



HAL
open science

Recognition memory decline is associated with the progression to prodromal Alzheimer's disease in asymptomatic at-risk individuals

Filipa Raposo Pereira, Maximilien Chaumon, Bruno Dubois, Hovagim Bakardjian, Mahsa Bahrami, Marie-Odile Habert, Katia Andrade, Nadjia Younsi, Valentina La Corte, Nathalie George

► To cite this version:

Filipa Raposo Pereira, Maximilien Chaumon, Bruno Dubois, Hovagim Bakardjian, Mahsa Bahrami, et al.. Recognition memory decline is associated with the progression to prodromal Alzheimer's disease in asymptomatic at-risk individuals. *Journal of Neurology*, 2024, 272 (70), <10.1007/s00415-024-12834-y>. <hal-05405398>

HAL Id: hal-05405398

<https://hal.science/hal-05405398v1>

Submitted on 8 Dec 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Recognition memory decline is associated with the progression to prodromal Alzheimer's disease in asymptomatic at-risk individuals

Filipa Raposo Pereira^{1,2} · Maximilien Chaumon¹ · Bruno Dubois^{1,2,3} · Hovagim Bakardjian^{1,2} · Mahsa Bahrami² · Marie-Odile Habert^{4,5,6} · Katia Andrade² · Nadjia Younsi² · Valentina La Corte^{2,7,8} · Nathalie George¹ on behalf of for the INSIGHT-preAD study group

Received: 25 September 2024 / Revised: 22 November 2024 / Accepted: 24 November 2024 / Published online: 16 December 2024
© Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

Episodic memory (EM) alterations are a hallmark of Alzheimer's disease (AD). We assessed EM longitudinally in cognitively normal elders at-risk for AD (with subjective memory complaints), as a function of amyloid- β ($A\beta$) burden, neurodegeneration (N), and progression to prodromal AD. We stratified 264 INSIGHT-preAD study subjects in controls ($A\beta$ -/N-), stable/N- or N+ ($A\beta$ +), and progressors/N- or N+ ($A\beta$ + groups (progressors were included only until AD-diagnosis). We used linear mixed-effect models with $A\beta$ and N status, or progression to AD as factors, to analyze behavioral performance in an old/new word-recognition task based on the free and cued selective reminding test (FCSRT). The controls and stable/N- groups showed near-ceiling accuracy and RT improvement across follow-up. The stable/N+ group showed accuracy reduction and no RT improvement, i.e., $A\beta$ +/N+ cumulative effect. The progressors showed a marked performance decline. EM alterations may constitute early preclinical markers of progression to prodromal AD, while individuals are cognitively normal according to neuropsychological standards.

Keywords Alzheimer's disease (AD) · Episodic memory · β -amyloid · Neurodegeneration · Free and cued selective reminding test (FCSRT) · Old/new effect

Introduction

Alzheimer's disease (AD) is the most dominant form of dementia in elderly people [1, 2]. It is characterized by progressive cognitive decline and is currently the major public health problem in elderly populations, with a prevalence

estimated in > 50 million people worldwide, expected to double at every 20 years [1, 2]

In typical forms, AD tends to be diagnosed at prodromal or dementia stages, with most existing knowledge on AD-linked cognitive decline, stemming from research performed at these symptomatic stages [2–6]. Therefore, the pathophysiological processes leading to the onset of cognitive

Valentina La Corte and Nathalie George contributed equally.

✉ Filipa Raposo Pereira
filipa.raposopereira@icm-institute.org

Marie-Odile Habert
http://www.cati-neuroimaging.com

¹ Sorbonne Université, Institut du Cerveau-Paris Brain Institute-ICM, INSERM, U 1127, CNRS, UMR 7225, AP-HP, CENIR, Centre MEG-EEG, Hôpital de La Pitié-Salpêtrière, 47 Boulevard de L'Hôpital, 75013 Paris, France

² Department of Neurology, Institute of Memory and Alzheimer's Disease (IM2A), Pitié-Salpêtrière University Hospital, AP-HP, Paris, France

³ Department of Neurology, Centre of Excellence of Neurodegenerative Disease (CoEN), ICM, CIC

Neurosciences, APHP, Hôpital Pitié-Salpêtrière, Sorbonne Université, Paris, France

⁴ Sorbonne Université, CNRS, INSERM, Laboratoire d'Imagerie Biomédicale, LIB, 75006 Paris, France

⁵ AP-HP, Hôpital de la Pitié-Salpêtrière, Médecine Nucléaire, 75013 Paris, France

⁶ Centre d'Acquisition et Traitement des Images (CATI), Paris, France

⁷ Laboratoire Mémoire Cerveau et Cognition (MC2Lab) EA 7536, Institut de Psychologie, Université Paris Cité, Boulogne-Billancourt, France

⁸ Institut Universitaire de France, Paris, France

decline in AD remain poorly understood, resulting in a shift in the focus of AD-research toward earlier preclinical stages preceding symptomatology —asymptomatic at-risk phase [2, 4, 6–9]. In this study, individuals were cognitively normal but considered at-risk of progression to prodromal-AD for reporting subjective memory complaints (SMC) for >6 months, with or without presenting AD neuropathophysiological alterations [2, 4, 5, 7, 9, 10]. However, not all at-risk individuals progress to AD, highlighting the need for an in-depth characterization of the preclinical asymptomatic phase, missing at present [2, 7, 11].

At the biologic level, the diagnosis of typical-AD requires the presence of neuropathophysiological markers such as β -amyloid [$A\beta$] burden (i.e., neocortical accumulation of $A\beta$ plaques in positron emission tomography [PET]) or neurodegeneration (as reflected by reduced neuro-metabolic glucose uptake) [7, 12–16]. These are considered AD-risk factors, which in-vivo evidence shows to silently appear several years before cognitive alterations are established, and to correlate with the appearance of SMC and the escalation of episodic memory (EM) deficits, a hallmark of AD [7, 9, 12, 16–18]. Moreover, when observed together, these markers seem to further exacerbate memory decline and progression to AD [7, 15, 16, 18–20]. Hence, such AD-markers might anticipate cognitive decline, as proposed in the 3-stage framework of Sperling et al.: (1) presence of cerebral amyloidosis (measured as either $A\beta$ burden in PET imaging or decreased $A\beta_{1-42}$ concentration in cerebrospinal fluid [CSF]); (2) presence of amyloidosis and neurodegeneration (considered as either hypometabolism in ^{18}F -fluorodeoxyglucose (^{18}F FDG) PET or increased tau level in CSF, and medial temporal lobe atrophy); (3) presence of amyloidosis, neurodegeneration, and subjective cognitive complaints [4, 7, 13, 18].

In typical AD, subjective cognitive complaints largely reflect the onset of EM deficits [2, 3, 9, 11, 21]. These are coined as SMC, an umbrella term for potential alterations in different EM processes, including encoding, retrieval, consolidation, and recognition [2, 3, 8, 11, 22]. However, SMC are common in both AD and normal cognitive aging populations, with the difference lying on the source of the encountered difficulties [8, 9, 21–25]. In normal cognitive aging, memory alterations are mostly related to inefficient encoding capacity or impairment of retrieval processes originated in executive decline [8, 9, 21–25]. It contrasts with typical AD, where memory alterations are mostly related to storage and consolidation decline, considered a signature of the amnesic hippocampal syndrome (AHS) characteristic of AD [8, 9, 21–25]. Preclinically, research shows that SMC are the first cognitive symptom to accelerate and the most predictive of future AD, starting 4–8 years before executive decline, and 7–10 years before other cognitive decline [8, 10, 25]. Nevertheless, it remains challenging to distinguish

normal cognitive aging vs. AD at the asymptomatic at-risk stage, when SMC are very subtle.

The Free and Cued Selective Reminding Test (FCSRT) is one of the most validated tools for the assessment of EM. When EM deficits start to be objectively perceived, the FCSRT has proven successful on distinguishing deficits linked to AD vs. normal cognitive aging [8, 24–27]. One strength of the test is the use of semantic cues at both encoding and recall stages—i.e., pairing a word to be remembered (e.g., grapes) with a corresponding semantic cue (e.g., fruit), which controls the learning process of the to-be-remembered words and consequently maximizes retrieval [8, 10, 24, 25, 27]. These play an important role on the distinction between cognitive aging participants, who can profit from cue presentation to improve their retrieval capacity, and individuals with AD-AHS, who fail to do so [8, 21, 25–30]. In practice, this can be observed on the main scores of the FCSRT, the free recall (FR) and the total recall (TR) [8, 24–26]. FR, without using semantic cues, evaluates the accessibility to stored material [8, 24–26, 31]. TR (the sum of FR and cued recall) evaluates the availability of stored material and is considered a signature of the AD-AHS [8, 11, 24–26, 31]. While FR tends to decline 6–7 years prior to AD-diagnosis, TR tends to parallel cognitive decline typically around 2–3 years prior diagnosis [8, 10, 11, 25, 31]. However, FR alterations in isolation are not AD-specific, they mostly reflect executive dysfunction, which can result from various deficits [2, 8, 11, 21, 31, 32]. It is only when concomitant TR alterations are detected that EM deficits can be conclusive of AD [2, 8, 21, 24]. Moreover, TR is prone to ceiling effects because the FCSRT uses only 1-list of words, restricting the comparison of its scores to cognitive aging equivalents [3, 8, 27, 31]. Hence, the detection of subtle variations in maximum recall, crucial to evaluate TR, tend to be delayed to a late preclinical phase when semantic cues start failing to overcome retrieval deficits and EM problems become evident [8, 10, 31]. Therefore, despite its validated capacity distinguishing the source of EM deficits, the FCSRT diagnosis value is restricted to the clinical phase and lacks the sensitivity to detect the onset of subtle memory alterations at an asymptomatic at-risk phase [3, 9].

Alternatively, computerized versions of old/new list-learning recognition tasks tend to be more sensitive to subtle EM deficits, allowing to record response accuracy and reaction time (RT) [2, 8, 27, 30, 33, 34]. More accurate and faster recognition of old vs. new words is the typical old/new effect, maintained even during cognitive aging [10, 27, 35]. In contrast, individuals with AHS tend to have trouble distinguishing old from new words even when semantic cues are used, often resulting from encoding deficits, which tend to impel the adoption of biased response strategies to support the old/new decision-making [2, 9, 21, 24, 30, 36–38]. Decision-making can be evaluated using signal detection

theory (SDT), with d' assessing the capacity to discriminate old and new words and response criterion assessing the bias toward predominance of either old or new responses [23, 30, 33, 36–38]. In AD, low d' and variable criterion are interpreted as a direct failure in profiting from cue-assisted learning and a sign of EM deficits of the AHS nature [9, 23, 30, 35, 36].

Here, we used an old/new word-recognition task based on the verbal content of the FCSRT. First, the task allowed profiting from semantic cues to maximize consolidation [8, 21, 24]. Second, to differentiate the nature of EM deficits introducing two types of new words, which were related to the old words (encoded during the FCSRT) either in their number of repetitions (New Repeated [NR] words), or in their semantic content (New Unique [NU] words). Research on the subtle onset of EM alterations, their development, and their association with AD-biomarkers such as A β burden or neurodegeneration, is currently insufficient. Within the INSIGHT-preAD 5-year longitudinal observational cohort with extensive multimodal data, we aimed at finding behavioral signatures (i.e., accuracy, RT, d' , and criterion) of the onset of subtle EM alterations, which can predict in cognitively normal individuals at-risk (SMC > 6-months) who will progress to an AD-diagnosis [2].

Our objective was threefold: (1) to retrospectively evaluate the behavioral performance to our old/new word-recognition task, to assess its efficacy at detecting the onset of subtle EM alterations in cognitively normal elder individuals with SMC, at baseline (i.e., M0) and longitudinally (i.e., 5 years follow-up period; M12, M24, M36, M48, M60); (2) to assess A β burden and neurodegeneration effects on the behavioral performance during this task; (3) to assess the progression to prodromal-AD effect on the behavioral performance of 15 individuals who progressed to prodromal-AD during the follow-up.

Methods

Participants

The INSIGHT-preAD study is a monocentric observational 5-year follow-up study, comprising 318 cognitively normal elderly human individuals with SMC recruited from the Institute for Memory and Alzheimer's disease (IM2A) at the Pitié-Salpêtrière Hospital in Paris [2, 39]. Inclusion criteria consisted in the following: age of 70–85 years; 6 months of SMC; normal global efficiency assessed with the minimal state [MMSE] ($\geq 27/30$ score); normal EM assessed with the FCSRT (TR $\geq 44/48$ score); clinical dementia rating (CDR) score of 0; good visual and auditory acuity [2, 8, 24, 39–41]. Participants were excluded from the study if under guardianship or in a nursing facility, diagnosed

with prodromal AD or any other neurologic disease, illiterate, or meeting any impossibility to perform a brain MRI. Drop-out from the study was possible at any time if participants wished so. Participation in the study was automatically ceased in case of diagnosis with another condition or death. In the case where the participant presented cognitive decline in two consecutive neuropsychological assessments (in MMSE, CDR and/or FCSRT-TR scores, performed at 6 month interval), an independent committee of two neurologists, a neuropsychologist, and a neuroimaging expert blind to the participant's A β burden status, reviewed their medical file and provided a diagnosis [42]. If prodromal or AD dementia was confirmed, the participant exited the INSIGHT-preAD cohort.

Procedure

The INSIGHT-preAD is a longitudinal 5-year follow-up initiative based on an in-depth multimodal approach with a battery of demographic, neuropsychological, neurobiologic, genetic, neuroimaging, and electrophysiological assessments, to fully characterize elderly subjects at-risk for developing AD (SMC > 6-months; see Online Resource 1, for details) [2, 39]. In this study, we used a subset of the data consisting of:

- ^{18}F -florbetapir PET scans (assessing deposition of A β plaques);
- ^{18}F FDG PET scans (assessing brain metabolism) collected at baseline (M0);
- Neuropsychologic assessments (MMSE assessing cognitive functioning; FCSRT assessing EM; Frontal Assessment Battery [FAB] assessing executive functioning; Trail Making Test [TMT] assessing visual attention and processing speed);
- Behavioral data from the old/new verbal recognition task, collected yearly from baseline. This task was performed with EEG recording, analyzed outside the scope of the present study and not presented here [2, 8, 24, 39–41].

Both neuropsychologic assessments and behavioral task were performed on the same day, with the neuropsychological evaluation in the morning and the behavioral task in the afternoon, within an interval of 4–6 h.

PET acquisitions

The standardized uptake value ratio (SUVr) of β -amyloid burden was measured with a Philips Gemini GXL CR-PET scanner at session M0, 50 min after an injection of 370 MBq (10 mCi) ^{18}F -florbetapir (neocortical marker of β -amyloid plaques deposition). A SUVr cut-off > 0.7918 distinguished A β + from A β - individual status, based on a

linear conversion of the INSIGHT-preAD CAEN method [2, 42]. The value of 0.7918 was obtained by mean averaging the SUVr of cortical regions of interest (i.e., left, and right precuneus, posterior cingulum, anterior cingulum, associative parietal cortex, associative temporal cortex and orbito-frontal cortex) with respect to a reference region (i.e., whole cerebellum plus pons) [42].

Since low glucose uptake (i.e., hypometabolism) is considered a marker of neurodegeneration, the ^{18}F -FDG uptake SUVr was measured 30 min after an injection of 2 MBq/kg ^{18}F -FDG. An SUVr cut-off < 2.27 was set to distinguish neurodegeneration-positive (N+) from neurodegeneration-negative (N-) individuals [2, 14]. Based on a similar method as above, the value of 2.27 was obtained by mean averaging the ^{18}F -FDG SUVr of regions of interest (i.e., posterior cingulate cortex, inferior parietal lobe, precuneus, inferior temporal gyrus) with respect to a reference region (here, the pons).

Neuropsychological assessments

Education level was assessed with the French socio-cultural (NSC) scale, which measures education in 8 levels, the level 1 corresponding to primary school and the level 8 to higher education—i.e., PhD level. Cognitive functioning was assessed with the MMSE [40]. The 30-points examination consists in the evaluation of the answers given to several questions assessing: orientation (time and place); memory; attention and problem solving; language; and visual-spatial skills [40]. A score below 27 is suggestive of cognitive impairment and probable dementia, requiring a more detailed cognitive examination [40]. Executive function was assessed with the FAB test, which allows discriminating frontal dementias from AD [43]. The examination consists of 6 categories scored from 0 to 3: conceptualization and abstract reasoning; mental flexibility; interference sensitivity; inhibitory control; motor programming; and environmental autonomy [43]. From a maximum score of 18, 12 is considered a cut-off score indicative of AD [43]. Visual attention and cognitive flexibility were assessed with the TMT (time B-A). The 2-part test (A and B) evaluates visual search, task monitoring, speed of processing, mental flexibility, and executive functions [44]. In part A, participants are presented with a paper with encircled numbers from 1 to 25, which must be sequentially connected by lines. The procedure is similar for part B, except that the numbers alternate with letters. The score measures the time required to complete the test, where an average score (time for B minus time for A) should be ± 46 s and an overall score > 190 s is considered deficient [44]. Subjective perception of memory deficits or SMC was assessed with the modified 15-item version of the McNair Frequency of Forgetting Questionnaire [45]. A list of 15 examples of memory deficits was

separately presented to the participant and a companion; the individuals were required to indicate the severity and regularity with which these deficits happened to the participant, from rarely to always. Scores ≥ 15 are associated with abnormally high levels of SMC [45].

As for the FCSRT (French version), it began by the consecutive presentation of four cards, each containing 4 different words, which amounts to a total of 16 words to be remembered [24, 25, 46, 47]. In each card, participants were asked to identify the item (e.g., grapes) that corresponds to a unique category cue (e.g., fruit) provided orally. After the presentation of the 4 cards (encoding and learning phase), 3 immediate FR trials of the same items were performed in alternation with an interference task of 20 s of backward counting. If some items were not retrieved after each FR trial, the same semantic cues as before were presented in immediate cued recall (CR) trials until the missing item was retrieved. Therefore, each immediate recall trial was composed by FR and CR if any item was not spontaneously retrieved. Then, delayed FR and CR trials were performed 20 min later. The two main FCSRT measures are: the FR score (i.e., the cumulative sum of all FR trials; range 0–48) and the TR score (i.e., the cumulative sum of all FR trials plus all CR trials, range 0–48) [24, 46]. A TR score ≤ 41 is typically linked to AD [24, 46]. Longitudinal analysis of the FCSRT can be found elsewhere [48].

Old/new word-recognition task

In the afternoon of every yearly visit, 4–6 h after the FCSRT performance, an old/new word-recognition task based on the verbal content of the FCSRT (French version) was performed while EEG data was recorded [2, 39, 46]. Here, only the behavioral data of this task is presented. The experimental session consisted in the following: 2 min of resting state; 1 min familiarization task; approximately 20 min of old/new word-recognition task; 2 min of resting state. The familiarization task aimed at getting the participant acquainted with the old/new recognition task; it consisted in the presentation of a sentence of 6 to-be-remembered words, followed by recognition of these words among confounders. The old/new word-recognition task followed immediately. It consisted in 4 runs of 52 words each, to allow short breaks for the participant (total of 208 words, 80 ‘Old’ and 128 ‘New’ words). The old words were the 16 words memorized during the FCSRT; each old word was pseudo-randomly presented 5-times across the runs (with 0 to 3 presentations in each run and no immediate repetition). The new words were: 16 ‘New Repeated’ (NR) words, which were New words semantically unrelated to the Old words and presented the same number of times as the Old words throughout the runs (with 0 to 3 presentations in each run and no immediate repetition); and 48 ‘New Unique’ (NU) words, which were semantically

related to the Old words and presented only once throughout the runs. Since the first occurrence of a NR word could be considered as a NU word, these first occurrences were coded as NU, resulting in a total of 80 Old, 64 NR, and 64 NU words. The different types of New words aimed at potentially distinguishing sources of EM deficits and was unbeknownst to the participants, whose task was to indicate whether each presented word had previously (during the morning session) been seen (old) or not (new) [26]. All the words were presented in pseudo-random order throughout the task runs; with the same order kept across the participants. In order to minimize repetition bias, there were two word-list versions used in alternation every 2 years. List-A was used at sessions M0/M24/ M48, and list-B was used at sessions M12/M36/M60. The words were displayed in white color on a black screen monitor at 1 m distance from the participant. Each trial consisted in the following appearance (Fig. 1): a fixation cross (1.508 s duration); a word (2.004 s duration); the written question “Have you already seen this word?”, where a yes or no response (i.e., old/new) was required by right hand button-press to initiate next question (no response time limit); a second written question “Are you sure of your previous response?”, where a yes/no response in the same fashion as before was required to start a new trial (no response time limit). The second response was out of the scope of this study and not analyzed here.

Behavioral data processing

Behavioral data were collected from baseline yearly over 5-years: M0, M12, M24, M36, M48, M60. Behavioral performance of the old/new word-recognition task was assessed through accuracy (i.e., percentage of correct responses) and

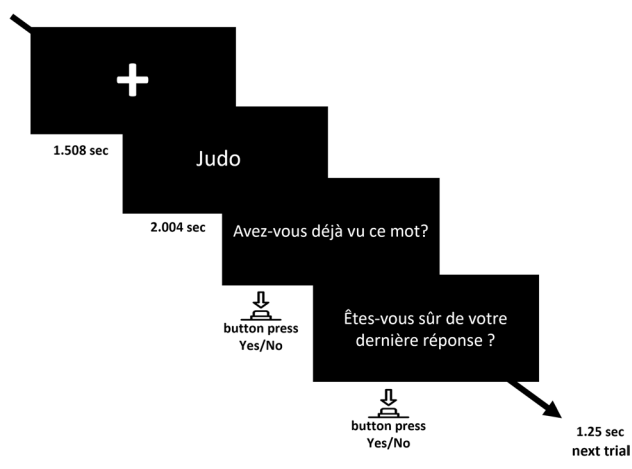


Fig. 1 Example trial of the old/new word-recognition task. Translation: 1) “Avez-vous déjà vu ce mot?” aka. Have you seen this word before? 2) “Êtes-vous sûr de votre dernière réponse?” aka. Are you sure of your response?

reaction time of correct responses (RT). Statistical analyses were performed using R studio (version 4.2.1). Exploratory examination of the behavioral data showed that in every session, RT of the first trial of every run tended to be slower than the remaining trials. Moreover, on the first run slower RTs were observed along the first 17 trials, likely corresponding to an adaptation period to the task (Online Resource 2). Therefore, we decided to dismiss the first 17 trials of the first run and the first trial of the following 3 runs; this also allowed us to balance the number of trials per word category. Then, we excluded trials with RT shorter than 150 ms <RT> or longer than 5000 ms. We concatenated the 4-runs data and computed the mean RT of correct responses per experimental condition (old, NR, NU), per session, per subject. The mean number of trials retained per subject at every session is shown in Table 1.

Accuracy was computed as the percentage of correct responses (i.e., [number of correct responses/total number of retained trials]*100) for each subject in each condition and on each session. Old/new responses were further analyzed in the SDT framework. We computed d' , that is, the discrimination sensitivity to old/new words, and criterion, that is, the bias to give either more ‘Old’ or more ‘New’ responses. For this, we first calculated hits (‘Old’ responses to old words presentation), misses (‘New’ responses to old words presentation), false alarms (FA; ‘Old’ responses to new words presentation), and correct rejections (CR; ‘New’ responses to new words presentation). Then, the Hit rate (HR; $number\ of\ hits/[number\ of\ hits + number\ of\ misses]$) and the False Alarm rate (FAr; $number\ of\ FA/[number\ of\ FA + number\ of\ CR]$) were transformed into quantiles of the normal distribution ($qnorm$) and we computed d' as $(qnorm\ HR - qnorm\ FAr)$ and criterion as $-(qnorm\ HR + qnorm\ FAr)/2$.

Table 1 Mean number of trials analyzed per subject, for each word category (old, New Repeated [NR], New Unique [NU]) on each session (M0 to M60).

Session	Old mean \pm SD	NR mean \pm SD	NU mean \pm SD
M0	78.3 \pm 2.1	62.4 \pm 2.1	61.9 \pm 2.8
M12	78.5 \pm 2.8	62.7 \pm 2.4	62.0 \pm 2.9
M24	78.4 \pm 3.0	62.5 \pm 2.8	62.1 \pm 2.9
M36	78.3 \pm 3.3	62.2 \pm 3.1	62.5 \pm 2.9
M48	77.8 \pm 3.8	61.9 \pm 4.1	62.0 \pm 3.1
M60	78.1 \pm 3.5	61.9 \pm 4.6	62.2 \pm 3.5

The average across subjects of the number of trials retained for analysis \pm standard deviation is given

NR new repeated, NU new unique, SD standard deviation

Statistical analyses

Demographic and clinical characterization of the groups

We tested if the demographic and clinical characteristics of our groups of interest (please see description below) differed at baseline (M0). We analyzed age, education level, sex, and SMCs self-reported by the participants using the five groups as a between-subject factor in a one-way Welch analyses of variance (ANOVA; ‘car’ R package, version 3.1.0). Welch ANOVA was chosen for its robustness to account for unequal group variances even in the case of different group sizes. Similarly, we tested A β SUVr differences considering only the A β + groups (excluding the controls/A β -/N- group), and FDG SUVr differences considering only the N + groups (i.e., the stable and progressors N + groups).

Group differences in the main neuropsychological scores included in this study (i.e., MMSE, FCSRT_FR, FCSRT_TR, FAB, TMT ListB-A) were evaluated with general linear models (glm; ‘stats’ R package, version 4.2.1). These models included the 5-groups as between-subject fixed-effect factor and age, sex, and education as fixed-effect covariates to account for their confounding effect. Post-hoc comparisons of the groups were performed with multiple comparison correction using Games-Howell tests (‘rstatix’ R package, version 0.7.0).

Behavioral performance in the recognition task

According to our research questions, we created two types of statistical models with different between-subject factors. First, we analyzed the effect of A β burden and neurodegeneration using A β _N model with the following 3 groups of stable participants: A β -/N- (controls), A β +/N- (stable/N-), and A β +/N+ (stable/N+). We tested the fixed effects of A β burden (contrasting controls vs. stable/N-), neurodegeneration (contrasting stable/N- vs. stable/N+), and their interaction with the word category-factor (Old, NR, and NU), across the 6-sessions (M0-M60), on the outcome variables (i.e., accuracy, d' , criterion, RT).

Second, considering that 15 A β + participants were diagnosed with prodromal-AD (progressors), we analyzed the effects associated with this diagnosis using stable_progressors model with only the A β + participants (that is, excluding controls) divided into a group of stable participants and a group of progressors. We further accounted for the potential cumulative effect of neurodegeneration by subdividing the stable and progressor groups as a function of their neurodegeneration status (stable/N- and /N+; progressors/N- and /N+). Thus, the stable_progressors model comprised two fixed-effect group factors (progression to prodromal-AD and neurodegeneration status).

The statistical models used were the following: binomial logistic regression mixed effects glm (glmer) for the analysis of accuracy, linear mixed-effects effects (lmer) for the analysis of d' and criterion, and log-transformed RT (log(RT)) (lme4 R package, version 1.1.29). These mixed effects models were chosen for their capacity to account for missing information and analyze longitudinal data. For all analyses we used full-model structures including all possible interactions among the within- and between-subjects fixed-effect factors. The following within-subject fixed-effect factors were considered: word-category (Old, NR, and NU) and session (M0-M60) for accuracy and RT analyses; word-category (Old, New) and session (M0-M60) for d' and criterion analyses. As stated above, two types of between-subject fixed-effect factors were used: the A β _N model comprised a between-subject factor with 3 levels (A β -/N- [controls], A β +/N- [stable/N-], and A β +/N+ [stable/N+]) and the stable_progressors model, focused on A β + participants, comprised two between-subject factors, each with 2 levels (stable/progressors and N- /N+). Participants were introduced as a random-effects factor and the following random-slopes were considered across them: word-category and session for accuracy and RT analysis; session for d' and criterion analysis, except for the stable_progressors model where random slopes could not be included because of the relatively small number of observations. Age, sex, and education were introduced as additional fixed-effect covariates, to account for their potential confounding effect. Age was mean-centered at baseline and considered a numerical variable. Sex and education were considered categorical variables. We reported significant effects based on Chi-square and p values for the glmer-based accuracy analysis, and F and p-values for the lmer-based d' , criterion, and RT analyses. Post-hoc analyses corrected for multiple comparisons were performed using the multivariate t-distribution with ‘emmeans’ (R package, version 1.7.5); we tested for differences within or between the following factors modalities: word-category, session, and group [49].

Additional characterization of the groups (control analyses)

Other outcome variables complementing the characterization of our five groups of interest (controls, stable/N-, stable/N+, progressors/N-, progressors/N+) were analyzed using one-way Welch ANOVA at baseline. These variables were: body mass index ([BMI] = weight/height²), licit substance use (alcohol, cigarettes), and comorbidities (arterial hypertension, atrial fibrillation, heart disease, dyslipidemia, diabetes, head trauma, mood disorders).

Complementary characterization of the groups considered longitudinal differences in neuropsychological scores (i.e., MMSE, FCSRT_IR, FCSRT_TR, FAB, TMT List B-A; from sessions M0 to M60) and in the SMC reported either

by the participants or by their companion (from sessions M12 to M60). For these analyses, we used lmer models for the neuropsychological scores including the following fixed-effects: the 5-groups as between-subject factor, the session as within-subject factor, and the 2-way interaction between them. Age, sex, and education were included as additional fixed-effect covariates. We included participants as a random-effect factor. Pairwise post-hoc analyses of group differences corresponding to the contrasts underlying our effects of interest (i.e., A β burden, neurodegeneration, or progression to prodromal AD) were performed with Games-Howell tests and corrected for multiple comparisons. The SMCs were analyzed with one-way Welch ANOVAs with groups-by-sessions as the fixed-effect factor.

Results

Demographic and neuropsychological characteristics of the groups and clinical relation with biomarkers

From 318 subjects initially included in the INSIGHT preAD cohort, we selected 263 participants for this study as detailed in Fig. 2. During follow-up, different reasons such as technical recording problems, quitting the study, death, or diagnosis with AD or other conditions led to the additional dismissal of several subjects (Fig. 2).

We began by stratifying the participants into groups according to A β burden, neurodegeneration, and progression to prodromal-AD status (Fig. 2). During follow-up, 249 participants did not progress to AD (stable (stable)

and were stratified as follows: 175 controls (without A β burden [A β -], SUVr < 0.7918, and without Neurodegeneration [N-], SUVr < 2.27); 57 (A β +) stable/N-; and 16 (A β +) stable/N+. Fifteen participants progressed to AD (eight progressors/N- and seven progressors/N+), and were considered here while cognitively normal but immediately removed once diagnosed (Fig. 2).

The mean A β SUVr at session M0 was significantly different among the 4 A β + groups (stable [N- and N+] and progressors [N- and N+]) (Table 2), with higher SUVr in the progressors than in the stable groups ($p = 0.0009$ in post-hoc tests). No other significant difference was found among the groups.

At baseline, the participants were on average 76 ± 3.5 years old (mean \pm standard deviation [SD]), mostly women ($n = 173$; 65.5%), with a higher education level (mean = 6.17 ± 2.0 ; i.e., bachelor undergraduate level; Table 2). Age and education were similar among the groups (Table 2).

Despite being cognitively normal, all participants were at-risk for AD by self-reporting SMC > 6 months. The neuropsychological scores considered at inclusion met the required mean thresholds and score-range for the cohort's age as follows: MMSE_total score = 29 (normal $\geq 27/30$), FAB_total score = 16 (normal $\geq 12/18$), and FCSRT_TR = 46 (normal $\geq 41/48$). On other neuropsychological scores included at baseline (Table 2), there was a main effect of progression-to-prodromal status on the FCSRT_FR score, which was lower in the progressors than the stable groups ($p = 0.003$); a main effect of group was also observed in the TMT List B-A score.

Fig. 2 Flowchart of the number of participants from INSIGHT-preAD cohort included in each session of this study

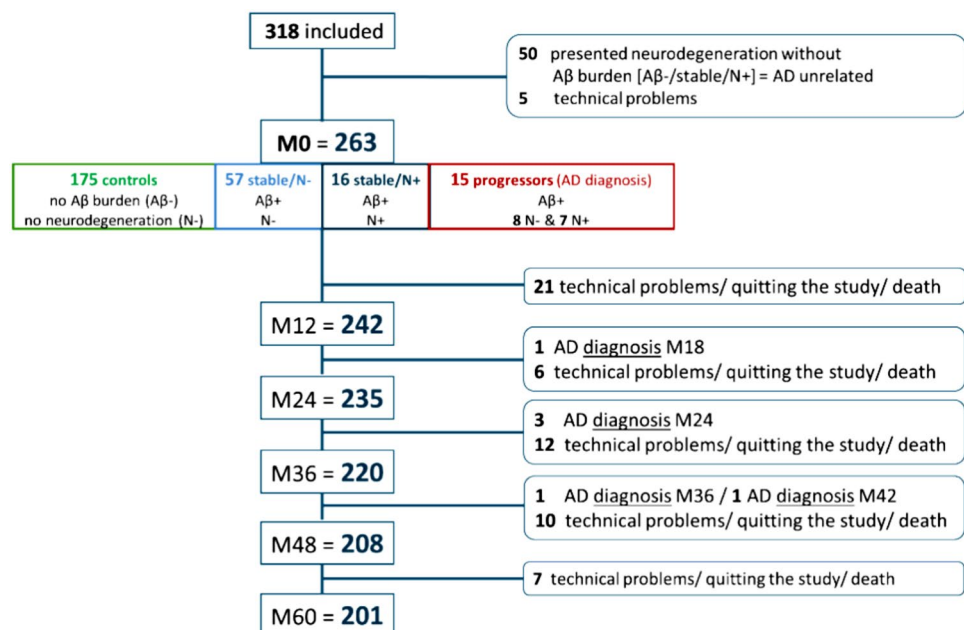


Table 2 Demographic and clinical characterization of the INSIGHT-preAD groups considered in this study at baseline (M0)

	Controls	Stable		Progressors		$F_{df1, df2} = \text{value}/$ $X^2_{df1} = \text{value}$	<i>p</i>
	A β -/N- (<i>n</i> =175)	A β +/N- (<i>n</i> =57)	A β +/N+ (<i>n</i> =16)	A β +/N- (<i>n</i> =8)	A β +/N+ (<i>n</i> =7)		
	Mean \pm SD	Mean \pm SD	Mean \pm SD	Mean \pm SD	Mean \pm SD		
Age	75.5 \pm 3.4	76.8 \pm 3.3	76.4 \pm 3.9	77.8 \pm 2.4	77.0 \pm 4.7	$F_{4, 22.2} = 2.6$	0.07 ^{a, c1}
Education level	6.3 \pm 2.0	5.7 \pm 2.0	6.9 \pm 1.7	6.3 \pm 2.2	6.6 \pm 2.4	$F_{4, 22.0} = 1.7$	0.18 ^{a, c2}
Sex (%)	F=117 (67) M=58 (33)	F=42 (73) M=15 (26)	F=6 (38) M=10 (63)	F=6 (75) M=2 (25)	F=2 (29) M=5 (71)	$X^2_4 = 11.9$	0.018 ^a
A β SUVR	0.7 \pm 0.1	1.0 \pm 0.1	1.1 \pm 0.2	1.2 \pm 0.2	1.2 \pm 0.2	$F_{3, 14.8} = 6.2$	0.006 ^{**, a, c3, d}
Neurodegeneration SUVR	2.6 \pm 0.2	2.5 \pm 0.2	2.1 \pm 0.1	2.5 \pm 0.2	2.1 \pm 0.1	$F_{1, 15.1} = 0.0009$	0.98 ^{a, e}
MMSE_total (\geq 27/30) [min-max]	28.7 \pm 1.0 [27-30]	28.5 \pm 0.9 [27-30]	28.6 \pm 1.0 [27-30]	28.1 \pm 0.8 [27-29]	28.3 \pm 0.5 [28-30]	$X^2_4 = 3.5$	0.48 ^b
FCSRT_free recall [min-max]	30.4 \pm 5.3 [13-16]	30.5 \pm 4.8 [17-43]	28.5 \pm 4.1 [21-39]	24.4 \pm 5.8 [18-34]	24.4 \pm 7.3 [14-36]	$X^2_4 = 12.7$	0.013 ^{**, b, c3}
FCSRT_total recall (\geq 41/48) [min-max]	46.2 \pm 1.9 [41-48]	46.4 \pm 1.8 [41-48]	46.1 \pm 1.5 [43-48]	44.6 \pm 1.9 [41-47]	44.7 \pm 2.6 [41-48]	$X^2_4 = 7.1$	0.13
FAB (12/18) [min-max]	16.5 \pm 1.7 [10-18]	16.0 \pm 1.7 [11-18]	16.3 \pm 1.2 [14-18]	16.3 \pm 1.5 [14-18]	15.7 \pm 2.7 [12-18]	$X^2_4 = 4.0$	0.40 ^b
TMT ListB-A (sec) [min-max]	46.0 \pm 34.9 [-14-205]	51.5 \pm 32.6 [-7 to 172]	70.2 \pm 54.7 [7-241]	51.4 \pm 38.2 [7-133]	77.1 \pm 34.0 [43-134]	$X^2_4 = 14.2$	0.007 ^{**, b}
M. complaints_participant (normal < 15)	13.3 \pm 6.6	12.9 \pm 5.3	10.6 \pm 4.9	14.0 \pm 6.0	9.3 \pm 5.2	$F_{4, 22.6} = 1.9$	0.16

The table represents demographic and clinical scores per group of participants at session M0 in mean (SD) or number (%), and the respective range of each measure score per group. The *p* value corresponds to the results of general linear models performed per measure

SD standard deviation, A β SUVR β -Amyloid Standardized Uptake Value ratio, MMSE Mini Mental State Examination, FCSRT Free and Cued Selective Reminding Test, FAB Frontal Assessment Battery, TMT Trail Making Test

The *p* value of the statistical analysis performed between the 5-groups at session M0 is reported for each outcome variable with ***: $p < 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$. It is followed by superscript letters indicating the type of analysis: a Welch ANOVA; b General linear model (GLM), *p* = main effect of group; c post-hoc Games-Howell tests yielding significant results when corrected for multiple comparisons for the effects of interest (c1 = effect of A β burden, c2 = effect of neurodegeneration, c3 = difference linked to progression to prodromal AD); d comparison performed between the A β + groups; e comparison performed between the neurodegeneration-positive groups. Significant differences were found on the following measures: c1) A β effect on age: $p = 0.04$ [controls < stable/N-]; c2) neurodegeneration effect on education: $p = 0.05$ [stable/N+ > stable/N-]; c3) progression-to-prodromal-AD effect on A β _SUVR: $p = 0.0009$ [progressors > stable], and on FCSRT_free recall: $p = 0.003$ [progressors < stable]

Behavioral results

Behavioral performance of the old/new recognition task in the stable participants (A β _N model)

To assess A β burden and neurodegeneration effect on behavioral performance, we used the A β _N model comprising the 3-groups of stable participants (i.e., controls, stable/N-, and stable/N+). The mean accuracy of word recognition (i.e., correct response rate) was maintained at nearly ceiling level in the 3-group across follow-up (0.97 \pm 0.04). Table 3 presents the main results of the Glmer analysis. It showed a main effect of word-category, reflecting better recognition of old words than NR ($p < 0.0001$) and NU ($p < 0.0001$) words. There was also a main effect of session (M0-M6) and an interaction between word-category and session. This reflected a better recognition of old than both NR and NU words at every session and of NR than NU words at session

M24, as shown by post-hoc tests. There were systematic interactions between group, word_category, and session, prompting investigation of performance by group (Fig. 2). Post-hoc analyses showed that the stable/N+ group recognized NU words significantly worse than the controls ($p = 0.02$) and stable/N- ($p = 0.03$) groups (Fig. 3). At M24, the stable/N+ participants were significantly worse in recognizing NR ($p = 0.05$) and NU ($p = 0.0005$) than the stable/N- group; and worse in recognizing NU words than the controls ($p = 0.02$; Fig. 3). Furthermore, there were main effects of age ($X^2(1) = 17.77$, $p < 0.0001$) and education level ($X^2(6) = 15.30$, $p = 0.02$), reflecting significant linear trends of decreasing accuracy with increasing age, and increasing accuracy with increasing level of education.

For the SDT analysis (d' and criterion), we pooled NR and NU words together as new words. The d' mean suggested similar sensitivity to old/new discrimination in the controls (4.62 \pm 0.65) and the stable/N- (4.59 \pm 0.54) groups, and

Table 3 Main effects and interactions relative to the behavioral performance (i.e., accuracy, dprime, criterion, and reaction time) of the groups included in the A β _N model (groups: controls, stable/N–, stable/N+)

Main effects and interactions	Accuracy		dprime		Criterion		Reaction time	
	X^2_{df} = value	p =	$F_{df1,df2}$ = value	p =	$F_{df1,df2}$ = value	p =	$F_{df1,df2}$ = value	p =
word category (wc)	$X^2_2 = 60.60$	<0.0001	NA	NA	NA	NA	$F_{2,478} = 32.73$	<0.0001
session	$X^2_5 = 25.03$	0.00014	$F_{5,34.2} = 4.94$	0.002	$F_{5,2} = 6.26$	0.14	$F_{5,220} = 6.72$	<0.0001
group	$X^2_2 = 0.60$	0.74	$F_{2,4.4} = 8.73$	0.03	$F_{2,2} = 1.31$	0.43	$F_{2,243} = 2.30$	0.10
wc * session	$X^2_{20} = 49.80$	0.00024	NA	NA	NA	NA	$F_{10,2332} = 2.51$	0.005
wc * group	$X^2_4 = 11.65$	0.02	NA	NA	NA	NA	$F_{4,471} = 0.29$	0.88
group * session	$X^2_{10} = 19.87$	0.03	$F_{10,2} = 2.23$	0.35	$F_{10,2} = 1.36$	0.50	$F_{10,217} = 2.23$	0.02
3-way interaction	$X^2_{20} = 49.80$	0.00024	NA	NA	NA	NA	$F_{20,2329} = 1.05$	0.39

The p values of the statistical analyses performed are reported for each experimental factor and interaction with significant values in bold df degrees of freedom of the X^2 ; df1, df2 degrees of freedom of the F statistics; NA not applicable

a somewhat decreased sensitivity in the stable/N+ group (4.11 ± 0.82 ; Fig. 4). This was confirmed by a main effect of group—(Table 3), linked to significantly higher d' in the stable/N– vs. N+ groups ($p = 0.0002$; neurodegeneration effect from post-hocs). There was also a main effect of session, where d' was significantly higher in the controls and stable/N– groups (i.e., neurodegeneration negative) than the stable/N+ group. Furthermore, there was a main effect of age ($F(1,1.8) = 53.36, p = 0.02$), where higher age was associated with significantly lower d' .

Criterion analysis showed that the mean criterion was close to 0 but slightly negative in all groups throughout sessions (controls = -0.18 ± 0.3 ; stable/N– = -0.14 ± 0.2 ; stable/N+ = -0.21 ± 0.3), suggesting a trend toward choosing more old than new responses, both when the response was the expected (hits) and when it was not (false alarms; Fig. 4). However, no significant differences were found between the groups.

In summary, SDT analysis suggested that lower performance in the stable/N+ group was not related to a change in response criterion but rather to a decrease in old/new discrimination sensitivity (d').

Regarding RT, there was an overall longitudinal reduction in all groups ($M0 = 937 \pm 335$; $M60 = 773 \pm 333$), suggesting a learning effect typical of these tasks. Accordingly, as can be seen in Table 3, there was a main effect of session, which reflected significant RT decrease from M0 to M60 ($p = 0.0001$), and a main effect of word-category, which reflected shorter RT to recognize Old vs. NR ($p < 0.0001$) and NU words ($p < 0.0001$) and NR vs. NU words ($p = 0.0003$). Also, there was a significant interaction between word-category and session, which reflected significant decrease in RT from session M0 to M60 in the recognition of NR ($p < 0.001$) and NU ($p = 0.0004$) words, while for Old words, there was only a trend to RT reduction from M36 to M60 ($p = 0.06$). An interaction between session and group reflected shorter RT in the controls vs.

the stable/N+ group at M36 ($p = 0.02$; Fig. 5). Finally, there was a significant RT increase with increasing age ($F(1,608.2) = 13.31, p = 0.0003$). Individual performances per group can be found in the Online Resource 3.

Behavioral performance of the old/new recognition task in the stable vs. progressor participants

Since fifteen participants progressed to prodromal AD during the 5-year follow-up period, we used a stable_progressors model (stable vs. progressors) including only A β + participants (AD-criteria), to assess behavioral performance differences associated with progression. The cumulative effect of neurodegeneration was also considered by subdividing the groups according to their N+/N– status. The main results of this analysis are summarized in Table 4.

As previously observed, the mean accuracy of word recognition was longitudinally maintained at nearly ceiling level in the stable participants ($M0 = 0.97 \pm 0.05$; $M60 = 0.98 \pm 0.02$). In contrast, there was a clear longitudinal decline in the progressors ($M0 = 0.93 \pm 0.10$; $M60 = 0.78 \pm 0.17$). There was no significant main effect of progression to prodromal AD on accuracy, but significant interactions between progression-status, session, and word-category (Table 4). Post-hoc tests showed that the progressors recognized NR and NU words with significantly lower accuracy than the stable group did at sessions M12 (both $p < 0.001$) and M60 ($p < 0.0001$ and $p = 0.004$, respectively). In addition, recognition of old words was less accurate in the progressors than the stable group at sessions M48 ($p < 0.001$) and M60 ($p < 0.001$). There was also an interaction between word-category and session, which reflected more accurate recognition of: old vs. NR words in all sessions but M48; old vs. NU words at sessions M12, M36, and M60; NU vs. NR words at session M24 (Fig. 6).

The SDT analysis further showed a higher d' in the stable (4.48 ± 0.64) vs. the progressor groups (3.28 ± 0.72),

Fig. 3 Accuracy performance in the old/new recognition task across the M0 to M60 sessions for each stable group (controls, stable/N-, and stable/N+). In each plot, the mean accuracy (in rate of correct responses) is represented for each group (comprising $n = 175$, 57, and 16 participants at M0, respectively), for each word category (NR words in blue, NU words in green, old words in red), and for every session (M0 to M60). Thin vertical lines represent the standard errors of the means

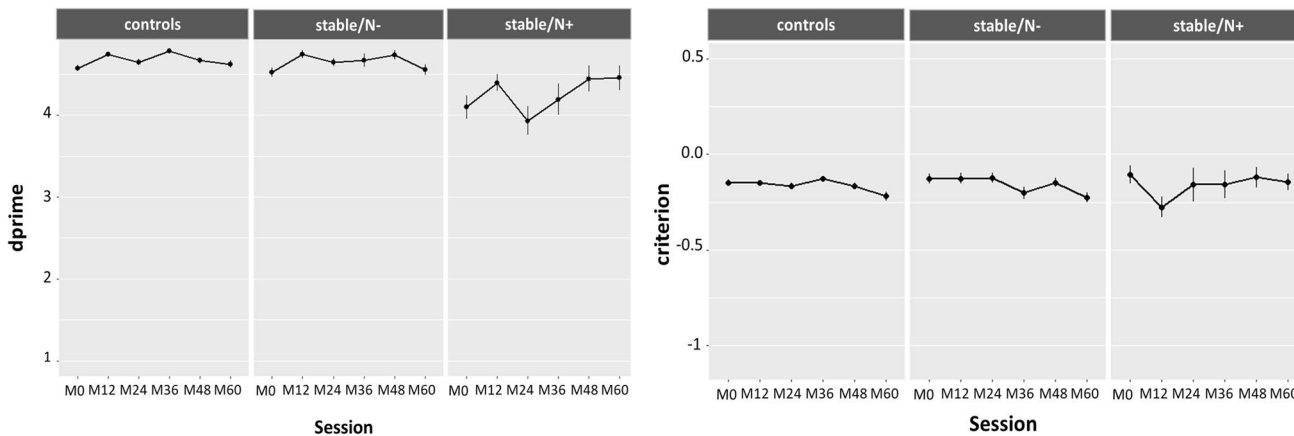
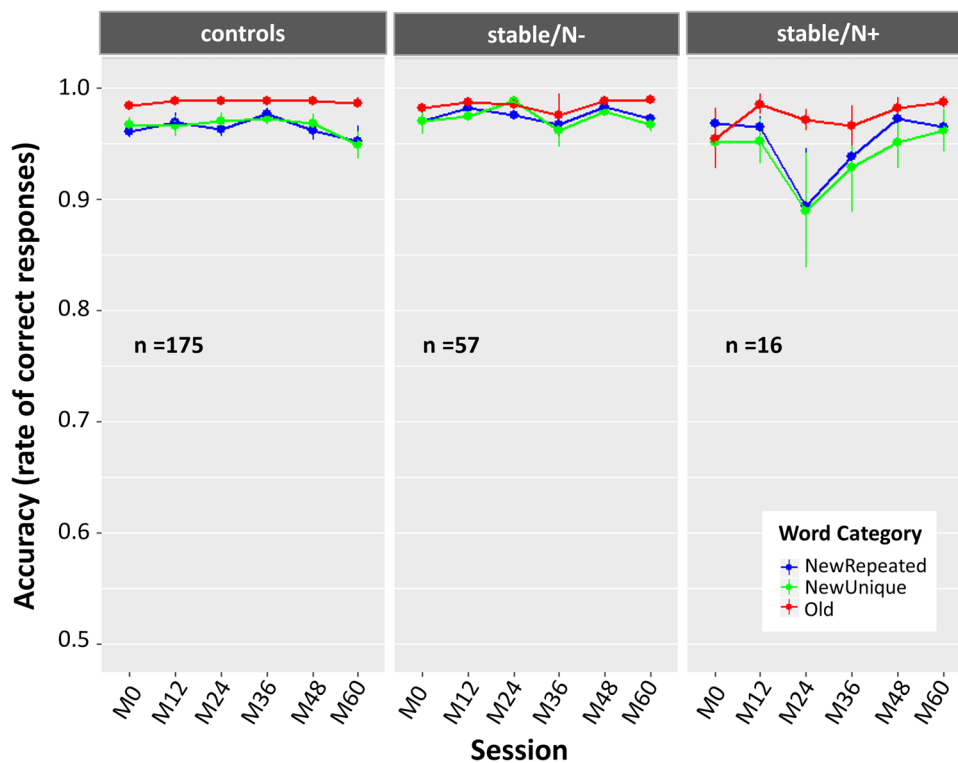
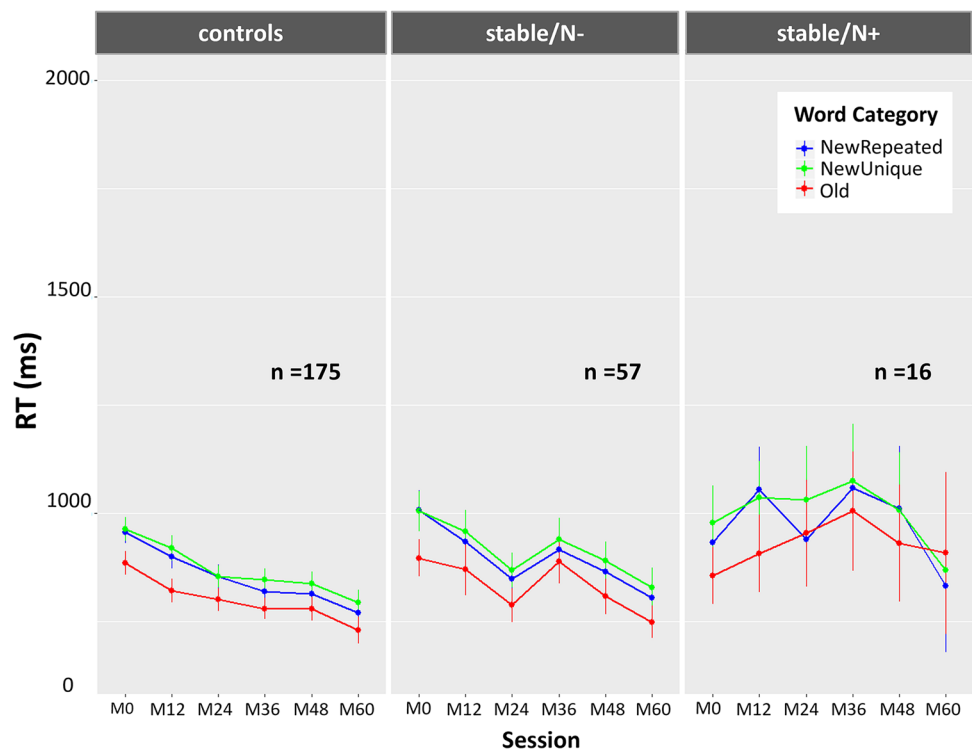


Fig. 4 dprime [d'] and response criterion across sessions in the old/new word-recognition task, in the controls, stable/N- and stable/N+ groups. The mean dprime on the left plots and the mean

criterion on the right plots (in ordinate) are represented for each group at every session (M0 to M60, in abscissa)

reflected by a main effect of progression-status and confirmed by post-hoc tests ($p < 0.001$). A main effect of session and an interaction between progression-status and session reflected significantly lower sensitivity to discriminate old/new words in the progressor than the stable groups at sessions M12, M24, M48, and M60. There were also interactions between session and neurodegeneration-status and session, neurodegeneration-status and progression-status, indicative of a cumulative effect of neurodegeneration. Response criterion was overall mostly negative indicating

Fig. 5 Reaction time [RT] to the old/new word-recognition task on each session (M0 to M60) in each stable group (controls, stable/N-, and stable/N+). In each plot, the mean RT (in ms, in ordinate) is presented per group ($n = 175, 57,$ and 16 participants at M0, respectively), for each word category (NR words in blue, NU words in green, Old words in red), at every session (M0 to M60). Vertical lines represent the standard errors of the means



a response bias toward choosing old (rather than new) responses, which although not significant, was on average greater in the progressors (-0.27 ± 0.47) than the stable group (-0.15 ± 0.23). Nevertheless, response criterion in the progressors was highly heterogeneous when compared with the stable group, which was confirmed by a significant main effect of session and an interaction between the progression-status and session (Table 4; Fig. 7). Post-hoc analyses showed a significantly lower criterion in the progressor

vs. the stable groups at M12 ($p = 0.03$) and significantly higher criterion at session M48 ($p = 0.02$). The interactions between session and neurodegeneration-status, and between session, neurodegeneration-status and progression-status were also significant, again indicative of a cumulative effect of neurodegeneration.

Overall, the progressors relative to the stable group showed an early and sustained decrement in word recognition, associated with a significant decrease in their

Table 4 Main effects and interactions relative to the behavioral performance (i.e., accuracy, dprime, criterion, and reaction time) of the groups included in the stable_progressors model (groups: stable/progressors and N+/N-)

Main effects and interactions	Accuracy		dprime		Criterion		Reaction time	
	$X^2_{df1} = \text{value}$	$p =$	$F_{df1,df2} = \text{value}$	$p =$	$F_{df1,df2} = \text{value}$	$p =$	$F_{df1,df2} = \text{value}$	$p =$
Word category (wc)	$X^2_2 = 2.41$	0.30	NA	NA	NA	NA	$F_{2,185} = 18.99$	<0.0001
Session	$X^2_5 = 43.45$	<0.0001	$F_{5,1208} = 26.14$	<0.0001	$F_{5,1216} = 11.95$	<0.0001	$F_{5,71} = 4.99$	0.0006
Progressor-status (progr-status)	$X^2_1 = 0.48$	0.49	$F_{1,71} = 38.28$	<0.0001	$F_{1,77} = 0.14$	0.71	$F_{1,76} = 5.67$	0.02
Neurodegeneration (N)	$X^2_1 = 9.93$	0.49	$F_{1,71} = 1.16$	0.28	$F_{1,77} = 0.14$	0.71	$F_{1,78} = 1.64$	0.20
wc * session	$X^2_{10} = 76.59$	<0.0001	NA	NA	NA	NA	$F_{10,753} = 2.33$	0.01
wc * progr-status	$X^2_2 = 0.51$	0.78	NA	NA	NA	NA	$F_{2,185} = 5.94$	0.003
progr-status * session	$X^2_5 = 27.28$	<0.0001	$F_{5,1209} = 31.86$	<0.0001	$F_{5,1217} = 9.07$	<0.0001	$F_{5,71} = 6.78$	<0.0001
progr-status * wc * session	$X^2_{10} = 26.70$	0.003	NA	NA	NA	NA	$F_{10,753} = 2.67$	0.003
N * session	$X^2_5 = 2.70$	<0.0001	$F_{1,1208} = 9.33$	<0.0001	$F_{1,1216} = 14.39$	<0.0001	$F_{5,71} = 0.32$	0.90
progr-status * N * session	$X^2_1 = 2.36$	0.0003	$F_{1,70} = 0.91$	0.34	$F_{1,75} = 0.02$	0.90	$F_{5,71} = 2.35$	0.05

The p values of the statistical analyses performed are reported for each experimental factor and interaction with significant values in bold df degrees of freedom of the X^2 ; df1, df2 degrees of freedom of the F statistics; NA not applicable

For the interactions involving the neurodegeneration factor, we only report the significant interactions as detailed in main text

Fig. 6 Accuracy of the old/new recognition task across the M0 to M60 sessions for each Aβ+ group (stable and progressors). The mean accuracy (in rate of correct responses) for the Old (in red), NR (in blue) and NU (in green) words at each session (M0 to M60) is represented for the stable group ($n=73$ at M0) and the progressor groups ($n=15$ at M0). The colored thin vertical lines represent the standard errors of the means

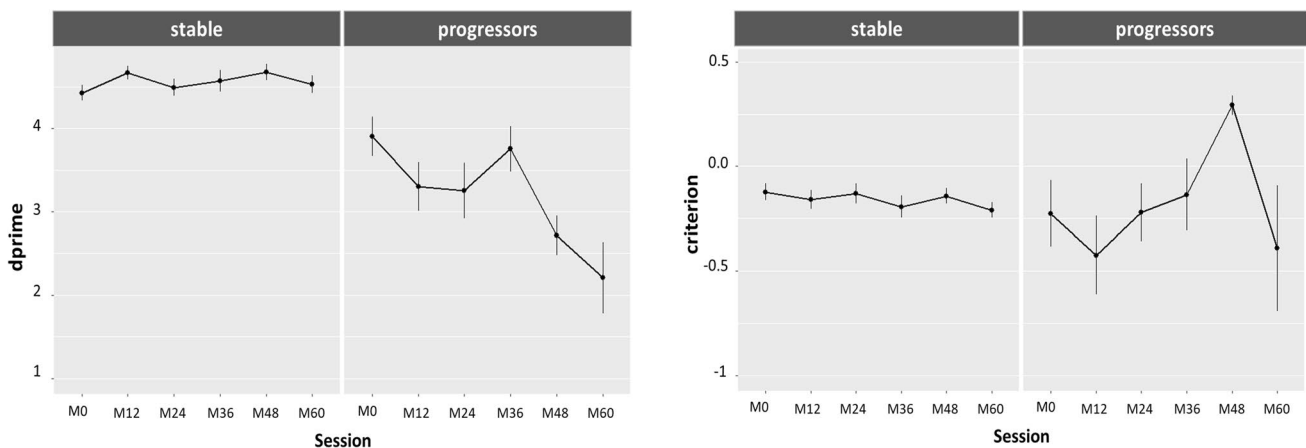
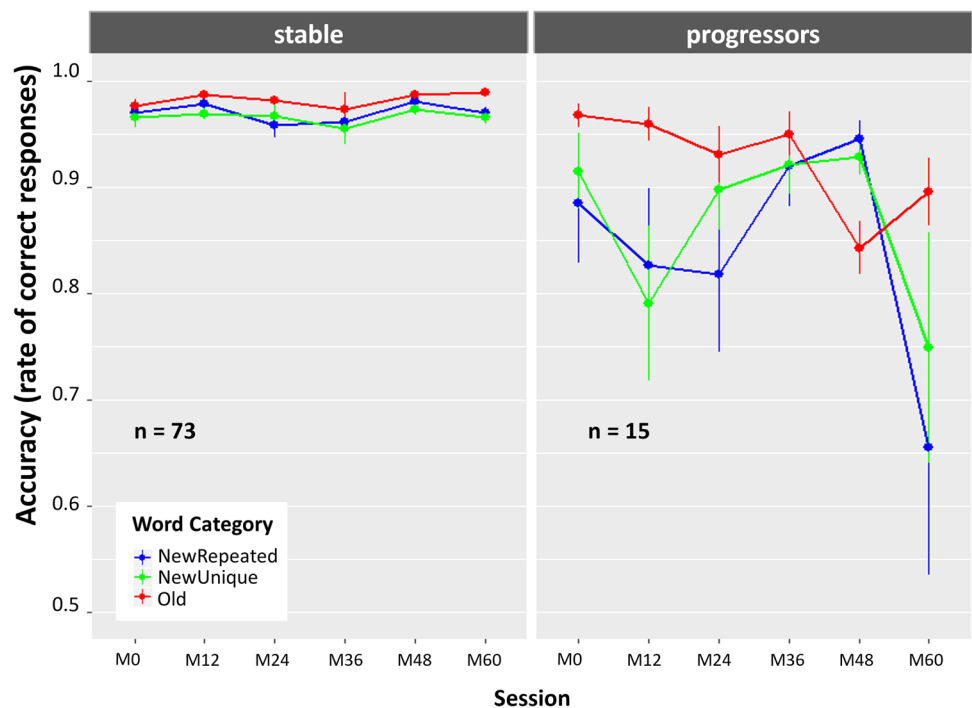


Fig. 7 d' and response criterion across sessions in the old/new word-recognition task in the stable and progressor groups. The mean d' on the left plots and the mean criterion on the right

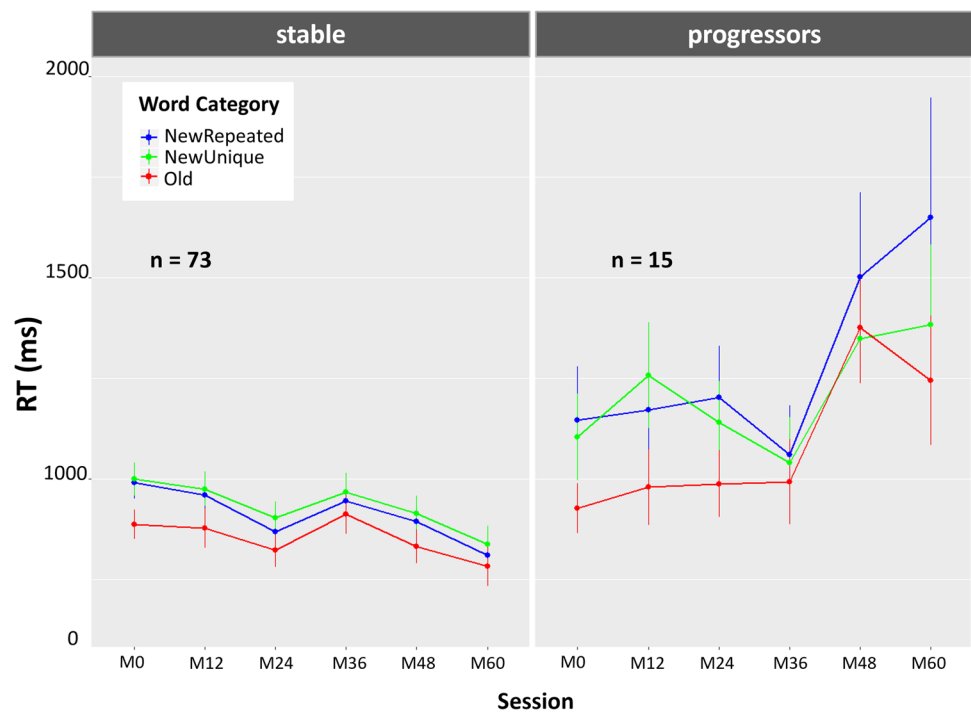
plots (in ordinate) are represented for each group at every session (M0 to M60; in abscissa)

sensitivity to discriminate old/new words. Alteration in the recognition of NR and NU words started already at session M12, while the recognition of old word started later, from session M48.

The analysis of RT showed a main effect of progression-status (Table 4; Fig. 8), with higher RT in the progressor (1148 ± 306 ms) than in the stable group (916 ± 292 ms). There were main effects of word-category, session, and an interaction between these factors. This reflected that the recognition was faster for the Old than NR words at

all sessions but M36 and M48 and it was faster for the old than the NU words at sessions M0, M12, and M24 (all $p \leq 0.001$). The significant interactions between progression-status, word-category, and session and associated post-hoc analyses showed that the progressors were significantly slower than the stable participants in recognizing all word categories at sessions M48 and M60 (all $p \leq 0.02$; Fig. 8). At both these sessions, the progressors showed significantly lower accuracy, d' , and longer RT than stable participants in recognizing old and new words.

Fig. 8 Reaction time [RT] to the old/new word-recognition task across M0 to M60 sessions for each A β + group (stable and progressors). The mean RT (in ms, in ordinate) for Old (in red), NR (in blue) and NU (in green) words at each session (M0 to M60, in abscissa) are represented for the stable group ($n = 73$ at M0) and the progressor group ($n = 15$ at M0). The colored thin vertical lines represent the standard errors of the means



Individual performances per group can be found in the Online Resource 3.

Complementary clinical and neuropsychological characterization of the groups

EM alterations are influenced by a myriad of physiologic and clinical characteristics, which were in part included in the INSIGHT-preAD study. Here we analyzed a number of these characteristics to complement the clinical characterization of our groups (Table 5).

First, comorbidities are common in aged populations. Here, hypertension (38% [$n = 101$]) and dyslipidemia (41% [$n = 107$]) were the most prevalent. The use of licit drugs, such as alcohol and cigarettes, have also been linked to alterations in cognitive performance. In our sample, the overall consumption average was within the official guidelines, without any group difference. However, though not significant, the percentage of weekly smoking was on average higher in the progressor/N- than the other groups (75% at M0) and the alcohol consumption was on average higher in the stable/N+ group than the other groups, with the male participants in this group reaching the status of heavy drinkers at M0 according to official guidelines (16.3 ± 13.7 units/week; normal < 14). Moreover, the stable/N+ participants drunk significantly more alcohol than the stable/N- participants did ($p = 0.03$), suggesting a link with neurodegeneration.

Considering participants with neurodegeneration (i.e., stable/N+ and progressor/N+ groups) vs. without (controls,

stable/N-, progressor/N- groups), the neurodegeneration-positive groups were mostly males (63% and 71%, respectively), contrasting with the female predominance of the overall sample, had higher body mass index (BMI, a proxy for body fat; 26% and 26%, respectively), classifying them as overweight [25–29.9]. They were also on average slightly more hypertensive (50% and 57%, respectively), which is often associated with overweight, and had lower TMT ListB-A scores (i.e., longer time to complete the test; 70.2 and 77.1, respectively).

M12 was the first session where SMC were reported from both the participants and their companions. Results show that SMC were higher when reported by the participants than by their companions in all groups except in the progressors/N+ group, when the companions reported higher SMC than the participants. The companions of the progressors/N+ reported the highest SMC, scoring normal (< 15) at session M12 but escalating from M24 onwards, while the progressors/N- companions reported abnormal SMC only at session M60 (19.0 ± 11.3). However, there was not any significant difference between the groups in SMC reported by either the participants or their companions. In the analysis of neuropsychological tests across sessions, significant group differences were found on the FCSRT_FR ($F(4,340.4) = 16.73$, $p < 0.0001$); on the FCSRT_TR ($F(4,321.5) = 26.70$, $p < 0.0001$), on the MMSE_total ($F(4,317.9) = 8.78$, $p < 0.0001$), on the FAB_total ($F(4,307.9) = 10.54$, $p < 0.0001$), and on the TMT List B-A scores ($F(4,317.6) = 6.82$, $p < 0.0001$). Post-hoc analyses on our effects of interest showed the following:

Table 5 Complementary physiologic and clinical characteristics of our study sample at baseline (M0, except for the SMC where the values from participants and companions are presented from M12)

	Controls	Stable		Progressors		$F_{df1, df2} = \text{value}$	p
	A β -/N- ($n = 175$)	A β + /N- ($n = 57$)	A β + /N+ ($n = 16$)	A β + /N- ($n = 8$)	A β + /N+ ($n = 7$)		
	mean \pm SD	mean \pm SD	mean \pm SD	mean \pm SD	mean \pm SD		
M. complaints_companion (normal < 15; from M12)	7.7 \pm 6.2	8.4 \pm 5.5	9.5 \pm 6.8	6.8 \pm 4.6	12.7 \pm 9.1	$F_{4, 17.8} = 0.7$	0.58
M. complaints_participant (normal < 15; from M12)	13.4 \pm 6.6	13.2 \pm 5.7	10.3 \pm 4.1	13.8 \pm 3.7	12.5 \pm 8.0	$F_{4, 16.1} = 1.0$	0.43
BMI (> 25 = overweight)	25 \pm 0.3	24 \pm 0.4	26 \pm 0.3	25 \pm 0.2	26 \pm 0.4	$F_{4, 22.5} = 1.2$	0.33
HTA n (%)	70 (40)	17 (29)	8 (50)	2 (25)	4 (57)	$F_{4, 22.0} = 1.0$	0.41
Atrial fibrillation n (%)	16 (9)	4 (7)	0 (0)	1 (13)	1 (14)	$F_{3, 16.2} = 0.2^a$	0.92
Heart disease n (%)	15 (9)	6 (11)	5 (31)	1 (13)	1 (14)	$F_{4, 21.4} = 0.8$	0.51
Dyslipidemia n (%)	65 (37)	25 (44)	9 (56)	5 (63)	3 (43)	$F_{4, 21.9} = 0.9$	0.48
Diabetes n (%)	6 (3)	4 (7)	2 (13)	0 (0)	0 (0)	$F_{2, 34.6} = 0.9^a$	0.40
Head trauma n (%)	15 (9)	5 (9)	3 (19)	0 (0)	0 (0)	$F_{2, 36.4} = 0.5^a$	0.62
Mood disorders n (%)	55 (31)	12 (21)	2 (13)	2 (25)	1 (14)	$F_{4, 22.5} = 1.4$	0.27
Smoking packs weekly (1 pack = 20 cigarettes)	0.5 \pm 0.6	0.5 \pm 0.6	0.6 \pm 0.8	0.8 \pm 0.7	0.3 \pm 0.5	$F_{4, 22.0} = 0.7$	0.61
Alcohol cups weekly (normal = F \leq 7; M \leq 14)	F: 4.5 \pm 5.3 M: 10.8 \pm 8.8	F: 5.1 \pm 6.7 M: 6.9 \pm 8.5	F: 4.0 \pm 4.9 M: 16.3 \pm 13.7	F: 5.7 \pm 4.9 M: 13.0 \pm 9.9	F: 5.0 \pm 7.1 M: 6.8 \pm 10.2	$F_{4, 5.8} = 0.1$ $F_{4, 6.0} = 1.0$	F: 0.97 M: 0.48

Mean scores \pm standard deviation (SD) or numbers (n) and percentage (%) of participants are given for each group. The p value of the statistical analysis (one-way Welch ANOVA) performed across the five groups is reported for each measure. Any significant difference was observed among the groups at baseline. The thresholds for normal alcohol consumption are different for female (F) and male (M) participants and are therefore reported separately for each gender in each group

SD standard deviation, M memory, BMI body mass index, HTA Arterial Hypertension, F female participants, M male participants

^aFor atrial fibrillation, diabetes, and head trauma, some groups had only 0 values (for all participants) and therefore could not be included in the ANOVA comparisons

an effect of neurodegeneration with lower scores in the stable/N- vs. control groups, in the FCSRT_TR ($p < 0.0001$) and the FAB_total ($p = 0.0008$); an effect of A β burden with worse scores in the stable/N+ vs. stable/N- groups, in the TMT List B-A (i.e., longer time to complete the task; $p = 0.03$); and an effect of progression to prodromal AD with lower scores in the progressors vs. the stable groups in the FCSRT_FR ($p < 0.0001$), the FCSRT_TR ($p < 0.0001$), the MMSE_total ($p < 0.0001$), the FAB_total ($p < 0.0001$), and the TMT List B-A score ($p = 0.0009$).

Discussion

Here we assessed the potential EM alterations appearing in elderly cognitively normal individuals at-risk for AD (> 6 months of SMC), by analyzing their behavioral performance in an old/new recognition task based on the verbal content of the FCSRT. We first assessed the effect of A β burden and neurodegeneration, which showed a cumulative detrimental effect on accuracy and RT. Then, we analyzed

the effect of progression to prodromal AD, which negatively impacted EM performance up to 3 years before diagnosis.

The faster and more accurate recognition of old vs. new words shown by the controls is expected in list-learning tasks and validated the capacity of our original old/new recognition task for assessing EM [23, 33, 34].

We started by assessing behavioral performance longitudinally in controls and stable participants (A β _N model), to examine the effects of A β burden and neurodegeneration. Accuracy was overall maintained at nearly 'ceiling level' and RT decreased across sessions, replicating the learning effect mentioned above expected from the repetition of similar cognitive tasks [11, 21, 23, 30, 33, 34]. Moreover, old words (i.e., previously studied) were on average recognized more accurately and faster than NR words (repeated the same number of times as the old words) and NU words (belonging to the same semantic category as the old words), with also faster RT for NR than NU words (inducing repetition and semantic interferences, respectively). Better and faster recognition of previously studied words vs. words seen at-new is expected as a reflect of the integrity of encoding and storage processes in cued list-learning tasks [8, 11, 21,

30, 34]. Cues are presented in these tasks during learning, to promote the encoding and storage of the to-be-remembered items (old words) and minimizing the interference of novel irrelevant items (new words), and during retrieval to enhance appropriate recognition [8, 11, 21, 25, 27, 35]. In normal cognitive aging, individuals typically benefit from semantic cues to recognize to-be-remembered items and overcome compromised encoding and retrieval strategies, which appears the case in the controls and stable/N- participants [3, 8, 21, 24, 25]. By contrast, in the typical AD with AHS, individuals struggle to benefit from semantic cues to overcome encoding and storage deficits, leading to a deficient consolidation and faulty retrieval of the memory trace [3, 8, 21, 24, 25]. We could detect this pattern in the stable/N+ group, who showed reduced recognition of novel items, particularly when these belonged to the same semantic category (NU), suggesting a misattribution of familiarity that thwarted the distinction between old and NU words [21, 23, 28]. The SDT analysis further supported this suggestion, showing a diminished sensitivity in old vs. new words discrimination in the stable/N+ group. Additionally, this group also showed overall longer RT than the other groups at session M36. In line with previous research, there was a cumulative effect of A β burden and neurodegeneration, which suggested a reduction in the cognitive resources vital for EM processing in these individuals, while they were stable and cognitively normal [7, 12, 15, 16, 18–20, 50].

However, despite A β burden being considered the main AD-risk factor, when in isolation, it did not appear to influence EM processing [6, 18, 51, 52]. This result was not entirely surprising, since the link between A β burden and the onset of memory deficits remains controversial, with some studies showing the absence of any memory or cognitive deficits for years, in individuals meeting A β -pathology criteria [6, 10, 18, 19, 27, 51–53]. Moreover, research has shown that the association between A β burden and cognitive decline should perhaps consider the accumulation rate of A β plaques instead of the A β status used here, as in preclinical AD the A β burden reaches a plateau, leading to ‘ceiling levels’ that might preclude the observation of memory deficits [7, 15, 16, 19, 20, 50].

The core part of our results was the observation of a very distinct pattern of behavioral performance in the progressors to prodromal AD, while they were still cognitively normal according with typical cut-offs and virtually without any apparent clinical manifestation. Here, we assessed the progressors in comparison with A β + stable participants, since A β burden is a main criterion in AD-diagnosis [2, 4, 7, 13, 14, 18]. We found that the progressors recognized less accurately all word-categories, with an irregular pattern of performance across sessions. The accuracy in NU and NR word recognition was particularly affected and started right from session M12—between 3 and 4 years prior to

diagnosis—, while recognition of old words started to be affected later, only up to 1 year prior to diagnosis. This aligns with research suggesting that the novel vs. previously studied words introduces an interference in the recognition process demanding higher cognitive resources, which seems to be independent of the interference source, either semantic relatedness or repetition number [8, 9, 18, 21, 25, 30, 34]. Adding to this evidence, the progressors showed an early decline in the sensitivity to discriminate between Old and New words (d'), which in typical AD (with AHS) has been interpreted as a performance tending toward chance guessing [9, 23, 30, 36–38]. Furthermore, evidence suggests the loss of this discriminatory capacity results from the failing in encoding, with negative impact on storage, and consequently on recognition [8, 11, 21, 24, 34, 51]. The encoding of similar word-list versions during follow-up might be considered a factor that contributed to source confusion (for the to-be-remembered words, ‘Old’), causing degraded representations that consequently prompted misattribution error and memory distortions [8, 21, 23, 30, 34]. Moreover, 2–3 years before diagnosis, the progressors became significantly slower at recognizing all word categories, contrasting with the reduction in RT associated with the typical learning effect observed in stable groups [33, 34]. The RT increase in the progressors may reflect a more conservative response strategy to attain the attended level of performance and potentially mask existing cognitive deficits [18, 30, 33, 34]. In sum, the recognition of old/new words was less accurate, with lower discrimination sensitivity, and longer in the progressors. These results suggest that although asymptomatic and cognitively normal, the progressors were likely already on the neurodegenerative path to AD, lacking the full resources to overcome the subtle EM deficits that seemed to start emerging at an at-risk phase, years before AD-diagnosis.

Finally, the supplementary demographic and clinical analyses complemented the characterization of the groups with interesting findings. For instance, self-reported SMC remained mostly within the norm (score < 15) longitudinally in the control and stable groups [45]. By contrast, when reported by the companions, SMC increased from the controls to the stable group and again to the progressor group. This was further exacerbated when considering the cumulative effect of neurodegeneration in the progressors, with abnormal levels reached at session M24 for the progressors/N+. In contrast, self-reported SMC did not significant difference. This is in line with research showing decrease in the awareness of the existence and magnitude of memory deficits by individuals at preclinical stages in contrast with the perception of their companions [22].

Furthermore, it was interesting to find characteristics common to the neurodegeneration-positive groups (i.e., the stable/N+ and progressor/N+ groups). They were in their

majority males (contrary to the female predominance of our overall sample), showed higher BMI, more hypertension and took longer to complete the TMT test (requiring physical performance). Although these differences were not significant when compared with the other groups and our data did not allow us to make formal inference on this topic, these results are in line with a growing body of evidence suggesting a link between weight, myelin, and neurodegeneration [54, 55].

A major advantage of our study was the use of an old/new word-recognition task based on the verbal content of the FCSRT. First, the same to-be-remembered words were encoded in the FCSRT (morning) and in the old/new recognition task (afternoon), strengthening the consolidation process [3, 34, 56]. Second, despite EM requiring between hours to weeks to consolidate, our experimental design granted a delay of 1 h/+ period between the morning encoding and the afternoon retrieval, well beyond the ≤ 20 min typically used in EM assessments. This certainly contributed to detecting the emergence of subtle EM alterations [3, 34, 56]. Moreover, our study included five different groups matched for age and education, which allowed us to isolate the effects of A β burden, neurodegeneration, and the progression to AD. That said, this study had some limitations. First, despite the groups being matched for age and education level at baseline, there were some group differences in neuropsychological variables. Also, the average high level of education of the cohort challenged the generalization of our findings to the overall population. We attempted to account for the effects of age, education level, and sex, by introducing these variables as additional covariates in the analysis, revealing some overall effects which were not investigated in detail and therefore, residual confounds cannot be excluded. Another limitation was the small sample size and relatively high inter-subject variability of the stable/N+ and progressor groups, which calls for caution in the generalizability of our results. Nevertheless, the uniqueness of this cohort needs to be considered in regard with the difficulty in finding participants at-risk for AD willing to commit to a study of this longevity.

Considering the increasing prevalence of AD worldwide, it becomes increasingly important to find accessible markers able to predict at an asymptomatic stage who will progress to a confirmed diagnosis. Within the INSIGHT-preAD study, the old/new recognition task based on the verbal content of the FCSRT presented here, seems a promising tool that may detect the emergence of subtle EM deficits occurring in cognitively normal aging individuals at-risk for AD, several years before symptomatology becomes apparent. In conclusion, our results highlight the emergence of subtle EM deficits as a potential marker of the preclinical stages of AD in a continuous process that runs in tandem with A β burden and neurodegeneration biomarkers. The anticipation of

diagnosis and a more refined characterization of AD-related memory deficits might greatly contribute to disease-modifying therapies that can ease or delay clinical symptoms, to extend the patient's quality of life and to relieve the burden of caregivers.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00415-024-12834-y>.

Author contributions Filipa Raposo Pereira wrote the original draft, edited, and was involved in the material preparation, project administration, methodology, formal analysis and investigation, data curation, and visualization. Maximilien Chaumon was involved in the formal analysis and investigation, methodology, and writing review. Bruno Dubois was responsible for the conceptualization, was involved in the funding acquisition, project administration, resources, and writing review. Hovagim Bakardjian was involved in the conceptualization and investigation. Mahsa Bahrami was involved in the resources, data collection, and writing review. Marie-Odile Habert was involved in the formal analysis and methodology of PET exams. Katia Andrade was involved in the writing review. Nadja Younsi was involved in the resources. Valentina La Corte was involved in the conceptualization, methodology, resources, supervision, project management, validation, funding acquisition, and writing paper review and editing. Nathalie George was involved in the formal analysis, project managing, methodology, resources, supervision, validation, funding acquisition, and writing review and editing. All authors read and approved the final manuscript.

Funding The authors would like to thank Christiane Metzinger for her help in getting access to the INSIGHT pre-AD study data, and Francis Nyasse for his involvement in earlier stages of the coordination of the INSIGHT pre-AD study. We thank all INSIGHT-preAD participants for their commitment to the study. This study was funded by a grant from the Fondation Recherche Alzheimer (FRA) to Nathalie George, Valentina La Corte, and Filipa Raposo Pereira [FRA, project: ePARAD]. The INSIGHT study [IDRCB: 2012-A01731-42] was legally sponsored by Inserm [protocol nb. C12-59]. It received funding from the program "Investissements d'avenir" [ANR-10-IAIHU-06] and from Pfizer. The Fondation Plan-Alzheimer further funded the MRI and the ^{18}F -fluorodeoxyglucose PET along with the MEMENTO study. Avid provided the ^{18}F -florbetapir ligand for amyloid β PET. The study was done in collaboration with the Centre Hospitalier Universitaire de Bordeaux [study number CIC EC7].

Code and data availability The current study was based on the INSIGHT-preAD cohort, whose principal investigator is Prof. Bruno Dubois. The data from the INSIGHT-preAD study cannot be publicly shared due to privacy and ethical restriction. For further details please find the following link: <https://www.gaaindata.org/partner/INSIGHT-preAD>. The code used for data analysis is available on the following link: https://osf.io/knfwb/?view_only=32c3a9e29da14ee1aa2f84ba48562573

Declarations

Conflict of interest The authors declare no competing interests that are relevant to the content of this article. The funding bodies had no role in the study design, in the data analysis and interpretation, or in the article writing.

Ethical approval All aspects of this study were designed in full compliance with the French law n° 2004–806 (9th August 2004), the Good Clinical Practice principles (I.C.H version 4 of 1st May 1996 and the decision of 24th November 2006), the guidelines of the Helsinki Decla-

ration (Ethical Principles for Medical Research involving Human Subjects, Tokyo, 2004), and the human ERP research technical guidelines [57]. The INSIGHT-PreAD study protocol was approved by the ethics committee of the Pitié-Salpêtrière Hospital (IDRCB: 2012-A01731-42) and the present project was approved by the INSIGHT-preAD scientific committee. All the clinicians and neuropsychologists involved in the acquisition of data from the INSGHT cohort had authorization for using all the tests include on the study.

Consent to participate After study conditions were presented, a signed informed consent and willingness to commit to the longevity of the project was obtained from all individual participants included in the study.

Consent to publish The authors confirm that human research participants provided informed consent for publication of the images present in the article.

References

- Gauthier S, Webster C, Servaes S, et al (2022) World Alzheimer report 2022—life after diagnosis: navigating treatment, care and support. London
- Dubois B, Epelbaum S, Nyasse F et al (2018) Cognitive and neuroimaging features and brain β -amyloidosis in individuals at risk of Alzheimer's disease (INSIGHT-preAD): a longitudinal observational study. *Lancet Neurol* 17:335–346
- De Simone MS, Perri R, Fadda L et al (2019) Predicting progression to Alzheimer's disease in subjects with amnesic mild cognitive impairment using performance on recall and recognition tests. *J Neurol* 266:102–111. <https://doi.org/10.1007/s00415-018-9108-0>
- Sperling RA, Aisen PS, Beckett LA et al (2011) Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dementia* 7:280–292. <https://doi.org/10.1016/j.jalz.2011.03.003>
- Dubois B, Feldman HH, Jacova C et al (2014) Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* 13:614–629. [https://doi.org/10.1016/S1474-4422\(14\)70090-0](https://doi.org/10.1016/S1474-4422(14)70090-0)
- Mormino EC, Papp KV (2018) Amyloid accumulation and cognitive decline in clinically normal older individuals: implications for aging and early Alzheimer's disease. *J Alzheimer's Disease* 64:S633–S646. <https://doi.org/10.3233/JAD-179928>
- Chipi E, Salvadori N, Farotti L, Parnetti L (2019) Biomarker-based signature of Alzheimer's disease in pre-MCI individuals. *Brain Sci* 9:213
- Grober E, Veroff AE, Lipton RB (2018) Temporal unfolding of declining episodic memory on the free and cued selective reminding test in the prodementia phase of Alzheimer's disease: implications for clinical trials. *Alzheimer's Dementia: Diagnosis Assess Disease Monit* 10:161–171. <https://doi.org/10.1016/j.dadm.2017.12.004>
- Backman L, Jones S, Berger A-K et al (2005) Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. *Neuropsychology* 19:520–531. <https://doi.org/10.1037/0894-4105.19.4.520.supp>
- Papp KV, Rentz DM, Mormino EC et al (2017) Cued memory decline in biomarker-defined preclinical Alzheimer disease. *Neurology* 88:1431–1438
- Gagliardi G, Epelbaum S, Houot M et al (2019) Which episodic memory performance is associated with Alzheimer's disease biomarkers in elderly cognitive complainers? Evidence from a longitudinal observational study with four episodic memory tests (Insight-PreAD). *J Alzheimer's Disease* 70:811–824. <https://doi.org/10.3233/JAD-180966>
- Hanseuw BJ, Betensky RA, Schultz AP et al (2017) Fluorodeoxyglucose metabolism associated with tau-amyloid interaction predicts memory decline. *Ann Neurol* 81:583–596. <https://doi.org/10.1002/ana.24910>
- Jack CR, Wiste HJ, Weigand SD et al (2017) Age-specific and sex-specific prevalence of cerebral β -amyloidosis, tauopathy, and neurodegeneration in cognitively unimpaired individuals aged 50–95 years: a cross-sectional study. *Lancet Neurol* 16:435–444. [https://doi.org/10.1016/S1474-4422\(17\)30077-7](https://doi.org/10.1016/S1474-4422(17)30077-7)
- Jack CR, Wiste HJ, Weigand SD et al (2017) Defining imaging biomarker cut points for brain aging and Alzheimer's disease. *Alzheimer's Dementia* 13:205–216. <https://doi.org/10.1016/j.jalz.2016.08.005>
- Tarawneh R, Holtzman DM (2010) Biomarkers in translational research of Alzheimer's disease. *Neuropharmacology* 59:310–322. <https://doi.org/10.1016/j.neuropharm.2010.04.006>
- Weise CM, Chen K, Chen Y et al (2018) Left lateralized cerebral glucose metabolism declines in amyloid- β positive persons with mild cognitive impairment. *Neuroimage Clin* 20:286–296. <https://doi.org/10.1016/j.nicl.2018.07.016>
- Staffaroni AM, Melrose RJ, Leskin LP et al (2017) The functional neuroanatomy of verbal memory in Alzheimer's disease: [18F]-Fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) correlates of recency and recognition memory. *J Clin Exp Neuropsychol* 39:682–693. <https://doi.org/10.1080/13803395.2016.1255312>
- Landau SM, Mintun MA, Joshi AD et al (2012) Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Ann Neurol* 72:578–586. <https://doi.org/10.1002/ana.23650>
- Carapelle E, Serra L, Modoni S et al (2017) How the cognitive reserve interacts with β -amyloid deposition in mitigating FDG metabolism: an observational study. *Medicine (Baltimore)* 96:e5876
- Nelson PT, Alafuzoff I, Bigio EH et al (2012) Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. *J Neuropathol Exp Neurol* 71:362–381
- Sarazin M, Berr C, De Rotrou J et al (2007) Amnesic syndrome of the medial temporal type identifies prodromal AD A longitudinal study. *Neurology* 69:1859–1867
- Teipel SJ, Cavedo E, Lista S et al (2018) Effect of Alzheimer's disease risk and protective factors on cognitive trajectories in subjective memory complainers: An INSIGHT-preAD study. *Alzheimer's Dementia* 14:1126–1136. <https://doi.org/10.1016/j.jalz.2018.04.004>
- Jardin EC (2018) Recognition Memory Revisited: An Aging and Electrophysiological Recognition Memory Revisited: An Aging and Electrophysiological Investigation Investigation. ETD Archive.
- Grober E, Lipton RB, Hall C, Crystal H (2000) Memory impairment on free and cued selective reminding predicts dementia. *Neurology* 54:827–832. <https://doi.org/10.1212/WNL.54.4.827>
- Auriacombe S, Helmer C, Amieva H et al (2010) Validity of the free and cued selective reminding test in predicting dementia: the 3C study. *Neurology* 74:1760–1767. <https://doi.org/10.1212/WNL.0b013e3181df0959>
- Lemos R, Marôco J, Simões MR et al (2017) The free and cued selective reminding test for predicting progression to Alzheimer's disease in patients with mild cognitive impairment: a prospective longitudinal study. *J Neuropsychol* 11:40–55. <https://doi.org/10.1111/jnp.12075>
- Loewenstein DA, Curiel RE, Duara R, Buschke H (2018) Novel cognitive paradigms for the detection of memory impairment

- in preclinical Alzheimer's Disease. *Assessment* 25:348–359. <https://doi.org/10.1177/1073191117691608>
28. Dubois B, Feldman HH, Jacova C et al (2007) Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 6:734–746. [https://doi.org/10.1016/S1474-4422\(07\)70178-3](https://doi.org/10.1016/S1474-4422(07)70178-3)
 29. Dubois B, Albert M (2004) Amnesic MCI or prodromal Alzheimer's disease? *Lancet Neurol* 3:246–248
 30. Schacter DL, Dodson CS (2001) Misattribution, false recognition and the sins of memory. *Philosophical Trans Royal Soc B: Biol Sci* 356:1385–1393. <https://doi.org/10.1098/rstb.2001.0938>
 31. Papp KV, Amariglio RE, Mormino EC et al (2015) Free and cued memory in relation to biomarker-defined abnormalities in clinically normal older adults and those at risk for Alzheimer's disease. *Neuropsychologia* 73:169–175. <https://doi.org/10.1016/j.neuropsychologia.2015.04.034>
 32. Grober E, Lipton RB, Sperling RA et al (2022) Associations of stages of objective memory impairment with Amyloid PET and structural MRI: the A4 study. *Neurology* 98:E1327–E1336. <https://doi.org/10.1212/WNL.000000000200046>
 33. Christ BU, Combrinck MI, Thomas KGF (2018) Both reaction time and accuracy measures of intraindividual variability predict cognitive performance in Alzheimer's disease. *Front Hum Neurosci* 12:124. <https://doi.org/10.3389/fnhum.2018.00124>
 34. Ramratan WS, Rabin LA, Wang C et al (2012) Level of recall, retrieval speed, and variability on the cued-recall retrieval speed task (CRRST) in individuals with amnesic mild cognitive impairment. *J Int Neuropsychol Soc* 18:260–268. <https://doi.org/10.1017/S1355617711001664>
 35. Loewenstein DA, Acevedo A, Agron J, Duara R (2007) Vulnerability to proactive semantic interference and progression to dementia among older adults with mild cognitive impairment. *Dement Geriatr Cogn Disord* 24:363–368. <https://doi.org/10.1159/000109151>
 36. Deason RG, Hussey EP, Ally BA, Budson AE (2012) Changes in response bias with different study-test delays: evidence from young adults, older adults, and patients with Alzheimer's disease. *Neuropsychology* 26:119–126. <https://doi.org/10.1037/a0026330>
 37. Deason RG, Tat MJ, Flannery S et al (2017) Response bias and response monitoring: evidence from healthy older adults and patients with mild Alzheimer's disease. *Brain Cogn* 119:17–24. <https://doi.org/10.1016/j.bandc.2017.09.002>
 38. Dunn JC (2004) Remember-know: a matter of confidence. *Psychol Rev* 111:524–542. <https://doi.org/10.1037/0033-295X.111.2.524>
 39. Raposo Pereira F, George N, Dalla Barba G et al (2024) The memory binding test detects early subtle episodic memory decline in preclinical Alzheimer's disease: a longitudinal study. *J Alzheimer's Disease* 98:465–479. <https://doi.org/10.3233/JAD-230921>
 40. Folstein MF, Folstein SE, Mchugh PR (1975) "Mini-Mental State" a practical method for grading the cognitive state of patients for the clinician*. *J psychiaf Res* 12:189–198
 41. Morris JC (1993) The clinical dementia rating (cdr): Current version and scoring rules. *Neurology* 43:2412–2414. <https://doi.org/10.1212/wnl.43.11.2412-a>
 42. Habert MO, Bertin H, Labit M et al (2017) Evaluation of amyloid status in a cohort of elderly individuals with memory complaints: validation of the method of quantification and determination of positivity thresholds. *Ann Nucl Med* 32:75–86. <https://doi.org/10.1007/s12149-017-1221-0>
 43. Dubois B, Slachevsky A, Litvan I, Pillon B (2000) The FAB: a frontal assessment battery at bedside. *Neurology* 55:1621–1626. <https://doi.org/10.1212/WNL.55.11.1621>
 44. Tombaugh TN (2004) Trail making test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol* 19:203–214. [https://doi.org/10.1016/S0887-6177\(03\)00039-8](https://doi.org/10.1016/S0887-6177(03)00039-8)
 45. McNair DM, Kahn RJ (1983) Self-assessment of cognitive deficits. In: Crook T, Ferris S, Bartus R (eds) *Assessment in geriatric psychopharmacology*. Mark Powley, New Canaan, CT, pp 119–136
 46. Grober E, Buschke H, Korey SR (1987) Genuine Memory Deficits in Dementia. *Dev Neuropsychol* 3:13–36. <https://doi.org/10.1080/87565648709540361>
 47. Van Der Linden M, Calicis F, Adam S, and members of the GREMEM (2004) L'épreuve de rappel libre/rappel indice à 16 items (RL/RI-16). In M. Van der Linden, S. Adam, A. Agniel, C. Baisset Mouly et les membres du GREMEM (Eds.), *L'évaluation des troubles de la mémoire: présentation de quatre tests de mémoire épisodique (avec leur étalonnage)*. Solar
 48. Raposo Pereira F, George N, Dalla Barba G et al (2024) The memory binding test can anticipate Alzheimer's disease diagnosis at an early preclinical stage: a longitudinal study in the INSIGHT-preAD cohort. *Front Aging Neurosci* 16:1414419. <https://doi.org/10.3389/fnagi.2024.1414419>
 49. Westfall PH, Troendle JF (2008) Multiple testing with minimal assumptions. *Biom J* 50:745–755. <https://doi.org/10.1002/bimj.200710456>
 50. Johnson SC, Christian BT, Okonkwo OC et al (2014) Amyloid burden and neural function in people at risk for Alzheimer's Disease. *Neurobiol Aging* 35:576–584. <https://doi.org/10.1016/j.neurobiolaging.2013.09.028>
 51. Dubois B, Picard G, Sarazin M (2009) Early detection of Alzheimer's disease: new diagnostic criteria. *Dialogues Clin Neurosci* 11:135–139
 52. Quenon L, Dricot L, Woodard JL et al (2016) Prediction of Free and Cued Selective Reminding Test Performance Using Volumetric and Amyloid-Based Biomarkers of Alzheimer's Disease. *J Int Neuropsychol Soc* 22:991–1004. <https://doi.org/10.1017/S1355617716000813>
 53. Morris GP, Clark IA, Vissel B (2018) Questions concerning the role of amyloid- β in the definition, aetiology and diagnosis of Alzheimer's disease. *Acta Neuropathol* 136:663–689. <https://doi.org/10.1007/s00401-018-1918-8>
 54. Iozzo P, Guzzardi MA (2019) Imaging of brain glucose uptake by PET in obesity and cognitive dysfunction: Life-course perspective. *Endocr Connect* 8:R169–R183. <https://doi.org/10.1530/EC-19-0348>
 55. Weise CM, Chen K, Chen Y et al (2022) Differential impact of body mass index and leptin on baseline and longitudinal positron emission tomography measurements of the cerebral metabolic rate for glucose in amnesic mild cognitive impairment. *Front Aging Neurosci* 14:1031189. <https://doi.org/10.3389/fnagi.2022.1031189>
 56. Lindbergh CA, Walker N, La Joie R et al (2021) Worth the Wait: Delayed Recall after 1 Week Predicts Cognitive and Medial Temporal Lobe Trajectories in Older Adults. *J Int Neuropsychol Soc* 27:382–388. <https://doi.org/10.1017/S1355617720001009>
 57. Picton TW, Bentin S, Berg P et al (2000) Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37:127–152. <https://doi.org/10.1111/1469-8986.3720127>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

SUPPLEMENTARY MATERIAL

Article title: Recognition memory decline is associated with the progression to prodromal Alzheimer's disease in asymptomatic at-risk individuals

Journal name: Journal of Neurology

Author names: Filipa Raposo Pereira¹, Maximilien Chaumon, Bruno Dubois, Hovagim Bakardjian, Mahsa Bahrami, Marie-Odile Habert, Katia Andrade, Nadja Younsi, Valentina La Corte, Nathalie George, for the INSIGHT-preAD study group

¹Corresponding author affiliations :

- Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, INSERM, U 1127, CNRS, UMR 7225, AP-HP, CENIR, Centre MEG-EEG, Hôpital de la Pitié-Salpêtrière, Paris, France
- Department of Neurology, Institute of Memory and Alzheimer's Disease (IM2A), Pitié-Salpêtrière University Hospital, AP-HP, Paris, France

¹Corresponding author email: filipa.raposopereira@icm-institute.org

ESM 1_List of the assessments and biological markers collected in the INSIGHT-PreAD study.

All the assessments were performed at baseline. The clinical, cognitive and psychobehavioural data were collected every 6-months. The neuropsychological, EEG, and actigraphy data were collected every 12-months. The blood sampling, MRI, ¹⁸F-FDG PET, and 18F-florbetapir PET data were collected every 24-months. Abbreviations: PET= Positron Emission Tomography; MRI= Magnetic resonance imaging; DTI= Diffusion Tensor Imaging; EEG= Electroencephalogram; CSF= Cerebrospinal Fluid; PCR= Polymerase Chain Reaction.

1.1 Subjective feelings about memory and cognition assessments

- 15-item version of the McNair Frequency of Forgetting Questionnaire (modified from16)
- Healthy Age Brain Care Monitor (HABC-M)
- INSIGHT-preAD Questionnaire of Cognitive Decline (IQCD)
- Self-Evaluation Assessment of Complaints (AC)
- Visual Analogic Scale for Complaints (ASC)
- AD-related Anxiety Questionnaire (AD-NOS)

1.2 Psycho-behavior, mood, autonomy, and quality of life assessments

- Neuropsychiatric Inventory (NPI)
- State-Trait-Anxiety Inventory Y (STAI-Y-B)
- Geriatric Depression Scale (GDS)
- Starkstein Apathy Scale
- Bristol Activities of Daily Living Scale (Bristol ADL)
- Amsterdam Instrumental Activity of Daily Living Questionnaire (IADL)
- EuroQol 5D test (EQ-5D)

1.3 Cognitive function assessments

1.3.1 Global cognitive function

- Mini Mental State Examination (MMSE)
- Clinical Dementia Rating (CDR)

1.3.2 Episodic memory:

- Free and Cued Selective Reminding (FCSRT)
- Delayed Matching to Sample Task (DMS-48; immediate and delayed)
- Rey-Osterrieth Complex Figure (3-min and 30-min recall)
- Memory Binding Test (MBT)

1.3.3 Working memory and executive function:

- Forward and backward Digit and Visuo-spatial span
- Frontal Assessment Battery (FAB)
- Trail Making Test (TMT; B–A time)
- Lexical Fluency (p-words in 2 minutes)

1.3.4 Instrumental function

- Semantic Fluency (animals in 2 minutes)
- Image Naming Test (DO-80)
- Praxis assessment
- Rey-Osterrieth Complex Figure (copy)

1.4 Brain imaging assessments

- PET: ¹⁸F-Florbetapir PET to test brain amyloid- β deposition
- PET: ¹⁸F-Fluorodeoxyglucose (FDG) PET to test brain glucose hypometabolism
- MRI: 3D T1-weighted magnetization-prepared rapid gradient echo (MP-RAGE) MRI
- MRI: 2D FLAIR T2*- fluid-attenuated inversion recovery MRI
- DTI & MRI: 2D T2* DTI and a T2*-weighted gradient-echo echo-planar MRI series
- MRI: Pulsed arterial spin labelling (ASL) MRI
- MRI: 3D T1 MRI sequences to measure hippocampal volume
- Cortical thickness.

1.5 Brain electrophysiological assessments (EEG)

- High-density EEG during rest
- High-density EEG during a cognitive task (memory recall of words)

1.6 Biomarker assessments

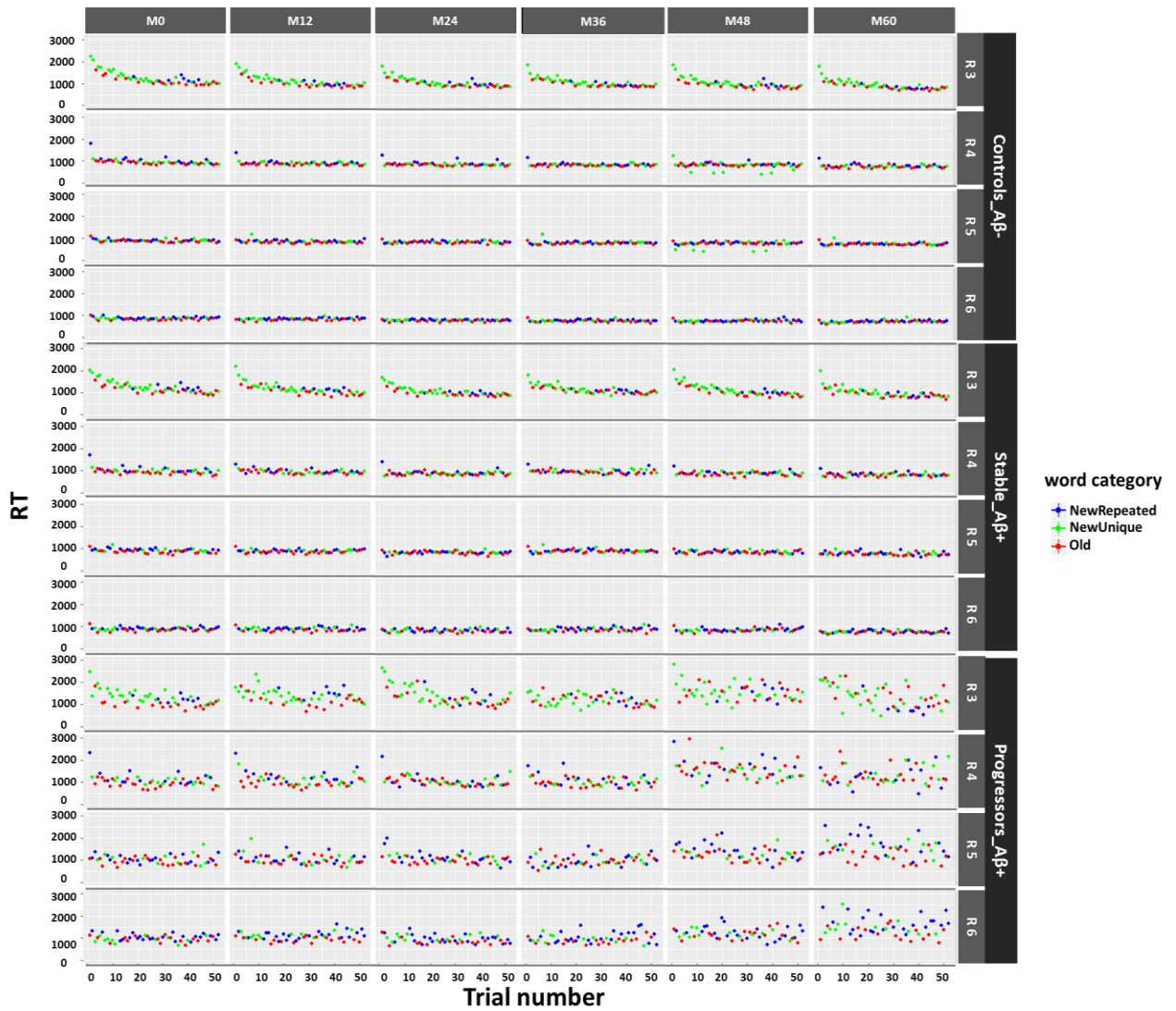
- CSF : total tau protein concentration (t-tau) in CSF
- CSF: tau protein phosphorylated concentration at threonine 181 (p-tau181) in CSF
- CSF: amyloid- β peptide 1-42 (A β 1-42) concentration in CSF
- CSF: ration of amyloid- β peptide (A β 1-42) to amyloid- β peptide (A β 1-40) in CSF
- Blood biomarkers

1.7 APOE ϵ 4 allele genotyping

- PCR-based Sanger sequencing

ESM 2_Overall distribution of data across trials in each run of the task, on each session, in all participants.

The plot represents the reaction time (RT; y-axis) of each individual for each word category (New Repeated words in blue; New Unique words in green; Old words in red) per trial number (x-axis). The columns correspond to every session (M0 to M60). The first 4 rows correspond to the performance by of the participants from the control group, the middle 4 rows correspond to the performance of the participants from the stable group, and the last 4 rows correspond to the performance of the participants from the progressor group. Each row represents one run of the task (with 4 runs in total: R3, R4, R5, R6).



ESM 3_ Individual reaction time (RT) and Accuracy performances

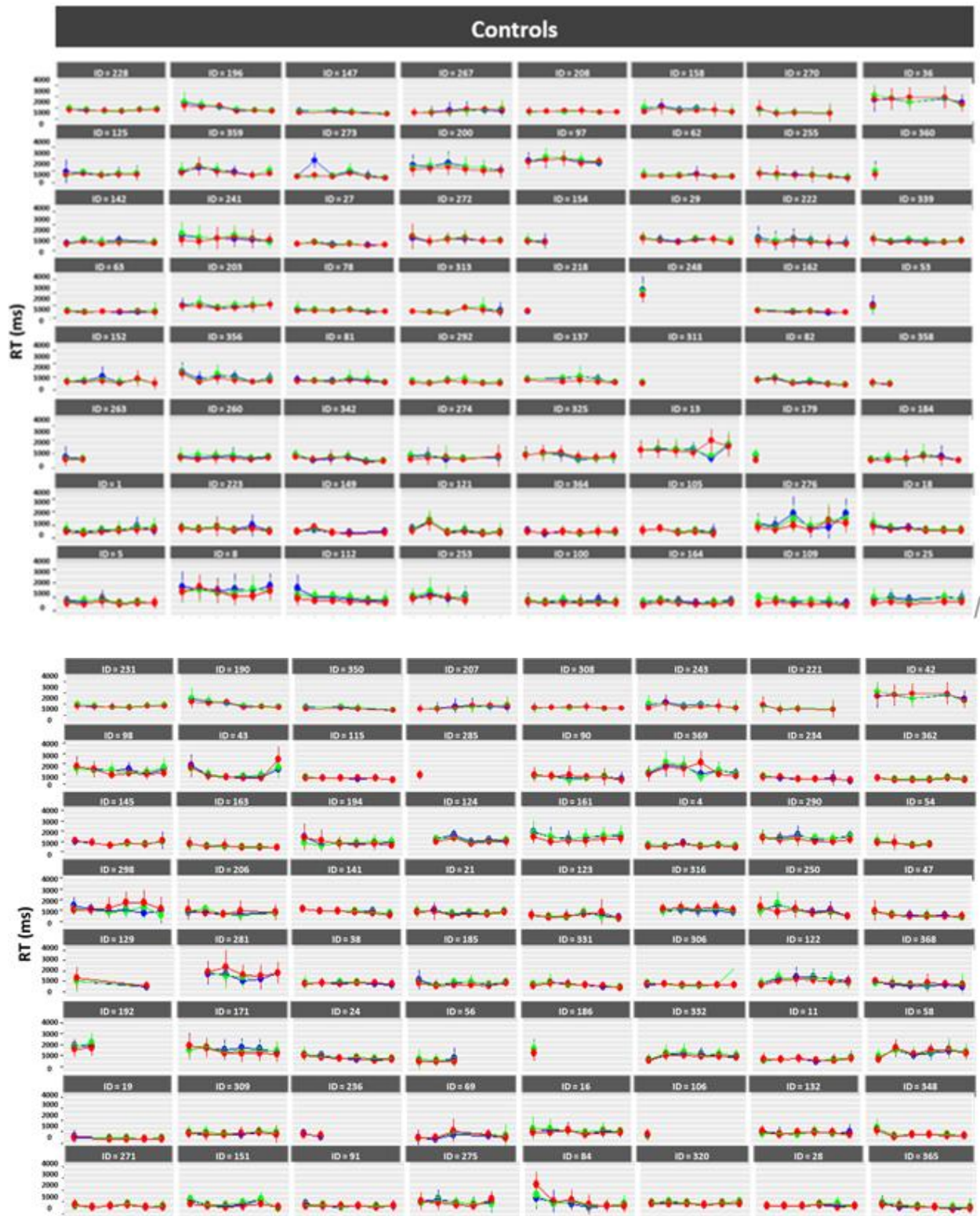
Each plot represents either the accuracy (unit: number of correct responses) or the reaction time (unit = ms; y-axis) of each individual (indicated by their ID number) for each word category (New Repeated words in blue; New Unique words in green; Old words in red) on each session (M0 to M60; x-axis). The individual data are presented successively for each group (control [$A\beta$ -/N-]; $A\beta$ +/stable N- and N+; $A\beta$ +/progressor N- and N+).

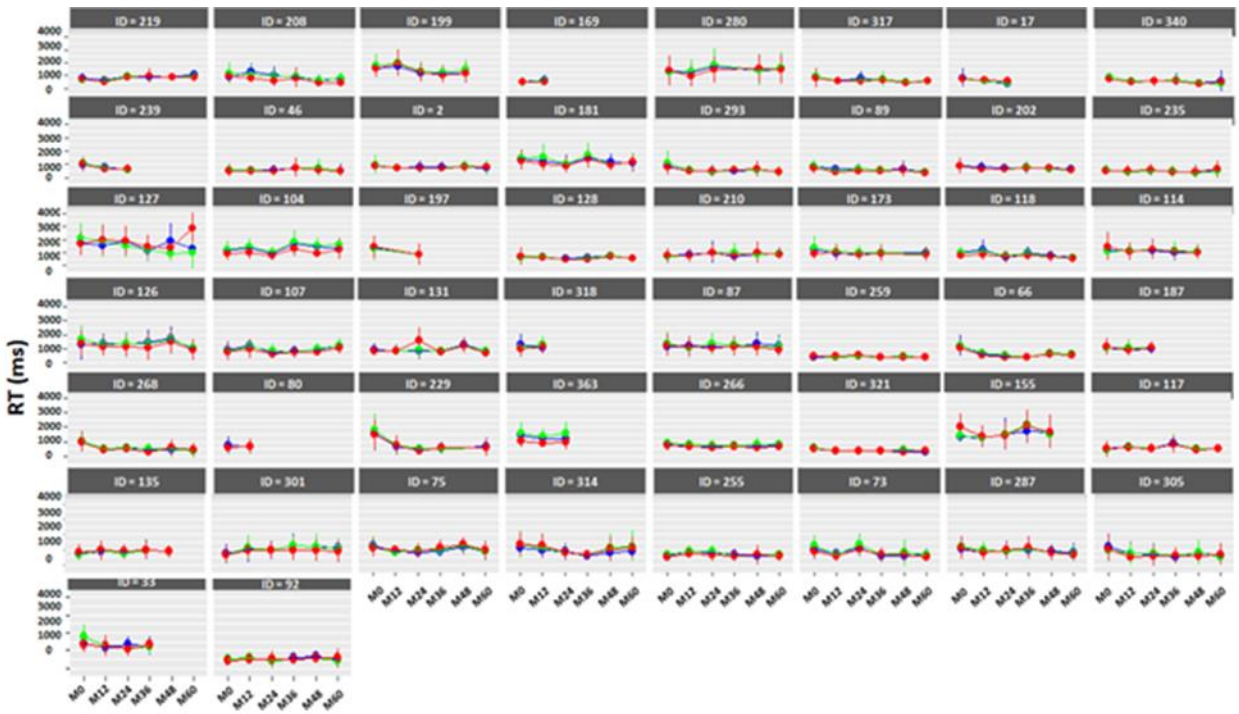
3.1 Accuracy of control participants ($A\beta$ -/N-)



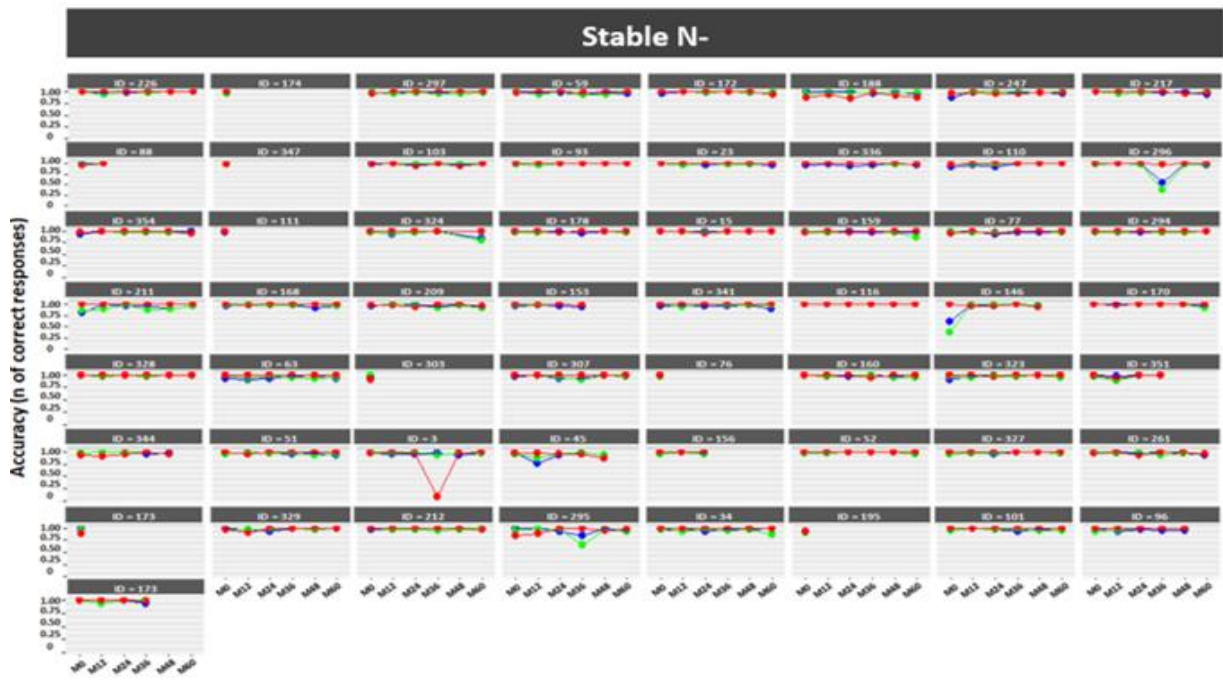


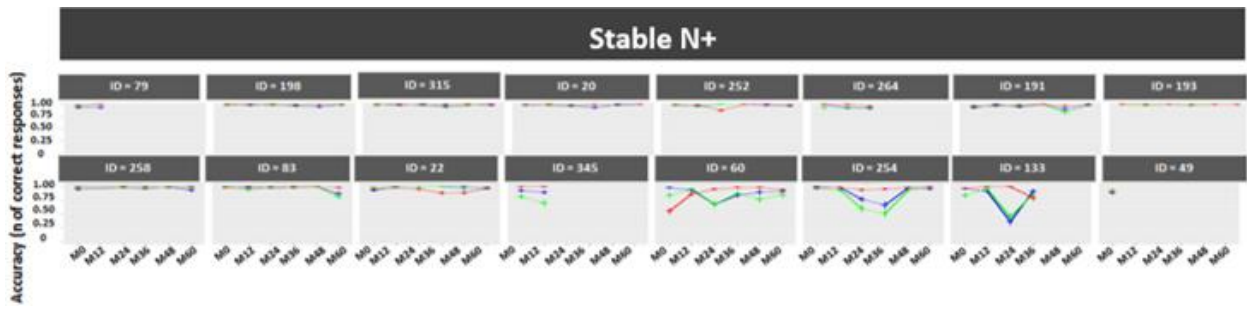
3.2 Reaction Time of control participants ($A\beta$ -/N-)



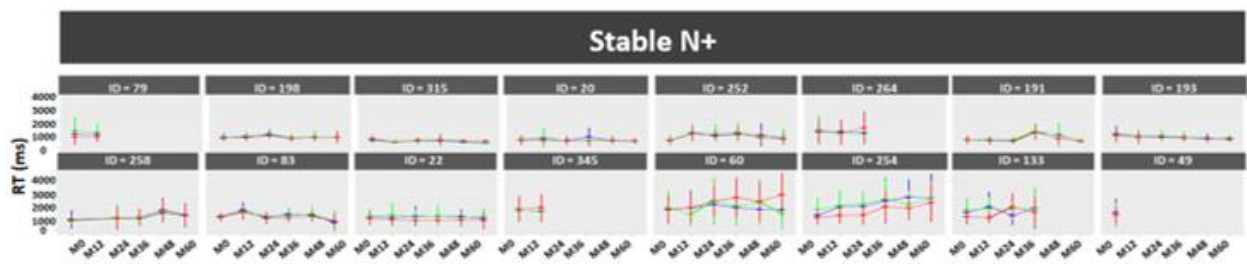
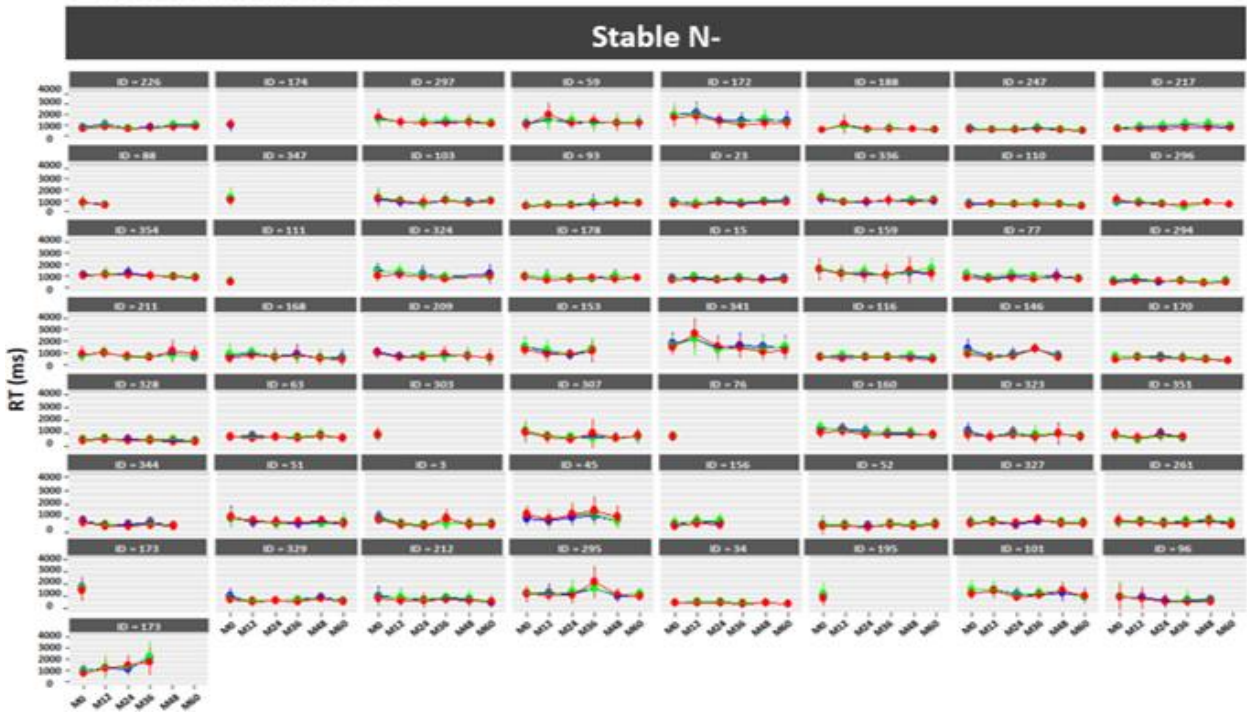


3.3 Accuracy of the stable participants





3.4 Reaction Time of the stable participants



3.5 Accuracy of the progressor participants



