our study—could facilitate precision care decisions.

Conclusion

Our findings demonstrate that a tiered diagnostic strategy—beginning with brief cognitive screening and progressing through clinical profiling, structural imaging, and plasma biomarker evaluation—can effectively detect both syndromic and biologically defined Alzheimer's disease, even in a primary care context. Plasma biomarkers such as GFAP and p-tau181 provide significant diagnostic value, and their integration with MRI and APOE ϵ 4 status further enhances accuracy. Although some subgroup analyses suggest excellent performance, these results should be interpreted with caution due to limited sample size and potential overfitting. Nevertheless, this study underscores the feasibility of implementing real-world, personalized diagnostic frameworks that accommodate varying resource levels and can support early intervention in neurodegenerative diseases.

Methods

Study design and population

This cross-sectional, observational study was based on data from the STOP ALZHEIMER – DEBA project, an epidemiological initiative conducted in 2015 in the municipality of Deba (Basque Country, Spain). The project aimed to assess the prevalence and clinical spectrum of cognitive impairment, including MCI and dementia, in individuals aged 60 years and older from a real-world, community-based population.

This screening involved the collection of sociodemographic data, cardiovascular risk assessment using the CAIDE score (Cardiovascular Risk Factors, Aging, and Incidence of Dementia)¹², and brief cognitive evaluation through the Mini-Mental State Examination (MMSE), the Memory Alteration Test (M@T), the AD8 questionnaire, and the Fototest. These brief cognitive tests were conducted in a primary care setting by professionals who had received prior training from clinical neuropsychologists. The screening tests were administered in the participant's preferred language—either Basque or Spanish—given that the study was conducted in a bilingual region.

Individuals screening positive on any test using traditional thresholds were invited to continue to the diagnostic phase, along with a matched sample of cognitively negative participants matched for age, sex, education, and CAIDE score. A total of 277 individuals completed a comprehensive diagnostic evaluation at the CITA-Alzheimer Foundation (Donostia-San Sebastián, Basque Country, Spain), including clinical syndromic diagnosis, full neuropsychological assessment, structural brain MRI, blood and plasma biomarker analysis, APOE genotyping, and optional lumbar puncture.

Of these 277 participants, 181 underwent lumbar puncture for CSF biomarker analysis. This subgroup was used for analyses involving CSF as a reference standard. Comparisons between participants with and without CSF data were performed to assess sample representativeness (see Table 1).

Clinical and neurological evaluation

Participants underwent a physical and neurological examination and a standardized clinical assessment that included documentation of vascular risk factors such as hypertension, diabetes, dyslipidemia, and smoking status, along with anthropometric measures including body mass index. Laboratory data included fasting glucose, HbA1c, and total cholesterol levels.

Neuropsychiatric symptoms were assessed using the Neuropsychiatric Inventory (NPI)³⁵, while affective symptoms were measured with the Hospital Anxiety and Depression Scale (HADS)³⁶, including separate anxiety and depression subscales. Global cognitive and functional status was evaluated using the Clinical Dementia Rating (CDR) and the CDR Sum of Boxes (CDR-SB)³⁷. Subtle motor symptoms were evaluated using Part III (motor examination) of the Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS)³⁸.

Neuropsychological assessment

A detailed neuropsychological battery was administered by trained professionals and covered major cognitive domains. Memory was assessed using the CERAD Word List (learning and delayed recall)³⁹, Face-Name Associative Memory Exam (FNAME)⁴⁰, and the Rey-Osterrieth Complex Figure recall⁴¹. Language abilities were measured using the Boston Naming Test (short form)⁴², semantic and phonemic verbal fluency tasks⁴³. Visuospatial and visuoconstructive skills were evaluated with the VOSP Number Location test⁴⁴, VOSP Object Decision test, Rey-Osterrieth Complex Figure copy, and the 15-Objects Test⁴⁵. Executive functioning and attention were assessed with the Trail Making Test (Parts A and B)⁴⁶, WAIS-III Digit Span (forward and backward). Results were interpreted using validated normative data adjusted for age, sex, and education level.

Syndromic diagnosis and cognitive screening

Diagnostic classification was established through multidisciplinary consensus meetings involving neurologists, neuropsychologists, and other trained clinicians. Final diagnoses were based on the integration of clinical history, functional status (assessed via the Clinical Dementia Rating and CDR Sum of Boxes), and performance on the neuropsychological battery described in Sect. 5.3, interpreted using age-, sex-, and education-adjusted norms. Participants were classified as cognitively unimpaired, MCI, or dementia according to internationally accepted clinical criteria. MCI was defined as objective cognitive decline in one or more domains without significant impairment of daily functioning, whereas dementia was diagnosed when cognitive deficits interfered with autonomy in daily life. All participants underwent brief cognitive testing using MMSE, T@M, AD8, and Fototest. Traditional cut-offs were defined as MMSE \leq 24, T@M \leq 37, Fototest \leq 29, and AD8 \geq 2. Optimized

cut-offs derived from recent research included $\text{MMSE} \leq 28$, $\text{T@M} \leq 40$, $\text{Fototest} \leq 35$, and $\text{AD8} \geq 1$ ¹³. Composite variables were created to indicate positive screening using either traditional (Brief Cognitive Test traditional cutoff, BCTt) or optimized (Brief Cognitive Test optimized cutoff, BCTo) thresholds. These variables were used in logistic regression models and ROC curve analyses.

Plasma and CSF biomarker analysis

Venous blood was collected and processed according to standardized protocols. Samples were processed within two hours, centrifuged, aliquoted, and frozen at -80°C until analysis. Plasma levels of phosphorylated tau at threonine 181 (p-tau181), amyloid $\beta 1-42$ and $\beta 1-40$ (used to calculate the $\text{A}\beta 42/40$ ratio), glial fibrillary acidic protein (GFAP), and neurofilament light chain (NfL) were quantified using the Single Molecule Array (SIMOA) HD-X platform (Quanterix, USA) at the Achucarro Basque Center for Neuroscience Foundation. The Neurology 4-Plex E kit was used for $\text{A}\beta 1-40$, $\text{A}\beta 1-42$, NfL, and GFAP, and the pTau181 V2 Advantage Kit was used for p-tau181. All laboratory personnel were blinded to clinical data.

CSF was collected via lumbar puncture under sterile conditions using atraumatic needles. Samples were processed within two hours, centrifuged, aliquoted, and frozen at -80°C until analysis. Biomarker quantification was performed using Elecsys[®] assays (Roche Diagnostics)⁴⁷, which included $\text{A}\beta 42$, total tau, and p-tau181. Amyloid positivity (A+) was defined by low $\text{A}\beta 42$ levels ($< 1000\text{ pg/ml}$) and the p-tau/ $\text{A}\beta 42$ ratio, with values > 0.024 indicating abnormal amyloid pathology. Tau positivity (T+) was defined by elevated p-tau181 levels ($> 27\text{ pg/ml}$). These thresholds were based on reference values provided by the laboratory and applied consistently to classify CSF biomarker status.

MRI acquisition and interpretation

Structural MRI was acquired using a Siemens Magnetom Trio Tim 3T scanner at the CITA-Alzheimer Foundation. Sequences included 3D T1-weighted, FLAIR, T2-weighted, diffusion-weighted imaging (DWI), and susceptibility-weighted imaging (SWI). All MRIs were interpreted by an experienced neuroradiologist blinded to clinical and biomarker data. Medial temporal atrophy (MTA)⁴⁸ was visually rated and considered abnormal when ≥ 2 . Cerebrovascular burden was defined as either a Fazekas score⁴⁹ ≥ 2 or the presence of ≥ 4 cerebral microbleeds.

APOE genotyping

APOE genotyping was conducted using polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) analysis. Genotypes were categorized based on the presence of at least one $\epsilon 4$ allele, and participants were classified as APOE $\epsilon 4$ carriers or non-carriers. Genotyping was performed blinded to clinical outcomes and biomarker status.

Statistical analysis

Descriptive statistics were computed for demographic, clinical, cognitive, metabolic, and biomarker variables. Continuous variables were expressed as means and standard deviations and compared using independent samples t-tests. Categorical variables were described as percentages and compared using chi-square tests.

Diagnostic performance of brief cognitive tests and plasma biomarkers was assessed using ROC curves and area under the curve (AUC) calculations. Optimized cut-offs were applied for MMSE (≤ 28), M@T (≤ 40), Fototest (≤ 35), and AD8 (≥ 1), and a binary composite variable (BCTo) was defined as positive if any test met its threshold. AUCs were calculated globally and stratified by sex, age group, and educational attainment.

Plasma biomarker thresholds were estimated using the Youden index, defined as the maximum value of sensitivity + specificity – 1. Sensitivity, specificity, and cut-offs for p-tau181, $\text{A}\beta 42/40$ ratio, GFAP, and NfL were reported. Given the lack of universally approved diagnostic cut-offs for plasma biomarkers measured on the SIMOA platform, and the variability in thresholds proposed across studies and cohorts, these cut-offs were derived in an exploratory fashion based on our cohort-specific distribution.

Multivariable logistic regression models were constructed to predict a range of binary outcomes: syndromic cognitive impairment, MCI, CSF amyloid positivity (A+), combined amyloid and tau positivity (A + T+), and hybrid outcomes (e.g., MCI with A + T + pathology). Predictors included BCTo, plasma biomarkers (p-tau181, $\text{A}\beta 42/40$), APOE $\epsilon 4$ status, demographic variables (age, sex), and MRI features (medial temporal atrophy ≥ 2 , vascular pathology). Models were progressively layered into three complexity levels: the first (interview-based) included BCTo, the total CAIDE score, and the total Neuropsychiatric Inventory (NPI) score (sum of frequency \times severity across all domains); the second level added visual MRI features—medial temporal atrophy (dichotomized as ≥ 2) and Fazekas scale (dichotomized as ≥ 2); the third (full multimodal) level incorporated plasma biomarkers (p-tau181, $\text{A}\beta 42/40$ ratio) and APOE $\epsilon 4$ status, all entered as binary variables based on internally derived cut-offs. AUCs were computed for each level and outcome, and subgroup analyses were conducted for sex, CAIDE score, and age.

Due to the limited number of participants with confirmed dementia and CSF biomarker data, predictive modeling in Sect. 2.6 and 2.7 focused on MCI combined with biological positivity (MCI + A+), which represents a clinically actionable pre-dementia state. This approach also aligns with our aim to support early identification strategies in primary care.

All analyses were performed using Python 3.10 and R 4.2.2 with appropriate packages (pandas, scikit-learn, pROC). AUC values were considered excellent if ≥ 0.80 , good if ≥ 0.70 , and acceptable if ≥ 0.60 . A p-value < 0.05 was considered statistically significant. No corrections were made for multiple comparisons. To evaluate whether the stepwise increases in AUCs between model levels were statistically significant, DeLong tests for correlated ROC curves were performed using the subset of participants with complete data across all levels.

Ethical considerations

The protocol and informed consent procedure of the STOP ALZHEIMER – DEBA project were approved by the Ethics Committee of the Basque Country, under reference number PI2015153. All participants gave written informed consent prior to enrollment. The study was conducted in accordance with the Declaration of Helsinki and Spanish data protection legislation.

Data availability

The original data and analysis scripts used in this study are available from the authors upon request. Requests for access should be directed to the corresponding and first author, Dr. Miren Altuna, via email at maltuna@cita-alzheimer.org.

Received: 29 May 2025; Accepted: 22 August 2025

Published online: 27 August 2025

References

- Jack, C. R. J. et al. Revised criteria for diagnosis and staging of alzheimer's disease: alzheimer's association workgroup. *Alzheimers Dement.* **20** (8), 5143–5169 (2024).
- Dubois, B. et al. Alzheimer Disease as a Clinical-Biological Construct—An International Working Group Recommendation. *JAMA Neurol* [Internet]. ;81(12):1304–11. (2024). Available from: <https://doi.org/10.1001/jamaneurol.2024.3770>
- Nichols, E. et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Heal* [Internet]. ;7(2):e105–25. (2022). Available from: [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8)
- Scheltens, P. et al. Alzheimer's disease. *Lancet (London England)*. **397** (10284), 1577–1590 (2021).
- Arevalo-Rodriguez, I. et al. Mini-Mental state examination (MMSE) for the detection of alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst. Rev.* **2015** (3), CD010783 (2015).
- Carnero-Pardo, C., Rego-García, I., Mené Llorente, M. & Alonso Ródenas, M. Vilchez Carrillo R. Diagnostic performance of brief cognitive tests in cognitive impairment screening. *Neurologia* ; (2019).
- Carnero-Pardo, C., Sáez-Zea, C., Montiel-Navarro, L., Feria-Vilar, I. & Gurpegui, M. Normative and reliability study of fototest. *Neurologia* **26** (1), 20–25 (2011).
- Tanwani, R. et al. Diagnostic accuracy of ascertain dementia 8-item questionnaire by participant and informant-A systematic review and meta-analysis. *PLoS One.* **18** (9), e0291291 (2023).
- Rami, L., Bosch, B., Sanchez-Valle, R. & Molinuevo, J. L. The memory alteration test (M@T) discriminates between subjective memory complaints, mild cognitive impairment and alzheimer's disease. *Arch. Gerontol. Geriatr.* **50** (2), 171–174 (2010).
- Arranz, J. et al. Diagnostic performance of plasma pTau217, pTau181, Aβ1–42 and Aβ1–40 in the LUMIPULSE automated platform for the detection of Alzheimer disease. *Alzheimers Res Ther* [Internet]. ;16(1):139. (2024). Available from: <https://doi.org/10.1186/s13195-024-01513-9>
- Palmqvist, S. et al. Plasma phospho-tau217 for Alzheimer's disease diagnosis in primary and secondary care using a fully automated platform. *Nat Med* [Internet]. ; (2025). Available from: <https://doi.org/10.1038/s41591-025-03622-w>
- Kivipelto, M. et al. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *Lancet Neurol.* **5** (9), 735–741 (2006).
- Tainta, M. et al. Brief cognitive tests as a decision-making tool in primary care. A population and validation study. *Neurologia* ; (2022).
- O'Connor, A. et al. Plasma GFAP in presymptomatic and symptomatic familial Alzheimer's disease: a longitudinal cohort study. Vol. 94, *Journal of neurology, neurosurgery, and psychiatry*. England; pp. 90–2. (2023).
- Benedet, A. L. et al. Differences between plasma and cerebrospinal fluid glial fibrillary acidic protein levels across the alzheimer disease continuum. *JAMA Neurol.* **78** (12), 1471–1483 (2021).
- Schindler, S. E. et al. Head-to-head comparison of leading blood tests for Alzheimer's disease pathology. *Alzheimer's Dement* [Internet]. ;20(11):8074–96. (2024). Available from: <https://doi.org/10.1002/alz.14315>
- 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* [Internet]. ;20(7):426–39. (2024). Available from: <https://doi.org/10.1038/s41582-024-00977-5>
- Sahrai, H. et al. SIMOA-based analysis of plasma NFL levels in MCI and AD patients: a systematic review and meta-analysis. *BMC Neurol.* **23** (1), 331 (2023).
- Warmenhoven, N. et al. A comprehensive head-to-head comparison of key plasma phosphorylated Tau 217 biomarker tests. *Brain* **148** (2), 416–431 (2025).
- U.S. Food and Drug Administration. Premarket Notification: K242706, Lumipulse G pTau217/β-Amyloid 1–42 Plasma Ratio [Internet]. (2025). Available from: <https://www.fda.gov/news-events/press-announcements/fda-clears-first-blood-test-used-diagnosing-alzheimers-disease>
- Chatterjee, P. et al. Plasma Aβ42/40 ratio, p-tau181, GFAP, and NfL across the alzheimer's disease continuum: A cross-sectional and longitudinal study in the AIBL cohort. *Alzheimers Dement.* **19** (4), 1117–1134 (2023).
- Baiardi, S. et al. Diagnostic value of plasma p-tau181, nfl, and GFAP in a clinical setting cohort of prevalent neurodegenerative dementias. *Alzheimers Res. Ther.* **14** (1), 153 (2022).
- Hampel, H. et al. Blood-based biomarkers for alzheimer's disease: current state and future use in a transformed global healthcare landscape. *Neuron* **111** (18), 2781–2799 (2023).
- Cano, A. et al. Clinical value of plasma pTau181 to predict Alzheimer's disease pathology in a large real-world cohort of a memory clinic. *eBioMedicine* [Internet]. ;108. (2024). Available from: <https://doi.org/10.1016/j.ebiom.2024.105345>
- De Meyer, S. et al. Serum biomarkers as prognostic markers for Alzheimer's disease in a clinical setting. *Alzheimer's Dement Diagnosis, Assess Dis Monit* [Internet]. ;17(1):e70071. (2025). Available from: <https://doi.org/10.1002/dad2.70071>
- Zetterberg, H. & Blennow, K. Moving fluid biomarkers for alzheimer's disease from research tools to routine clinical diagnostics. *Mol. Neurodegener.* **16** (1), 10 (2021).
- Chandra, A., Dervenoulas, G. & Politis, M. Magnetic resonance imaging in alzheimer's disease and mild cognitive impairment. *J. Neurol.* **266** (6), 1293–1302 (2019).
- Ottoy, J. et al. Vascular burden and cognition: Mediating roles of neurodegeneration and amyloid PET. *Alzheimer's & Dement* [Internet]. ;19(4):1503–17. Available from: <https://alz-journals.onlinelibrary.wiley.com/doi/abs/> (2023). <https://doi.org/10.1002/alz.12750>
- Livingston, G. et al. Dementia prevention, intervention, and care: 2024 report of the lancet standing commission. *Lancet (London England)*. **404** (10452), 572–628 (2024).

30. Ecay-Torres, M. et al. Increased CAIDE dementia risk, cognition, CSF biomarkers, and vascular burden in healthy adults. *Neurology* **91** (3), e217–e226 (2018).
31. Belloli, L. et al. *Multi-modal AI Screening for MCI and Alzheimer's Disease: Results from an Argentine Cohort* Vol. 20 (Alzheimer's & Dementia, 2024).
32. Tascedda, S. et al. Advanced AI techniques for classifying alzheimer's disease and mild cognitive impairment. *Front. Aging Neurosci.* **16**, 1488050 (2024).
33. van Dyck, C. H. et al. Lecanemab in Early Alzheimer's Disease. *N Engl J Med* [Internet]. ;388(1):9–21. (2022). Available from: <https://doi.org/10.1056/NEJMoa2212948>
34. Sims, J. R. et al. Donanemab in Early Symptomatic Alzheimer Disease: The TRAILBLAZER-ALZ 2 Randomized Clinical Trial. *JAMA* [Internet]. ;330(6):512–27. (2023). Available from: <https://doi.org/10.1001/jama.2023.13239>
35. Cummings, J. The neuropsychiatric inventory: development and applications. *J. Geriatr. Psychiatry Neurol.* **33** (2), 73–84 (2020).
36. Zigmond, A. S. & Snaith, R. P. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* **67** (6), 361–370 (1983).
37. O'Bryant, S. E. et al. Staging dementia using clinical dementia rating scale sum of boxes scores: a Texas alzheimer's research consortium study. *Arch. Neurol.* **65** (8), 1091–1095 (2008).
38. Goetz, C. G. et al. Movement disorder Society-sponsored revision of the unified parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* **23** (15), 2129–2170 (2008).
39. Morris, J. C. et al. The consortium to Establish a registry for alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of alzheimer's disease. *Neurology* **39** (9), 1159–1165 (1989).
40. Rentz, D. M. et al. Face-name associative memory performance is related to amyloid burden in normal elderly. *Neuropsychologia* **49** (9), 2776–2783 (2011).
41. Peña-Casanova, J. et al. Spanish multicenter normative studies (NEURONORMA Project): norms for the Rey-Osterrieth complex figure (copy and memory), and free and cued selective reminding test. *Arch. Clin. Neuropsychol. Off J. Natl. Acad. Neuropsychol.* **24** (4), 371–393 (2009).
42. Casals-Coll, M. et al. Spanish multicenter normative studies (NEURONORMA project): normative data and equivalence of four BNT short-form versions. *Arch. Clin. Neuropsychol. Off J. Natl. Acad. Neuropsychol.* **29** (1), 60–74 (2014).
43. Peña-Casanova, J. et al. Spanish multicenter normative studies (NEURONORMA Project): norms for verbal fluency tests. *Arch. Clin. Neuropsychol. Off J. Natl. Acad. Neuropsychol.* **24** (4), 395–411 (2009).
44. Peña-Casanova, J. et al. Spanish multicenter normative studies (NEURONORMA Project): norms for the visual object and space perception battery-abbreviated, and judgment of line orientation. *Arch. Clin. Neuropsychol. Off J. Natl. Acad. Neuropsychol.* **24** (4), 355–370 (2009).
45. Alegret, M. et al. Detection of visuo-perceptual deficits in preclinical and mild alzheimer's disease. *J. Clin. Exp. Neuropsychol.* **31** (7), 860–867 (2009).
46. Peña-Casanova, J. et al. Spanish multicenter normative studies (NEURONORMA Project): norms for verbal span, visuospatial span, letter and number sequencing, trail making test, and symbol digit modalities test. *Arch. Clin. Neuropsychol. Off J. Natl. Acad. Neuropsychol.* **24** (4), 321–341 (2009).
47. Blennow, K. et al. Second-generation Elecsys cerebrospinal fluid immunoassays aid diagnosis of early alzheimer's disease. *Clin. Chem. Lab. Med.* **61** (2), 234–244 (2023).
48. Scheltens, P. et al. Atrophy of medial Temporal lobes on MRI in probable alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J. Neurol. Neurosurg. Psychiatry.* **55** (10), 967–972 (1992).
49. Fazekas, F., Chawluk, J. B., Alavi, A., Hurtig, H. I. & Zimmerman, R. A. MR signal abnormalities at 1.5 T in alzheimer's dementia and normal aging. *AJR Am. J. Roentgenol.* **149** (2), 351–356 (1987).

Acknowledgements

We appreciate the collaboration of the entire team at the CITA-Alzheimer Foundation.

Author contributions

Conceptualization, M.A. and PML; methodology, M.A. and MGS.; formal analysis, M.A. and MGS; investigation, M.A., MGS., R.C., ECS, EA., J.S., M.C., M.E., A.I., C.L, M.T. and PML; resources, M.A. and PML; data curation, M.A. and M.G.S.; writing—original draft preparation, M.A.; writing—review and editing, M. A., MGS, R.C., ECS, EA., AE, J.S., M.C., M.E., A.I., C.L, M.T. and PML; resources, M.A. and PML.; visualization, M.A.; supervision, M.A.; project administration, M.A. and MGS; funding acquisition, PML. All authors have read and agreed to the published version of the manuscript.

Funding

This study has been partially funded by Instituto de Salud Carlos III through the project “PI21/00718”(Co-funded by European Regional Development Fund; “A way to make Europe”) and by BIOEF (Convocatoria de Ayudas a la Investigación en Alzheimer de la Fundación Vasca de Innovación e Investigación Sanitarias) through the project “BIO22/ALZ/014”. The DEBA project was supported by the Department of Health of the Basque Government.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-17394-3>.

Correspondence and requests for materials should be addressed to M.A. or P.M.-L.

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

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025