Disease progression modeling in Alzheimer’s disease: insights from the shape of cognitive decline

Lars Lau Raket, PhD¹ ², for the Alzheimer’s Disease Neuroimaging Initiative*

¹H. Lundbeck A/S, Denmark
²Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Lund, Sweden

*Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background: The characterizing symptom of Alzheimer disease (AD) is cognitive deterioration. While much recent work has focused on defining AD as a biological construct, most patients are still diagnosed, staged, and treated based on their cognitive symptoms. But the cognitive capability of a patient at any time throughout this deterioration will not directly reflect the disease state, but rather the effect of the cognitive decline on the patient’s predisease cognitive capability. Patients with high predisease cognitive capabilities tend to score better on cognitive tests relative to patients with low predisease cognitive capabilities at the same disease stage. Thus, a single assessment with a cognitive test is not adequate for determining the stage of an AD patient.

Methods and Findings: I developed a joint statistical model that explicitly modeled disease stage, baseline cognition, and the patients’ individual changes in cognitive ability as latent variables. The developed model takes the form of a nonlinear mixed-effects model. Maximum-likelihood estimation in this model induces a data-driven criterion for separating disease progression and baseline cognition. Applied to data from the Alzheimer’s Disease Neuroimaging Initiative, the model estimated a timeline of cognitive decline in AD that spans approximately 15 years from the earliest subjective cognitive deficits to severe AD dementia. It was demonstrated how direct modeling of latent factors that modify the observed data patterns provide a scaffold for understanding disease progression, biomarkers and treatment effects along the continuous time progression of disease.

Conclusions: The suggested framework enables direct interpretations of factors that modify cognitive decline. The results give new insights to the value of biomarkers for staging patients and suggest alternative explanations for previous findings related to accelerated cognitive decline among highly educated patients and patients on symptomatic treatments.
1. Background

Alzheimer disease (AD) is slowly progressing with preclinical and prodromal phases lasting many years before the onset of dementia. The stage of the underlying disease process of an AD patient entering a clinical trial is largely unknown, but may be estimated by a combination of, for example, cognitive testing, clinical evaluation, and biomarker results. While these procedures for evaluating disease severity are useful for creating coarse groupings of patients, the factors used to create groupings may be systematically affected by a wealth of factors not directly tied to the disease process, for example, intelligence, level of education, comorbidities, and genetics.

So far, efforts to develop therapies that delay or halt the progression of AD have generally been unsuccessful, and the vast majority of trials testing symptomatic agents in AD have also failed. These failures may be due to wrong therapeutic targets or non-efficacious therapies, but it is possible that a proportion of trial failures could be attributed to other factors such as study design, endpoints, and non-optimal patient population selection. For disease-modifying drugs, for example, the current standard durations for interventional studies may not be adequate. Simulations based on cohort studies suggest that prevention of disease in cognitively normal individuals may require study lengths far beyond the current standard to achieve high statistical power for detecting an effect of even very efficacious drugs [1,2], and this may also be the case for secondary prevention studies.
1.1 Cognitive decline and symptom onset

The characterizing symptom of AD is cognitive deterioration. The cognitive capability of a patient at any time throughout this deterioration will not directly reflect the disease state, but the effect of the cognitive decline on the patient’s predisease cognitive capability.

Many factors influence instantaneous cognitive ability, and low cognitive ability at a single time point is not necessarily an indication of cognitive decline. Cognitive decline can only be established by repeated evaluations of patients’ cognition over time. Longitudinal assessments of patient cognition also offer the benefit of hindsight – once cognitive decline is established; one can traverse back in time along the cognitive trajectory and predict when the decline started and search for patterns that are indicative of future cognitive decline. If done properly, one can synchronize individual observed trajectories to one long-term timeline representative of the full span of cognitive decline over the disease.

Age is typically considered the major risk factor for developing AD, but age of first diagnosis of AD can vary by decades between patients, and because this span is much greater than the entire course of cognitive decline associated with AD, patient age is not an appropriate scale for understanding the pattern of cognitive decline in AD. The natural scale for studying the patterns of cognitive decline is time since symptom onset. However, self- or caregiver-reported age of symptom onset may be imprecise due to the patient’s memory problems; recall bias, where early sporadic cognitive issues are believed to be symptoms of the disease; or personal differences in sensitivity and interpretation of the earliest cognitive problems.

1.2 Disease progression modeling

Alzheimer disease typically presents in a sporadic late-onset form. The autosomal dominant forms of AD (ADAD) caused by rare genetic mutations have earlier onset than sporadic AD, but otherwise the pathogenesis is largely similar [3]. In ADAD, age at symptom onset is strongly affected by mutation type,
parental age at symptom onset, APOE genotype and sex [4]. These factors can be used to calculate
expected patient age at symptom onset for ADAD patients, which can be used to construct a more
synchronized time scale for studying biomarkers and the pathological cascade of the disease [5].
Furthermore, this makes it possible to do primary prevention studies in a highly efficient manner [6].
In sporadic AD, age at onset cannot be predicted accurately from demographic or genetic factors.
Assessment of biomarkers such as amyloid and tau load in cerebrospinal fluid or by positron emission
tomography may be used to diagnose the disease even in the earliest stages [7], but such assessments
can be both invasive and expensive, and data is sparse. There are however rich datasets with
longitudinal cognitive measurements that span different parts of the disease. An appealing use of this
data is to assemble the individual observed short-term trajectories to one long-term timeline
representative of the full span of cognitive decline over the disease.

Different approaches to construct disease progression models for AD have been taken. A classic
approach is to formulate the changes in cognitive scores using differential equations [8,9,10,11]. One
major drawback of this type of modeling is that covariate effects and different sources of random
variation should be formulated in the differential equation framework and may be very difficult to
handle and interpret. An alternative class of disease progression models is based on direct modeling of
the observed longitudinal trajectories and explicit modeling the patient-level disease stage. An
important example of this type of approach is the model by Donohue et al. [12] which simultaneously
models multiple observations of cognitive measures and biomarkers. This modeling approach has been
powerful in illuminating the multivariate nature of Alzheimer disease progression, but it does have some
drawbacks. In particular, the model assumes that all included outcomes are synchronized over a
common disease time scale; it does not model correlations between outcomes; it does not include
covariates; and finally, it uses a heuristic estimation procedure that does not simultaneously account for
the assumed random variability in both timing and measurements. The approach was recently
generalized to a wider class of Bayesian latent-time joint mixed-effects models [13]. This class of models
can in principle take advantage of the dependence between different outcomes and allow inclusion of
covariates, but covariates can only model variation in outcomes and not disease stage or progression
rate. Furthermore, the Bayesian element of the model requires somewhat arbitrary choices of priors for
different model parameters. Progression is modeled on an age scale and, as discussed above, this choice
may amplify the a-priori dis-synchronization between patients by orders of magnitude. This in turn
negatively affects the results of the model as can be seen in Figure 1 in the paper [13] where patient-
level trajectories go from minimal to maximal severity over 10-15 years, while variation of when
maximal severity is reached between patients is spread out over 20-year periods.

In this paper, we propose a new approach to disease progression modeling that overcomes several of
the shortcomings of the above-mentioned methods. In the presented form, the model is estimating a
disease timeline from repeated assessments of a univariate measure such as a cognitive scale. The
model is inspired by the statistical framework presented by Raket et al. [14] where systematic patterns
of variation in both vertical (observed cognitive score) and horizontal (disease-timing) directions are
modeled simultaneously on both the population and individual level. The model allows covariate effects
on both outcomes and disease progression and all model parameters are estimated simultaneously
using maximum likelihood estimation.

The goal of this work was to explore whether the proposed disease progression model could align
observed cognitive trajectories to a precise timeline of cognitive decline associated with AD and to
evaluate if this modeling would shed new light on aspects related to disease progression and
biomarkers. When the model was fitted to cognitive scores from ADNI, the presented model aligned the
cognitive trajectories of patients to a consistent shape of cognitive decline with a span of approximately
15 years from the earliest subjective cognitive deficits to severe AD dementia. It was shown that the model’s predictions of patients’ disease stages based on their longitudinal cognitive scores could predict time since symptom onset and diagnosis. It was further demonstrated that the predicted disease stages provided a more suitable time scale for modeling the evolution of biomarkers over the course of disease than group-wise modeling based on patient symptoms at baseline. The model was used to estimate the effects of sex, age, and education on cognitive decline and to evaluate the effects of cholinesterase inhibitor treatment on cognitive decline. Finally, the model was fitted to the cognitive trajectories of a subset of patients with a rich set of biomarkers available at baseline to estimate if baseline biomarker profile could predict disease stage. The results of the model in an independent held-out validation data set confirmed that baseline biomarker profiles could predict the disease stage of unseen individuals – even in the preclinical phases of disease where no clinically detectable cognitive impairment is present.

2. Methods

2.1. Data

Data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

Patients included in the current study were required to have a valid classification at baseline (cognitively normal, significant memory concern, mild cognitive impairment [early], mild cognitive impairment [late], or dementia).
Outcomes

The main outcome measure considered was the total score of the 13-item Alzheimer’s Disease Assessment Scale-cognitive subscale (ADAS-cog; range: 0-85; lower score indicates less impairment) [15]. Included patients were required to have at least one valid ADAS-cog total score to be included in the present study.

Other outcomes were reported onset of various symptoms related to cognitive impairment and AD; Clinical Dementia Rating scale – sum of boxes [16]; Functional Activities Questionnaire [17], FDG-PET (meta-ROI) [18]; cross-sectional hippocampal volume extracted from MRI using FreeSurfer [19]; Florbetapir PET SUVr [20]; Aβ_{1-42}, total tau, and p-tau_{181} concentrations in CSF as measured using the Roche Elecsys® immunoassay [21]; the ratio of Aβ_{1-42}/Aβ_{1-40} concentrations in CSF as measured by 2D-UPLC tandem mass spectrometry; and Neurofilament Light chain (NFL) levels in plasma measured using a single molecule array platform [22].

2.2. Disease progression model

Let $y_{ij}$ represent the observed cognitive score of patient $i$ at time $t_{ij}$ ($i = 1, ..., n, j = 1, ..., m_i$). We assume that $y_{ij}$ is generated by a model of the form

$$ y_{ij} = \theta(v_i(t_{ij})) + x_i(t_{ij}) + \varepsilon_{ij} $$

where $\theta$ is a function that represents the shape of cognitive decline; $v_i$ is a warping function that transforms observation time $t_{ij}$ to a disease time scale $v_i(t_{ij})$ that is aligned across patients; $x_i$ is the idiosyncratic patient-level deviation from the mean shape that represents consistent deviations over time; and $\varepsilon_{ij}$ is independent measurement noise.
Cognitive scores can be extremely noisy due to many different sources of variation, and to accurately infer the shape of the disease timeline of cognitive decline $\theta_i$, to predict patient-level disease stage $v_i$, and to predict the entire patient-level course of decline $\gamma$, one will have to make suitable model choices. In the following, we describe the basic model choices and their motivations.

Because we are modeling cognitive decline in a pathological aging, it is natural to assume that the representative shape of decline $\theta$ is a function that has a stable left asymptote (predisease cognitive normality) and a monotone decline. In this paper, we focus on ADAS-cog scores that show a distinct exponential decline in dementia [23], and thus we will work with a parametrized family of exponential functions to model the mean progression pattern

$$\theta(t) = t \cdot \exp \left( \frac{t - s}{\exp(g)} \right) + v,$$

where $v$ is the left asymptote representing the average stable predisease cognitive score, and where the remaining parameters determine the shape of the decline.

The mean progression pattern $\theta$ can be modeled differently to achieve other properties, for example as a generalized logistic function or as a monotone spline [24].

The mapping of observed time to disease time $v_i$ should allow the model to assemble short-term longitudinal observations to a long-term timeline of cognitive decline. The major source of variation can likely be ascribed to differences in how long the patient has had the disease before we begin observing them, so we model $v_i$ as a shift of study time

$$v_i(t) = t + s_i.$$

Random effects
When modeling longitudinal data for groups of individuals it is often natural to describe systematic differences between individuals using random effects. The proposed disease progression model has three types of random effects.

- $s_i$: random patient-level shift that models the disease stage of patient $i$. Assumed to follow a zero-mean normal distribution with unknown variance.

- $x_i$: random patient-level systematic deviation from the mean curve. Assumed to be a discrete-time observation of a Brownian motion with a zero-mean normally distributed starting level with unknown variance. The Brownian motion has an unknown parameter controlling variance scale.

- $\epsilon_{ij}$: random observation noise. Assumed to be independent zero-mean normally distributed with unknown variance.

A free correlation between $s_i$ and the normally distributed starting level of $x_i$ is included in the model, the remaining effects are assumed independent.

**Fixed effects**

The basic model parameters $l$, $g$, $s$, and $v$ that describe the shape of $\theta$ are modeled as fixed effects.

- $l$ is a scaling parameter of the exponential function. Since a goal of disease progression modeling is to find a common pattern of decline, $l$ will be modeled as a single free parameter.

- $g$ is a scaling parameter of time. Patient-level differences in rate of decline that can be ascribed to a covariate or factor can be modeled as a regression-type model on $g$. Initially this parameter will be modeled as a single free parameter.

- $s$ is a shift of observed time. Patient-level differences in disease stage that can be ascribed to a covariate or factor can be modeled as fixed effects. Since the present study includes several
cohorts at different disease stages (e.g. cognitively normal patients, patients with dementia) the initial modeling will have different parameters for non-cognitively normal cohorts, thus modeling disease time since the baseline stage of the cognitively normal patients.

• $v$ is an intercept parameter describing the left asymptote. Patient-level differences in predisease cognition that can be ascribed to a covariate or factor can be modeled as a regression-type model on $v$. Initially this parameter will be modeled as a single free parameter.

### 2.3. Statistical analysis

To investigate the effect of covariates on the pattern of disease progression, forward selection was used to evaluate models with all combinations of covariate effects on rate of decline $g$, disease stage $s$, and predisease cognition $v$. The search was continued as long as the Akaike Information Criterion [25] improved, but the model selection was based on the more conservative Schwarz' Bayesian Information Criterion (BIC) [26].

To investigate if predicted disease time was predictive of time since reported symptom onset, linear regression was done on time since reported symptom onset (at baseline) using predicted disease time as a covariate. P-values were computed using T-tests.

To investigate if predicted disease time offered a better time scale for modeling other longitudinal outcomes (e.g. biomarkers) than time since baseline for the five baseline groups, linear mixed effects modeling was used. To allow for nonlinear trends in the mean pattern, the outcome was modeled using a cubic B-spline function with 3 degrees of freedom across predicted disease time and time since baseline (one pattern per baseline group), respectively. Patient-level random slopes and intercepts were included to model longitudinal deviations within an individual. P-values were computed using likelihood ratio tests with maximum likelihood estimation.
Comparisons of quantitative outcomes between groups with two levels were done using Wilcoxon rank sum tests and correlations were evaluated with Spearman’s rank correlation coefficients.

**Software**

All analyses were done using R version 3.5.2 [27]. Maximum likelihood estimation in the disease progression models was done using the method of Lindstrom and Bates [28] using the nlme and covBM R packages [29].

### 3. Results

#### 3.1 Basic model

The basic model described in Section 2.2 was fitted on longitudinal ADAS-cog data from ADNI. The data comprised 9,830 ADAS-cog scores across 2,142 individuals. The ADAS-cog scores plotted against study time and predicted time-scales (“disease month”) are shown in Figure 1. Relative to the average baseline disease stage of the Cognitively normal group the model estimated that the Significant memory concern group were 29 months later into the trajectory of cognitive decline, while the early and late MCI groups were respectively 42 and 88 months later, and that the dementia group was 136 months later.

The model had 12 degrees of freedom, and the log likelihood of the fitted model was -29734.26. AIC and BIC were 59492.52 and 59578.84 respectively.
Figure 1. Observed longitudinal ADAS-cog trajectories for 2,142 ADNI participants plotted against time in study (top), predicted patient baseline-status group-level disease time (middle), and predicted individual disease time.
3.2 Validation of the basic model

The presented disease progression model aggregates the information in baseline status groups and the longitudinal trajectories of participants to a single number, the predicted disease month. For this continuous disease progression scale to be relevant to AD, it should also hold information that describe other aspects of the disease than the cognitive deterioration observed on ADAS-cog that the model was fitted on.

To first evaluate whether the disease progression model captured milestones of cognitive deterioration, we evaluated the model’s ability to predict self-reported onset of cognitive symptoms, mild cognitive impairment symptoms, Alzheimer’s disease symptoms or diagnosis of Alzheimer’s disease. There were 1.142 participants who had at least one entry of these data during the study follow-up. Age of symptom onset or diagnosis plotted against the age at the model’s predicted disease time 0 is shown in Figure 2.

Ideally (based on perfectly aligned trajectories and perfect self-reported onsets/diagnosis between individuals) the results of each measure would lie on a line with slope 1 and the intercept would represent the difference in years between age at disease time 0 and the age at onset/diagnosis time. For the age of onset of cognitive symptoms, there seems to be different intercepts for the different baseline groups, where more severe baseline groups tend to report symptom onset later relative to the model prediction of the less severe groups. This may be an effect of different subjective definitions of onset of cognitive symptoms across baseline groups; it may be because of biased model estimates of the staging of the baseline groups; or a combination.

Based on linear regression, predicted disease month was predictive of time since cognitive symptom onset ($p < 0.0001$), time since Alzheimer’s disease symptoms onset ($p < 0.0001$) and time since Alzheimer’s diagnosis ($p < 0.0001$) – all times relative to study baseline. Predicted disease month was not significantly predictive for time since mild cognitive impairment symptom onset ($p = 0.558$).
Figure 2. Reported age of onset of cognitive symptoms (top left), cognitive impairment symptoms (top right), Alzheimer’s disease symptoms (bottom left) and Alzheimer’s disease diagnosis (bottom right) as a function of age at predicted disease month 0.
Secondly, to validate that the predicted disease time also synchronizes other independently captured aspects of the disease than cognition as measured by ADAS-cog, we analyzed if this continuous disease scale based on baseline patient statuses and ADAS-cog trajectories better captured patterns of variation in other clinical scales and biomarkers than separate modeling of the different baseline groups. We found that the predicted disease time better described the patterns of variation compared to allowing separate patterns per baseline group in 7 of the 10 cases when measured by log likelihood (Table 1), even though the latter model had 16 degrees of freedom more than the former. When measured using AIC and BIC that both adjust for additional degrees of freedom to compare model quality, the predicted disease time model was better in 8 of the 10 cases for AIC and 10 of the 10 cases for BIC. Interestingly, the three biomarkers where group-wise modeling was better as measured by log likelihood were all measuring to beta amyloid (CSF Aβ₁₋₄₂ and Aβ₁₋₄₂/Aβ₁₋₄₀ ratio, Florbetapir PET). These biomarkers are known to have a bimodal distribution [30] and are thus poorly modeled by a single trajectory. Figure 3 shows the estimated trajectories of the two models for hippocampal volume (MRI), Aβ₁₋₄₂ (CSF), and NfL (plasma).
Table 1. Comparison of longitudinal modeling of clinical scales and biomarkers based on patient baseline group versus continuous disease time. Comparison in terms of log likelihood (larger is better), AIC and BIC (smaller is better). Bold numbers indicate the best fitting model for a given measure.

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Number of observations (number of patients)</th>
<th>One trajectory per baseline group (df = 24)</th>
<th>One trajectory across predicted disease time (df = 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Dementia Rating scale – sum of boxes</td>
<td>9,712 (2,142)</td>
<td>-14792.0</td>
<td>29632.0</td>
</tr>
<tr>
<td>Functional Activities Questionnaire</td>
<td>9,715 (2,126)</td>
<td>-25664.1</td>
<td>51376.2</td>
</tr>
<tr>
<td>FDG-PET (meta-ROI)</td>
<td>3,461 (1,454)</td>
<td>3592.7</td>
<td>-7137.3</td>
</tr>
<tr>
<td>Hippocampal volume (MRI)</td>
<td>6,052 (1,675)</td>
<td>-44202.4</td>
<td>88452.7</td>
</tr>
<tr>
<td>Florbetapir PET SUVr</td>
<td>2,568 (1,224)</td>
<td>1733.2</td>
<td>-3418.3</td>
</tr>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>mean</td>
<td>std. dev</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td><strong>Aβ1-42 (CSF)</strong></td>
<td>2,342 (1,252)</td>
<td>-16979.2</td>
<td>34006.5</td>
</tr>
<tr>
<td><strong>Aβ1-42/Aβ1-40 (CSF)</strong></td>
<td>1,425 (867)</td>
<td>2919.1</td>
<td>-5790.2</td>
</tr>
<tr>
<td><strong>Total tau (CSF)</strong></td>
<td>2,334 (1,247)</td>
<td>-1320.5</td>
<td>26489.1</td>
</tr>
<tr>
<td><strong>p-tau (CSF)</strong></td>
<td>2,330 (1,246)</td>
<td>-7924.9</td>
<td>15897.8</td>
</tr>
<tr>
<td><strong>NfL (plasma)</strong></td>
<td>4,219 (1,576)</td>
<td>-18792.2</td>
<td>37632.5</td>
</tr>
</tbody>
</table>

*df* degrees of freedom

---

It is made available under a CC-BY-NC 4.0 International license.
Figure 3. Data and estimated biomarker trajectories for hippocampal volume (MRI), $\text{A}\beta_{1-42}$ (CSF), and NfL (plasma). Left column shows results when allowing different trajectories for the five baseline groups, right column shows the results when requiring a single trajectory over predicted disease time.
3.2 Age, sex, education and cognitive decline

There were systematic differences in follow-up time, age at baseline and length of education between male and female participants (Supplementary Table 1). Compared to female participants, male participants on average had 3.2 months longer follow-up (Wilcoxon \(p = 0.0085\)), were on average 2.0 years older at baseline (Wilcoxon \(p < 0.0001\)) and had 0.89 years more education (Wilcoxon \(p < 0.0001\)).

Age at baseline and years of education were not significantly correlated (Spearman's \(\rho = -0.04, p = 0.0792\)).

To explore whether age at baseline, sex and length of education affected the pattern of cognitive decline, stepwise forward model selection was done to include these factors in the model. The best model included fixed effects of age and sex on \(g\), and \(v\), and fixed effects of years of education on \(g\) and \(v\). While there were some substantial differences in marginal parameter estimates due to age, sex and length of education (e.g. men are predicted to be 57 months later in disease compared to women in the same baseline groups; Supplementary Table 2), the estimates should not be interpreted in isolation since all parameters simultaneously affect the shape of the disease trajectory and may counteract each other. Figure 4 shows how age, sex and education differences systematically affected the mean trajectories. From the figure we see that male participants consistently scored lower on ADAS-cog throughout the disease (3.1 points), but that they remained more stable in the initial 100 months where female participants had a more gradual decline. Lower age at baseline and longer education were both associated with higher cognitive scores, but also slightly increased rates of decline as evident in the stages of overt dementia (predicted disease time >120 months).
Figure 4. Estimated trajectories for different combination of patient age, sex and length of education. The trajectories are aligned at predicted month 0 that corresponds to the average cognitive stage of cognitively normal individuals at baseline.

3.3 Cholinesterase inhibitors and cognitive decline

Using the search terms described in the supplementary material, we identified 1,347 individuals that were treated with cholinesterase inhibitors (ChEIs), but only 64 of these initiated or discontinued treatment during the observation time (total of 9 initiations and 60 discontinuations).

To explore if treatment with ChEIs affected the shape of the decline trajectories, stepwise forward model selection from the basic model was done to include ChEI treatment in the model. The best model included fixed effects of treatment on $s$ and $v$, but not on rate of decline $g$ (14 degrees of freedom, log likelihood = -29638.79, AIC = 59305.59, BIC = 59406.29). The model found that patients treated with ChEIs generally had worse level of cognition (effect on $v$ was 5.50 ADAS-cog points for treated
individuals, \( p < 0.0001 \) and a delayed progression within baseline groups (effect on was 7.53 months, \( p < 0.0001 \)). The average trajectories and distribution of data across treatment are shown in Figure 5.
Figure 5. Estimated trajectories for patients with and without cholinesterase inhibitor treatment (top) and corresponding distribution of observed ADAS-cog scores. The trajectories are aligned at predicted month 0 that corresponds to the average cognitive stage of cognitively normal individuals that are not treated with cholinesterase inhibitors at baseline.
3.4 Biomarkers for disease staging

The disease progression model relies on observing patients longitudinally and uses the temporal pattern of cognitive scores to predict the patient’s status at baseline. This is valuable for increasing disease understanding for cohorts or for doing retrospective analyses, but the models presented thus far offer only little insight into the disease stage of a patient entering, for example, a clinical trial. In this setting, only the baseline classification of the patient, the cognitive score and possibly other demographic data would be able to inform the stage of the patient. However, as shown in Section 3.2, several biomarkers have temporal patterns that follow the trajectory of cognitive decline. Biomarker data collected at baseline may thus enable a better assessment of the stage of an individual.

688 individuals had complete biomarker data at baseline for the eight biomarkers considered in Section 3.2. These individuals had 3,301 visits with valid ADAS-cog scores.

Training and validation data

540 individuals (80%) were randomly selected for the training cohort and the remaining 148 (20%) comprised the validation cohort.

Model development

Using the BIC-based model selection procedure described previously, we searched for the best model among models that included adjustment for sex, baseline age and education (on parameters $g, s, v$) as well as adjustment for the eight baseline biomarkers on disease stage (parameter $s$). The model selection was done on the training data. The best model included the biomarkers FDG-PET (meta-ROI), hippocampal volume (MRI), Florbetapir PET SUVr, $A\beta_{1-42}/A\beta_{1-40}$ (CSF), and NfL (plasma) (22 degrees of freedom, log likelihood = -7591.17, AIC = 15226.34, BIC = 15355.45). The parameter estimates for the model are given in Supplementary Table 3.
Model validation

To validate the biomarker model, the model fitted on training data was used to predict disease stage in two different scenarios. The first used baseline biomarker data and the second used baseline biomarkers in combination with baseline ADAS-cog total score. Visual inspection of the longitudinal ADAS-cog trajectories suggests that the baseline data does hold information that improves prediction of disease stage in the test data (Figure 6). To quantify this, the biomarker model was compared to the baseline model on its predictive accuracy for the longitudinal ADAS-cog total score trajectories (Table 2). Inclusion of biomarker data clearly reduced the mean squared error (MSE) and median absolute error (MAE) of predictions on both test and training data (MSE/MAE 65.1/4.21 vs. 100.0/4.98 on test data). Including the baseline ADAS-cog total score improved the post-baseline predictive accuracy of the biomarker model further (MSE/MAE 55.1/3.48 for baseline biomarkers + ADAS-cog model vs. 69.8/4.31 for biomarker model on test data).
Figure 6. Predicted disease month for training and test datasets. Top row displays predicted disease-time alignment of observed ADAS-cog total score trajectories based on baseline biomarker data; bottom row displays predicted disease-time alignment of trajectories based on baseline ADAS-cog total score and baseline biomarker data.
Table 2. Predictive accuracies of predicted ADAS-cog total score trajectories for the basic model and the biomarker model both with and without the baseline ADAS-cog total score. Predictions were censored to the interval [0, 85] to respect the range of the ADAS-cog scores.

<table>
<thead>
<tr>
<th>Model and data</th>
<th>Mean squared error</th>
<th>Median absolute error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Test</td>
</tr>
<tr>
<td>Basic model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(baseline status)</td>
<td>90.0</td>
<td>100.0 (110.6&lt;sup&gt;1&lt;/sup&gt;)</td>
</tr>
<tr>
<td>BM model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(baseline status + BMs)</td>
<td>58.3</td>
<td>65.1 (69.8&lt;sup&gt;1&lt;/sup&gt;)</td>
</tr>
<tr>
<td>Basic model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(baseline status + ADAS-cog)</td>
<td>78.2&lt;sup&gt;1&lt;/sup&gt;</td>
<td>102.7&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>BM model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(baseline status + ADAS-cog + BMs)</td>
<td>53.5&lt;sup&gt;1&lt;/sup&gt;</td>
<td>55.1&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>1</sup>Baseline ADAS-cog measurements excluded in computation of prediction errors
4. Discussion

Disease progression modeling

In this paper we presented a model for progression of dementia based on longitudinal cognitive assessments. Disease stages of individual patients were modeled using a latent variable approach. As opposed to conventional latent variable models, for example, in item response theory for modeling cognitive tests [31,32], the proposed model imposes explicit structures to ensure that the longitudinal modeling respects the known course of disease (e.g. that disease progression is an increasing function of elapsed time and that cognition on average declines with disease progression). By imposing these structures, the model provides a scaffold for understanding disease progression in pathological aging in terms of three continuous measures, disease stage, rate of decline, and cognitive deviation from the mean.

The proposed model aligned trajectories of cognitive decline, the characterizing symptom of dementia, and to demonstrate that this approach provided valid insights about other aspects of the disease, it was shown that predicted disease stage was predictive of various measures of disease onset. Furthermore, the use of ADAS-cog trajectories to map patients to a one-dimensional disease timeline consistently provided a better explanation of other clinical scales and biomarker trajectories than a conventional approach that grouped patients based on baseline symptoms.

Age, sex, education and cognitive decline

The effect of demographic and socioeconomic factors on disease risk and manifestation in Alzheimer’s disease has been the subject of much study. In this work we focused on the combined effects of age, sex and length of education.
When considered individually, these factors have been observed to result in differences in disease progression. While age is typically considered the major risk factor for developing AD, higher age at AD onset has been observed to be associated with a slower rate of cognitive decline [33,34]. Similarly, female sex has been identified as a major risk factor with almost two-thirds of AD cases being women [35]. While this difference has been known for a long time, it has only become apparent more recently that there are sex differences in symptomatology, rate of decline, and possibly in neural anatomy [36,37]. Finally, the effects of cognitive reserve on age-related cognitive decline has been the subject of much study [38]. Cognitive reserve is often studied using educational attainment as an operational proxy for cognitive reserve. It has consistently been found that higher education is associated with increased rate of cognitive decline in incident Alzheimer’s disease [39,40,41,42,43,44,45]. Several of these studies also report that education is associated with higher baseline cognition.

Differences in cognitive decline is often studied by comparing slopes in statistical models that assume that cognitive decline follows a linear pattern. The argumentation and interpretation around the cognitive reserve model is somewhat more sophisticated, but still largely centered on an assumption of a linear rate of decline (e.g. illustrated in Figure 1 in [46]). The prevailing hypothesis within the field of cognitive reserve research is that, compared to individuals with low cognitive reserve, individuals with high cognitive reserve have higher predisease cognitive scores and that their brains tolerate a higher load of neuropathology before cognitive decline is seen. At a sufficiently high level of neuropathology, cognitive ability reaches its floor for all participants. If the timescale of neuropathological buildup is similar across individuals, this suggests that individuals with high cognitive reserve will have to decline a wider range of cognitive scores in a shorter time, thus leading to an accelerated rate of decline [46].

The analyses in the present paper clearly illustrate that rate of cognitive decline as measured on ADAS-cog is not constant but increases over the course of AD. Thus, findings of an increased rate of decline in
a certain group of patients using slope models could either be because the group of patients has accelerated decline, because they are at a later disease stage, or a combination. The proposed disease progression model seeks to align cognitive trajectories on a disease timeline, and thus it allows one to separate the hypothesized mechanisms of cognitive decline. The best model that adjusted for effects of age at baseline, sex, and length of education on, respectively, disease stage, rate of decline and cognitive deviation, found that that all three factors affected all three disease measures except for disease stage which was not affected by length of education.

When considering the combination of effects (Figure 4), the results suggested that higher age at baseline was associated with lower cognition throughout disease time and a slightly reduced rate of decline. Women tended to have better predisease cognition but also an accelerated decline. Finally, longer education was associated with slightly faster rate of decline and a systematically better cognition throughout the disease.

While these findings are largely consistent with previous findings, they also illustrate that previous results that do not take the long-term disease trajectories into account may be systematically biased. In particular, the fact that highly educated patients tend to have above-mean cognition throughout the disease means that they will meet cognitive cut-offs used for inclusion criteria in clinical studies longer in to their disease than patients with less education. Because of the accelerated cognitive decline in the later stages of disease, these patients will have a much faster rate of decline when using conventional slope models, but this difference will primarily be due to their later disease stage.

Symptomatic medications for Alzheimer disease and cognitive decline

ChEIs have consistently shown a symptomatic benefit in mild to severe dementia due to AD in randomized, double blind, placebo-controlled trials [47]. It has however been questioned whether long-term treatment with ChEIs could be harmful [48]. A recent meta-analysis found that AD patients treated
with symptomatic treatments had a faster rate of cognitive decline [49]. This could be interpreted as a harmful side effect, but since the included studies were not randomized with respect to symptomatic treatments, such causal link cannot be made. An alternative explanation is simply that ChEIs work – that patients that are being treated at study inclusion have a cognitive benefit that, similarly to higher levels of education, means that they meet inclusion criteria for clinical studies further in to their disease. The optimal disease progression model identified in the model search did not include effects of ChEI treatment on rate of decline. Instead, the results of this model showed that patients treated generally had lower cognition compared to untreated patients (confounding by indication) and that their progression was slightly delayed.

**Biomarker-based disease staging**

The final application of the model examined how a patient’s biomarker profile at study entry could be used to predict their disease stage. Based on training data used for model development a set of 5 biomarkers were included in the model. Inclusion of biomarker profiles considerably improved prediction of future ADAS-cog trajectories in the unseen validation dataset and inclusion of baseline ADAS-cog score further improved the prediction. Among the biomarkers, FDG-PET explained most variation followed by CSF Aβ1-42/Aβ1-40 and Florbetapir SUVr. Hippocampal volume and plasma NfL explained the least.

This modeling of baseline biomarkers for patients in the earliest stages of disease takes advantage of the long-term follow-up that is unique to ADNI. The modeling essentially relies on hindsight because the patients’ disease stage can only be predicted with high reliability once a systematic pattern of cognitive decline has been observed. By using these patterns, the model identified how combinations of biomarkers could be used to predict disease stage. The results of the model suggest that biomarker...
Profiles at a single time point may be used to predict the disease stage of an individual even in the preclinical phases of disease where no clinically detectable cognitive impairment is present.

With further validation, these results can be used to define a space of permissible biomarker profiles to use as inclusion criteria in clinical trials. Such biomarker-based synchronization of patient’s disease stage would enable testing a drug in a more homogeneous population. This would in turn greatly increase the power of clinical trials in AD where it is common to see extreme levels of variability in patient trajectories [50,51].

Acknowledgement

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). 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.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmmun; 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 Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.
References


2(10): 1-10.


27. R Core Team. R: A language and environment for statistical computing. 2015.


**Supplementary material**

**Cholinesterase inhibitor treatment data extraction and processing**

Treatment status with cholinesterase inhibitors was extracted from data on key background medications and the concurrent medications log. The latter was searched for the following common names and misspellings `aricept`, `aricpet`, `donepezil`, `exelon`, `galantamine`, `razadyne`, `remenyl`, `reminyl`, `rivastigmine`.

Missing month and day in dates were interpolated as follows: missing day was set to the 1st day of the month. Missing month was set to July.
### Supplementary Table 1 Sex differences.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants, n</td>
<td>1,002</td>
<td>1,140</td>
<td>—</td>
</tr>
<tr>
<td>ADAS-cog scores, n</td>
<td>4,364</td>
<td>5,466</td>
<td>—</td>
</tr>
<tr>
<td>Follow-up time (months),</td>
<td>24.3 [5.5, 59.5]</td>
<td>25.4 [11.9, 60.6]</td>
<td>0.0085</td>
</tr>
<tr>
<td>median [IQR]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at baseline (years),</td>
<td>72.0 [57.4, 77.5]</td>
<td>74.3 [69.5, 79.3]</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>median [IQR]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of education (years),</td>
<td>15 [14, 18]</td>
<td>16 [15, 18]</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>median [IQR]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p-values computed using Wilcoxon rank sum tests

IQR: Interquartile range
**Supplementary Table 2** Parameter estimates in model adjusting for age, sex and education. *p*-values are computed using likelihood ratio tests.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>l</em> (intercept)</td>
<td>0.01337</td>
<td>—</td>
</tr>
<tr>
<td><em>s</em> (significant memory concern)</td>
<td>-35.11</td>
<td>0.0106</td>
</tr>
<tr>
<td><em>s</em> (early MCI)</td>
<td>-44.58</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>s</em> (late MCI)</td>
<td>-87.73</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>s</em> (dementia)</td>
<td>-136.1</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>s</em> (age at baseline)</td>
<td>-1.878</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>s</em> (male)</td>
<td>57.15</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>g</em> (intercept)</td>
<td>3.169</td>
<td>—</td>
</tr>
<tr>
<td><em>g</em> (age at baseline)</td>
<td>0.007061</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>g</em> (male)</td>
<td>-0.1895</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>g</em> (years of education)</td>
<td>-0.003337</td>
<td>0.0011</td>
</tr>
<tr>
<td><em>v</em> (intercept)</td>
<td>7.623</td>
<td>—</td>
</tr>
<tr>
<td><em>v</em> (age at baseline)</td>
<td>0.1234</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>v</em> (male)</td>
<td>3.113</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td><em>v</em> (years of education)</td>
<td>-0.4285</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>
**Supplementary Table 3** Parameter estimates in model adjusting for age, sex, education and biomarkers. *p*-values are computed using likelihood ratio tests.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th></th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>l (intercept)</td>
<td>17.92</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>s (significant memory concern)</td>
<td>-18.20</td>
<td>0.0663</td>
<td></td>
</tr>
<tr>
<td>s (early MCI)</td>
<td>-13.14</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (late MCI)</td>
<td>-40.00</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (dementia)</td>
<td>-64.12</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (FDG-PET)</td>
<td>73.52</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (hippocampal volume)</td>
<td>0.005980</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (Florbetapir SUVr)</td>
<td>-34.57</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (CSF Aβ1-42/Aβ1-40)</td>
<td>137.1</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>s (plasma NfL)</td>
<td>-0.2579</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>g (intercept)</td>
<td>3.280</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>g (age at baseline)</td>
<td>0.01276</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>g (male)</td>
<td>-0.3764</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>g (years of education)</td>
<td>-0.03897</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>v (intercept)</td>
<td>11.71</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>v (male)</td>
<td>4.026</td>
<td>&lt; 0.0001</td>
<td></td>
</tr>
<tr>
<td>v (years of education)</td>
<td>-0.2133</td>
<td>0.217</td>
<td></td>
</tr>
</tbody>
</table>