Received: 11 September 2025 | Accepted: 31 October 2025

DOI: 10.1002/dad2.70223

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Amyloid-negative neuropsychological norms: Added value in the era of biomarkers and disease-modifying therapies

Sara Rubio-Guerra[1,2,3] | Isabel Sala[1,2] | María Belén Sánchez-Saudinós[1,2] | Laura Videla[1,2] | Alexandre Bejanin[1,2] | Ainara Estanga[4] | Mirian Ecay-Torres[4] | Carolina Lopez de Luis[4] | Lorena Rami[2,5] | Adrià Tort-Merino[2,5] | Magdalena Castellví[5] | Ana Pozueta[2,6] | María García-Martínez[2,6] | David Gómez-Andrés[7] | Carmen Lage[2,6] | Sara López-García[2,6] | Pascual Sánchez-Juan[2,6] | Mircea Balasa[2,5] | Albert Lladó[2,5] | Miren Altuna[4,8] | Mikel Tainta[4] | Javier Arranz[1,2,3] | Nuole Zhu[1,2,3] | Daniel Alcolea[1,2] | Alberto Lleó[1,2,3] | Juan Fortea[1,2] | Eloy Rodríguez Rodríguez[2,6,9] | Raquel Sánchez-Valle[2,5] | Pablo Martínez-Lage[4] | Ignacio Illán-Gala[1,2,3]

[1]Sant Pau Memory Unit, Department of Neurology, IR SANT PAU, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[2]Center for Biomedical Network Research on Neurodegenerative Diseases (CIBERNED), Instituto de Salud Carlos III, Madrid, Spain

[3]Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

[4]Fundación CITA-Alzhéimer Fundazioa, Centro de Investigación y Terapias Avanzadas, San Sebastián, Gipuzkoa, Spain

[5]Alzheimer's disease and other cognitive disorders Unit, Hospital Clínic de Barcelona, Fundació de Recerca Clínic Barcelona–Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[6]Hospital Universitario Marqués de Valdecilla, Instituto de Investigación Marqués de Valdecilla (IDIVAL), Santander, Spain

[7]Hospital Universitari Vall d'Hebron, Vall d'Hebron Institut de Recerca (VHIR), Barcelona, Spain

[8]Debabarrena Integrated Health Organization, Osakidetza Basque Health Service, Elgoibar, Gipuzkoa, Spain

[9]Department of Medicine and Psychiatry, Universidad de Cantabria, Santander, Spain

Correspondence
Ignacio Illán-Gala, Sant Pau Memory Unit, Neurology Department, Hospital de la Santa Creu i Sant Pau, Building 18, Floor 2, Office 44, C/ Sant Antoni Maria Claret, 167, Barcelona 08041, Spain.
Email: iillan@santpau.cat

Funding information
Ajuntament de Barcelona; Alzheimer's Association; Alzheimer's Society; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas; Department of Economic Promotion of the Provincial Government of Gipuzkoa;

Abstract

INTRODUCTION: We previously applied generalized additive models for location, scale, and shape to derive amyloid β–negative next-generation norms (NGN) for a comprehensive neuropsychological battery. Here, we evaluated the accuracy of NGN in detecting cognitive impairment compared to traditional norms (TN).

METHODS: This multicenter study included $N = 2405$ participants classified as cognitively normal (CN, $n = 987$) or with mild cognitive impairment (MCI, $n = 1418$) using conventional criteria. All participants underwent neuropsychological testing and cerebrospinal fluid Alzheimer's disease (AD) biomarker assessment. We used actuarial neuropsychological criteria to reclassify all participants using TN and NGN.

Department of Health of the Basque Government; European Union; Fondo Europeo de Desarrollo Regional; Fundació Catalana Síndrome de Down; Fundació 'la Caixa'; Fundació La Marató de TV3; Fundación Tatiana Pérez de Guzmán el Bueno; Fundació Víctor Grífols i Lucas; Generalitat de Catalunya; Global Brain Health Institute; Instituto de Salud Carlos III; Jérôme Lejeune Foundation; National Institutes of Health

Diagnostic groups were compared on cognitive performance, AD biomarker positivity, and longitudinal cognitive trajectories.

RESULTS: Nineteen percent of TN-classified CN participants were diagnosed with MCI by NGN, whereas 3% of TN-classified MCI were identified as CN by NGN. NGN demonstrated stronger associations with neuropsychological performance, AD biomarkers, and progression than TN.

DISCUSSION: NGN enhance the detection of objective cognitive impairment, with direct implications for clinical practice and research.

KEYWORDS

Alzheimer's disease, biomarkers, clinical trials, cognitive trajectories, diagnosis, disease-modifying therapies, MCI, neuropsychological norms, normative data

Highlights

- Next-generation norms (NGN) reclassify one of every five cases from cognitively normal (CN) to mild cognitive impairment (MCI).
- This group shows poor cognitive performance and a high prevalence of amyloid β positivity.
- NGN-based diagnosis of MCI predicts cognitive progression on follow-up.
- Results indicate that NGN improve the detection of objective cognitive impairment.
- NGN can inform biomarker use, therapy indication, and clinical trial design.

1 | BACKGROUND

Alzheimer's disease (AD) is the leading cause of dementia;[1] yet, the majority of individuals on the AD continuum are in the predementia stages, including the preclinical AD (i.e., positive AD biomarkers, but performance in the normal range on objective cognitive assessment) and prodromal AD (i.e., mild cognitive impairment [MCI] due to AD) stages.[2] The recent approval of the first disease-modifying drugs for AD, targeting amyloid β (Aβ) plaques in individuals at early symptomatic stages,[3–5] underscores the importance of timely, biomarker-confirmed AD diagnosis, as these therapies show the greatest benefit in earlier disease stages.[6,7] At the same time, in the absence of effective preventive treatments, biomarkers are not recommended for use in cognitively normal (CN) individuals in clinical practice.[8,9] Accurate neuropsychological norms are therefore critical to identify the earliest objectively measurable cognitive deficits that inform the use of biomarkers and the selection of patients for therapeutic interventions.

We recently developed comprehensive neuropsychological normative data, referred to as next-generation norms (NGN).[10] NGN incorporated two methodological advancements: (1) the selection of a normative sample consisting of CN individuals without biomarker evidence of Aβ accumulation, and (2) the modeling of normative data under the generalized additive models for location, scale, and shape (GAMLSS). Norms were adjusted for age, education, and sex.

In the present follow-up study, we aim to evaluate the accuracy of NGN in detecting cognitive impairment compared to traditional norms (TN). However, direct comparison of neuropsychological norms is challenged by the lack of both uniform test batteries and an operationalized definition of MCI.[11,12] Conventional MCI criteria rely on subjective cognitive decline, clinical judgment based on cognitive and functional screening tools, and impairment in one or more neuropsychological tests,[13,14] which may result in false-positive and false-negative diagnostic errors.[15,16] In contrast, actuarial MCI criteria such as the Jak/Bondi approach, which require impairment on at least two measures within a cognitive domain,[17] have demonstrated greater diagnostic stability and stronger associations with AD biomarkers and progression.[18,19]

Accordingly, to compare NGN and TN, we selected a large multicenter dataset of individuals without dementia and applied actuarial neuropsychological criteria to cognitive scores based on both NGN and TN from the same test battery. We hypothesized that NGN would yield stronger associations with neuropsychological performance, AD biomarker positivity, and clinical progression than TN.

2 | METHODS

2.1 | Participants and design

Data were prospectively collected as part of the longitudinal multicenter SIGNAL study, which harmonized core neuropsychological procedures across several Spanish centers.[20] For the present study, we retrospectively selected $N = 2405$ individuals without dementia (CN: $n = 987$; MCI: $n = 1418$) recruited between March 2006 and

RESEARCH IN CONTEXT

1. Systematic Review: Current Spanish neuropsychological norms do not account for biomarker status and are based on traditional norming methods. Building upon accumulated knowledge from prior studies, we recently derived next-generation norms (NGN) from a large sample of amyloid β–negative cognitively normal individuals using generalized additive models for location, scale, and shape (GAMLSS).
2. Interpretation: Despite the theoretical advantages of NGN, evidence is needed to establish their actual added value. NGN-based cognitive classification demonstrated enhanced associations with neuropsychological performance, Alzheimer's disease biomarker positivity, and longitudinal cognitive trajectories compared to current normative standards. These results suggest that NGN enhance the detection of cognitive impairment and may therefore improve clinical and research outcomes.
3. Future Directions: Future research should provide further evidence to confirm the utility of NGN.

July 2022. Participants were recruited from both outpatient memory clinics and community-based volunteer populations. The inclusion criteria were: (1) absence of dementia, (2) age ≥ 30 years, and baseline availability of both (3) a neuropsychological evaluation and (4) cerebrospinal fluid (CSF) AD biomarker assessment, including Aβ status. The exclusion criterion was a history of neurological, psychiatric, or systemic conditions that might affect cognitive performance (Figure S1). Participants were classified at baseline as CN or MCI using conventional criteria[13,14] based only on clinical and neuropsychological data at the time of CSF collection, without regard to biomarker results or follow-up information. We previously derived NGN from a normative dataset comprising all Aβ-negative (Aβ–) CN individuals (n = 774).[10] This study adhered to the 1964 Declaration of Helsinki and was approved by the ethics committees of all participating centers. Written informed consent was obtained from all participants before enrollment.

2.2 | Materials and procedure

2.2.1 | Neuropsychological evaluation

All participants underwent a baseline neuropsychological evaluation according to their center's protocol (Table S1). The assessment incorporated the Mini-Mental State Examination (MMSE; range 0–30, with lower scores reflecting greater impairment) as a measure of general cognitive functioning.[21] The neuropsychological measures included in the present validation study were those for which NGN were derived in our prior study, and included at least two measures across the domains of: visuospatial skills (the Visual Object and Space Perception Battery [VOSP] number location subtest and the Rey-Osterrieth Complex Figure [ROCF] copy), memory (the Free and Cued Selective Reminding Test [FCSRT] total free recall, FCSRT total recall, FCSRT delayed free recall, FCSRT delayed total recall, and the ROCF delayed recall), attention/executive function (the Trail Making Test Part A [TMT-A], Trail Making Test Part B [TMT-B], Digit Span Forward, Digit Span Backward, and phonetic fluency [words beginning with P]), and language (semantic fluency [animal names] and the Boston Naming Test [BNT]).

Raw neuropsychological scores of all participants (N = 2405) were converted into demographically adjusted normed scores according to NGN and published Spanish normative data (the Neuronorma Project[22,23]—from this point forward, "TN" will specifically refer to this set of traditional neuropsychological norms). Data were missing for 10% of the observations, with copy and delayed recall of the ROCF each accounting for 37% of the overall missing data.

2.2.2 | CSF biomarkers of AD

All participants underwent a baseline lumbar puncture to analyze CSF AD biomarkers, of which Aβ 1–42 (Aβ42) was the most widely available Aβ biomarker across centers. Individuals were classified as either Aβ-positive (Aβ+) or Aβ– based on each center's cutoff for Aβ42 (Supplementary Methods).

2.2.3 | Cognitive trajectories

In a subset of participants from Hospital de la Santa Creu i Sant Pau for whom at least a 1-year longitudinal neuropsychological assessment was available (n = 812), follow-up changes in the MMSE score were used to track their cognitive performance over time.

2.3 | Added value of NGN

To explore the potential added value of NGN over TN, we examined if NGN-derived cognitive status classification (i.e., CN or MCI) improved the associations with neuropsychological performance, CSF AD biomarkers, and cognitive trajectories compared to TN.

To compare the norms, we applied the same actuarial neuropsychological criteria, adapted from Jak/Bondi criteria,[17] to both NGN- and TN-based scores. Adaptation of Jak/Bondi criteria involved defining impairment as a performance < −1.5 standard deviations (SD) from the normative mean (corresponding to a z-score < −1.5 and a scaled score ≤ 5), as opposed to < −1 SD put forward by Jak/Bondi. This midpoint cutoff (−1.5 vs. −1 or −2 SD) aligns with common practices in the field and offers a good compromise between sensitivity and specificity. Participants were actuarially diagnosed with MCI if they had two or more impaired normed scores within at least one cognitive domain across memory, attention/executive function, or language. Participants were considered CN only if they had at least two cognitive measures available across all the examined domains, and the criteria for MCI were not

met. In making actuarial diagnoses, we did not account for visuospatial function, as 47% of the participants had insufficient data available for classification in this domain; thus, considering visuospatial skills would have precluded many otherwise CN individuals from receiving a neuropsychological diagnosis.

Four possible combinations of actuarial neuropsychological diagnoses were expected: individuals with a normal performance according to both TN and NGN (CN_{tn}/CN_{ngn}); individuals performing abnormally by both norms (MCI_{tn}/MCI_{ngn}); individuals classified as CN based on TN, but diagnosed with MCI under NGN (CN_{tn}/MCI_{ngn}); and those classified as MCI according to TN, but performing normally as per NGN (MCI_{tn}/CN_{ngn}). We explored the frequencies of these diagnostic groups and conducted between-group comparisons on their demographic characteristics, neuropsychological performance, and rates of AD-related biomarkers. We also compared cognitive trajectories across diagnostic groups, as measured by changes in MMSE scores over time: we first analyzed the complete longitudinal dataset; subsequently, we performed separate subanalyses for Aβ+ and Aβ– participants to capture cognitive progression driven by AD and non-AD underlying etiologies, respectively.

2.4 | Statistical analyses

Statistical analyses were conducted using R, version 4.2.3.[24] Between-group differences in continuous variables were assessed using *t*-tests for two-group comparisons and one-way analysis of variance (ANOVA) for three or more groups. For unequal variances, we applied Welch's *t*-test and Welch's ANOVA. Group differences in categorical variables were examined using chi-squared tests. Statistical significance for all tests was set at p value < 0.05 ($\alpha = 0.05$). Standardized effect size statistics were calculated to quantify the magnitudes of between-group differences: we reported Cohen's d (d) for *t*-tests, eta-squared (η^2) for ANOVA, and Cramér's V (φc) for chi-squared tests. The magnitudes of the differences were interpreted as small, medium, or large based on generally accepted thresholds.[25] To account for demographic influences when comparing neuropsychological performance across diagnostic groups, we calculated *w*-scores (standardized residuals) using a fully adjusted regression approach. For each neuropsychological measure, a multiple linear regression model was fitted in the CN_{tn}/CN_{ngn} reference group. Based on clinical knowledge, age, education, and sex were included as covariates in all models, irrespective of their univariate or multivariate significance. The predicted value for each participant was subtracted from their observed raw score and divided by the model's residual SD, yielding a *w*-score that represents individual deviation from demographically adjusted expected performance. Post hoc analyses were corrected for multiple comparisons using the Bonferroni method. To conduct difference and equivalence tests, we prespecified thresholds for a meaningful difference (δ): ≥ 0.20 for d, ≥ 0.01 for η^2, and ≥ 0.10 for φc, according to conventions.[25] We concluded that differences in between-group comparisons existed whenever a statistically significant and meaningful difference was found. Equivalence between groups was established when the upper limit of the 90% (1–2α) confidence interval for the effect size fell entirely below the predefined equivalence threshold δ.[26] We applied linear mixed-effects (LME) models for the longitudinal analyses to predict annual changes in MMSE scores across diagnostic groups. The models incorporated fixed effects for the group (up to four levels: CN_{tn}/CN_{ngn}, MCI_{tn}/MCI_{ngn}, CN_{tn}/MCI_{ngn}, and MCI_{tn}/CN_{ngn}), time, and their interaction. Age at baseline, education, and sex were also incorporated as fixed effects to account for their potential influence on cognitive trajectories. To consider individual variability in baseline scores and rates of change over time, a random intercept, as well as a random slope for time, were included for each participant. We used the restricted maximum likelihood method for model estimation. Akaike Information Criterion was used to assess our final models' fit.

3 | RESULTS

3.1 | Demographics and baseline characteristics

Table 1 presents baseline characteristics of all participants, stratified by conventional diagnosis. The MCI group was older, had fewer years of education, lower MMSE scores, and a higher proportion of CSF Aβ+ participants than the CN group. The sex distribution was equivalent between the MCI and CN groups. Raw neuropsychological scores of MCI individuals were significantly worse than those of CN individuals on all 14 cognitive measures, with large effect sizes.

3.2 | Added value of NGN

Of the 2405 participants, 2185 had enough baseline neuropsychological data to receive an actuarial cognitive classification based on both TN and NGN. Among these, 81% of the individuals actuarially identified as CN by TN, and 97% of those diagnosed with MCI by TN, were classified the same by NGN (CN_{tn}/CN_{ngn} and MCI_{tn}/MCI_{ngn} groups, respectively). Conversely, 19% of TN-classified CN participants were diagnosed with MCI by NGN (CN_{tn}/MCI_{ngn}). Finally, 3% of TN-classified MCI individuals were considered CN by NGN (MCI_{tn}/CN_{ngn}; Figure 1). Table 2 displays baseline demographics, neuropsychological performance, and proportions of CSF Aβ+ individuals across the four actuarially reclassified diagnostic groups. Comparisons of TN- and NGN-derived actuarial classifications versus conventional diagnoses made at each center are provided in Tables S2 and S3.

3.2.1 | Demographic characteristics

Across the four diagnostic groups, statistically significant overall differences were observed for age and years of education, with effect sizes indicating a medium-to-large difference for age and a small-to-medium difference for education. No significant differences were observed for sex distribution (Table 2).

TABLE 1 Baseline demographic, neuropsychological, and biomarker characteristics of participants.

Characteristics	All participants $N = 2405$*	CN $n = 987$*	MCI $n = 1418$*	p value†	Effect size‡
Demographics					
Age (years)	67 (9.5)	62 (8.9)	70 (8.4)	**< 0.001**	$d = 0.98$
Education (years)	12 (4.8)	13 (4.3)	11 (4.8)	**< 0.001**	$d = 0.50$
Sex (female)	1367 (57%)	594 (60%)	773 (55%)	**0.007**	$\varphi c = 0.05$
Clinical measures					
MMSE (/30)	27.0 (2.9)	28.7 (1.5)	25.8 (3.0)	**< 0.001**	$d = 1.17$
Neuropsychological performance (raw scores)					
VOSP number location (/10)	8.0 (2.3)	9.1 (1.2)	7.4 (2.5)	**< 0.001**	$d = 0.80$
ROCF copy (/36)	29.8 (6.2)	32.0 (3.6)	26.9 (7.5)	**< 0.001**	$d = 0.90$
FCSRT total free recall (/48)	17.2 (10.3)	26.7 (6.3)	11.4 (7.6)	**< 0.001**	$d = 2.15$
FCSRT total recall (/48)	33.0 (12.8)	43.3 (4.7)	26.6 (12.1)	**< 0.001**	$d = 1.67$
FCSRT delayed free recall (/16)	6.4 (4.8)	10.7 (2.9)	3.7 (3.6)	**< 0.001**	$d = 2.10$
FCSRT delayed total recall (/16)	11.3 (4.9)	15.1 (1.4)	8.9 (4.8)	**< 0.001**	$d = 1.64$
ROCF delayed recall (/36)	13.2 (7.0)	16.1 (5.9)	9.2 (6.4)	**< 0.001**	$d = 1.14$
TMT-A (s)	62.6 (36.1)	41.3 (18.4)	77.5 (37.9)	**< 0.001**	$d = 1.16$
TMT-B (s)	176.1 (98.8)	103.0 (63.5)	227.3 (86.1)	**< 0.001**	$d = 1.60$
Digit Span Forward (/9)	5.2 (1.2)	5.7 (1.2)	4.9 (1.1)	**< 0.001**	$d = 0.75$
Digit Span Backward (/8)	3.9 (1.2)	4.4 (1.2)	3.5 (1.1)	**< 0.001**	$d = 0.88$
Phonetic fluency (n)	12.2 (5.6)	15.8 (4.9)	9.6 (4.6)	**< 0.001**	$d = 1.31$
Semantic fluency (n)	16.3 (6.8)	21.1 (6.0)	12.9 (5.0)	**< 0.001**	$d = 1.51$
BNT (%)	81.7 (14.5)	90.5 (7.8)	75.6 (15.0)	**< 0.001**	$d = 1.18$
CSF Aβ status					
Aβ+	1008 (42%)	213 (22%)	795 (56%)	**< 0.001**	$\varphi c = 0.34$

Note: Descriptive and inferential statistics on baseline demographics, MMSE scores, raw neuropsychological scores, and frequencies of CSF Aβ+ participants, stratified by conventional diagnosis. Bold p values indicate statistical significance (< 0.05).

Abbreviations: Aβ, amyloid β; Aβ+, Aβ–positive; AD, Alzheimer's disease; BNT, Boston Naming Test; CN, cognitively normal; CSF, cerebrospinal fluid; FCSRT, Free and Cued Selective Reminding Test; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; ROCF, Rey–Osterrieth Complex Figure; SD, standard deviation; TMT-A/B, Trail Making Test Part A/Part B; VOSP, Visual Object and Space Perception Battery.

*Mean (SD); n (%).

†Welch's two sample t-test; Pearson's chi-squared test.

‡d = Cohen's d, φc = Cramér's V.

3.2.2 | Neuropsychological performance

For most neuropsychological measures, visual inspection of results indicated a stepwise cognitive performance across the four diagnostic groups: participants in the CN_{tn}/CN_{ngn} group performed better than those in the MCI_{tn}/CN_{ngn} group, who in turn outperformed the CN_{tn}/MCI_{ngn} group; finally, the MCI_{tn}/MCI_{ngn} group showed the lowest performance. Inferential analyses confirmed significant and meaningful between-group differences across all neuropsychological measures after adjusting for relevant demographic factors, with the largest differences in FCSRT subtests (i.e., verbal memory measures; Table 2).

3.2.3 | CSF biomarkers of AD

A chi-squared test revealed significant between-group differences in CSF Aβ+ prevalence. The largest difference was observed between the CN_{tn}/CN_{ngn} (22%) and MCI_{tn}/MCI_{ngn} (65%) groups. Notably, the MCI_{tn}/CN_{ngn} group showed an Aβ+ prevalence (8.7%) statistically equivalent to that in the CN_{tn}/CN_{ngn} group, while the proportion of Aβ+ individuals in the CN_{tn}/MCI_{ngn} group (48%) more closely resembled that in the MCI_{tn}/MCI_{ngn} group (Table 2; Figure 2).

3.2.4 | Cognitive trajectories

Longitudinal data were available for 812 participants (CN_{tn}/CN_{ngn}, n = 374; MCI_{tn}/CN_{ngn}, n = 8; CN_{tn}/MCI_{ngn}, n = 105; MCI_{tn}/MCI_{ngn}, n = 325), representing 34% of the total sample. The longitudinal cohort consisted of 322 Aβ+ and 490 Aβ– participants. Significant between-group differences in baseline MMSE scores were observed, with a large effect size. The differences in baseline MMSE scores followed this pattern: $CN_{tn}/CN_{ngn} > MCI_{tn}/CN_{ngn} > CN_{tn}/MCI_{ngn} > MCI_{tn}/MCI_{ngn}$. Participants underwent a mean (SD) of 3.6 (1.9) MMSE assessments,

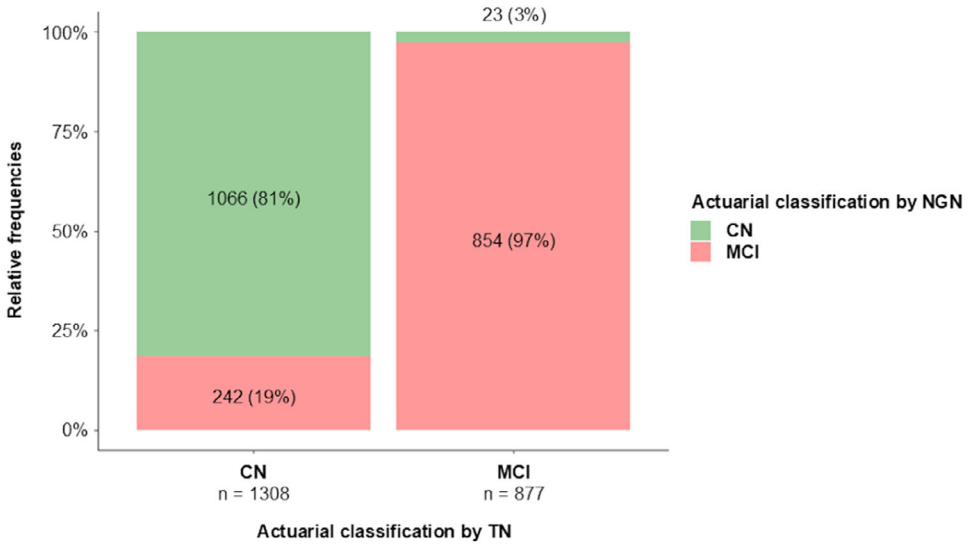

FIGURE 1 Cognitive classification by NGN versus TN. Absolute and relative frequencies of actuarial cognitive classifications based on NGN compared to those based on TN. CN, cognitively normal; MCI, mild cognitive impairment; NGN, next-generation norms; TN, traditional norms.

with a median (range) of 3 (2–11) evaluations, over a mean (SD) follow-up period of 4.2 (2.7) years, with a median (range) of 3.5 (1.0–14.1) years. Importantly, no statistically significant between-group differences were found in terms of follow-up number of visits or duration, enhancing the comparability of longitudinal cognitive trajectories across groups (Table S4).

We first applied an LME model to the complete longitudinal dataset to explore general trends. Next, we conducted subanalyses on the Aβ+ and Aβ− cohorts. The MCI_{tn}/CN_{ngn} group was not included in the models due to insufficient sample size ($n = 8$) for valid estimation.

When analyzing the complete longitudinal dataset, the LME model revealed statistically significant differences in baseline MMSE scores and their annual changes for the MCI_{tn}/MCI_{ngn} and the CN_{tn}/MCI_{ngn} groups compared to the reference CN_{tn}/CN_{ngn} group. Specifically, the MCI_{tn}/MCI_{ngn} group exhibited a mean MMSE score that was 1.97 points lower at baseline and decreased by an additional 0.80 points per year relative to the CN_{tn}/CN_{ngn} group. The cognitive trajectory of the CN_{tn}/MCI_{ngn} group was also significantly different from that of the CN_{tn}/CN_{ngn} group, with a mean MMSE score that was 0.72 points lower at baseline and an annual decline that exceeded by 0.32 points that of the CN_{tn}/CN_{ngn} group (Figure 3A; Table S5).

Subsequently, we conducted separate LME models of the Aβ+ and Aβ− subsamples. The diagnostic groups exhibited similar tendencies in cognitive trajectories as observed in the complete dataset, although not all comparisons reached statistical significance (Figure 3B,C; Tables S6 and S7).

4 | DISCUSSION

We previously derived Aβ− NGN for a comprehensive neuropsychological battery using GAMLSS.[10] In this study, we evaluated the added value of NGN in detecting cognitive impairment compared to TN. Our results indicate that NGN enhance diagnostic sensitivity, while preliminary analyses do not reveal any overt loss of specificity.

We observed notable discrepancies between TN- and NGN-based cognitive classifications. Specifically, one of every five (19%) participants considered CN under TN was diagnosed with MCI by NGN, while 3% of those diagnosed with MCI under TN were classified as CN by NGN. These reclassifications were supported by patterns of neuropsychological performance and AD biomarker positivity. Participants consistently classified as CN (CN_{tn}/CN_{ngn}) performed best, followed by those TN-diagnosed with MCI but NGN-classified CN (MCI_{tn}/CN_{ngn}); these, in turn, outperformed participants TN-classified CN but NGN-diagnosed with MCI (CN_{tn}/MCI_{ngn}); finally, participants consistently diagnosed with MCI (MCI_{tn}/MCI_{ngn}) showed the poorest neuropsychological performance. Regarding AD biomarkers, the MCI_{tn}/CN_{ngn} group had the lowest proportion of Aβ+ individuals (8.7%), statistically equivalent to that in the CN_{tn}/CN_{ngn} group (22%). In turn, the proportion of Aβ+ participants in the CN_{tn}/MCI_{ngn} group (48%) more closely resembled that in the MCI_{tn}/MCI_{ngn} group (65%).

These findings suggest that CN_{tn}/MCI_{ngn} individuals likely represent false-negative diagnostic errors under TN. The enhanced sensitivity achieved by NGN was primarily driven by impairments in the FCSRT subtests, detected by NGN but not by TN. This aligns with established evidence that episodic memory deficits are the hallmark of MCI due to AD.[27–29] Conversely, MCI_{tn}/CN_{ngn} individuals might represent false-positive diagnostic errors by TN, possibly related to alternative, more benign etiologies.

Longitudinal analyses further validated the utility of NGN. The CN_{tn}/MCI_{ngn} group demonstrated baseline MMSE scores and annual decline rates that were intermediate between those observed in the CN_{tn}/CN_{ngn} and MCI_{tn}/MCI_{ngn} groups. These trends remained consistent across Aβ+ and Aβ− subgroups, although some comparisons did not reach statistical significance due to limited statistical power.

TABLE 2 Baseline demographic, neuropsychological, and biomarker characteristics across actuarially reclassified diagnostic groups.

Characteristics	CN_{tn}/CN_{ngn} $n = 1066$*	MCI_{tn}/CN_{ngn} $n = 23$*	CN_{tn}/MCI_{ngn} $n = 242$*	MCI_{tn}/MCI_{ngn} $n = 854$*	*p* value†	Effect size‡
Demographics						
Age (years)	63.0 (9.7)	57.4 (11.7)	70.3 (8.8)	68.8 (8.4)	**< 0.001**	$\eta^2 = 0.11$
Education (years)	13.0 (4.7)	12.2 (5.2)	10.0 (4.2)	11.7 (4.8)	**< 0.001**	$\eta^2 = 0.04$
Sex (female)	625 (59%)	14 (61%)	146 (60%)	466 (55%)	0.2	$\varphi c = 0.03$
Clinical measures						
MMSE (/30)	28.5 (1.7)	27.2 (2.8)	26.5 (2.8)	25.2 (3.1)	**< 0.001**	$\eta^2 = 0.28$
Neuropsychological performance (*w*-scores)						
VOSP number location	0.0 (1.0)	−0.5 (1.2)	−0.4 (1.3)	−0.7 (1.5)	**< 0.001**	$\eta^2 = 0.07$
ROCF copy	0.0 (1.0)	−0.4 (1.3)	−0.3 (1.4)	−1.0 (1.8)	**< 0.001**	$\eta^2 = 0.1$
FCSRT total free recall	0.0 (1.0)	−0.9 (0.9)	−1.2 (0.9)	−2.5 (1.0)	**< 0.001**	$\eta^2 = 0.59$
FCSRT total recall	0.0 (1.0)	−0.7 (1.1)	−2.0 (1.5)	−4.9 (2.5)	**< 0.001**	$\eta^2 = 0.62$
FCSRT delayed free recall	0.0 (1.0)	−0.6 (1.0)	−1.4 (1.0)	−2.5 (1.0)	**< 0.001**	$\eta^2 = 0.58$
FCSRT delayed total recall	0.0 (1.0)	−0.5 (1.0)	−2.1 (1.7)	−5.4 (3.0)	**< 0.001**	$\eta^2 = 0.59$
ROCF delayed recall	0.0 (1.0)	−0.4 (1.2)	−0.7 (1.0)	−1.3 (1.0)	**< 0.001**	$\eta^2 = 0.25$
TMT-A	0.0 (1.0)	−0.9 (1.4)	−1.0 (1.8)	−1.6 (2.1)	**< 0.001**	$\eta^2 = 0.18$
TMT-B	0.0 (1.0)	−0.8 (1.0)	−1.0 (1.2)	−1.5 (1.4)	**< 0.001**	$\eta^2 = 0.27$
Digit Span Forward	0.0 (1.0)	−1.0 (0.9)	−0.3 (0.9)	−0.5 (1.0)	**< 0.001**	$\eta^2 = 0.05$
Digit Span Backward	0.0 (1.0)	−1.1 (0.8)	−0.4 (0.9)	−0.6 (1.0)	**< 0.001**	$\eta^2 = 0.09$
Phonetic fluency	0.0 (1.0)	−0.9 (1.0)	−0.6 (0.8)	−1.0 (1.0)	**< 0.001**	$\eta^2 = 0.18$
Semantic fluency	0.0 (1.0)	−0.6 (1.1)	−0.7 (0.7)	−1.1 (0.9)	**< 0.001**	$\eta^2 = 0.25$
BNT	0.0 (1.0)	−0.7 (0.8)	−1.1 (1.5)	−1.9 (2.1)	**< 0.001**	$\eta^2 = 0.24$
CSF Aβ status						
Aβ+	234 (22%)	2 (8.7%)	115 (48%)	558 (65%)	**< 0.001**	$\varphi c = 0.42$

Note: Descriptive and inferential statistics on baseline demographics, MMSE scores, neuropsychological performance, and frequencies of CSF Aβ+ participants across actuarially reclassified diagnostic groups. Raw neuropsychological scores were converted to demographically adjusted *w*-scores (reference group: CN_{tn}/CN_{ngn}) to account for the potential influence of demographic variables on neuropsychological performance and to facilitate comparison across measures. Bold *p* values indicate statistical significance (< 0.05).

Abbreviations: Aβ, amyloid β; Aβ+, Aβ–positive; AD, Alzheimer's disease; ANOVA, analysis of variance; BNT, Boston Naming Test; CN, cognitively normal; CSF, cerebrospinal fluid; FCSRT, Free and Cued Selective Reminding Test; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; ngn, next-generation norms; ROCF, Rey–Osterrieth Complex Figure; SD, standard deviation; TMT-A/B, Trail Making Test Part A/Part B; tn, traditional norms; VOSP, Visual Object and Space Perception Battery.

*Mean (SD); *n* (%).

†One-way ANOVA; Welch's ANOVA; Pearson's chi-squared test.

‡η^2 = eta-squared, φc = Cramér's V.

These results support the utility of NGN in detecting early cognitive impairment and reflect the accelerating nature of cognitive decline observed in previous longitudinal studies.[30] Additionally, consistent with the progressive nature of AD, the Aβ+ CN_{tn}/CN_{ngn} subgroup exhibited a steeper cognitive decline than the Aβ− CN_{tn}/CN_{ngn} subgroup, which was likely composed primarily of individuals without neurodegenerative pathology.

The diagnostic advantage of NGN over TN likely stems from a combination of factors. First, our approach of excluding individuals with biomarker evidence of Aβ accumulation from the normative sample is endorsed by prior research,[31–34] and was further validated by our previous findings.[10] This approach may thus have reduced the risk of attributing to normal aging the cognitive decline due to undetected pathology, thereby enabling a more accurate assessment of neuropsychological performance and AD-related cognitive decline.[35] Second, the GAMLSS framework enables the inclusion of continuous predictors (e.g., age and education) and modeling of all distribution parameters (not just the mean), leading to more precise and realistic normative estimates.[36,37] Finally, adjusting NGN for sex, along with age and education, has likely contributed to their improved accuracy.[38]

These results have relevant implications for clinical practice and research. Current diagnostic criteria and expert consensus recommend that, in clinical settings, AD biomarkers be reserved for cognitively impaired individuals.[8,9] Our findings suggest that NGN could support clinical decision-making by enhancing the identification of individuals who qualify for biomarker testing. Additionally, because

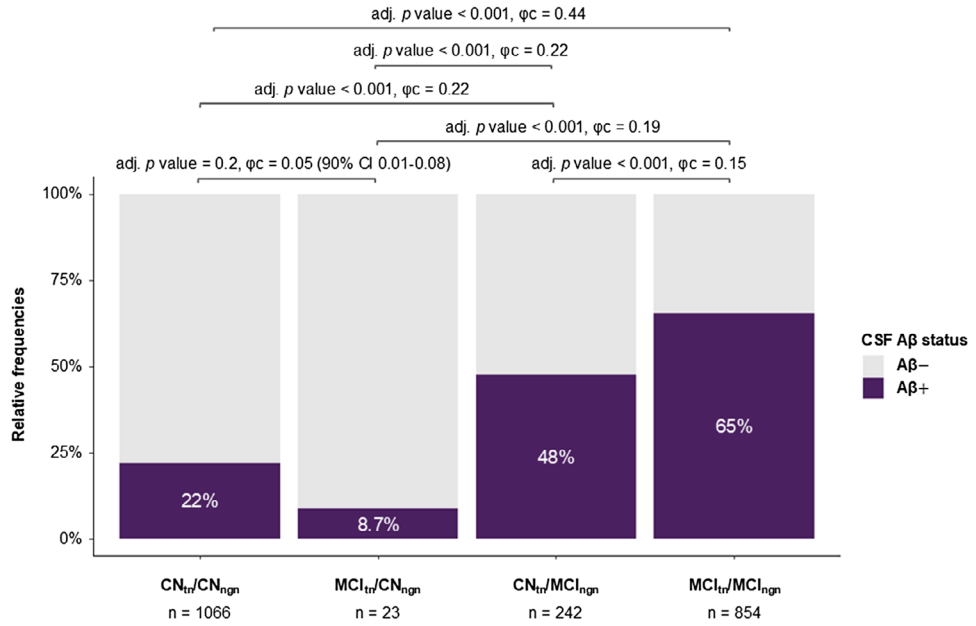

FIGURE 2 Proportions of CSF Aβ+ participants across actuarially reclassified diagnostic groups and Bonferroni-corrected pairwise comparisons. Aβ, amyloid β; Aβ−, Aβ-negative; Aβ+, Aβ-positive; adj. *p* value, Bonferroni-adjusted *p* value; CI, confidence interval; CN, cognitively normal; CSF, cerebrospinal fluid; MCI, mild cognitive impairment; ngn, next-generation norms; tn, traditional norms; φc, Cramér's V.

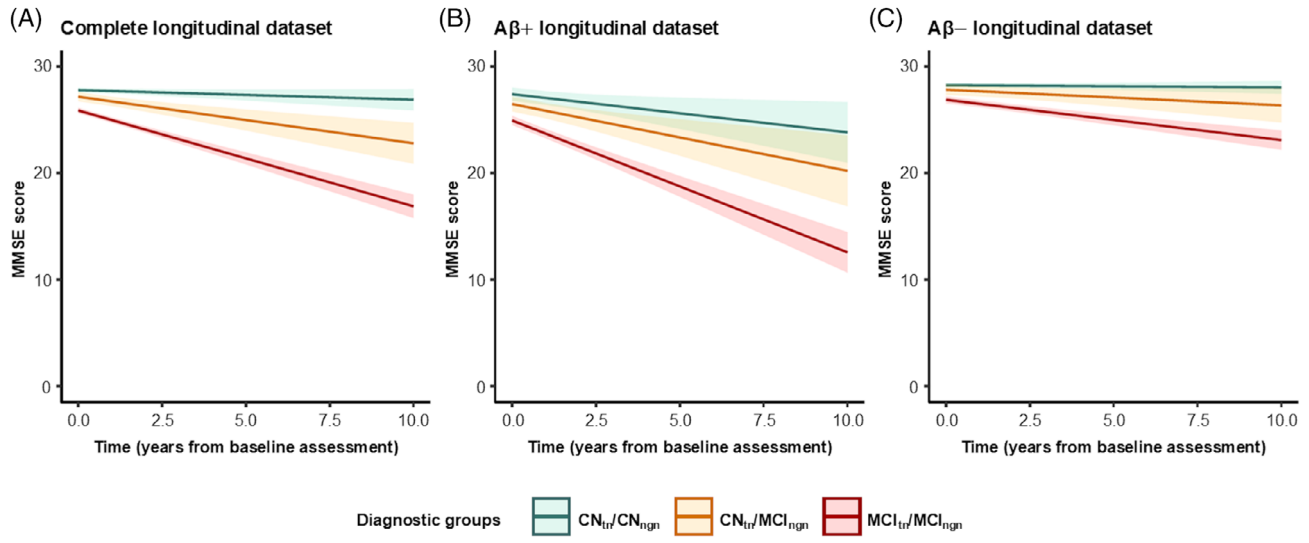

FIGURE 3 Longitudinal trajectories of MMSE scores across actuarially reclassified diagnostic groups. LME model estimates of temporal changes in MMSE scores are shown for the complete longitudinal dataset (A), and for the Aβ+ (B) and Aβ− (C) subsamples. The MCI_{tn}/CN_{ngn} group was not included in the models due to insufficient sample size for reliable analysis. The shaded areas represent 95% CI. Aβ, amyloid β; Aβ−, Aβ-negative; Aβ+, Aβ-positive; CI, confidence interval; CN, cognitively normal; LME, linear mixed-effects; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; ngn, next-generation norms; tn, traditional norms.

copathology occurs frequently, it is recommended that a positive biomarker result should not lead to a diagnosis of AD unless backed by a suggestive clinical phenotype, including a consistent profile of cognitive impairment.[6,8,9,39] In this regard, a comprehensive neuropsychological evaluation covering the main cognitive domains (i.e., memory, language, attention/executive, and visuospatial functions) is essential to capture the nature and extent of the deficits.[40] NGN contribute to the characterization of cognitive impairment and the assessment of severity across these primary domains, thereby informing the interpretation of biomarkers in routine clinical practice. Importantly, timely and accurate detection of AD-related cognitive impairment enables earlier treatment, which may improve patient outcomes and boost the ability of clinical trials to detect drug efficacy.[7,41]

This study has limitations. The first limitation is that the validation sample is not completely independent of the normative dataset. The preceding effort to maximize the sample size of the normative dataset made it unfeasible to obtain a contemporaneous, entirely independent validation sample. However, more than two thirds of the participants (1507 out of 2185, 69%) involved in assessing the utility of NGN were independent of the normative sample. We judged that maintaining the overlapping normative participants in the validation sample (678 out of 2185, 31%) was the best available option, as excluding them would likely have introduced other limitations. Notably, this would have biased the Aβ+ prevalence in the validation sample, particularly in the CN_{tn}/CN_{ngn} group, by enriching it with Aβ+ participants above the expected prevalence in the general population.[42] This overrepresentation of Aβ+ CN_{tn}/CN_{ngn} individuals could dilute the comparative results. Thus, although the lack of entirely independent samples is a limitation, we believe our approach offers the most reliable validation given the constraints. Second, the absence of autopsy or biomarker confirmation for non-AD neurodegenerative diseases limits the interpretability of our findings. To address this limitation, we used longitudinal analyses to capture cognitive deterioration over time as a surrogate marker of neurodegeneration. Finally, because the MCI_{tn}/CN_{ngn} subgroup—providing the most direct test of specificity—was small, and only a minority had at least 1-year follow-up data, our study cannot definitively quantify specificity. Future research should aim to replicate these findings in an independent sample. In addition, future studies could incorporate base-rate approaches that explicitly model false-positive rates[43] and longitudinal clinical outcomes to provide a more rigorous estimate of specificity.

In conclusion, our findings indicate that, by integrating methodological strengths, NGN improve the detection of cognitive impairment, particularly AD-related cognitive impairment. Implementing NGN in clinical and research settings could refine decision-making, enhance clinical trial design, and ultimately improve patient outcomes.

ACKNOWLEDGMENTS

We want to express our gratitude to all participants of this study and all the clinical and research teams involved in their assessment. Alexandre Bejanin acknowledges support from Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union through the Miguel Servet grant (CP20/00038), and the Alzheimer's Association (AARG-22-923680). Mircea Balasa reported receiving support from ISCIII (PI22/00343). Javier Arranz was supported by ISCIII through the Río Hortega Fellowship CM21/00243 and co-funded by the European Union. Nuole Zhu was supported by contract CM21/00113 from ISCIII, co-funded by the European Union. Juan Fortea reports grants from Fundación Tatiana Pérez de Guzmán el Bueno (IIBSP-DOW-2020-151). He was also supported by Horizon2020–Research and Innovation Framework Programme from the European Union (H2020-SC1-BHC-2018-2020). Raquel Sánchez-Valle reported receiving funding from ISCIII (20/00448). Pablo Martínez-Lage reported receiving funding from the Department of Economic Promotion of the Provincial Government of Gipuzkoa (124/16), the Department of Health of the Basque Government (2016111096, S-PR12CH001, S-PR13ZH001), and ISCIII (PI12/02262, PI15/00919). Ignacio Illán-Gala was supported by the Juan Rodés Contract (JR20/0018) from ISCIII, partly funded by the European Social Fund. He is a senior Atlantic Fellow for Equity in Brain Health at the Global Brain Health Institute (GBHI) and receives funding from the Alzheimer's Association (AACSF-21-850193), and the Alzheimer Society (GBHI ALZ UK-21-72097). The SPIN cohort received funding from Fondo de Investigaciones Sanitario, ISCIII (PI13/01532, PI14/01126, PI16/01825, PI17/01019, PI17/01896, PI18/00335, PI18/00435, PI19/00882, PI20/01473, PI20/00836, PI21/01395, PI21/00063, PI21/00791, PI22/00611, PI22/00307, PI23/01786, PI24/00598, PI24/00968, PI24/01087, INT19/00016, INT21/00073, INT23/00048, AC19/00103, DTS22/00111) and the CIBERNED program (Program 1, Alzheimer Disease), jointly funded by Fondo Europeo de Desarrollo Regional, Unioń Europea, "Una manera de hacer Europa". The SPIN cohort was also supported by the National Institutes of Health grants (1R01AG056850-01A1, R21AG056974, R01AG061566, 1R01AG081394-01, 1R61AG066543-01, and R01AG080470), by Generalitat de Catalunya (2017-SGR-547, 2021-SGR-00979, SLT006/17/125, SLT006/17/119, SLT002/16/408), Fundació La Marató de TV3 (20141210, 044412, and 20142610), Ajuntament de Barcelona in collaboration with Fundació 'la Caixa' (23S06157-001), Fundació Catalana Síndrome de Down, and Fundació Víctor Grífols i Lucas. Horizon 21 Consortium is partly funded by Jérôme Lejeune Foundation. The GAP study was supported by CITA-Alzheimer Foundation, ISCIII (PI12/02262 and PI1500919), and the Department of Health of the Basque Government (S-PR12CH001 and S-PR13ZH001). The DEBA project was partially funded by the Department of Economic Promotion of the Provincial Government of Gipuzkoa (2016 Support Programme for the Gipuzkoa Network for Science, Technology, and Innovation) and the Department of Health of the Basque Government (2016111096).

CONFLICT OF INTEREST STATEMENT

Sara Rubio-Guerra reported honoraria for educational events from Esteve Pharmaceuticals. María Belén Sánchez-Saudinós reported receiving travel support for attending meetings from Lilly and serving on an advisory board for Novo Nordisk. Miren Altuna reported receiving honoraria for educational events from Lilly, Esteve, ITAMAR, Almirall, and KERN Pharma; as well as support for attending meetings or travel from Lilly, Esteve, and Almirall. Javier Arranz reported receiving honoraria for educational events from Roche, Esteve, and Lilly, and support for attending meetings or travel from Esteve Pharmaceuticals, S.A. Daniel Alcolea reported receiving consulting fees from Grifols S.A., Lilly, Fujirebio-Europe, and Roche Diagnostics; honoraria for educational events from Fujirebio-Europe, Roche Diagnostics, Nutricia, Krka Farmacéutica S.L., Zambon S.A.U., Esteve Pharmaceuticals, S.A., Neuraxpharm, Alter, and Lilly; and support for attending meetings or travel from Fujirebio-Europe, Lilly, and Nutricia. Daniel Alcolea also participated on advisory boards for Grifols S.A., Lilly, Fujirebio-Europe, and Roche Diagnostics. Alberto Lleó reported receiving consulting fees from Eisai, Esteve, Fujirebio-Europe, Roche, Grifols S.A., and Lilly. Juan Fortea reported receiving consulting fees from Lundbeck, Ionis, and AC Immune; honoraria for educational events from Roche, Esteve,

Biogen, Laboratorios Carnot, Adamed, Life Molecular Imaging, Eisai, and Lilly. Juan Fortea has participated on advisory boards for AC Immune, Alzheon, Zambon, Lilly, Roche, Eisai, and Perha. Juan Fortea has held leadership or fiduciary roles in the Spanish Neurological Society, T21 Research Society, Lumind Foundation, Jérôme-Lejeune Foundation, Alzheimer's Association, Health Research Board, and Dementia Trials Ireland (all without payments). Juan Fortea also participated in Study Sections for the European Commission, National Institutes of Health, and Instituto de Salud Carlos III, with payments received for these activities. Juan Fortea reports receipt of equipment, materials, or other services from Life Molecular Imaging, provided to his institution. Daniel Alcolea, Alberto Lleó, and Juan Fortea declare a filed patent application (WO2019175379 A1 Markers of synaptopathy in neurodegenerative disease). Raquel Sánchez-Valle reported receiving consulting fees from Lilly (paid to her institution) and UCB (paid personally). Pablo Martínez-Lage reported receiving consulting fees from Lilly, Roche, and Eisai; honoraria for educational events from Lilly, Roche, Eisai, Nutricia, and Almirall; and support for attending meetings or travel from Lilly, Roche, Eisai, and Nutricia. Ignacio Illán-Gala reported receiving honoraria for educational events from Esteve, Sociedad Española de Neurología, Societat Catalana de Neurologia, Lilly, Kern Pharma, and Almirall; support for attending meetings or travel from Esteve and Almirall; and participation on advisory boards for UCB and Nutricia. The remaining authors declare no conflicts of interest. Author disclosures are available in the Supporting Information.

CONSENT STATEMENT

Written informed consent was obtained from all human subjects before enrollment in the study.

ORCID

Sara Rubio-Guerra https://orcid.org/0000-0001-7652-8029
Ignacio Illán-Gala https://orcid.org/0000-0002-5418-2052

REFERENCES

1. World Health Organization. *Dementia n.d*. World Health Organization. (Accessed September 5, 2025). https://www.who.int/news-room/fact-sheets/detail/dementia
2. Gustavsson A, Norton N, Fast T, et al. Global estimates on the number of persons across the Alzheimer's disease continuum. *Alzheimers Dement*. 2023;19:658-670. doi:10.1002/alz.12694
3. Cummings J, Rabinovici GD, Atri A, et al. Aducanumab: appropriate use recommendations update. *J Prev Alzheimers Dis*. 2022;9:221-230. doi:10.14283/jpad.2022.34
4. Cummings J, Apostolova L, Rabinovici GD, et al. Lecanemab: appropriate use recommendations. *J Prev Alzheimers Dis*. 2023;10:362-377. doi:10.14283/jpad.2023.30
5. Rabinovici GD, Selkoe DJ, Schindler SE, et al. Donanemab: appropriate use recommendations. *J Prev Alzheimers Dis*. 2025;12:100150. doi:10.1016/j.tjpad.2025.100150
6. Dubois B, Von Arnim CAF, Burnie N, Bozeat S, Cummings J. Biomarkers in Alzheimer's disease: role in early and differential diagnosis and recognition of atypical variants. *Alzheimers Res Ther*. 2023;15:175. doi:10.1186/s13195-023-01314-6
7. Cummings J, Ritter A, Zhong K. Clinical trials for disease-modifying therapies in Alzheimer's disease: a primer, lessons learned, and a blueprint for the future. *J Alzheimers Dis*. 2018;64:S3-S22. doi:10.3233/JAD-179901
8. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement*. 2024;20(8):5143-5169. doi:10.1002/alz.13859
9. Dubois B, Villain N, Schneider L, et al. Alzheimer disease as a clinical-biological construct—an International Working Group Recommendation. *JAMA Neurol*. 2024;81:1304. doi:10.1001/jamaneurol.2024.3770
10. Rubio-Guerra S, Sánchez-Saudinós MB, Sala I, et al. Development of amyloid-negative neuropsychological norms using GAMLSS. *Alzheimer's Dement*. 2025; e70224. doi:10.1002/dad2.70224
11. Clark LR, Delano-Wood L, Libon DJ, et al. Are empirically-derived subtypes of mild cognitive impairment consistent with conventional subtypes? *J Int Neuropsychol Soc*. 2013;19:635-645. doi:10.1017/S1355617713000313
12. Schinka JA, Loewenstein DA, Raj A, et al. Defining mild cognitive impairment: impact of varying decision criteria on neuropsychological diagnostic frequencies and correlates. *Am J Geriatr Psychiatry*. 2010;18:684-691. doi:10.1097/JGP.0b013e3181e56d5a
13. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256:183-194. doi:10.1111/j.1365-2796.2004.01388.x
14. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on mild cognitive impairment. *J Intern Med*. 2004;256:240-246. doi:10.1111/j.1365-2796.2004.01380.x
15. Edmonds EC, Delano-Wood L, Clark LR, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimers Dement*. 2015;11:415-424. doi:10.1016/j.jalz.2014.03.005
16. Edmonds EC, Delano-Wood L, Jak AJ, Galasko DR, Salmon DP, Bondi MW. "Missed" mild cognitive impairment: high false-negative error rate based on conventional diagnostic criteria. *J Alzheimers Dis*. 2016;52:685-691. doi:10.3233/JAD-150986
17. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry*. 2009;17:368-375. doi:10.1097/JGP.0b013e31819431d5
18. Bondi MW, Edmonds EC, Jak AJ, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimers Dis*. 2014;42:275-289. doi:10.3233/JAD-140276
19. Thomas KR, Edmonds EC, Eppig JS, et al. MCI-to-normal reversion using neuropsychological criteria in the Alzheimer's disease neuroimaging initiative. *Alzheimers Dement*. 2019;15:1322-1332. doi:10.1016/j.jalz.2019.06.4948
20. Alcolea D, Martínez-Lage P, Sánchez-Juan P, et al. Amyloid precursor protein metabolism and inflammation markers in preclinical Alzheimer disease. *Neurology*. 2015;85:626-633. doi:10.1212/WNL.0000000000001859
21. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatric Res*. 1975;12:189-198. doi:10.1016/0022-3956(75)90026-6
22. Peña-Casanova J, Blesa R, Aguilar M, et al. Spanish multicenter normative studies (NEURONORMA Project): methods and sample characteristics. *Arch Clin Neuropsychol*. 2009;24:307-319. doi:10.1093/arclin/acp027
23. Peña-Casanova J, Casals-Coll M, Quintana M, et al. Spanish normative studies in a young adult population (NEURONORMA young adults project): methods and characteristics of the sample. *Neurología*. 2012;27:253-260. doi:10.1016/j.nrleng.2011.12.008

24. R Core Team. *R: A Language and Environment for Statistical Computing*. R Core Team; 2023.
25. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed.. L. Erlbaum Associates; 1988.
26. Shishkina T, Farmus L, Cribbie RA. Testing for a lack of relationship among categorical variables. *Quant Meth Psych*. 2018;14:167-179. doi:10.20982/tqmp.14.3.p167
27. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7:270-279. doi:10.1016/j.jalz.2011.03.008
28. Sarazin M, Berr C, Fabrigoule C, et al. Amnestic syndrome of the medial temporal type identifies prodromal AD. *Neurology*. 2007;69(19):1859-1867.
29. Wagner M, Wolf S, Reischies FM, et al. Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology*. 2012;78:379-386. doi:10.1212/WNL.0b013e318245f447
30. Van Der Veere PJ, Hoogland J, Visser LNC, et al. Predicting cognitive decline in amyloid-positive patients with mild cognitive impairment or mild dementia. *Neurology*. 2024;103:e209605. doi:10.1212/WNL.0000000000209605
31. Baker JE, Lim YY, Pietrzak RH, et al. Cognitive impairment and decline in cognitively normal older adults with high amyloid-β: a meta-analysis. *Alzheimer's Dementia*. 2017;6:108-121. doi:10.1016/j.dadm.2016.09.002
32. Insel PS, Donohue MC, Sperling R, Hansson O, Mattsson-Carlgren N. The A4 study: β-amyloid and cognition in 4432 cognitively unimpaired adults. *Ann Clin Transl Neurol*. 2020;7:776-785. doi:10.1002/acn3.51048
33. Hahn A, Kim YJ, Kim HJ, et al. The preclinical amyloid sensitive composite to determine subtle cognitive differences in preclinical Alzheimer's disease. *Sci Rep*. 2020;10:13583. doi:10.1038/s41598-020-70386-3
34. Bos I, Vos SJB, Jansen WJ, et al. Amyloid-β, tau, and cognition in cognitively normal older individuals: examining the necessity to adjust for biomarker status in normative data. *Front Aging Neurosci*. 2018;10:193. doi:10.3389/fnagi.2018.00193
35. Boyle PA, Wang T, Yu L, et al. To what degree is late life cognitive decline driven by age-related neuropathologies? *Brain*. 2021;144:2166-2175. doi:10.1093/brain/awab092
36. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *J R Stat Soc C: Appl Stat*. 2005;54:507-554.
37. Timmerman ME, Voncken L, Albers CJ. A tutorial on regression-based norming of psychological tests with GAMLSS. *Psychol Methods*. 2021;26:357-373. doi:10.1037/met0000348
38. Sundermann EE, Maki P, Biegon A, et al. Sex-specific norms for verbal memory tests may improve diagnostic accuracy of amnestic MCI. *Neurology*. 2019;93:e1881-e1889. doi:10.1212/WNL.0000000000008467
39. Dubois B, Villain N, Frisoni GB, et al. Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol*. 2021;20:484-496. doi:10.1016/S1474-4422(21)00066-1
40. Weintraub S. Neuropsychological assessment in dementia diagnosis. *Continuum*. 2022;28:781-799. doi:10.1212/CON.0000000000001135
41. Dubois B, Padovani A, Scheltens P, Rossi A, Dell'Agnello G. Timely diagnosis for Alzheimer's disease: a literature review on benefits and challenges. *J Alzheimers Dis*. 2015;49:617-631. doi:10.3233/JAD-150692
42. Jansen WJ, Ossenkoppele R, Knol DL, et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA*. 2015;313:1924. doi:10.1001/jama.2015.4668
43. Oltra-Cucarella J, Sánchez-SanSegundo M, Rubio-Aparicio M, Arango-Lasprilla JC, Ferrer-Cascales R. The association between the number of neuropsychological measures and the base rate of low scores. *Assessment*. 2021;28:955-963. doi:10.1177/1073191119864646

SUPPORTING INFORMATION

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

How to cite this article: Rubio-Guerra S, Sala I, Sánchez-Saudinós MB, et al. Amyloid-negative neuropsychological norms: Added value in the era of biomarkers and disease-modifying therapies. *Alzheimer's Dement*. 2025;17:e70223. https://doi.org/10.1002/dad2.70223