

## Measuring cognitive change in subjects with prodromal Alzheimer's disease

T Mura\*, MD, PhD <sup>1, 2, 3, 4</sup>, C Proust-Lima, PhD <sup>5, 6</sup>, H Jacqmin-Gadda PhD <sup>5, 6</sup>, TN Akbaraly PhD <sup>1,2,7</sup>, J. Touchon, MD, PhD<sup>1,2,8</sup>, B Dubois, MD, PhD <sup>9</sup>, C Berr, MD, PhD <sup>1,2,8</sup>.

<sup>1</sup> INSERM, U1061, Neuropsychiatrie : Recherche Epidémiologique et Clinique, 34093 Montpellier, Cedex 5, France

<sup>2</sup> Université Montpellier I, 34095 Montpellier, Cedex 5, France

<sup>3</sup> Département d'Information Médicale & Centre d'Investigation Clinique, CHRU Montpellier, 34093 Montpellier, France

<sup>4</sup> INSERM, CIC 1001, Montpellier, France

<sup>5</sup> INSERM U897, Equipe de Biostatistique, Centre de Recherche en Epidémiologie et Biostatistique, F-33076 Bordeaux, France

<sup>6</sup> Université Bordeaux Segalen, ISPED, F-33076 Bordeaux, France

<sup>7</sup> Department of Epidemiology and Public Health, University College London, London, United Kingdom.

<sup>8</sup> CMRR Languedoc Roussillon, service de Neurologie, CHRU Montpellier, 34093 Montpellier, France

<sup>9</sup> INSERM-UPMC UMRS 975, Institut de la Mémoire et de la Maladie d'Alzheimer, ICM, APHP, Salpêtrière Hospital, University Paris 6, Paris, France

**Corresponding author:** Dr Thibault Mura, Inserm U1061 - Hôpital La Colombière, Pavillon 42  
39 av. Charles Flahault - 34493 Montpellier - Cedex 5 - France

**Phone:** 33 (0)4 67 33 23 28, **Fax:** 33 (0)4 67 33 23 35 **email:** [t-mura@chu-montpellier.fr](mailto:t-mura@chu-montpellier.fr)

**Word count:** Text: 3 483 Abstract: 248 Figure: 4 Table: 3 References: 40

**Key words:** Mild Cognitive Impairment; Alzheimer's disease; Cognitive neuropsychology in dementia; cohort study; neuropsychological assessment.

## **Structured Abstract:**

**Objective:** To investigate the sensitivity of a large set of neuropsychological tests to detect cognitive changes due to prodromal Alzheimer's Disease(AD); to compare their metrological properties in order to select a restricted number of these tests for the longitudinal follow-up of subjects with prodromal-AD.

**Participants:** 212 patients with mild cognitive impairment(MCI) were tested at baseline by a standardized neuropsychological battery which included: the Free-and-Cued-Selective-Recall-Reminding test (FCSRT), the Benton-Visual-Retention test, the DENO100, verbal-fluency, a serial-digit-learning test, the double-task-of-Baddeley, the WAIS similarities, the Trail-Making-Test, and the WAIS-digit-symbol test. Patients were monitored every 6 months for up to 3 years in order to identify those who converted to AD (retrospectively classified as prodromal-AD). Statistical analyses were performed using a nonlinear multivariate mixed model involving a latent process. This model assumes that the psychometric tests are nonlinear transformations of a common latent cognitive process, and captures the metrological properties of tests.

**Results:** 57 patients converted to AD. The most sensitive tests in the detection of cognitive changes due to prodromal-AD were the FCSRT, the semantic verbal-fluency, and the DENO100. Some tests exhibited a higher sensitivity to cognitive changes for subjects with high levels of cognition: the free-recall, free-delayed-recall scores (FCSRT) and the semantic verbal-fluency, whereas others showed a higher sensitivity at low levels of cognition: the total-recall score (FCSRT).

**Conclusions:** Tests used for the follow-up of prodromal-AD subjects should be chosen among those which actually decline in this stage of the disease, and should be selected according to the subject's initial scores.

## **GLOSSARY**

AD = Alzheimer's disease; BVRT = Benton Visual Retention Test; DSM-III-R = Diagnostic and Statistical Manual of Mental Disorders, 3rd ed., revised; DST = Digit Substitution Test; FCSRT = Free and Cued Selective Recall Reminding Test; IADL = Instrumental Activities of Daily Living; LCP = Latent Cognitive Process; MCI = Mild Cognitive Impairment; MMSE = Mini-Mental State Examination; SDOT = Serial Digit Ordering Test; TMT = Trail Making Test; WAIS = Wechsler Adult Intelligence Scale.

## Background

To describe patients with cognitive difficulties without criteria diagnosis for dementia, Flicker *et al.*<sup>1</sup> coined the term Mild Cognitive Impairment (MCI) with the idea that this concept could help in the early prediction of the development of dementia. However, MCI defines a syndrome – whose definition has been revised and changed several times since 1991<sup>2-4</sup> – and may be the consequence of different diseases with distinct aetiologies. Consequently, some authors have emphasized the need to clearly characterize the underlying disorder (i.e. Alzheimer's Disease (AD), vascular dementia...)<sup>5, 6</sup>. This characterization could facilitate predicting the clinical progression of the disease and promote the development of specific therapeutic approaches at an early stage of the disease<sup>6</sup>.

In this context, the measurement of cognitive changes has two main uses. First, this measure could help to identify subjects with prodromal-AD (i.e. symptomatic pre-dementia phase of AD or "MCI due to AD"<sup>6</sup>). Indeed, a working group has recently highlighted the importance of measuring cognitive change for the diagnosis of "MCI due to AD" (i.e. prodromal-AD) stating as a diagnostic criterion that it “provides evidence of longitudinal decline in cognition”<sup>7</sup>. The second use is to enable the follow-up of patients with prodromal-AD, either in a clinical care context, or in intervention trials with cognitive change as a primary outcome to assess the effects of drugs at an early stage of AD<sup>8-10</sup>.

Yet, in all of these situations, there is currently no recommendation concerning the choice of neuropsychological tests to be used. In order to be informative for the identification or the follow-up of subjects with prodromal-AD, a neuropsychological test must fulfill some requirements. First, selected tests have to be markers of the evolution of the disease, and thus have to explore the cognitive domains which are affected in the course of the different stages of the disease. Secondly, the tests have to be able to measure cognitive changes in the range of cognitive levels of the targeted population<sup>8</sup>. This capacity can be limited by the metrological

properties of neuropsychological tests (i.e. floor and ceiling effects, curvilinearity: sensitivity to detect cognitive changes varying according to the initial cognitive level). These points can be analyzed using a recently published statistical model which allows the study of metrological properties by modeling a latent cognitive factor which is the common cognitive part measured by the neuropsychological tests<sup>11</sup>.

This work aims to investigate the sensitivity of a large set of neuropsychological tests frequently used in subjects with MCI, to measure cognitive changes due to prodromal-AD; and to describe and compare the metrological properties of these tests. This comparison will aid in selecting a restricted number of psychometric tests for the longitudinal identification and follow-up of subjects with prodromal-AD.

## Methods

### Study population

Data came from the PRE-AL Study (PREdiction of ALzheimer's disease)<sup>12</sup>. Briefly, in 2001 and 2002, 251 patients with MCI<sup>2</sup> were recruited and followed up semiannually over a period of 3 years. Subjects were recruited from memory clinics of 14 expert centers in the field of dementia across France (see Acknowledgments). Patients were enrolled according to the following criteria: 1) a subjective memory complaint; 2) an objective memory impairment documented by at least one word missing in the three-word recall of the Mini-Mental State Examination<sup>13</sup> (MMSE), or a score under 29 on the Isaacs-set test<sup>14</sup>, or both; 3) a preservation of general cognitive functioning (MMSE between 25 and 29/30); 4) a normal score or only one item impaired at the first level in the four Instrumental Activities of Daily Living (IADL)<sup>15</sup>; and 5) the absence of the Diagnostic and Statistical Manual of Mental Disorders, 3<sup>rd</sup> ed., revised (DSM-III-R) criteria for dementia<sup>16</sup>. Patients with focal lesions in cerebral imaging, or documented depressive symptoms or with other medical conditions, which could interfere with memory performance or follow-up, were excluded. Each subject signed an informed consent form once the nature of the procedures had been fully explained. The study was approved by the Ethics Committee of the "La Salpêtrière" Hospital.

Out of the 251 participants included in the PREAL study, patients who had had no follow-up (n=17), those with only one visit at 6 months (n=13), and patients with no measurements for any of the neuropsychological tests during their follow-up (n=3), were excluded from our analyses. As AD was the primary outcome of the study, patients who converted to a non-AD dementia (n=6) were excluded from further statistical analyses. Finally, the present analyses were carried out on a total of 212 participants.

### Data collection

#### *Dementia diagnosis and definition of prodromal-AD:*

Patients were seen every 6 months for 3 years. At each visit, clinical assessment included the recording of all medical events, current treatment, and complete neurologic examination. Activities of daily life were rated with the IADL scale during an interview with the patient and a knowledgeable collateral source (a spouse or a child)<sup>15</sup>. Memory complaint was assessed by a specific questionnaire<sup>17</sup>. The Clinical Dementia Rating scale (CDR) was completed at each visit during follow-up<sup>18</sup>. During the follow-up, when patient meets the clinical DSM-III-R<sup>16</sup> criteria for dementia and/or the score of 1 at the CDR scale, he was considered as "converted" to dementia. In order to minimize intercenter variability, all clinical records were reviewed by an Expert Committee composed of neurologists (n=3), neuropsychologists (n=3), geriatricians (n=3), and psychiatrists (n=3). They re-examined whether clinical criteria for dementia were satisfied using DSM-III-R criteria<sup>16</sup> and classified the patient as AD using the *NINCDS-ADRDA* Criteria<sup>19</sup>. Patients who converted to AD during the 3-year follow-up were retrospectively classified as patients with prodromal-AD, other subjects were designated as MCI-non AD.

#### *Assessment of neuropsychological performance:*

All subjects were tested at inclusion and annually using a standardized neuropsychological battery. In the event of a suspected conversion, the patient underwent an additional neuropsychological evaluation 6 months later. Six cognitive domains commonly affected by ageing and AD were assessed using 13 scores<sup>20-27</sup> described in Table 1. *The mean duration of the administration of the complete predictive battery was 91 minutes (SD = 15).*

### Statistics

We used a nonlinear mixed model for multivariate longitudinal data involving a latent process (illustrated in [figure 1](#)) to analyze the cognitive trajectory over time of the MCI subjects during the follow-up. The statistical model assumes that the correlation between the neuropsychological tests is induced by a *latent cognitive process (LCP)* representing the common cognitive part measured by the neuropsychological tests. The model is divided into two parts: (a) a linear mixed model describes the change over time in the LCP and evaluates the common effects of covariates on this latent cognitive trajectory, and (b) test-specific nonlinear measurement models relate each administration of the neuropsychological tests with the LCP, by describing and accounting for the metrological properties of the tests, and evaluating test-specific associations with covariates (here specifically prodromal-AD status). Therefore the impact of a given covariate can be explored on the LCP and on each test.

(a) Effect of the covariates on latent cognitive change with time: The LCP trajectory was modeled using a linear mixed model<sup>28</sup> which evaluates changes of a repeated outcome over time (i.e. the LCP) and accounts for correlations between the repeated measures of each subject. The linear mixed model included a subject-specific random intercept and a slope for time after inclusion (in years of follow-up), as well as the following covariates: age, sex, educational level, prodromal-AD status and their interaction with the time elapsed since inclusion. To define the LCP dimension, we assumed it followed a Gaussian distribution  $N(0,1)$  at baseline for the reference state of the linear mixed regression model (i.e. women, with a low baseline level, MCI non AD of 55 years old).

(b) Metrological properties and test specific effects: The test-specific measurement models consisted in flexible transformations linking the neuropsychological tests with the latent cognitive process. These transformations which capture the metrological properties of the tests<sup>29</sup>, are covariate-and-time-invariant parametric functions depending on parameters that are estimated



simultaneously with the other parameters of the model. Beta cumulative distribution functions were chosen as flexible and parsimonious transformations. Indeed, these functions offered a large variety of shapes (concave, convex, sigmoid, or simply linear) and thus modeled the curvilinearity of the tests. The complete methodology was developed elsewhere<sup>11, 30</sup>. This part of the model also evaluated the test-specific effects of covariates on each test after adjustment for their effect on the LCP.

(c) Mean annual change for each neuropsychological test according to the occurrence of AD during the follow-up (in latent cognitive process units) was calculated by adding together the common effect of time on the LCP (a), the test-specific effect of time on each test (b) and the test-specific effect of the interaction between time and prodromal-AD status on each test (b).

Statistical tests were performed at the conventional two-tailed  $\alpha$  level of 0.05. Data were analyzed using SAS Enterprise Guide version 4.3 (SAS Institute, Cary, North Carolina) and a FORTRAN90 executable for the nonlinear mixed model with the latent process (program HETMIXSURV available at <http://biostat.isped.u-bordeaux2.fr/>).

## Results

Out of the 212 MCI participants included in the present analyses, 57 converted to AD during the 3-year follow-up (figure 2) and were retrospectively classified as patients with prodromal-AD. The characteristics of these 212 participants are detailed in [Table 2](#).

### *Sensitivity of the tests to detect cognitive changes due to prodromal-AD*

As shown in [Table 3](#), older age was associated with both a lower latent cognitive level at baseline ( $\beta=-0.081$  units of LCP at baseline per year of age  $p<0.01$ ) and a steeper latent cognitive decline (steeper decline of  $\beta=-0.010$  units of LCP per years of follow-up per year of age  $p<0.01$ ). A higher educational level was associated with a higher latent cognitive level at baseline, but not with a slower latent cognitive decline. Gender was neither associated with latent cognitive level at baseline, nor with a steeper latent cognitive decline. As expected, cognitive change over time was different between prodromal-AD and MCI non-AD (steeper decline of  $\beta=-0.615$  units of LCP per year of follow-up in prodromal-AD,  $p\text{-value}<0.01$ ). Possibly due to a practice effect, MCI non-AD patients showed an improvement in their scores during the study ( $\beta=0.25$  units of LCP per years of follow-up,  $p\text{-value}<0.01$ ), whereas the score of prodromal-AD patients declined during the follow-up ( $\beta=-0.35$  units of LCP per year,  $p\text{-value}<0.01$ ).

[Figure 3](#) presents the *mean test-specific annual change* in both groups (for the reference state of regression model (i.e. women, with a low baseline level, MCI non AD) at the age of 71.8 years old). In subjects with prodromal-AD, the three FCSRT scores, the DENO100, the verbal fluencies and the SDOT significantly decreased over time. For MCI non-AD patients, none of the scores significantly decreased during the follow up, and four tests showed a significant improvement over time (FCSRT-free-recall and delayed-free-recall, BVRT, DENO100 and WAIS-DST).

Finally, when comparing the annual variations of tests between patients with prodromal-AD and MCI non-AD ([Figure 3](#)), change over time in all of the tests except three (Baddeley, Weis-Similarities and TMT-B) differed between the two groups. Tests displaying the greatest differences between groups were the three FCSRT scores, the DENO100 and the verbal fluencies.

*Comparison of the metrological properties of the neuropsychological tests.*

The estimated transformations between each test and the common **LCP** are shown in [Figure 4](#). Except for the verbal-fluency (category) and the WAIS-DST, all the tests showed a nonlinear transformation. This indicates that a given loss of points for a given cognitive test does not correspond to the same intensity of decline of the LCP over the whole range of the test. Different curve shapes could be identified between the transformations: (1) *concave* for the FCSRT (total-recall), the DENO100, the SDOT and the WAIS-similarities; (2) *convex* for verbal-fluency (letter-R) and the TMT-A and TMT-B; (3) *sigmoid* for the other FCRST scores (free-recall and free-delayed-recall) and the BVRT; (4) close to *horizontal* for the Baddeley; (5) almost *linear* for the verbal-fluency (category) and the WAIS-DST.

Each of these shapes highlighted a different sensitivity pattern for cognitive change over time. Tests with a *concave shape* are more sensitive to cognitive change at **low** rather than at high levels of cognition. For example, a change of 1.0 unit of LCP in the medium-**high** cognitive range (0.0;5.0) represents a loss of 2.8 points for the SDOT, whereas the same change in the medium-**low** range (-5.0;0.0) represents a loss of 7.4 points. Similarly, tests with a *convex shape* are more sensitive to cognitive change at **high** rather than at low cognitive levels. Tests with a *sigmoid shape* are more sensitive to cognitive change in the medium range of the LCP. The *horizontal shape* of the Baddeley-Mü indicates that this test does not vary with changes in the LCP and, therefore, does not reflect any cognitive change in this population. A *linear shape*

indicates that a loss of points in the scoring scale used in the test corresponds to the same intensity of decline of the LCP over the whole range of the test.

Exploring the low levels of cognition and the floor effect: Despite their high sensitivity to cognitive change at low levels of cognition, the four tests with a concave shape do not present the same ability to detect changes at low levels of cognition. While the FCSRT (total-recall score) can explore low and very low levels of cognition, the shape of the SDOT is quite vertical below -5 units of LCP, revealing the inability of the SDOT to discriminate major cognitive impairment. Among other tests, the two TMTs showed horizontal shapes in the lowest range of the LCP. This illustrates the floor effect of these tests: minimum values of the tests are reached for a medium LCP value.

Exploring the high levels of cognition and the ceiling effect: The horizontal shape of total-recall (FCSRT) in high cognitive levels underlines the ceiling effect of this test, and suggests that it is not appropriate to assess cognitive changes in subjects with medium and high cognitive levels. Among the other tests, only the free-recall (FCSRT) and the verbal-fluency (category) cover the highest values of the LCP (+5.0 to +10.0) and can actually discriminate changes in these high cognitive levels.

## Discussion

In this clinical prospective study of 212 MCI individuals, we provide rational explanations for selecting neuropsychological tests for the longitudinal identification and follow-up of subjects with prodromal-AD. These explanations are based on the analysis of the sensitivity of neuropsychological tests to detect cognitive changes due to prodromal-AD; and on the description of the metrological properties of these tests (i.e. variable sensitivity to cognitive change, floor and ceiling effect).

A major result of this study is the identification of tests that are sensitive to the cognitive change due to prodromal-AD. Some tests present a poor sensitivity to cognitive change and seem, therefore, of limited interest in the context of research or clinical care for a longitudinal follow-up, namely Baddeley's double task, the TMT-B and the WAIS-Similarities. Conversely, among the thirteen scores analyzed, six showed a significant decrease over time in the prodromal-AD patients, and a substantial difference in their evolution compared with MCI non-AD subjects: the 3 FCSRT scores, the semantic verbal-fluency, the DENO100 and the SDOT. These results are consistent with the neuroanatomic distribution of histopathological abnormalities reported in the mildest stages of AD which overlaps, at least partially, with the regions implicated in these tests. For example, the neuropathological changes in the early stage of AD begin primarily in medial temporal regions (hippocampus, entorhinal cortex)<sup>31</sup> which are known to be critical for episodic memory function. Consequently, an impaired ability to learn and retain new information (i.e. an episodic memory deficit) is usually the earliest and most prominent feature of AD<sup>32</sup>. The decrease over time of both free-recall and total-recall scores (FCSRT) is therefore consistent with the aggravation of the amnesic syndrome of the medial temporal type<sup>12</sup>. Among other tests, the observed decline of the verbal-fluency (category) test and the DENO100 illustrates the early impairment of access to semantic memory, due to the temporal neocortical damage that occurs in AD<sup>33</sup>. The greater impairment in semantic fluency rather than in letter fluency has been

previously reported in Mild AD<sup>34, 35</sup>. A potential limitation of this study is that it included subjects with MCI defined according criteria dating from 1999<sup>2</sup>. The use of other criteria to define MCI, or the carrying out of a similar study in the general population, could have led to different results. Another point is that subjects were followed over a limited period of three years, and it is possible that some AD occurred after the third years of follow-up. However, considering the decreasing incidence of AD during the follow-up, with only 5 cases diagnosed during the last visit (Figure 2), it is likely that this number is low and that this information bias is minimal.

With regards to the metrological properties of the tests, our results suggest a potential interest of selecting neuropsychological tests for a longitudinal follow-up according to the initial cognitive level of the target population (in a research setting) or of the patient (in a clinical setting). Some tests should be chosen for evaluating cognitive changes at high levels of cognition (FCSRT free-recall score/ delayed-free-recall score and verbal-fluency "*category*"), whereas others should be avoided because of their ceiling effect (FCSRT total-recall, DENO100). Conversely, some tests seem more useful for evaluating cognitive changes at low levels of cognition (FCSRT total-recall score), while some should be avoided (WAIS-DST, TMT-A, SDOT). Another result of this study is that a majority of the scores exhibited strong curvilinearity. This curvilinearity could have consequences on the results of intervention trials or in epidemiological research<sup>36</sup>, and thus has to be taken into account in the analysis. In addition, this curvilinearity partly explains the remaining difficulties in providing a cut-off value for cognitive decline in the diagnosis of Prodromal-AD. Indeed, for most of the tests, a decrease of a given number of points will not have the same meaning according to the initial value of the score. Thus it could be interesting to create an algorithm which would include different scores and/or different cut-off values to provide a prediction of the probability of developing AD rather than use a single test with a single cut-off value.

The improvement observed for some tests in MCI non-AD patients (FCRST, BVRT, DENO100) is probably due to practice and/or learning effects<sup>37</sup>. This kind of effect has been extensively documented in healthy participants for tasks assessing different cognitive functions including verbal episodic memory<sup>38</sup>. Contrary to some studies<sup>37, 39</sup>, the observed improvement was not limited to the second testing but was observed during the entire study (data not shown). This effect is generally considered as an interfering variable complicating the interpretation of results<sup>40</sup>. However, it could also be considered as an interesting property which could help to differentiate the normal ability of an individual to learn and adjust with practice, from the pathological process of an individual with prodromal-AD.

Our statistical methodology has several advantages. It can handle unbalanced repeated measurements and bounded quantitative outcomes, and it can take into account and describe the metrological properties of neuropsychological tests<sup>11</sup>. However, it is worth noting that the latent cognitive process in this model is necessarily defined according to the pool of psychometric tests used in the analysis. In this study we used a large battery of neuropsychological tests which assessed several cognitive domains. Therefore, the modeled LCP could reasonably be interpreted as a global cognitive factor. However, a limitation of this study is that it was not designed to compare numerous different measures of a specific cognitive domain (like episodic memory), and hence cannot reasonably highlight one measure over others that were not tested.

Our study provides rational explanations for the selection of neuropsychological tests for the cognitive follow-up of patients with MCI in a clinical care context (for diagnostic or prognostic purposes of prodromal-AD) or research context (to identify a target population, or as an outcome for the effect of an early intervention in prodromal-AD). The tests that can be recommended are those which actually demonstrate a decline at the stage of prodromal-AD (the three FCRST scores, the semantic verbal-fluency, the SDOT and the DENO100), and which are able to measure cognitive changes in the range of cognitive levels of the targeted population. In

the current study, the free-recall, the free-delayed-recall score (FCSRT) and the semantic verbal-fluency test cover a wide cognitive range and seem to be adapted for the follow-up of subjects with an initial medium or high level of cognition. Conversely, the total-recall score (FCSRT) suffered from a very considerable ceiling effect, but appeared to be the best score for following up subjects with an initial low cognitive level. For future research, comparing several neuropsychological tests (or score) within a specific cognitive domain (particularly verbal episodic memory and language) using the same methodology could be of great interest for refining the choice of cognitive tests to be used.



**Table 1:** Assessment of neuropsychological performances

<b>Cognitive domain assessed</b>	<b>Name of the Neuropsychological test</b>	<b>Scores (units/max)</b>
<b>Verbal episodic memory</b>	Free and Cued Selective Reminding Test (FCSRT) <sup>20,21</sup>	- Free-recall score (number of words at the free recall /48) - Total-recall score (number of words at the total free and cued recall /48) - Delayed-free-recall score (number of words at the delayed free recall /16)
<b>Immediate visual memory</b>	Benton Visual Retention Test <sup>22</sup>	- BVRT (number of accurate recognitions /15)
<b>Language</b>	DENO100 <sup>23</sup>	- DENO1000 (number of accurate recognitions /100)
	Verbal-fluency test	- Letter "R" (number of words in 2 minutes) - Category "fruits" (number of words in 2 minutes)
<b>Working memory</b>	Serial Digit Ordering Test <sup>24</sup>	- SDOT (number of figures until the order was broken in the series /105)
	Double task of Baddeley <sup>25</sup>	- Baddeley-Mü (decline of the visual-motor task and the serial recall task in single and dual condition /150)
<b>Conceptual elaboration</b>	Similarities of the Wechsler Adult Intelligence Scale <sup>27</sup>	- WAIS-Similarities (2 points for abstract or fundamental aspect, 1 point for concrete or functional aspect /28)
<b>Executive function</b>	Trail Making Test <sup>26</sup>	- TMT-A ((number of correct items/total time)*10) - TMT-B ((number of correct items/total time)*10)
	Digit Symbol Test of the WAIS <sup>27</sup>	- WAIS-DST (number of symbols correctly associated in 90 seconds)

**Table 2:** Baseline characteristics of the 212 MCI patients included in the analysis

	Mean (std) or		
	N(%)	Interquartile	
	N=212	min-max	range (P25;P75)
<b>Age at baseline</b>	71,8(5,3)	58,0;81,0	68,0;76,0
<b>Gender : Male</b>	87(41%)		
<b>Educational level</b>			
8 years and under	122 (57%)		
9 to 12 years	40(19%)		
12 years or over	50(24%)		
<b>Prodromal-AD</b>	57(27%)		
<b>Neuropsychological tests :</b>			
FCSRT: free-recall-score	22,2(7,9)	1,0;38,0	16,5;28,0
FCSRT: total-recall-score	40,8(8,2)	11,0;48,0	39,0;46,5
FCSRT: delayed-free-call	8,0(3,9)	0,0;16,0	5,0;11,0
BVRT	11,6(2,1)	6,0;15,0	10,0;13,0
Deno100	88,2(8,3)	58,0;100,0	84,5;94,0
Verbal Fluency: letter-R	17,7(6,5)	3,0;38,0	13,0;21,0
Verbal Fluency: Fruit	15,9 (4,7)	2,0-32,0	12,0 ; 19,0
SDOT	79,8(11,6)	11,0;97,0	75,0;87,0
Baddley Mü	94,3(12,2)	49,5;122,7	88,2;102,6
WAIS similarities	18,0(5,3)	1,0;27,0	15,0;22,0
TMT-B	2,1(1,1)	0,4;6,1	1,2;2,7
TMT-A	5,2(2,1)	1,4;13,2	3,8;6,8
WAIS-DST	35,9(10,9)	9,0; 65,0	28,0;43,0

**Table 3:** Parameter estimates of the linear mixed model for the latent cognitive process (LCP): effect of age, sex, educational status and prodromal-AD status and their interaction with time during the follow-up

	Cross-sectional effects <i>at baseline</i>			Longitudinal effects <i>Interaction with Time</i>		
	Estimate	SE	P-value	Estimate	SE	P-value
<b>Time*</b>				0,258*	0,072	<0.01
<b>Prodromal-AD</b>	-1,811	0,312	<0.01	-0,615	0,136	<0.01
<b>Age<sup>†</sup> at baseline</b>	-0,081	0,016	<0.01	-0,010	0,003	<0.01
<b>Gender (male)</b>	0,162	0,161	0.32	-0,062	0,037	0.09
<b>Educational level</b>						
8 years and under	0,000	---	---	0,000	---	---
9 to 12 years	0,993	0,208	<0.01	0,000	0,004	0.99
12 years or over	1,158	0,203	<0.01	0,005	0,043	0.91

\* Mean change each year after inclusion in the reference group (MCI non-AD)

† Age in year after 55 years of age

## **Acknowledgements**

The authors thank all of the investigators of the PreAl study—Serge Belliard (Rennes), Didier Hannequin (Rouen), Marie Pierre Hervy (Kremlin Bicetre), Bernard Laurent (Saint-Etienne), Sylvie Legrain (Paris), Bernard Michel (Marseille), Florence Pasquier (Lille), Michelle Puel (Toulouse), Anne Sophie Rigaud (Paris), Philippe Robert (Nice), Martine Vercelletto (Nantes), Marc Verny (Paris) —for clinical evaluations and their help during this study.

Thanks to Joana Norton for proofreading.

## **Competing Interests Statement:**

The authors have no conflict of interest to declare.

## **Funding Sources**

This work was funded by the "French Ministry of Health"[grant PHRC national 2000]. TNA is funded by the “*Chercheur d'Avenir*” grant from the Languedoc-Roussillon region (France).

## **Authorship Statement:**

All authors have made substantial contribution to this work: T.M. performed statistical analyses and interpretation of the data and wrote the first draft of the manuscript. J.T., B.D. and C.B. were involved in data acquisition, and revised the article for important intellectual content. C.P. H.J. and T.A. have made substantial contribution to the conception of the article, helped to interpret the statistical analysis, and revised the article for important intellectual content. All authors read and approved the final version of the article. This article contains original data. Thibault MURA had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## References

1. Flicker C, Ferris Sh, Reisberg B. Mild cognitive impairment in the elderly: predictors of dementia. *Neurology* 1991;41:1006–9.
2. Petersen Rc, Smith Ge, Waring Sc, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56:303–8.
3. Petersen Rc, Doody R, Kurz A, et al. Current concepts in mild cognitive impairment. *Arch Neurol* 2001;58:1985–92.
4. Artero S, Petersen R, Touchon J, et al. Revised criteria for mild cognitive impairment: validation within a longitudinal population study. *Dement Geriatr Cogn Disord* 2006;22:465–70.
5. Molinuevo JI, Valls-Pedret C, Rami L. From mild cognitive impairment to prodromal Alzheimer disease: A nosological evolution. *European Geriatric Medicine* 2010;1:146–154.
6. Dubois B, Albert MI. Amnesic MCI or prodromal Alzheimer's disease? *Lancet Neurol* 2004;3:246–8.
7. Albert MS, Dekosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:270–9.
8. Morris MC, Evans DA, Hebert LE, et al. Methodological issues in the study of cognitive decline. *Am J Epidemiol* 1999;149:789–93.
9. Siemers E. Designing clinical trials for early (pre-dementia) Alzheimer's disease: determining the appropriate population for treatment. *J Nutr Health Aging* 2011;15:22–4.

10. Dubois B. Interest of the new criteria for drug trials in AD. *J Nutr Health Aging* 2009;13:356–7.
11. Proust C, Jacqmin-Gadda H, Taylor Jm, et al. A nonlinear model with latent process for cognitive evolution using multivariate longitudinal data. *Biometrics* 2006;62:1014–24.
12. Sarazin M, Berr C, De Rotrou J, et al. Amnesic syndrome of the medial temporal type identifies prodromal AD: a longitudinal study. *Neurology* 2007;69:1859–67.
13. Folstein Mf, Folstein Se, Mchugh Pr. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
14. Isaacs B, Kennie At. The Set test as an aid to the detection of dementia in old people. *Br J Psychiatry* 1973;123:467–70.
15. Barberger-Gateau P, Dartigues Jf, Letenneur L. Four Instrumental Activities of Daily Living Score as a predictor of one-year incident dementia. *Age Ageing* 1993;22:457–63.
16. American Psychiatrique Association. Diagnostic and statistical manual of mental disorders. Washington, DC: American Psychiatrique Association, 1987.
17. Al Aloucy J, Dalla Barba G, Roudier M. Metamemory, self-awareness and anosognosia in patient with mild cognitive impairment, mild Alzheimer's disease and control healthy older adults. *neurology* 2006;66(suppl 2):A129. Abstract S16.005.
18. Morris Jc. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–4.
19. Mckhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of

- Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–44.
20. Grober E, Buschke H. Genuine memory deficits in dementia. *Dev Neuropsychol* 1987;3:13–36.
21. Van Der Linden M, Coyette F, Poitrenaud J, et al. L'épreuve de rappel libre/rappel indicé à 16 items (RL/RI – 16). In: Van der Linden M AS, Agniel A et les membres du GREMEM, ed. L'évaluation des troubles de la mémoire. Présentation de quatre tests de mémoire épisodique (avec leur étalonnage). Marseille, France: Solal, 2004.
22. Benton A. The revised visual retention test: clinical and experimental applications. New York: Psychological corporation, 1963.
23. Kremin H, Perrier D, De Wilde M. DENO-100 Paradigme experimental et test clinique de dénomination contrôlée: effet relatif de 7 variables expérimentales sur les performances de 16 sujets atteints de maladies dégénératives. *Rev Neuropsychol* 1999;9:439–440.
24. Cooper Ja, Sagar Hj, Jordan N, et al. Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain* 1991;114 ( Pt 5):2095–122.
25. Baddeley Ad, Baddeley Ha, Bucks Rs, et al. Attentional control in Alzheimer's disease. *Brain* 2001;124:1492–508.
26. Reitan R. Validity of the Trail making Test as an indicator of organic brain damage. *Percept Mot Skills* 1958;8:271–276.
27. Weschler D. The Wechsler Adult Intelligence Scale-Revised. San Antonio: Psychological Corporation, 1981.
28. Laird Nm, Ware Jh. Random-effects models for longitudinal data. *Biometrics* 1982;38:963–74.

29. Proust-Lima C, Amieva H, Dartigues Jf, et al. Sensitivity of four psychometric tests to measure cognitive changes in brain aging-population-based studies. *Am J Epidemiol* 2007;165:344–50.
30. Proust-Lima C, Amieva H, Letenneur L, et al. Gender and education impact on brain aging: a general cognitive factor approach. *Psychol Aging* 2008;23:608–20.
31. Delacourte A, David Jp, Sergeant N, et al. The biochemical pathway of neurofibrillary degeneration in aging and Alzheimer's disease. *Neurology* 1999;52:1158–65.
32. Salmon Dp. Neuropsychological Features of Mild Cognitive Impairment and Preclinical Alzheimer's Disease. *Curr Top Behav Neurosci* 2011.
33. Carter Sf, Caine D, Burns A, et al. Staging of the cognitive decline in Alzheimer's disease: insights from a detailed neuropsychological investigation of mild cognitive impairment and mild Alzheimer's disease. *Int J Geriatr Psychiatry* 2011.
34. Biundo R, Gardini S, Caffarra P, et al. Influence of APOE Status on Lexical-Semantic Skills in Mild Cognitive Impairment. *J Int Neuropsychol Soc* 2011:1–8.
35. Adlam Al, Bozeat S, Arnold R, et al. Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex* 2006;42:675–84.
36. Galasko Dr, Gould Rl, Abramson Is, et al. Measuring cognitive change in a cohort of patients with Alzheimer's disease. *Stat Med* 2000;19:1421–32.
37. Collie A, Maruff P, Darby Dg, et al. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. *J Int Neuropsychol Soc* 2003;9:419–28.
38. Blasi S, Zehnder Ae, Berres M, et al. Norms for change in episodic memory as a prerequisite for the diagnosis of mild cognitive impairment (MCI). *Neuropsychology* 2009;23:189–200.



39. Jacqmin-Gadda H, Fabrigoule C, Commenges D, et al. A 5-year longitudinal study of the Mini-Mental State Examination in normal aging. *Am J Epidemiol* 1997;145:498–506.
40. Bartels C, Wegrzyn M, Wiedl A, et al. Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neurosci* 2010;11:118.

**Figure 1:** Conceptualization of the nonlinear mixed model involving a latent process to model cognition from several neuropsychological tests.

**Figure 2:** Diagram mapping the administration of the neuropsychological tests and the occurrence of AD during the three-year follow-up (FU) of the study.

**Figure 3:** Mean annual change for each neuropsychological test according to the occurrence of AD during the follow-up (in latent cognitive process units).

**Figure 4:** Metrological properties of the thirteen neuropsychological scores used in the study