Estimating Dementia Onset: AT(N) Profiles and Predictive Modeling in Mild Cognitive Impairment Patients

Carlos Platero[a,e], Jussi Tohka[b], Bryan Strange[c,d]

February 6, 2024

[a]Health Science Technology Group, Technical University of Madrid, 28012 Madrid, Spain. [b]Biomedical Image Analysis Group, University of Eastern Finland, FI-70211 Kuopio, Finland. [c]Laboratory for Clinical Neuroscience, CTB, Technical University of Madrid, IdISSC, Madrid, Spain. [d]Alzheimer Disease Research Unit, CIEN Foundation, Queen Sofia Foundation Alzheimer Center, Madrid, Spain.

[e]Corresponding author: Carlos Platero at Health Science Technology Group, Technical University of Madrid, Ronda de Valencia 3, 28012, Madrid, Spain, carlos.platero@upm.es, phone number: +34910677576.

Platero C, Tohka J, Strange B. Estimating Dementia Onset: AT(N) Profiles and Predictive Modeling in Mild Cognitive Impairment Patients. Curr Alzheimer Res. 2024;20(11):778-790. doi: 10.2174/0115672050295317240223162312.

Abstract

Background: Mild cognitive impairment (MCI) usually precedes the symptomatic phase of dementia and constitutes a window of opportunities for preventive therapies.

Objectives: Predict the time a MCI patient has left to reach dementia and obtain a most likely natural history in the progression of MCI towards dementia.

Methods: This is a study with 633 MCI patients and 145 subjects with dementia through 4726 visits over 15 years from Alzheimer Disease Neuroimaging Initiative (ADNI) cohort. A combination of data from AT(N) profiles at baseline and longitudinal predictive modeling was applied. A data-driven approach was proposed for categorical diagnosis prediction and timeline estimation of cognitive decline progression, which combined supervised and unsupervised learning techniques.

Results: A reduced vector of only neuropsychological measures was selected for training the models. At baseline, this approach had high performance in detecting subjects at high risk of converting from MCI to dementia in the coming years. Furthermore, a disease progression model (DPM) was built and also verified using three metrics. As a result of the DPM focus on the studied population, it was inferred that amyloid pathology (A+) appears about 7 years before dementia, tau pathology (T+) and neurodegeneration (N+) occur almost simultaneously, between 3 and 4 years before dementia. In addition, MCI-A+ subjects were shown to progress more rapidly to dementia compared to MCI-A- subjects.

Conclusions: Based on proposed natural histories and cross-sectional and longitudinal analysis of AD markers, the results indicate that only a single cerebrospinal fluid sample is necessary during the prodromal phase of AD. Prediction from MCI into dementia and its timeline can be achieved exclusively through neuropsychological measures.

Keywords: Mild cognitive impairment, Alzheimer's disease, AT(N) biomarkers, Predictive models, Disease Progression Modeling.

1 INTRODUCTION

Alzheimer's disease (AD), the most common dementia, is marked by cognitive loss and specific neuropathological changes. Accumulation of amyloid plaques, neurofibrillary tangles (pathological tau) and neurodegeneration lead to impairments in memory and other cognitive domains and subsequently a dementia syndrome [1]. The AT(N) framework defines AD by neuropathologic change, with cognitive impairment viewed as a symptom [2]. Biomarkers associated with amyloidosis are denoted by the label A, those indicative of tau pathology are marked with T, and those measuring neurodegeneration carry the designation N. These measures can be binary, leading to eight biomarker profiles. Amyloid pathology biomarkers determine if an individual is on the Alzheimer's continuum. Tau biomarkers confirm an AD diagnosis. Neurodegenerative/neuronal injury biomarkers and cognitive symptoms are used to stage severity.

Mild cognitive impairment (MCI) is often considered to be a transitional phase between healthy cognitive aging and dementia [3, 4]. MCI is a key factor in managing AD, but it varies in clinical outcomes. Despite many MCI cases converting to AD (MCI progressing to dementia (pMCI)), many remain stable or regain normal cognitive function [5]. Understanding AD and effectively identifying individuals with pMCI at early stages is crucial. A prognosis with high sensitivity and specificity for MCI to AD progression holds significant value in clinical practice and medical research [6, 7]. The risk of progression from MCI to dementia depends on a number of factors, including age, sex, and apolipoprotein E (ApoE) status [8]. AT(N) profiles may be of particular value in patients with MCI by providing prognostic information on the risk of progression from MCI to dementia as demonstrated by recent studies [9, 10]. Studies have examined these risks at group and individual levels [11, 12, 9, 13, 14, 15]. Automated techniques for predicting diagnosis using varied modalities have been reviewed [16, 17, 18, 19]. A meta-analysis found neuropsychological measures (NMs), cerebrospinal fluid (CSF) biomarkers or FDG-PET variables improved predictive performance, whereas T1-MRI did not significantly affect prediction [19]. On the other hand, gene sequencing technology has been employed for analyzing the pathogenesis and achieving precise diagnoses of MCI [20]. Recently, a new field named imaging genetics emerged to comprehensively study pathogeny with fusion data [21].

A new machine learning framework called disease progression models (DPMs) regards AD progression as a continuous process and derives long-term pathological trajectories from short-term clinical data [22, 23, 24, 25, 26]. Beyond giving a data-driven description of the natural evolution of AD, DPMs provide automatic diagnosis by explicitly ordering biomarkers from normal to pathological stages along the disease time axis in a multivariate manner. DPM is based on the analysis of longitudinal samples at different stages of the disease. A class of DPM relies on the ordering of the observed longitudinal trajectories and extracting the modeling of the temporal progress of the disease. Donohue et al. [22] modeled these trajectories based on self-modeling regression. Long-term progression curves for the multiple outcomes and subject-specific random effects and time shifts were estimated iteratively until the convergence of the proposed algorithm, which is called growth models by alternating conditional expectation (GRACE).

GRACE addresses the limitations of using clinical diagnosis in predictive models, which can overestimate performance and lead to errors in identifying clinical events. GRACE is an algorithm that uses unsupervised learning and does not rely on clinical information. However, each set of selected markers used in building a predictive model with GRACE represents a specific timeline in the natural progression of the disease. This study proposes combining supervised and unsupervised techniques to enhance predictions of progression to dementia.

Here, two different types of information on the progression from MCI to dementia were considered: 1) AT(N) profiles and 2) the evolution of cognitive symptoms. This article proposes to use the measurements of the AT(N) biomarkers at baseline and to estimate the cognitive evolution by means of the GRACE algorithm.

2 MATERIAL AND METHODS

2.1 Participants and measurements

The Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset was used [27]. Inclusion, exclusion, and assessment details are at http://adni.loni.usc.edu/. The study included MCI or dementia patients with a cognitive battery and at least one T1-weighted MRI. Data from the ADNIMERGE R package were used [28], focusing on neuropsychological measures like Alzheimer's Disease Assessment-Cognitive 13-item scale (ADAS13), Clinical Dementia Rating-Sum of Boxes (CDRSB), Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MOCA), Rey Auditory Verbal Learning Test (RAVLT), Everyday Cognition (ECog)-total by participant (ECogPtTotal) and study partner (ECogSPTotal) and Functional Assessment Questionnaire (FAQ). MRI measures of hippocampal, ventricular, entorhinal, and whole brain volumes were also included. A total of 633 MCI patients met our inclusion criteria, and their baseline diagnoses were early MCI (EMCI, n=307) and late MCI (LMCI, n=326) [29].

A subgroup of patients with available AD neuropathology biomarkers was identified. This included those who underwent baseline CSF Aβ_{42}, CSF pTau, CSF Tau, and/or FDG-PET examinations. In total, 501 MCI had AT(N) biomarker data. Aβ_{42}, pTau, and Tau concentrations were quantified using Elecsys assays [30,31]. For FDG-PET, an FDG composite score was calculated as the mean uptake in left and right angular, temporal, and posterior cingulate regions [32].

A+ subjects had CSF Aβ_{42} concentration levels below 880 pg/ml or $pTau/A\beta_{42}$ over 0.028 [31]. T+ and N+ were defined by CSF pTau over 27 pg/ml and CSF Tau over 300 pg/ml respectively [31]. N+ in FDG PET was identified with levels below 1.21 [32].

2.2 Convert and censuring times

Clinical diagnoses were made at each visit to identify MCI patients who progressed to probable dementia, according to ADNI clinical assessments. The conversion time for those who progressed (pMCI) was set between the baseline and their first dementia diagnosis, confirmed in subsequent visits. The stable MCI (sMCI) group consisted of MCI patients who did not develop dementia. The last visits of the sMCI subjects established the censorship times. From the selected population, 403 subjects were classified as sMCI and 230 as pMCI based on clinical follow-ups. In the subgroup that included CSF markers, MRI measurements, and neuropsychological tests, there were 333 sMCI and 168 pMCI. Additionally, dementia subjects were included to characterize the progression from MCI to AD using the selected biomarkers, adding 145 and 128 patients in the first and second analyses, respectively (see supplementary material, S1). About 30% (n = 182/633 or n = 134/501) of the MCI population converted in the first four years, which represented roughly 80% (n = 182/230 or n = 134/168) of all pMCI subjects. Conversion to dementia within 4 years was defined as fast pMCI converters, while later converters were termed as slow. The same categories were applied to sMCI subjects based on a 48-month frame.

2.3 Data processing

Two types of participant data were considered: AT(N) profiles and cognitive impairment severity. Only normal or abnormal AT(N) values were considered, with profiles obtained at baseline and longitudinally. Progression from MCI to dementia was assessed clinically. Both supervised and unsupervised learning techniques were used for diagnosis prediction and AD progression timeline estimation. Survival analysis was applied to longitudinal data to detect high-risk dementia progression markers. These markers were used to develop DPMs. The most reliable DPM was selected based on three measures of the suggested natural history.

2.4 Categorical predictive models and DPMs

A two-stage data-driven approach combining both supervised and unsupervised learning techniques was employed to build the categorical predict models and DPMs [33]. We analyzed the link between longitudinal data and AD progression using survival models that consider conversion times and finite follow-up [34, 35, 36]. Briefly, within the population of individuals with MCI, longitudinal measurements were obtained from the studied cohort. The modeling of these

outcomes was accomplished through Linear Mixed Effects (LME) models, incorporating both fixed effects and subject-specific random effects to capture the longitudinal trajectories [37, 38]. Consequently, this modeling approach allowed the estimation of marker values for each subject over time. Furthermore, information regarding the progression of the disease to dementia during the follow-up period was available, i.e. conversion and censorship times. An extended Cox model was constructed for each significant discrete time [36]. To build each of these predictive models, the hazard ratios were calculated and transformed into probabilistic terms of conversion from MCI to dementia using a logistic regression model. Utilizing a set of markers derived from NMs and T1-MRI data, feature selection and model building were carried out using a nested cross-validation (CV) procedure, preventing overfitting and biased performance estimates [39, 11]. The procedure consisted of two nested CV loops: an inner loop, designed to select the optimal feature subsets for the proposed predictive models, and an outer loop, designed to obtain an unbiased estimate of model performance. Within each inner CV loop, diverse combinations of markers with varying dimensions were proposed and subsequently evaluated in the outer CV loop. A feature ordering stage that uses the minimal-redundancy-maximal-relevance (mRMR) algorithm [40] to propose good subsets of markers for the prediction of MCI conversion into dementia was used. A resampling method was employed to identify the top subsets of features based on mutual information difference in mRMR. Then, predictive models were constructed using training data with the identified feature subsets, and the highest-performing combinations were selected based on classification measures using withheld test data [41].

Given proposed marker subsets from a study population, which were chosen among the better combinations of classification scores obtained by the categorical predictive models, GRACE has been shown to reveal the natural history of cognitive decline by ordering the short-term trajectories of the subjects. Let p markers be measured from n individuals at different follow-up times. We denote the measured outcome k for individual i at time j as y_{ijk}, where i=1,...,n, k = 1,...,p and j = 1,...,q_{ik}. The proposal of Donohue et al. [22] is expressed as:

$$y_{ijk} = g_k\left(t^c_{ijk} + \delta_i\right) + x'_{t^c_{ijk}}\beta_k + \alpha_{0_{ik}} + \alpha_{1_{ik}}t^c_{ijk} + e_{ijk},$$

where g_k is a continuously differentiable monotone function and δ_i is the unknown subject-specific time shift, which follows a normal distribution with mean zero and variance σ_δ^2. Short-term observation time is represented by t^c_{ijk}, which indicates the centered years in relation to the temporal evolution of the visits, $t^c_{ij} = t_{ij} - \left(t_{i_1} + t_{i_{end}}\right)/2$, where t_{ij} is the visit time and i_1 and i_{end} represented the first and last visit index of i–subject. $x'_{t^c_{ijk}}$ is the row vector for the fixed effects (including variables such as age and scan time), and β_k are the fixed effects coefficients. The long-term progression time is computed from $t^c_{ijk} + \delta_i$. The parameters $\alpha_{0_{ik}}$ and $\alpha_{1_{ik}}$ are the subject- and outcome-specific random intercept and slope. These values reflect how the subset of regression parameters for the i-th subject deviates from those of the population. e_{ijk} is a measurement error term that follows a zero-mean Gaussian distribution with variance σ_k. A self-modeling regression model was applied with linear subject-level effects and long-term features with nonparametric monotone smoothing [22]. The goal of GRACE is to estimate both the time shift parameters and the short-term and long-term curves. Importantly, follow-up clinical diagnostic information was not used in the development of the DPM. Before fitting DPMs to the data, the outcomes were transformed into percentiles using a weighted empirical cumulative distribution function to ensure a common scale. The resulting scale ranged from 0 to 1, with 0 representing the least severe observed value and 1 representing the most severe observed value. Finally, bootstrapping was applied to obtain confidence intervals for the DPM evaluation parameters and to analyze the long-term trajectories of the markers in patients with and without amyloid pathology.

2.5 The onset or zero time

We observe that δ_i represents a relative measure of disease progression, taking into account the variabilities of biomarkers observed in the studied population. However, the onset of dementia may be biased due to the inclusion of more MCI subjects than dementia subjects in our populations. Given an initial year of dementia or zero time, t_{onset}, the short-term trajectories of sMCI subjects are expected to be to the left of t_{onset}. In contrast, for pMCI subjects, the marker trajectories should cross t_{onset} and move towards the right. Therefore, when estimating a temporal ordering using GRACE with a subset of proposed markers, certain scores can be suggested. The first score we define is the percentage of sMCI subjects whose last visits, $t^c_{i_{end}} + \delta_i$, occur before t_{onset} in relation to the total number of sMCI subjects:

$$SCORE_1 = \frac{\#\left\{i|(i \in sMCI) \cap \left(\left(t^c_{i_{end}} + \delta_i\right) < t_{onset}\right)\right\}}{\#sMCI}$$

Regarding pMCI subjects, two measures were established: a) the proportion of converters who at baseline did not have dementia and $\left(t^c_{i_1} + \delta_i\right) < t_{onset}$, compared to the total number of pMCI subjects:

$$SCORE_2 = \frac{\#\left\{i|(i \in pMCI) \cap \left(\left(t^c_{i_1} + \delta_i\right) < t_{onset}\right)\right\}}{\#pMCI}$$

and b) the ratio of pMCI subjects whose last visit was greater than t_{onset} with respect to the total number of pMCI subjects:

$$SCORE_3 = \frac{\#\left\{i|(i \in pMCI) \cap \left(\left(t^c_{i_{end}} + \delta_i\right) > t_{onset}\right)\right\}}{\#pMCI}$$

In addition, the classification of subjects with dementia at baseline should also be considered. On their first visit, these patients should be on a time greater than t_{onset},

$$SCORE_4 = \frac{\#\left\{i|(i \in Dementia) \cap \left(\left(t^c_{i_{end}} + \delta_i\right) < t_{onset}\right)\right\}}{\#Dementia}$$

Therefore, a zero time was estimated by maximizing the scores above. As t_{onset} is moved to higher values, $SCORE_1$ and $SCORE_2$ increase, while $SCORE_3$ and $SCORE_4$ decrease. An optimal time will be the one that achieves the maximum score of all the $SCORE_r$ and that these values are similar. It is proposed to obtain the optimum by minimizing the differences in absolute value of the $SCORE_r$:

$$\hat{t}_{onset} = \arg\min_{t_{onset}} \sum_{r=1}^{3} \sum_{s=r+1}^{4} |SCORE_r(t_{onset}) - SCORE_s(t_{onset})|$$

2.6 Criteria for validation of the proposed DPM

For each vector used to train GRACE, a timeline of the disease is proposed. To select the most likely natural history, three criteria are proposed to measure the reliability of the suggested DPM: i) Consider the four scores from the previous subsection, $SCORE_r$. ii) Perform a linear regression analysis with the pMCI subjects between their conversion times and the times estimated by the proposed DPM, $t_{onset} - \left(t^c_{i_1} + \delta_i\right)$, for all $i \in pMCI$. iii) Compare the values of the long-term trajectories of the markers at t_{onset} with their cutoffs used in clinical practice to diagnose patients with dementia. Therefore, among the vectors that showed good performance in predicting a high risk of progressing into dementia at baseline, these three criteria were used to select the optimal combination of markers to train GRACE.

3 RESULTS

3.1 AT profiles in MCI/Dementia subjects

At baseline, 64% of the sMCI patients had an A-T- profile and 21% with A+T+. However, the pMCI and dementia groups had a tendency of 70% with an A+T+ profile and around 15% A-T- at baseline. The percentages of the AT profiles in the longitudinal analysis were similar with respect to the values at baseline. In the longitudinal study, the maximum value of $pTau/A\beta_{42}$ among the subject's visits was selected. For that particular visit, the value of pTau was also chosen. Considering all clinical groups, over time the group A-T- decreased slightly while the percentages of A+T- and A+T+ increased somewhat. This result indicated that abnormal amyloid plaques or fibrillary tau (A + or T +) were or were not present, in general, in the patient at baseline, so their temporal evolutions did not provide more information for the AT binary classification. In the case of dementia, the longitudinal percentages of AT profiles were those at baseline. Only 15 subjects of the 128 patients with dementia had a second CSF evaluation in month 24. Note that these patients were only followed up for 2 years in the ADNI study (see supplementary material, S2).

3.2 MCI to dementia conversion prediction

For the progression from the MCI stage to dementia, the proposed methodology between the categorical predictive models and the DPMs selected the following vector (see supplementary material, S4):

$$\{ADAS13, FAQ, MMSE, CDRSB\}$$

At baseline, the predictive model showed good classification between sMCI versus pMCI patients (sensitivity or correct identification rate of positive cases or pMCI, SEN = 72.2% (70.6%-73.9%), specificity as a measure of correct detection of sMCI, SPE = 79.1% (77.9%-80.3%), accuracy of success of positive and negative cases, ACC = 76.6% (75.6%-77.5%), and area under the curve that measures the discriminative capacity of the classifier, AUC = 0.828 (0.816-0.838)). The classification scores were higher with the fast converters (pMCI-fast), *SEN* = 80.8% (79.9%−81.7%), and in those sMCI subjects who were followed for a longer time (sMCI-slow), *SPE* = 86.2% (85.4%−88.4%). On the contrary, the slow converters, *SEN* = 43.8% (40.9%−45.2%), or sMCI censored in less than four years, *SPE* = 71.5% (69.7%−73.5%), had worse scores. Clearly, it is easier to detect conversion into dementia, at baseline, the subjects who convert before the four years than those who will do so in a longer time. Likewise, those sMCI patients with censoring times greater than four years are easier to classify as non-converters than those sMCI with shorter time, since the latter could convert shortly after their censoring times.

Table 1 provides, for each clinical group, the percentages of the population with positivity in the three most important risk factors for conversion to dementia at baseline [13]: late MCI (LMCI) diagnosis, amyloid pathology (A +) and APOE4 status. The risk factor that most discriminates was amyloid pathology at baseline. The pMCI-fast population had the highest risk to conversion with the worst cognitive diagnosis LMCI (76.9 − 82.4%), the highest amyloid positivity (84.3%) and APOE4 carrier rate between 69.4% or 69.8%. In the opposite direction were the sMCI-slow subjects. The proposed predictive model better discriminated between sMCI vs pMCI in relation to these three risk factors. Using baseline data, the algorithm estimated that around 80% of pMCI-fast subjects progress to dementia and only 14% of sMCI-slow subjects were misclassified by this approach.

[Table 1 about here.]

3.3 DPM with GRACE

With the studied population, the proposed vector was used to train the GRACE algorithm. Fig 1 shows both the number of the subjects as a function of the time shift, δ_i, as well as the relationship between the conversion times of the pMCI subjects, $t_{convert}$, and their estimates by GRACE, $\hat{t}_{convert}$. For the conversion times of the pMCI patients, the regression line was an unit slope and intercept is near zero, i.e. the estimated times and the AD timeline were very similar ($\hat{t}_{convert} = 1.08 \cdot t_{convert} - 0.16$) and with a Pearson correlation coefficient of R = 0.724(0.718−0.731). By optimizing

the classification scores of the subjects based on $\left(t_{ij}^{c}+\delta_i\right)$ and the clinical groups, the time to onset of dementia was estimated, $\hat{t}_{onset} \approx 3.4 years$.

[Figure 1 about here.]

Fig. 2 shows the individual observed trajectories for each marker, in the original scale, ordered in the progression timeline using the DPM approach. The long-term trajectories are also highlighted for each marker. Note that the GRACE approach was blind to the information of the clinical group or AT profiles of the subjects. Even so, the short-term trajectories were painted with different colors depending on the clinical group or AT profiles of the subjects. This visualization shows that the ordering proposed by estimating δ_i was consistent both in the clinical classification and in the AT profiles. For the estimated time $\hat{t}_{onset} \approx 3.4 years$, $SCORE_1$ = 95.3%(95.1% − 95.6%) for the sMCI group $\left(\left(t_{i_{end}}^{c}+\delta_i\right) < t_{onset}\right)$, $SCORE_2$ = 80.6%(79.7% − 81.5%) for the pMCI subjects at baseline $\left(\left(t_{i_1}^{c}+\delta_i\right) < t_{onset}\right)$, $SCORE_3$ = 74.5%(73.8% − 75.2%) at the end of the visits $\left(\left(t_{i_{end}}^{c}+\delta_i\right) > t_{onset}\right)$ of the pMCI patients were correctly classified and $SCORE_4$ = 81.5%(80.6%−82.4%) of the dementia subjects $\left(\left(t_{i_1}^{c}+\delta_i\right) > t_{onset}\right)$were also correctly classified. Most pMCI and dementia subjects had an AT type A+T+ pattern. In contrast, sMCI subjects typically displayed amyloid negativity (A-). On the other hand, the short-term trajectories tend from A- towards A+T- and then A+T+ as a function of the time of progression towards dementia.

[Figure 2 about here.]

Once the progression timeline from MCI to dementia was defined from the δ_i values of the subjects, it was possible to infer the long-term trajectories of any marker acquired from the studied population (see supplementary material, S5). From the long-term trajectories of $A\beta_{42}$, pTau and FDG, together with the cut-off values to define the positivities of the AT(N) profiles, the following times were obtained in the progression of dementia: a) $t_{A\beta_{42}\,=\,880} = -3.7 years$ or $t_{pTau/A\beta 42\,=\,0.028} = -3.9 years$, b) $t_{pTau\,=\,27\,pg/ml} = -0.8 years$ and c) $t_{FDG=1.21} = -0.9 years$. These times suggested that the amyloid pathology (A+) appeared first and that later both the tau pathology (T+) and the neurodegenerative process (N+) occurred at a similar time. This suggestion is also supported by the high linear correlation between Tau and pTau ($R > 0.9$ for any of the three clinical groups analyzed, see supplementary material, S2), as a diagnosis with T+ ($pTau > 27 pg/ml$) would also imply N+ ($Tau > 290 pg/ml$). On the other hand, both FDG PET and Tau in CSF reflect neurodegeneration. Tau in CSF likely indicates the intensity of neuronal injury at a given time point, while hypometabolism on FDG likely indicates both cumulative loss of neuropil and functional impairment of neurons [2]. There is no correlation between Tau and FDG ($R = -0.03 to -0.10$, see supplementary material, S2). However, there is a reasonable agreement of positivity (N +) and negativity (N-) on neurodegenerative processes between Tau and FDG. At baseline, the highest agreement of positivity and negativity between Tau and FDG was in dementia, around 70%. While the positive agreement at sMCI falls below 30%, although it has a negative agreement of 76%. The pMCI subjects were in intermediate values of agreement (65% in positivity and 40% in negativity).

3.4 Validation of long-term trajectories

Table 2 shows a comparison among the mean values of the NMs, at baseline, of the sMCI, pMCI and dementia populations, with respect to the values estimated by the long-term trajectories at the onset of the dementia. The calculated NM values at $\hat{t}_{onset} \approx 3.4 years$ were similar to the cut-off points used in clinical practice for determining the progression to dementia. Furthermore, these NM scores were between the mean values of the pMCI and dementia population at baseline.

On the other hand, the values of the long-term trajectories of AT(N) markers at dementia onset were closer to the mean scores of subjects with pMCI and dementia than to clinical practice cut-off points, indicating that the majority of patients at $\hat{t}_{onset}$ had both amyloid and tau pathology. This result was consistent with the fact that the pMCI and dementia populations had around 70% of the A+T+ profile at baseline (see supplementary material, S2).

[Table 2 about here.]

Fig. 3 shows the application of bootstrap techniques on the long trajectories of the NMs between the subjects with A+ and A-. It was observed how cognitive decline accelerates among A+ subjects compared to A-. Furthermore, the trajectories of subjects A+ showed less dispersion. These subjects correspond to Alzheimer's continuum.

Using the cut-off points from MCI to dementia in the scores of the neuropsychological measures, it was observed that A+ subjects converted around year 0 while subjects A- progressed to dementia around year 10, i.e. there was a difference of approximately a decade of conversion between the subjects who had amyloid pathology and those who did not present it at baseline.

[Figure 3 about here.]

4 DISCUSSIONS

4.1 Conversion rates from MCI to dementia

MCI subjects are at high risk of progressing to dementia [42]. Around 39% of those diagnosed with MCI in specialist settings and 22% in population studies, develop dementia over the subsequent three years, compared to 3% of the non-MCI population of the same age [43]. In our study using the ADNI database, 30% of the MCI population converted in the first four years, which represented 80% of all pMCI patients (see supplementary material, S3). This is in line with previous observations from ADNI [12, 14] and other cohorts with similar ages [44]. Studies have shown that 32% of MCI patients over 65 years of age developed AD after 5 years of follow up [45]. However, the conversion ratios depend on the criteria for patient selection from these cohorts [46].

4.2 Approach

Clinical practice requires not only the detection of subjects at high risk of progression to AD, but also which pathology an MCI individual is likely to develop [19]. With both objectives in mind, two types of information were applied: a) the AT(N) profile and b) the cognitive evolution towards dementia. While other authors mixed markers from different modalities, here, CSF biomarkers and FDG PET measurements were exclusively used for AT(N) profiles and evolution of dementia syndrome only used cognitive tests. The AT(N) markers were binarized through cut off points for abnormal values, both at baseline and longitudinally. Moreover, the AT(N) profiles allowed us to consider whether the patient evolves into the Alzheimer's continuum or towards a dementia not suspected of being Alzheimer's. For the severity of cognitive impairment, a group of NMs were used for the cognitive and functional progression via predictive models and DPMs.

4.2.1 AT profiles

The percentages of the MCI/dementia subjects using the AT profiles (A-T-, A-T+, A+T-, A+T+) remained similar between baseline and longitudinal analyses. A+ subjects are at higher risk of developing dementia [47]. As described previously [10, 13], more than 45% of the MCI population with A+ at baseline converted before the first four years. Furthermore, 80% of the fast converters had amyloid pathology (A+) at baseline. Previous studies have shown that 40– 60% of MCI patients were A+ on PET, with evidence of 40% to 80% risk of conversion to AD dementia within 3 years, a level that was 4 to 9 fold higher than their A- counterparts [47, 48, 49]. The MCI subjects that did not progress to dementia over time correspond mainly to A- subjects, especially true in the sMCI-slow subjects, i.e. with censoring times greater than 48 months.

The fast converters with A+T+ profile represented 70% and, considering the cut-off $FDG < 1.21$ for N+ at baseline, decreased to 53% with an A+T+(N+) profile. In dementia, 68% of the subjects had an A+T+ profile and 63% had A+T+(N+), consistent with the fact that between 60-80% of dementia cases are AD [50]. By contrast, 22% of sMCI slow subjects, were A+, 14% A+T+ and 3% A+T+(N+). A greater agreement was observed between CSF Tau and FDG for (N) in dementia (see supplementary material, S2). Previous work has shown that the combination of abnormal CSF pTau, FDG PET and abnormal amyloid biomarker provides more powerful prediction of future dementia than an abnormal amyloid study alone [2]. FDG PET can be useful to differentiate between AD and other types of dementia and to measure disease progression [51]. FDG-PET is not recommended for diagnosing patients with early phase of AD, as there is no way to ascertain whether the hypometabolism is directly related to AD pathology [52]; however, clinicians may refer patients with more established symptomatology for an FDG-PET scan to identify regions of glucose hypometabolism and neurodegeneration that could be indicative of AD [53].

4.2.2 Risk factors

Advanced MCI stage (LMCI), abnormal levels of amyloid, and the presence of APOE4 carrier were associated with fast clinical decline (see Table 1). Previous studies demonstrated that EMCI patients take longer to progress to dementia compared with LMCI [54] and LMCI individuals have a 15-fold higher risk of being fast converters [13]. Multiple

observational studies have shown that abnormal levels of $A\beta$ were associated with cognitive decline and functional progression [55, 56]. The bootstrap analysis on the long trajectories of the NMs showed greater cognitive impairment among subjects with amyloid pathology (see Fig. 3). A+ was associated with poorer performance in the cognitive domains of global cognitive function, memory, language, visuospatial ability, processing speed, and attention/working memory/executive functions when compared to A- subjects [57]. The presence of an APOE4 allele was significantly more frequent in the pMCI-fast group among MCI patients. A strong association between APOE4 genotype and amyloid accumulation in the brain is well described in both in vivo [58] and neuropathology studies [59].

4.2.3 Predictive models

Predictive models of progression to dementia that combine measures of multiple domains simultaneously have been proposed previously. A combination of supervised and unsupervised machine learning approach was employed to avoid circular contamination of data with the clinical diagnostic information. The proposed methodology was applied to the ADNI cohort, specifically for individuals diagnosed with MCI at their baseline assessments, and whether they converted to dementia during the follow-up was determined. A vector of cognitive measures was proposed for estimating the conversion probabilities from the MCI stage to dementia and to describe the natural history of the temporal progression of MCI to dementia:

$$\{ADAS13, FAQ, MMSE, CDRSB\}.$$

These markers are easily interpretable, generating a robust, verifiable and reliable predictive model. The results reported here indicate that the proposed approach significantly captures the underlying disease burden, which could be transferable to other cohorts, and can be applied across wide age ranges. At baseline, the proposed battery of cognitive measures better discriminated, between sMCI and pMCI subjects with respect to other suggested risk factors (LMCI, A +, APOE4), particularly between the pMCI fast and sMCI slow subjects. The predictive performance was in line with the classification scores of previously presented algorithms [19]. In addition, it was observed that adding other marker modalities to cognitive measures did not produce significant improvements in the performance of the predictive models. The cost of collecting cognitive variables compared to performing a MRI or a amyloid PET is quite low. Therefore, the non-significant improvement in performance might not be worth the cost, logistics and patient inconvenience arising from the collection of other modalities. Methods that focus solely on cognition, as proposed here, should be further explored [19]. In this direction, AD treatment effectiveness may be inferred by examining impairment measured solely by longitudinal NMs [60]. Recent FDA preliminary guidance for industry similarly raises this possibility, suggesting that it will "consider strongly justified arguments that a persuasive effect on sensitive measures of neuropsychological performance may provide adequate support for a marketing approval" [61].

4.2.4 DPM

The proposed natural history of the progression from MCI to dementia was verified at three different levels: a) The high linear correlation (R = 0.724(0.718−0.731)) between the conversion times of the pMCI subjects and their values estimated by GRACE. b) The good performance on clinical classification (sMCI, pMCI and dementia) of subjects with the proposed temporal ordering (see fig. 2, $\hat{t}_{onset} \approx 3.4 years$, $SCORE_1$ = 95.3%(95.1%−95.6%), $SCORE_2$ = 80.6%(79.7% − 81.5%), $SCORE_3$ = 74.5%(73.8% − 75.2%) and $SCORE_4$ = 81.5%(80.6% − 82.4%)) and c) The similarities of the cut-off points of the cognitive and functional measures, used in clinical practice to diagnose dementia, and the values defined in these NMs at the time of onset of dementia given by the proposed natural history of the dementia (see table 2).

The long-term trajectories of AT(N) markers, given by the proposed temporal ordering of the studied MCI/dementia population, indicated that amyloid pathology (A+) was present around 7 years before the onset of dementia (($t_{onset} - t_{A\beta_{42}} = 880$) or ($t_{onset} - t_{pTau/A\beta 42} = 0.028$)). Furthermore, the time to reach T+ ($t_{onset} - t_{pTau} = 27$) and N+ ($t_{onset} - t_{FDG=1.21}$) were obtained, simultaneously, about 3 or 4 years before the onset of dementia. The cut-off values of N+ given by FDG and CSF-Tau markers were also observed that occur practically at the same time of the proposed natural history. This proposed temporal ordering is in line with previous observations. Positive β-amyloid detectable by both CSF and amyloid PET has been shown to precede substantial tau deposition [62]. Levels of $A\beta_{42}$ in CSF are already decreased at least 5 to 10 years before conversion to AD dementia, whereas Tau and pTau became progressively more abnormal as the time to diagnosis

of dementia [63, 64, 65]. These conclusions represent sample rather than individual effects. Furthermore, these estimated times were consistent with the observation that around 70% of pMCI-fast and dementia subjects show an A+T+ profile at baseline.

In order to detect MCI subjects at high risk of conversion into dementia, the long-term trajectories of the AT(N) markers and NMs on a common scale (see supplementary material, S5) indicated that the progressions of AT(N) markers were slower, and occur prior to, cognitive and functional decline measures. This supports the dual use of the AT(N) profile and the predictive models of the clinical symptom progression. The AT(N) profile, at baseline, allows the discrimination between the various stages of the Alzheimer's continuum and other dementias. On the other hand, the predictive models, presented here, estimated the time to convert to dementia. Furthermore, it was suggested that estimating the progression to dementia only requires the exclusive use of neuropsychological measures.

4.3 Limitations

This study has several limitations. Participants were selected from a single database. Moreover, ADNI patients are mostly well-educated and Caucasian. Further validation studies are needed with larger and more generalizable populations. However, the ADNI cohort has been used in 84.6% of articles studying the progression of subjects with MCI [19].

On the other hand, although the importance of amyloid pathology was highlighted by the current results, further suggested risk factors, such as sex and its interaction with the APOE4 genotype [66], require further interrogation.

Although GRACE has demonstrated superiority in estimating disease progressions compared to alternative approaches [33], new DPM training algorithms have emerged in recent years [67, 68, 69], including proposals based on recurrent neural networks [70]. Thus, a comparison between GRACE and these new approaches remains pending.

5 CONCLUSIONS

A combination of data from AT(N) profiles at baseline and longitudinal predictive modeling is proposed. The AT(N) profiles of the MCI subjects at baseline were not substantially modified longitudinally, especially true of the pMCI subjects. This is consistent with amyloid pathology being present years before the onset of dementia.

With a selected vector of cognitive and functional measures, high performance is achieved for detecting subjects at high risk of converting to dementia in the first four years and a natural history of the progression of MCI into dementia is also proposed. The proposed temporal evolution towards dementia was validated through three different metrics, highlighting the high linear correlation between the conversion times of the pMCI subjects and the estimated times with the proposed approach (R=0.72).

By combing the AT(N) profile with predictive modeling, the risk of converting from MCI to dementia was better estimated, as well as whether patients progressed within the Alzheimer's continuum or into a suspected non-Alzheimer's dementia. Thus, the MCI fast converters had an A+T+ profile of almost 70% at baseline, which coincided with the percentage of subjects with dementia at baseline and A+T+ profile. These percentages are in the line with evidence that between 60-80% of dementia have AD [50]. In addition, the amyloid pathology (A+) was estimated in 7 years before dementia and that tau pathology (T+) and neurodegeneration (N+) were between 3-4 years before dementia.

Finally, MCI subjects with amyloid pathology (A+) compared to those without (A-) demonstrated a difference of one decade in reaching the onset of dementia. In addition, the longitudinal trajectories of the cognitive and functional markers showed a faster and less dispersed deterioration among the A+ subjects compared to the A-.

6 ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson& Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

7 FUNDING

This work was supported by the Comunidad de Madrid, grant number MINA CM P2022 BMD 7236.

8 DECLARATIONS OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9 SUPPLEMENTARY MATERIAL

Below is the link to the electronic supplementary material.

References

[1] Hyman, B.T., Phelps, C.H., Beach, T.G., Bigio, E.H., Cairns, N.J., Carrillo, M.C., Dickson, D.W., Duyckaerts, C., Frosch, M.P., Masliah, E., et al.: National institute on aging–alzheimer's association guidelines for the neuropathologic assessment of Alzheimer's disease. Alzheimer's & dementia **8** (2012) 1–13

[2] Jack Jr, C.R., Bennett, D.A., Blennow, K., Carrillo, M.C., Dunn, B., Haeberlein, S.B., Holtzman, D.M., Jagust, W., Jessen, F., Karlawish, J., et al.: NIA-AA research framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia **14** (2018) 535–562

[3] Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R.C., Ritchie, K., Broich, K., Belleville, S., Brodaty, H., Bennett, D., Chertkow, H., et al.: Mild cognitive impairment. The lancet **367** (2006) 1262–1270

[4] Petersen, R.C., Roberts, R.O., Knopman, D.S., Boeve, B.F., Geda, Y.E., Ivnik, R.J., Smith, G.E., Jack, C.R.: Mild cognitive impairment: ten years later. Archives of neurology **66** (2009) 1447–1455

[5] Manly, J.J., Tang, M.X., Schupf, N., Stern, Y., Vonsattel, J.P.G., Mayeux, R.: Frequency and course of mild cognitive impairment in a multiethnic community. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society **63** (2008) 494–506

[6] Sperling, R.A., Rentz, D.M., Johnson, K.A., Karlawish, J., Donohue, M., Salmon, D.P., Aisen, P.: The A4 study: stopping AD before symptoms begin? Science translational medicine **6** (2014) 228fs13–228fs13

[7] Falahati, F., Westman, E., Simmons, A.: Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. Journal of Alzheimer's Disease **41** (2014) 685–708

[8] Vermunt, L., Sikkes, S.A., Van Den Hout, A., Handels, R., Bos, I., Van Der Flier, W.M., Kern, S., Ousset, P.J., Maruff, P., Skoog, I., et al.: Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. Alzheimer's & Dementia **15** (2019) 888–898

[9] van Maurik, I.S., Vos, S.J., Bos, I., Bouwman, F.H., Teunissen, C.E., Scheltens, P., Barkhof, F., Frolich, L., Kornhuber, J., Wiltfang, J., et al.: Biomarker-based prognosis for people with mild cognitive impairment (ABIDE): a modelling study. The Lancet Neurology **18** (2019) 1034–1044

[10] Blazhenets, G., Frings, L., Ma, Y., S¨orensen, A., Eidelberg, D., Wiltfang, J., Meyer, P.T., Initiative, A.D.N., et al.: Validation of the Alzheimer disease dementia conversion-related pattern as an ATN biomarker of neurodegeneration. Neurology **96** (2021) e1358–e1368

[11] Korolev, I.O., Symonds, L.L., Bozoki, A.C., A.D.N.I., et al.: Predicting progression from mild cognitive impairment to Alzheimer's dementia using clinical, MRI, and plasma biomarkers via probabilistic pattern classification. PloS one **11** (2016) e0138866

[12] Steenland, K., Zhao, L., John, S.E., Goldstein, F.C., Levey, A., Alvaro, A., Initiative, A.D.N., et al.: A 'framingham-like' algorithm for predicting 4-year risk of progression to amnestic mild cognitive impairment or Alzheimer's disease using multidomain information. Journal of Alzheimer's Disease **63** (2018) 1383–1393

[13] Jang, H., Park, J., Woo, S., Kim, S., Kim, H.J., Na, D.L., Lockhart, S.N., Kim, Y., Kim, K.W., Cho, S.H., et al.: Prediction of fast decline in amyloid positive mild cognitive impairment patients using multimodal biomarkers. NeuroImage: Clinical **24** (2019) 101941

[14] Cullen, N.C., Leuzy, A., Palmqvist, S., Janelidze, S., Stomrud, E., Pesini, P., Sarasa, L., Allué, J.A., Proctor, N.K., Zetterberg, H., et al.: Individualized prognosis of cognitive decline and dementia in mild cognitive impairment based on plasma biomarker combinations. Nature Aging **1** (2021) 114–123

[15] Palmqvist, S., Tideman, P., Cullen, N., Zetterberg, H., Blennow, K., Dage, J.L., Stomrud, E., Janelidze, S., Mattsson-Carlgren, N., Hansson, O.: Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. Nature Medicine **27** (2021) 1034–1042

[16] Zheng, C., Xia, Y., Pan, Y., Chen, J.: Automated identification of dementia using medical imaging: a survey from a pattern classification perspective. Brain informatics **3** (2016) 17–27

[17] Rathore, S., Habes, M., Iftikhar, M.A., Shacklett, A., Davatzikos, C.: A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. NeuroImage **155** (2017) 530–548

[18] Grueso, S., Viejo-Sobera, R.: Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimer's research & therapy **13** (2021) 1–29

[19] Ansart, M., Epelbaum, S., Bassignana, G., Bône, A., Bottani, S., Cattai, T., Couronné, R., Faouzi, J., Koval, I., Louis, M., et al.: Predicting the progression of mild cognitive impairment using machine learning: a systematic, quantitative and critical review. Medical Image Analysis **67** (2021) 101848

[20] Lee, E., Giovanello, K.S., Saykin, A.J., Xie, F., Kong, D., Wang, Y., Yang, L., Ibrahim, J.G., Doraiswamy, P.M., Zhu, H., et al.: Single-nucleotide polymorphisms are associated with cognitive decline at Alzheimer's disease conversion within mild cognitive impairment patients. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring **8** (2017) 86–95

[21] Bi, X.a., Xing, Z., Zhou, W., Li, L., Xu, L.: Pathogeny detection for mild cognitive impairment via weighted evolutionary random forest with brain imaging and genetic data. IEEE Journal of Biomedical and Health Informatics **26** (2022) 3068–3079

[22] Donohue, M.C., Jacqmin-Gadda, H., Le Goff, M., Thomas, R.G., Raman, R., Gamst, A.C., Beckett, L.A., Jack Jr, C.R., Weiner, M.W., Dartigues, J.F., et al.: Estimating long-term multivariate progression from short-term data. Alzheimer's & Dementia **10** (2014) S400–S410

[23] Guerrero, R., Schmidt-Richberg, A., Ledig, C., Tong, T., Wolz, R., Rueckert, D., (ADNI, A.D.N.I., et al.: Instantiated mixed effects modeling of Alzheimer's disease markers. NeuroImage **142** (2016) 113–125

[24] Schmidt-Richberg, A., Ledig, C., Guerrero, R., Molina-Abril, H., Frangi, A., Rueckert, D., Initiative, A.D.N., et al.: Learning biomarker models for progression estimation of Alzheimer's disease. PloS one **11** (2016)

[25] Li, D., Iddi, S., Thompson, W.K., Donohue, M.C., Initiative, A.D.N.: Bayesian latent time joint mixed effect models for multicohort longitudinal data. Statistical methods in medical research **28** (2019) 835–845

[26] Lorenzi, M., Filippone, M., Frisoni, G.B., Alexander, D.C., Ourselin, S., Initiative, A.D.N., et al.: Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. NeuroImage **190** (2019) 56–68

[27] Wyman, B.T., Harvey, D.J., Crawford, K., Bernstein, M.A., Carmichael, O., Cole, P.E., Crane, P.K., DeCarli, C., Fox, N.C., Gunter, J.L., et al.: Standardization of analysis sets for reporting results from ADNI MRI data. Alzheimer's & Dementia **9** (2013) 332–337

[28] The ADNI team: ADNIMERGE: Alzheimer's Disease Neuroimaging Initiative. (2021) R package version 0.0.1.

[29] Petersen, R.C., Aisen, P., Beckett, L.A., Donohue, M., Gamst, A., Harvey, D.J., Jack, C., Jagust, W., Shaw, L., Toga, A., et al.: Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. Neurology **74** (2010) 201–209

[30] Bittner, T., Zetterberg, H., Teunissen, C.E., Ostlund Jr, R.E., Militello, M., Andreasson, U., Hubeek, I., Gibson, D., Chu, D.C., Eichenlaub, U., et al.: Technical performance of a novel, fully automated electrochemiluminescence immunoassay for the quantitation of β-amyloid (1–42) in human cerebrospinal fluid. Alzheimer's & dementia **12** (2016) 517–526

[31] Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J.Q., Bittner, T., Lifke, V., Corradini, V., Eichenlaub, U., Batrla, R., et al.: CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: a study of fully automated immunoassays in BioFINDER and ADNI cohorts. Alzheimer's & Dementia **14** (2018) 1470–1481

[32] Landau, S.M., Mintun, M.A., Joshi, A.D., Koeppe, R.A., Petersen, R.C., Aisen, P.S., Weiner, M.W., Jagust, W.J., Initiative, A.D.N.: Amyloid deposition, hypometabolism, and longitudinal cognitive decline. Annals of neurology **72** (2012) 578–586

[33] Platero, C.: Categorical predictive and disease progression modeling in the early stage of Alzheimer's disease. Journal of Neuroscience Methods **374** (2022) 109581

[34] Kleinbaum, D.G., Klein, M.: Survival analysis. Volume 3. Springer (2010)

[35] Sabuncu, M.R., Bernal-Rusiel, J.L., Reuter, M., Greve, D.N., Fischl, B., A.D.N.I., et al.: Event time analysis of longitudinal neuroimage data. NeuroImage **97** (2014) 9–18

[36] Platero, C., Tobar, M.C.: Longitudinal survival analysis and two-group comparison for predicting the progression of mild cognitive impairment to Alzheimer's disease. Journal of Neuroscience Methods **341** (2020) 108698

[37] Bernal-Rusiel, J.L., Greve, D.N., Reuter, M., Fischl, B., Sabuncu, M.R., A.D.N.I., et al.: Statistical analysis of longitudinal neuroimage data with linear mixed effects models. Neuroimage **66** (2013) 249–260

[38] Bernal-Rusiel, J.L., Reuter, M., Greve, D.N., Fischl, B., Sabuncu, M.R., A.D.N.I., et al.: Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. Neuroimage **81** (2013) 358–370

[39] Kuhn, M., Johnson, K.: Applied predictive modeling. Volume 26. Springer (2013)

[40] Peng, H., Long, F., Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence **27** (2005) 1226–1238

[41] Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M.O., Chupin, M., Benali, H., Colliot, O., Initiative, A.D.N., et al.: Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. neuroimage **56** (2011) 766–781

[42] Vos, S.J., Verhey, F., Frölich, L., Kornhuber, J., Wiltfang, J., Maier, W., Peters, O., Rüther, E., Nobili, F., Morbelli, S., et al.: Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. Brain **138** (2015) 1327–1338

[43] Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S.G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., et al.: Dementia prevention, intervention, and care. The Lancet **390** (2017) 2673–2734

[44] Roberts, R.O., Knopman, D.S., Mielke, M.M., Cha, R.H., Pankratz, V.S., Christianson, T.J., Geda, Y.E., Boeve, B.F., Ivnik, R.J., Tangalos, E.G., et al.: Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. Neurology **82** (2014) 317–325

[45] Ward, A., Tardiff, S., Dye, C., Arrighi, H.M.: Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature. Dementia and geriatric cognitive disorders extra **3** (2013) 320–332

[46] Chen, Y., Denny, K.G., Harvey, D., Farias, S.T., Mungas, D., DeCarli, C., Beckett, L.: Progression from normal cognition to mild cognitive impairment in a diverse clinic-based and community-based elderly cohort. Alzheimer's & Dementia **13** (2017) 399–405

[47] Doraiswamy, P.M., Sperling, R., Johnson, K., Reiman, E.M., Wong, T., Sabbagh, M., Sadowsky, C.H., Fleisher, A., Carpenter, A., Joshi, A., et al.: Florbetapir F 18 amyloid PET and 36-month cognitive decline: a prospective multicenter study. Molecular psychiatry **19** (2014) 1044–1051

[48] Okello, A., Koivunen, J., Edison, P., Archer, H., Turkheimer, F., Nagren, K.u., Bullock, R., Walker, Z., Kennedy, A., Fox, N., et al.: Conversion of amyloid positive and negative MCI to AD over 3 years: an 11C-PIB PET study. Neurology **73** (2009) 754–760

[49] van Rossum, I.A., Vos, S.J., Burns, L., Knol, D.L., Scheltens, P., Soininen, H., Wahlund, L.O., Hampel, H., Tsolaki, M., Minthon, L., et al.: Injury markers predict time to dementia in subjects with mci and amyloid pathology. Neurology **79** (2012) 1809–1816

[50] Sosa-Ortiz, A.L., Acosta-Castillo, I., Prince, M.J.: Epidemiology of dementias and Alzheimer's disease. Archives of medical research **43** (2012) 600–608

[51] Dubois, B., Feldman, H.H., Jacova, C., Hampel, H., Molinuevo, J.L., Blennow, K., DeKosky, S.T., Gauthier, S., Selkoe, D., Bateman, R., et al.: Advancing research diagnostic criteria for Alzheimer's disease: the iwg-2 criteria. The Lancet Neurology **13** (2014) 614–629

[52] Dubois, B., Hampel, H., Feldman, H.H., Scheltens, P., Aisen, P., Andrieu, S., Bakardjian, H., Benali, H., Bertram, L., Blennow, K., et al.: Preclinical Alzheimer's disease: definition, natural history, and diagnostic criteria. Alzheimer's & Dementia **12** (2016) 292–323

[53] Frisoni, G.B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., D´emonet, J.F., Garibotto, V., Giannakopoulos, P., Gietl, A., et al.: Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. The Lancet Neurology **16** (2017) 661–676

[54] Jessen, F., Wolfsgruber, S., Wiese, B., Bickel, H., Mösch, E., Kaduszkiewicz, H., Pentzek, M., Riedel-Heller, S.G., Luck, T., Fuchs, A., et al.: AD dementia risk in late MCI, in early MCI, and in subjective memory impairment. Alzheimer's & Dementia **10** (2014) 76–83

[55] Mormino, E.C., Betensky, R.A., Hedden, T., Schultz, A.P., Amariglio, R.E., Rentz, D.M., Johnson, K.A., Sperling, R.A.: Synergistic effect of β-amyloid and neurodegeneration on cognitive decline in clinically normal individuals. JAMA neurology **71** (2014) 1379–1385

[56] Lim, Y.Y., Maruff, P., Pietrzak, R.H., Ames, D., Ellis, K.A., Harrington, K., Lautenschlager, N.T., Szoeke, C., Martins, R.N., Masters, C.L., et al.: Effect of amyloid on memory and non-memory decline from preclinical to clinical Alzheimer's disease. Brain **137** (2014) 221–231

[57] Han, S.D., Nguyen, C.P., Stricker, N.H., Nation, D.A.: Detectable neuropsychological differences in early preclinical Alzheimer's disease: A meta-analysis. Neuropsychology review **27** (2017) 305–325

[58] Morris, J.C., Roe, C.M., Xiong, C., Fagan, A.M., Goate, A.M., Holtzman, D.M., Mintun, M.A.: APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. Annals of neurology **67** (2010) 122–131

[59] Gomez-Isla, T., West, H.L., Rebeck, G.W., Harr, S.D., Growdon, J.H., Locascio, J.J., Perls, T.T., Lipsitz, L.A., Hyman, B.T.: Clinical and pathological correlates of apolipoprotein e ε4 in Alzheimer's disease. Annals of neurology **39** (1996) 62–70

[60] Papp, K.V., Buckley, R., Mormino, E., Maruff, P., Villemagne, V.L., Masters, C.L., Johnson, K.A., Rentz, D.M., Sperling, R.A., Amariglio, R.E., et al.: Clinical meaningfulness of subtle cognitive decline on longitudinal testing in preclinical AD. Alzheimer's & Dementia (2019)

[61] Sabbagh, M.N., Hendrix, S., Harrison, J.E.: FDA position statement "Early Alzheimer's disease: Developing drugs for treatment, Guidance for Industry". Alzheimer's & Dementia: Translational Research & Clinical Interventions **5** (2019) 13–19

[62] Reimand, J., Collij, L., Scheltens, P., Bouwman, F., Ossenkoppele, R., Initiative, A.D.N., et al.: Association of amyloid-β CSF/PET discordance and tau load 5 years later. Neurology **95** (2020) e2648–e2657

[63] Fagan, A.M., Mintun, M.A., Shah, A.R., Aldea, P., Roe, C.M., Mach, R.H., Marcus, D., Morris, J.C., Holtzman, D.M.: Cerebrospinal fluid tau and ptau181 increase with cortical amyloid deposition in cognitively normal individuals: implications for future clinical trials of Alzheimer's disease. EMBO molecular medicine **1** (2009) 371–380

[64] Buchhave, P., Minthon, L., Zetterberg, H., Wallin, ˚A.K., Blennow, K., Hansson, O.: Cerebrospinal fluid levels of β-amyloid 1-42, but not of tau, are fully changed already 5 to 10 years before the onset of Alzheimer dementia. Archives of general psychiatry **69** (2012) 98–106

[65] Tan, M.S., Ji, X., Li, J.Q., Xu, W., Wang, H.F., Tan, C.C., Dong, Q., Zuo, C.T., Tan, L., Suckling, J., et al.: Longitudinal trajectories of Alzheimer's ATN biomarkers in elderly persons without dementia. Alzheimer's research & therapy **12** (2020) 1–7

[66] Ishida, T., Tokuda, K., Hisaka, A., Honma, M., Kijima, S., Takatoku, H., Iwatsubo, T., Moritoyo, T., Suzuki, H., Initiative, A.D.N.: A novel method to estimate long-term chronological changes from fragmented observations in disease progression. Clinical Pharmacology & Therapeutics **105** (2019) 436–447

[67] Ghazi, M.M., Nielsen, M., Pai, A., Modat, M., Cardoso, M.J., Ourselin, S., Sørensen, L.: Robust parametric modeling of Alzheimer's disease progression. NeuroImage **225** (2021) 117460

[68] Koval, I., Bˆone, A., Louis, M., Lartigue, T., Bottani, S., Marcoux, A., Samper-Gonzalez, J., Burgos, N., Charlier, B., Bertrand, A., et al.: Ad course map charts Alzheimer's disease progression. Scientific Reports **11** (2021) 8020

[69] Garbarino, S., Lorenzi, M., Initiative, A.D.N., et al.: Investigating hypotheses of neurodegeneration by learning dynamical systems of protein propagation in the brain. Neuroimage **235** (2021) 117980

[70] Ghazi, M.M., Sørensen, L., Ourselin, S., Nielsen, M.: Carrnn: A continuous autoregressive recurrent neural network for deep representation learning from sporadic temporal data. IEEE Transactions on Neural Networks and Learning Systems (2022)

[71] Cho, S.H., Woo, S., Kim, C., Kim, H.J., Jang, H., Kim, B.C., Kim, S.E., Kim, S.J., Kim, J.P., Jung, Y.H., et al.: Disease progression modelling from preclinical Alzheimer's disease (AD) to AD dementia. Scientific reports **11** (2021) 1–10

[72] Teng, E., Becker, B.W., Woo, E., Knopman, D.S., Cummings, J.L., Lu, P.H.: Utility of the Functional Activities Questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer's disease. Alzheimer disease and associated disorders **24** (2010) 348

[73] Kukull, W., Larson, E., Teri, L., Bowen, J., McCormick, W., Pfanschmidt, M.: The mini-mental state examination score and the clinical diagnosis of dementia. Journal of clinical epidemiology **47** (1994) 1061–1067

[74] Yang, Y.W., Hsu, K.C., Wei, C.Y., Tzeng, R.C., Chiu, P.Y.: Operational determination of subjective cognitive decline, mild cognitive impairment, and dementia using sum of boxes of the clinical dementia rating scale. Frontiers in Aging Neuroscience **13** (2021)

[75] Ou, Y.N., Xu, W., Li, J.Q., Guo, Y., Cui, M., Chen, K.L., Huang, Y.Y., Dong, Q., Tan, L., Yu, J.T.: FDGPET as an independent biomarker for Alzheimer's biological diagnosis: a longitudinal study. Alzheimer's research & therapy **11** (2019) 1–11

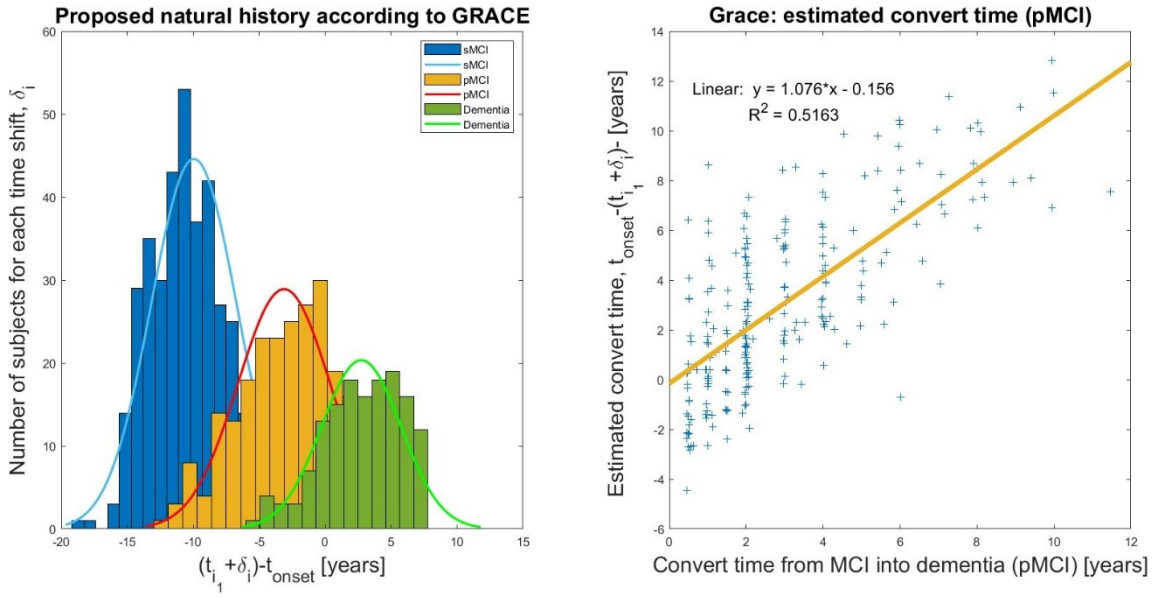

Figure 1: Classification of the subjects according to the time shift δ_i. The first column shows the histogram of the number of subjects as a function of the temporal ordering proposed by GRACE, $\left(t_{ij}^{c}+\delta_i\right)-\hat{t}_{onset}$. The second column shows the conversion times of the pMCI subjects with respect to their estimations by GRACE.

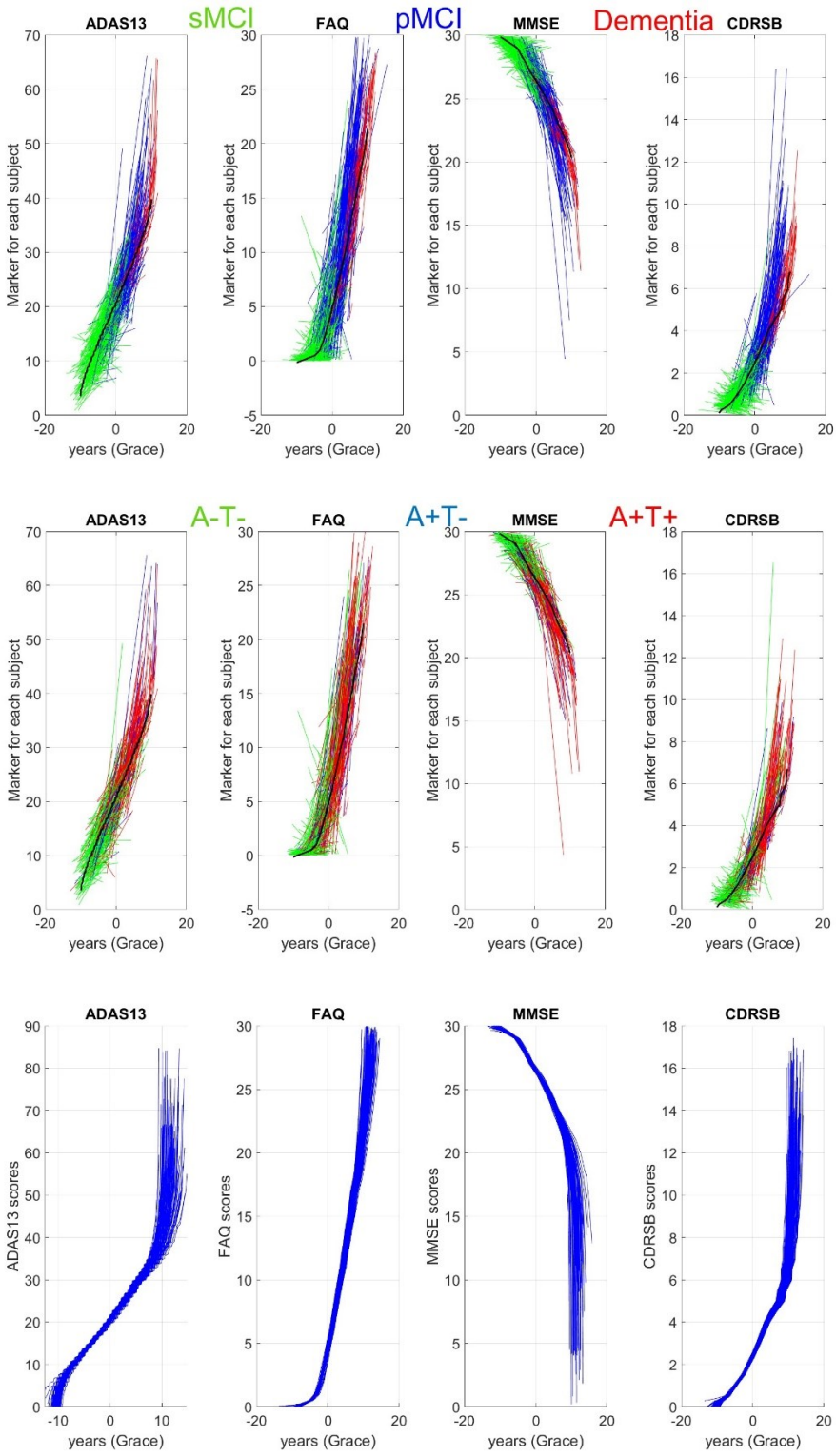

Figure 2: For each selected marker, the long-term trajectories (black lines) are superimposed over the subject-level observations in the original scale and colored according to diagnosis in the first row (sMCI in green, pMCI in blue and dementia in red) or AT profiles in the second row (A- in green, A+T- in blue and A+T+ in red). In the third panel, in blue, the long term trajectories of the selected markers from the bootstrapping technique.

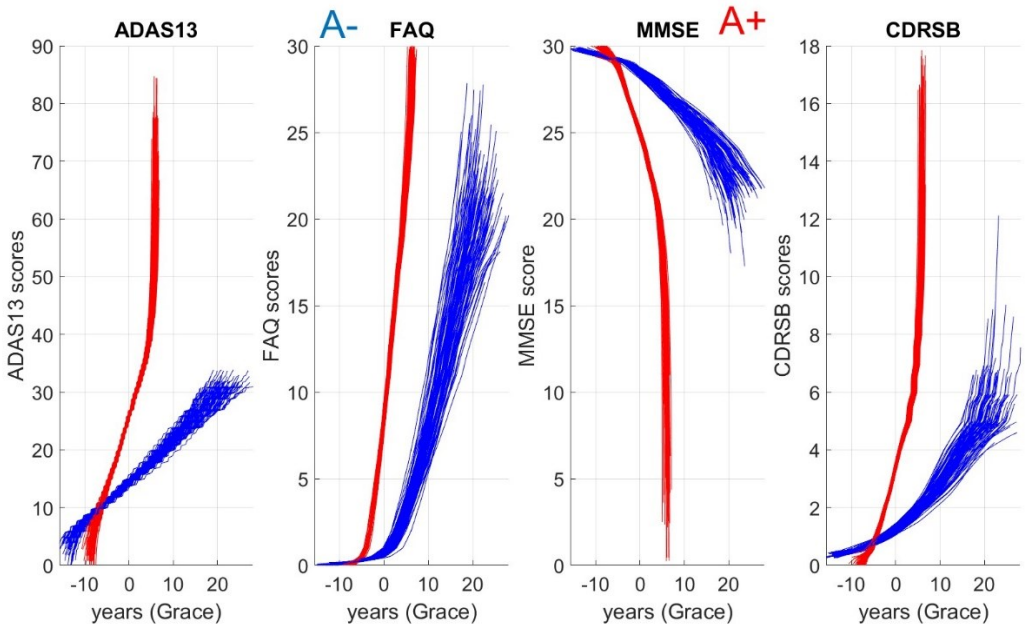

Figure 3: Long trajectories between the subjects with A+ and A- of the selected NM vector. Use of the bootstrap technique to infer the confidence intervals of the long trajectories.

Table 1: Comparison between the results of the proposed predictive model for conversion to dementia (using a binary classification, convert or not) with respect to the three highest risk factors (LMCI, A +, APOE4) at baseline. N= number of subjects.

		General population				Sub-group with CSF markers at baseline				
Clinical	*time* ≤ *month*48?	n	LMCI	APOE4	model	n	LMCI	APOE4	A+	model
pMCI	fast	182	82.4%	69.8%	80.8%	134	76.9%	69.4%	84.3%	82.1%
	slow	48	68.6%	50%	43.8%	34	58.8%	52.9%	64.7%	47.1%
sMCI	fast	207	32.9%	43.5%	28.5%	176	32.4%	44.3%	40.3%	26.7%
	slow	196	38.3%	33.7%	13.8%	157	28.7%	34.4%	22.3%	14.0%

Table 2: For each selected marker, a comparison among the mean values by clinical group (sMCI, pMCI, dementia) at baseline, with respect to the values estimated by the long trajectories at the onset of dementia and the cut-offs used in the clinical practice for determining progression to dementia.

Demographic and clinical characteristics					Model
Subjects	sMCI	pMCI	Dem	clinical cutoff	$t_{onset} = 3.4 years$
Cognitive outcomes					
ADAS13	12.9	19.9	31.1	26.5 [71]	25.9
CDRSB	1.2	1.9	4.5	3 [74]	3.7
FAQ	1.7	4.8	13.4	6 [72]	9.6
MMSE	28.3	27.2	23.1	26 [73]	24.8
CSF biomarkers & FDG					
$A\beta_{42}$	1117.42	752.04	690.21	880 [31]	605
pTau	23.08	34.54	37.14	27 [31]	31
$pTau/A\beta_{42}$	0.03	0.05	0.06	0.028 [31]	0.052
FDG	1.29	1.19	1.06	1.21 [75]	1.13