Neuropsychiatric Disease and Treatment

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Disease progression model for Clinical Dementia Rating–Sum of Boxes in mild cognitive impairment and Alzheimer's subjects from the Alzheimer's Disease Neuroimaging Initiative

This article was published in the following Dove Press journal:
Neuropsychiatric Disease and Treatment
24 May 2014
Number of times this article has been viewed

Mahesh N Samtani
Nandini Raghavan
Gerald Novak
Partha Nandy
Vaibhav A Narayan

On behalf of the Alzheimer's Disease Neuroimaging Initiative

Janssen Research and Development, LLC, Raritan, New Jersey, USA

Video abstract

Point your SmartPhone at the code above. If you have a QR code reader the video abstract will appear. Or use: http://dvpr.es/1jsXAGx

Correspondence: Mahesh N Samtani
Model Based Drug Development, Janssen Research and Development, LLC, Pharmaceutical Companies of Johnson and Johnson, 920 Route 202, PRD 2723, Raritan, NJ 08869, USA
Tel +1 908 704 5367
Fax +1 908 927 2573
Email msamtani@its.jnj.com

Background: The objective of this analysis was to develop a nonlinear disease progression model, using an expanded set of covariates that captures the longitudinal Clinical Dementia Rating Scale–Sum of Boxes (CDR–SB) scores. These were derived from the Alzheimer's Disease Neuroimaging Initiative ADNI-1 study, of 301 Alzheimer's disease and mild cognitive impairment patients who were followed for 2–3 years.

Methods: The model describes progression rate and baseline disease score as a function of covariates. The covariates that were tested fell into five groups: a) hippocampal volume; b) serum and cerebrospinal fluid (CSF) biomarkers; c) demographics and apolipoprotein Epsilon 4 (ApoE4) allele status; d) baseline cognitive tests; and e) disease state and comedications.

Results: Covariates associated with baseline disease severity were disease state, hippocampal volume, and comedication use. Disease progression rate was influenced by baseline CSF biomarkers, Trail-Making Test part A score, delayed logical memory test score, and current level of impairment as measured by CDR–SB. The rate of disease progression was dependent on disease severity, with intermediate scores around the inflection point score of 10 exhibiting high disease progression rate. The CDR–SB disease progression rate in a typical patient, with late mild cognitive impairment and mild Alzheimer's disease, was estimated to be approximately 0.5 and 1.4 points/year, respectively.

Conclusions: In conclusion, this model describes disease progression in terms of CDR–SB changes in patients and its dependency on novel covariates. The CSF biomarkers included in the model discriminate mild cognitive impairment subjects as progressors and nonprogressors. Therefore, the model may be utilized for optimizing study designs, through patient population enrichment and clinical trial simulations.

Keywords: NONMEM®, beta-regression, CSF $A\beta_{1-42}$, hippocampal volume, trial enrichment

Introduction

Alzheimer's disease (AD) is a terminal neurodegenerative disorder characterized by progressive memory decline and loss of other cognitive functions, which develops with neuropathological changes, such as neurofibrillary tangles and senile plaques. The predominant component of senile plaques is an amyloid beta (Aβ) peptide, in particular, a 42 amino acid isoform ($A\beta_{1-42}$) that is derived from a larger amyloid precursor protein.[1] In the brain, $A\beta_{1-42}$ can form soluble neurotoxic oligomers, fibrillar parenchymal plaques closely associated with neuritic dystrophy and gliosis, and fibrillar cerebral amyloid angiopathy. Currently available treatments for AD focus on temporary enhancement of impaired neurotransmitter systems to maximize the remaining activity in neurons

affected by the disease, but these treatments do not slow the progression of cognitive and functional decline.[2] Efforts are underway to develop disease-modifying AD therapeutics that can alter the underlying progressive pathology of AD, thereby slowing clinical decline.

The amyloid cascade hypothesis likely captures the key elements of the pathophysiology of AD.[1] However, none of the potential disease-modifying agents based on this hypothesis have been successful in clinical studies.[1] One possible explanation for the failure of antiamyloid treatments could be that therapy was started too late for patients who already had a diagnosis of AD. Thus, a key learning from these failed trials could be to shift disease interception to an earlier stage of disease ie, mild cognitive impairment (MCI). However, a major challenge is that the syndrome of MCI represents a heterogeneous group of subjects. Postmortem studies suggest that a subset of MCI subjects develop other disorders, including vascular dementia, dementia with Lewy bodies, frontotemporal dementia, and semantic dementia.[3] In this regard, a new draft guidance document released by the US Food and Drug Administration (FDA) seeks to provide industry sponsors with guidance on how to develop drugs for the treatment of early-stage AD.[4] The draft guidance supports "the concept of enriching the trial populations with patients most likely to progress to more overt dementia using both clinical and biomarker based criteria".[4] The present study therefore tested the hypothesis that biomarkers can identify MCI subjects with a higher probability of exhibiting disease progression and eventually developing AD. These MCI patients with pathologic biomarkers could be selected for clinical trials of new candidate therapies, which could serve three purposes highlighted in the FDA's recent draft guidance on Enrichment Strategies for Clinical Trials:[5] a) decreasing heterogeneity – studying MCI patients with pathologic biomarkers has the potential to reduce inter- and intraindividual (residual) variability; b) prognostic enrichment – MCI patients with pathologic biomarkers are more likely to progress to AD; and c) predictive enrichment – amyloid-positive MCI patients are more likely to respond to antiamyloid therapies.

Successful disease interception trials in the MCI populations, in addition to early detection and enrichment strategies, would also require an understanding of disease progression. This could ensure that early-pathology patients recruited in clinical trials would progress measurably within the specified duration of the study. However, the large intersubject variability in disease progression rate for an enriched MCI population could hinder the assessment of drug effects unless additional covariates influencing disease progression are identified. For patients with MCI, the FDA considers the Clinical Dementia Rating Scale–Sum of Boxes (CDR–SB) as an example of a tool to assess disease progression and as a candidate for a single primary efficacy end point.[4] The CDR–SB was also chosen for disease progression modeling in this analysis, based on its long history of usage in AD and MCI trials, the availability of publicly available data for modeling, and its clinical meaningfulness.[6,7] One of the goals of this research was to assess the ability of the CDR–SB to track progression for late MCI (LMCI) and mild-AD patients. The current analysis focused on modeling CDR–SB disease progression, while taking into consideration baseline biomarker status (cerebrospinal fluid [CSF] biomarkers and/or hippocampal volume). The aim was to understand patient heterogeneity by dichotomizing the population into predicted progressors and nonprogressors, based on baseline biomarker status. This modeling approach is analogous to the exercise performed as part of the Alzheimer's Disease Assessment Scale–Cognitive 11-item Subscale (ADAS-cog/11) quantitative clinical trial simulation model that was recently qualified by the FDA and European Medical Agency (EMA), for mild to moderate AD.[8–10] This recently qualified clinical trial simulation model was deemed fit for purpose as a drug development tool for aiding future clinical trial design, in the study of patients with mild to moderate AD. Thus, another goal of this exercise was to extend the drug development tool to the MCI population, to facilitate trial design in earlier stages of the disease.

The EMA has qualified three biomarkers to enrich recruitment into clinical trials in predementia AD.[11–13] These include, a) magnetic resonance imaging (MRI) to measure hippocampal volume; b) CSF tau and Aβ biomarkers; and c) amyloid-related positron emission tomography (PET) signal to image brain Aβ plaques. The current analysis focused on 102 AD patients that were followed for 2 years and 199 MCI patients who were followed for 3 years in the Alzheimer's Disease Neuroimaging Initiative-1 (ADNI-1) study, who had both baseline hippocampal volume and CSF biomarker data. In the current analysis, amyloid-related Pittsburgh compound B (PiB) PET was not considered because it was included as an "add-on" substudy in the ADNI-1.[14] The PiB-PET substudy was started towards the end of the first year of the ADNI-1 study, and most subjects did not have a PiB scan at baseline but rather, at month 12 or 24. This analysis aimed to assess the impact of baseline covariates on disease progression parameters in mild AD and LMCI, and since PiB information was missing at baseline for most subjects, it could not be considered here. Moreover, it has been shown that CSF $A\beta_{1-42}$ and PiB-PET, when used as categorical covariates to

classify subjects as biomarker-positive or -negative, exhibit almost perfect agreement for the dichotomous categorization (91% agreement).[14] PiB-PET does not offer any substantial information as a covariate when compared with CSF $A\beta_{1-42}$ and was therefore not considered further.

The focus was on both mild AD and LMCI because these two states may actually represent a disease continuum. During the 3-year follow up of the ADNI-1 study, around 40% of LMCI subjects converted to AD,[15] and therefore, this analysis focused on a joint analysis of LMCI and mild-AD subjects. The other advantage of utilizing the scores from both mild-AD and LMCI subjects was that it provides a wider dynamic range for the disease scores. This allows the model-based approach to check the hypothesis regarding the underlying progression process (linear versus [vs] nonlinear and saturable) over the course of the disease.

The model building proceeded as follows to fulfill the objectives the analysis highlighted above: a) model disease progression, taking into account baseline disease status and change in cognition as a function of time; b) assess whether disease progression is nonlinear, and check the MCI/AD continuum assumption ie, assess whether LMCI and mild AD reside at different places on the disease progression curve and after accounting for disease severity, assess whether these two populations have the same disease progression rate parameter; c) identify a subset of LMCI and mild AD subjects that carry the underlying AD pathology and have a higher risk of disease progression, using biomarkers; d) quantify the different sources of variability in disease progression, such as demographics, ApoE4 carrier status, and an expanded set of covariates that affect disease progression parameters; and e) assess missing data mechanisms. The model-building process is described in detail in the "Methods" section; an overview of the process is provided in a flow diagram in Figure S1.

Methods

Study details

The ADNI study is registered at ClinicalTrials.gov.[16] The data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI was launched in 2003, by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the FDA, private pharmaceutical companies, and nonprofit organizations, as a $60 million, 5-year public–private partnership. The primary goal of ADNI has been to test whether serial MRI, PET, biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials.

The principal investigator of this initiative is Michael W Weiner, VA Medical Center and University of California, San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and reflects the findings of subjects recruited from over 50 sites across the US and Canada. The initial goal, of ADNI-1, was to recruit just 800 subjects, but ADNI-1 has been followed by ADNI-Grand Opportunity (GO) and ADNI-2. To date, these three protocols have recruited over 1,500 adults, ages 55 to 90, to participate in the research, consisting of cognitively normal (CN) older individuals, people with early or late MCI, and people with early AD. The follow-up duration of each group is specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO (http://www.adni-info.org/Scientists/ADNIStudyProcedures.aspx). Subjects originally recruited for ADNI-1 and ADNI-GO had the option to be followed in ADNI-2. For up-to-date information, interested readers can refer to http://www.adni-info.org.

All AD subjects had clinical/neuropsychological assessments and MRI measurements. AD subjects were assessed at 0, 0.5, 1, and 2 years, while MCI subjects were assessed at 0, 0.5, 1, 1.5, 2, and 3 years. ADNI allows public access to all accumulating data. The data set utilized in the current analysis was downloaded on Apr 11, 2013 from the ADNI-1 database. The unique IDs (roster IDs) for 415 subjects whose data were used in the preparation of this article can be found in data file UPENNBIOMK.csv (one screening-failed subject [roster ID =975] in this file, was not included in the analysis[17]). The CDR–SB data set consisted of 301 AD and LMCI subjects recruited in the ADNI-1 study. CDR–SB data from elderly control CN subjects were not employed in the disease progression model because 80% of CDR–SB observations in CN subjects were zeros. Thus the majority of the CDR–SB data from CN subjects do not carry information about disease progression due to this ceiling effect. However, the baseline biomarker data from 114 CN subjects was utilized to determine the threshold for abnormal biomarker values. The availability of the individual-level data made it possible to evaluate covariate effects on progression rate and baseline disease score, along with an assessment of interindividual variability (IIV) in the disease progression model.

Computer software

Data set preparation was performed using SAS® 9.2 (SAS Institute Inc., Cary, NC, USA). Data set exploration and

visualization were performed using S Plus® 6.0 Professional Edition, Release 2 software (Insightful Corp, Seattle, WA, USA) and R 2.15.1 (R Foundation for Statistical Computing, Vienna, Austria). CDR–SB and biomarker data were used for nonlinear mixed-effects modeling using NONMEM® version 7.1 or higher (ICON Development Solutions, Hanover, Maryland, USA).[18] The covariate submodel was established using the stepwise covariate modeling (SCM) procedure in Perl-speaks-NONMEM (PsN; Department of Pharmaceutical Biosciences Uppsala University, Sweden) 3.4.2.[19] The estimation for the missing data analysis was performed using the generalized linear model (glm) function in R version 2.15.1.

Selection of the structural model with additive residual error

A sequence of models was tested and the results compared, in order to select the best model. Ten possible disease progression structural models were assessed. In these models R(t) referred to CDR–SB response as a function of time; R_0 characterized baseline disease score or severity; α was the progression rate parameter; β was a shape factor that also controls the inflection point; bCDRSB was the baseline CDR–SB score; θ, θ_1, and θ_2 were fixed-effects parameters depicting dependency of α on bCDRSB, and the subscript "i" referred to the "i"th subject. All random effects in the model (including residual variability) were additive in nature except the R_0 IIV, which was described using an exponential error model, to prevent individual baseline scores from trending negative. The ten possible disease progression models were the following:

- Linear progression without shape parameter[20]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (1)$$

- Linear progression with shape parameter θ[20]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (2a)$$

$$\alpha_i = \alpha\left(\frac{bCDRSB_i}{2} \cdot \frac{18 - bCDRSB_i}{16}\right)^{\theta} \quad (2b)$$

- Linear progression with shape parameter θ_1[21]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (3a)$$

$$\alpha_i = \alpha\left(\frac{bCDRSB_i}{2}\right)^{\theta_1}\left(\frac{18 - bCDRSB_i}{16}\right) \quad (3b)$$

- Linear progression with shape parameter θ_2[21]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (4a)$$

$$\alpha_i = \alpha\left(\frac{bCDRSB_i}{2}\right)\left(\frac{18 - bCDRSB_i}{16}\right)^{\theta_2} \quad (4b)$$

- Linear progression with shape parameters θ_1 and θ_2[21]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (5a)$$

$$\alpha_i = \alpha\left(\frac{bCDRSB_i}{2}\right)^{\theta_1}\left(\frac{18 - bCDRSB_i}{16}\right)^{\theta_2} \quad (5b)$$

- Linear progression with α_i dependent on $bCDRSB_i$ via a power relationship[20]

$$R_i(t) = R_{0i} + \alpha_i \cdot t \quad (6a)$$

$$\alpha_i = \alpha\left(\frac{bCDRSB_i}{2}\right)^{\theta} \quad (6b)$$

- Exponential progression[22]

$$R_i(t) = R_{0i} \cdot \exp(\alpha_i \cdot t) \quad (7)$$

- Verhulst logistic model[23,24]

$$R_i(t) = \frac{18 \cdot R_{0i}}{(18 - R_{0i})e^{-\alpha i t} + R_{0i}} \quad (8)$$

This is the explicit solution of the differential equation: $dR_i/dt = \alpha_i \cdot R_i \cdot (1 - [R_i/18])$; $R_i(0) = R_{0i}$

- Richard's logistic model[24,25]

$$R_i(t) = \frac{18 \cdot R_{0i}}{\left[(18^{\beta} - R_{0i}^{\beta})e^{-\alpha i \beta t} + R_{0i}^{\beta}\right]^{\frac{1}{\beta}}} \quad (9)$$

This is the explicit solution of the differential equation: $dR_i/dt = \alpha_i \cdot R_i \cdot (1 - [R_i/18]^{\beta})$; $R_i(0) = R_{0i}$

- Three-parameter logistic model (this equation does not have an explicit solution)[24,25]

$$\frac{dR_i}{dt} = \alpha_i \cdot R_i^{\beta}\left[1 - \frac{R_i}{18}\right]; \; R_i(0) = R_{0i} \quad (10)$$

These ten structural models were chosen because they are used in the AD literature to describe disease progression of related end points, such as the ADAS-cog/11, functional activities questionnaire, and Mini-Mental State Exam (MMSE).[20–26] The simplest linear progression model (Equation 1) was initially chosen for model fitting. This was followed by a series of linear models (Equations 2–5), with slope dependent on bCDRSB, to characterize the inflection point characteristic of disease progression ie, the disease progression is characterized by an initially increasing rate of progression

with increasing scores, followed by slowing of the disease progression rate as scores approach an asymptote of 18. In addition, a model with a power relationship between α and bCDRSB (Equation 6) and the exponential progression model (Equation 7) were also tested. Finally, logistic models that represent nonlinear/saturable disease progression (Equations 8–10) were also assessed. The mathematical relationships tested for the structural model were often nonnested and had the same number of parameters. Therefore, the selection of the best structural model was guided by Akaike information criteria (AIC), which is equal to the NONMEM® objective function value (OFV) plus twice the number of parameters in a given model.[24,25] A few other criteria that were used to compare different candidate models were a) decrease in the residual error; b) precision of the parameter estimates; c) goodness-of-fit diagnostics, including eta [η] shrinkage; d) ill conditioning and overparameterization, which was checked by inspecting the eigenvalues of the covariance matrix – ideally, the ratio of the largest eigenvalue to the smallest eigenvalue, referred to as the condition number, was not to exceed 1,000; and e) biological plausibility and prior knowledge about the structural model. Recent work indicates that disease progression for various AD end points follows a nonlinear and saturable trajectory,[24,25] suggesting that Equations 9 and 10 may exhibit better performance as compared with Equations 1–8.

Beta regression

Since CDR–SB is a bounded outcome, the function for disease progression was nonlinear, which was accounted for by models 2–10 in the previous section, either through the structural component or the fixed effects. However, these models assume that the residual error is normally distributed (additive residual error model). The additive model for residual variability does not account for the fact that the underlying distribution is heteroskedastic because with time, the variance in the CDR–SB scores decreases (approaches 0) as the scores approach the upper or lower boundaries of the scale.

Recently, the use of beta-distributed residuals was proposed (as part of the qualified ADAS-cog/11 quantitative clinical trial simulation model)[8–10] to constrain predictions to the range of the original instrument and to model the heteroskedasticity. The limitation of the standard beta regression is that the data should be contained within a certain interval with no data at the boundaries. A survey of the CDR–SB data from ADNI-1 LMCI and mild AD subjects suggested that <2% of the data resided at the boundaries; these data were not used for beta regression initially until a sensitivity analysis was performed for inclusion of boundary data.

To compare the model with normally distributed error versus a beta regression model, the model with the lowest AIC from the previous section was rerun by deleting the data at the boundaries. In addition, because the models with normally distributed residual errors versus beta residuals used different probability distributions, the method of complete likelihood functions (see the following section, "Optimized residual error structure") was used to compare the appropriate model from the previous section versus beta regression. Five different beta regression models were tested that had the property of producing individual predictions within the boundaries, which was a requirement for beta regression. These models were compared against each other using AIC values, η shrinkage, and condition numbers. The models assume that the scaled response (CDRSB/18) follows a beta distribution ie, $R(t)/18 \sim Beta(\mu\tau,(1-\mu)\tau)$; where μ is the scaled conditional expectation resulting from the structural model [restricted within the interval (0–1)] and τ is the precision parameter.

- Linear progression on the logit scale:[27]

$$\mu(t,\ r_{0i},\ \alpha_i) = \frac{\exp(r_{0i} + \alpha_i\, t)}{1 + \exp(r_{0i} + \alpha_i\, t)} \tag{11}$$

- Linear progression on the logit scale with $bCDRSB_i$ covariate on progression rate:[10]

$$\mu(t,\ r_{0i},\ \alpha_i) = \frac{\exp(r_{0i} + \alpha_i t)}{1 + \exp(r_{0i} + \alpha_i t)} \tag{12a}$$

$$\alpha_i = \alpha + \theta \cdot (bCDRSB_i - 2) \tag{12b}$$

- Verhulst logistic model:[23,24]

$$\mu(t, R_{0i}, \alpha_i) = \frac{R_{0i}}{(18 - R_{0i})e^{-\alpha i t} + R_{0i}} \tag{13}$$

- Richard's logistic model:[24,25]

$$\mu(t, R_{0i}, \alpha_i, \beta) = \frac{R_{0i}}{\left[(18^{\beta} - R_{0i}^{\beta})e^{-\alpha i \beta t} + R_{0i}^{\beta}\right]^{\frac{1}{\beta}}} \tag{14}$$

- Three-parameter logistic model:[24,25]

$$\frac{dR_i}{dt} = \alpha_i \cdot R_i^{\beta}\left[1 - \frac{R_i}{18}\right];\ R_i(0) = R_{0i} \tag{15a}$$

$$\mu(t, R_{0i}, \alpha_i, \beta) = \frac{R_i}{18} \tag{15b}$$

Optimized residual error structure

Published AD disease progression models have used different types of residual error structures for disease progression

modeling.[10,20–22,24,25] Therefore, models that were based on different probability distributions ie, normal, log-normal, logit-normal, and beta residuals were compared. Log-normal, logit-normal, and beta residuals are not capable of handling data at the lower and/or upper boundaries, and hence these data were not considered here. The complete log-likelihood function ($L = \sum_{i}^{n} l_i$) for each distribution is shown below, including all component parts and normalizing constants arising from the complete probability distributions.[28] The model with the lowest value for minus twice the log likelihood (–2LL) was chosen for further development. In the equation below, σ is the standard deviation, y_i are the scaled observed data, and μ_i is the scaled prediction [to the interval (0–1)] from the model:

- Normal distribution:[28]

$$l_i = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{(y_i - \mu_i)^2}{2\sigma^2} \quad (16)$$

- Log-normal distribution:[28]

$$l_i = \log\left(\frac{1}{y_i}\right) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{(\log(y_i) - \log(\mu_i))^2}{2\sigma^2} \quad (17)$$

- Logit-normal distribution:[29]

$$l_i = \log\left(\frac{1}{y_i(1-y_i)}\right) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{(\mathrm{logit}(y_i) - \mathrm{logit}(\mu_i))^2}{2\sigma^2} \quad (18)$$

- Beta distribution:[30]

$$l_i = \log\Gamma(\tau) - \log\Gamma(\mu_i\tau) - \log\Gamma((1-\mu_i)\tau) + (\mu_i\tau - 1)\log(y_i) + ((1-\mu_i)\tau - 1)\log(1-y_i) \quad (19)$$

Handling boundary data

Some of the residual error models assumed that the dependent variable was continuous in nature and that the data resided within the boundaries of the scale. If log-normal, logit-normal, or beta residuals were found to be a more suitable distribution, then the boundary data were also utilized in the subsequent analysis – all the data were rescaled by introducing a small noise (δ). This δ moved the observations slightly within the edges and rescaled all the data to lie within the boundaries. It has been shown that beta regression models that use this simple and practical approach are less sensitive to data at the boundaries and the choice of δ value. The scaling method reported by Verkuilen and Smithson[31] was implemented using the equation: $y_{new} = y_{original} \cdot (1–\delta) + (\delta/2)$, where $y_{original}$ is the score transformed from the original scale to the unit interval [0, 1] by first normalizing CDR–SB (CDRSB/18). The appropriateness of this choice of δ value was evaluated by assessing the change in parameter estimates with different δ values for the best structural model/residual variability submodel combination (which will be called the base model from here onwards).

Biomarker mixture models and use of optimal baseline biomarker as covariates

In order to establish thresholds for dichotomizing subjects, at baseline, into progressors and nonprogressors, mixture models were developed for each of the biomarkers. A mixture of two normal distributions was fitted separately for each of the baseline biomarker data. These baseline biomarker data were pooled from CN, LMCI, and mild AD subjects and included age and head size–adjusted average hippocampal volume, the CSF biomarkers $A\beta_{1–42}$, total tau protein (t-tau), tau phosphorylated at position threonine 181 (p-tau_{181P}), the t-tau/$A\beta_{1–42}$ ratio, and the p-tau_{181P}/$A\beta_{1–42}$ ratio. Age and head size–adjusted average hippocampal volume was computed using the covariance method, as described by Jack et al.[32] The distribution of tau-related markers was skewed to the right, and CSF t-tau, p-tau_{181P}, t-tau/$A\beta_{1–42}$ ratio, and p-tau_{181P}/$A\beta_{1–42}$ ratio were log-transformed before analysis, to reduce skewness. The models involved fitting the probability density function for the mixture of two normal components, which is as follows: $f(x) = \xi \cdot f(x;\mu_1,\sigma_1) + (1\text{-}\xi)\cdot f(x;\mu_2,\sigma_2)$, where μ and σ are the mean and standard deviations of the two distributions, while ξ and 1–ξ are the mixing proportions. In addition, mixing proportions in the subject groups clinically diagnosed as CN or LMCI or AD was assessed by testing the disease group label as a covariate on the ξ parameter. A threshold for normal vs abnormal biomarker levels was computed as the crossover point between the two normal distributions for the overall data set.

The first step of covariate analysis in the disease progression model was driven by prior knowledge about covariate effects on disease progression.[25] The early AD population is characterized by heterogeneity in the rate of disease progression. The ability to predict the likelihood of disease progression can be guided by the bimodal pattern of key baseline biomarkers. Previous work has shown that pathologic levels of baseline AD biomarkers quantitatively distinguish progressors from nonprogressors.[25] Therefore, based on the estimated threshold from the mixture model, the baseline biomarkers were also dichotomized and tested as categorical covariates

on the slope and baseline severity parameters in the base model of CDR–SB progression. The use of biomarkers as dichotomous predictors, as opposed to a continuous predictor, provides a useful way of selecting subjects for potential enrolment in a clinical study. However, for completeness, the optimal biomarker was also tested as a continuous covariate. The choice of the biomarker and its functional form, ie, continuous vs dichotomous parameterization for covariate modeling, was based on AIC. The model with the optimal biomarker was referred to as the base reference model.

Covariate screening

A list of 37 covariates was considered in this analysis to find additional covariates that could influence disease progression. These covariates were of five types: a) hippocampal volume; b) serum biomarkers; c) demographics and number of ApoE4 alleles; d) cognitive tests at baseline; and e) disease status (MCI vs AD) and AD comedication use. Since there was an insufficient number (<4%) of subjects on memantine alone at baseline, comedication use was categorized into three groups: a) subjects taking no AD comedications; b) subjects taking one AD comedication (either acetylcholinesterase inhibitor or memantine); and c) subjects taking two AD comedications (acetylcholinesterase inhibitor and memantine).

The demographic covariates that were tested included years of education, body mass index, age, sex, and family history of dementia. The serum biomarkers included total cholesterol, triglycerides, vitamin B12, mean corpuscular volume, fasting glucose, homocysteine, 8-iso prostaglandin F2 alpha, and 8,12-iso-iPF2alpha isoprostanes. Finally, the baseline cognitive tests included the American National Adult Reading test, Auditory Verbal Learning recognition, immediate and delayed recall, the Boston Naming Test, Category Fluency test (animal names and vegetable names), Clock-Drawing Test, Digit Symbol Substitution test, Digit Span forwards and backwards, Functional Activities Questionnaire, Logical Memory (delayed and immediate), MMSE, the Neuropsychiatric Inventory, ADAS-cog/13, Trail Making Test parts A and B, and bCDRSB. Cognitive test scores were not tested as covariates on bCDRSB score since the objective of this analysis was not to find whether CDR–SB is correlated with other scales. However, baseline cognitive test scores were tested as covariates on the progression rate parameter, to assess whether baseline performance on other cognitive and functional measures could be predictive of future decline on the CDR–SB scale.

This second stage of the covariate search involved model building that was performed using PsN in a stepwise manner, starting with the base reference model. Covariate effects were parameterized as follows: a) categorical covariates were coded using a linear function (with one added parameter for each level above the reference category); b) continuous covariates were generally parameterized using the power function; c) continuous covariates that could have a value of zero (eg, certain neuropsychological tests) were modeled using the exponential function since a power function is undefined for such a covariate; d) continuous covariates were centered on the median value ([covariate–median] for the exponential function and [covariate/median] for the power function); and e) the effect of different covariates on each parameter was described using a multiplicative relationship. During the step-up process of covariate model building, the relationship that gave the largest improvement in OFV was retained in the model, given that inclusion resulted in an OFV decrease of >3.84 points ($P<0.05$). The full model was derived when no more covariates could be added according to the prespecified criterion. The included covariates were then left out of the full model one at a time during backward elimination and tested using the stricter criterion of OFV change of >6.63 ($P<0.01$).

Missing data

Missing at random (MAR) was assumed as the mechanism for missing data, because likelihood-based mixed-effects modeling was conducted. This assumption is consistent with the primary statisticalanalysis generally used for clinical trials, where the mixed-effect model with repeated measures is commonly employed.[33] The primary goal of the missing data analysis was to explore the missing data mechanism (ie, why do certain subjects drop out from clinical trials). Also, the missing data analysis provided a dropout probability to facilitate visual predictive check (VPC) simulations for data with dropouts. For ADNI-1, scores were measured at prescheduled times (ie, at years 0, 0.5, 1, 1.5, 2, and 3). The grouped-time survival model[34] was a natural choice to evaluate the missing data for this type of study design as follows:

$$\begin{aligned}&\log\left(-\log(1-\mathrm{P}[\mathrm{Dropout}=1])\right)\\&\quad=[\alpha_1 D_{0-0.5}+\alpha_2 D_{0.5-1}+\alpha_3 D_{1-1.5}+\alpha_4 D_{1.5-2}\\&\qquad+\alpha_5 D_{2-3}]+\beta_1 PD_{ij-1}+\beta_2(\mathrm{bAge}_i-75)\end{aligned}\tag{20}$$

where $\alpha_1 \ldots \alpha_5$ are intercept parameters that characterize the baseline hazards at different time periods (eg, 0 to 0.5 year) and PD_{ij-1} is the pharmacodynamic CDR–SB score for the "i"th subject at the "j–1"th time period, ie, the relationship between the probability of dropout and the CDR–SB score prior to dropout was evaluated. The D (eg, $D_{0-0.5}$) terms

represent a dummy variable coded as "1" if the observation represented the corresponding interval (eg, 0 to 0.5 year), and otherwise coded as 0. Previous analysis of dropout data in AD studies has indicated that baseline age is a predictor of subject dropout.[22] The influence of individual baseline age_i (centered on the median baseline age of 75 years) on subject dropout was therefore tested. The model employs a complementary log–log link, which has the practical application of model parameters being interpreted in terms of hazard ratios.[34]

Model qualification

Once the disease progression, covariates, and dropout submodels were finalized, model qualification was carried out by a) the use of diagnostic plots; b) checking the plausibility of the model parameter values; and c) further evaluating the model using VPCs. VPC, in the presence of missing data,[35] was performed as follows: The observed data was grouped at prescheduled sampling times and compared with the simulated data at the respective time points. The event of missing data (dropout) was simulated based on the conditional probability at each time point, assuming a binomial distribution for the missing event. The conditional probability was calculated based on the baseline hazards estimated in the dropout model and the simulated CDR–SB scores at the previous visit. A percentile VPC with confidence intervals[35] was used to assess the extent to which the median prediction and extremes of the 90% prediction interval reflected the median and the fifth and 95th percentiles of the observed data.

Results

Subject characteristics

The characteristics of the ADNI-1 subjects have been published extensively.[20,21,24,25,36] Briefly, the mild AD subjects with CSF information (102 subjects with 383 CDR–SB observations) were between ages of 57 to 89 years (mean ± standard deviation [SD]: 75±8 years), and the LMCI subjects with CSF biomarker data (n=199, 1,219 CDR–SB observations) were between 55 to 89 years (74±8 years). The subjects were well educated, with an average of 15 years (±3 years) of education for the mild AD subjects and 16 years (±3 years) for LMCI subjects. A total of 47% and 49% of mild AD and LMCI subjects, respectively, had a family history of dementia with at least one parent having the disease. A total of 70% of mild AD subjects and 54% of LMCI subjects, respectively, were ApoE4 carriers; this confirmed that ApoE4 is a significant risk factor for AD.[24,25] Among all subjects, 58% of the mild AD subjects and 67% of LMCI subjects were male. In all, 93% of the mild AD subjects took AD comedications, with 40% taking acetylcholinesterase inhibitors and memantine, and 53% taking acetylcholinesterase inhibitors or memantine. Similarly, 55% of the LMCI subjects took AD comedications, with 12% taking two AD comedications and 43% taking at least one AD comedication. The mean baseline MMSE and CDR–SB scores were 24±2 and 4.3±1.6, respectively for mild AD subjects, while the mean baseline MMSE and CDR–SB scores were 27±2 and 1.6±0.9, respectively for LMCI subjects. While the subjects with LMCI and mild AD subjects both had memory complaints, on the MMSE, the range for LMCI and mild AD subjects was 24–30 and 20–26, respectively. The CDR global score for LMCI was 0.5, with a memory box score of $\geq$0.5, while the CDR global score for mild AD subjects was 0.5 or 1. The subjects with LMCI were generally intact with regard to cognitive and functional performance and did not qualify for dementia diagnosis.[36] The mild AD subjects met the National Institute of Neurological and Communicative Disorders and Stroke-AD and Related Disorders Association criteria for probable AD.[36]

Development of the disease progression structural model with additive residual error

A simple diagnostic plot, based on the 12-month interval method, was constructed to decipher how CDR–SB progresses over time (ie, linear vs nonlinear progression). The 12-month interval method involves computing the change in CDR–SB between pairs of visits separated by 1 year, by subtracting the earlier CDR–SB measurement from the later one. Since most subjects have more than one pair of CDR–SB measurements, individuals contributed more than one value for this diagnostic plot. The results of this plot are shown in Figure S2, which indicates that the rate of progression was slower at low and high CDR–SB scores, with intermediate scores of about 10 exhibiting faster rates of deterioration. This inverted U-shaped relationship for progression rate vs baseline score is a characteristic of an S-shaped progression curve, often described using the logistic models.[23–26] The expectation that a logistic model could describe the CDR–SB progression curve was realized when the different structural models were compared against one another using AIC values (Table 1).

The structural models that were tested had two key characteristics: a) utilized all the data, including observations at the boundaries of the CDR–SB scale; and b) used a simple additive residual error model. The structural model

Table 1 Summary of structural models

Model description	Akaike information criterion
Linear progression	2,883.8
Linear progression with shape parameter θ	2,814.1
Linear progression with shape parameter θ_1	2,814.3
Linear progression with shape parameter θ_2	2,814.8
Linear progression with two shape parameters	2,814.8
Linear progression with power relationship	2,812.8
Exponential progression	2,551.4
Verhulst logistic model	2,571.5
Richard's logistic model	2,496.7
Three-parameter logistic model	2,496.2

that had the largest AIC value was the two-parameter linear progression model (Equation 1). This model places a strong constraint on the progression rate being constant across the entire range of CDR–SB scores, and this restriction could lead to poor model performance.

The linear progression models, with the various shape parameters (Equations 2–6), performed quite similarly. These models are more flexible than the two-parameter linear progression model since the progression rate can vary across subjects based on their bCDRSB score. However, the progression rate within a subject stays constant across the 2- to 3-year study duration. This constraint could explain why the nonlinear progression models (exponential progression and logistic models) perform better than the linear progression models, in terms of the AIC (Table 1).

The exponential progression model (Equation 7) and the logistic model (Equation 8), which both have only two structural parameters, did not do as well as the Richard's function (Equation 9) and the three-parameter logistic model (Equation 10). The three-parameter logistic model had the best performance for CDR–SB, in terms of AIC.

Comparison of models employing beta regression

Five different models employing this technique were tested. The AIC values for models represented by Equations 11 to 15 were –4,054, –4,080, –4,050, –4,089, and –4,094. Equations 11 and 12 assumed that the CDR–SB scores progress linearly on the logit scale. Among these two models, Equation 12, which allowed the progression rate to depend on baseline disease status, performed better (ie, had a lower AIC), and this is expected based on both statistical and biological considerations. However, the three-parameter logistic model performed the best, with the lowest AIC among all five models. Thus, the same equation (Equation 10 and 15) was chosen for the structural model with additive residual error and beta regression. This suggests that the preservation of the inverted U-shaped relationship between progression rate and disease status is an important feature that should be incorporated when analyzing the CDR–SB data.

Selection of regular model vs model with beta regression

The choice between the regular parameterization with additive residual error versus beta regression is limited to the selection of only the residual error component. To compare the three-parameter logistic model with normally distributed residual error versus beta residuals, the model with additive residual error was rerun by deleting the data at the boundaries of the scale. These models along with two more models comparing different residual error parameterization are presented next.

Optimal residual error model

A comparison of the –2LL values for models represented by Equation 16–19 is as follows: –3,743, –4,002, –4,121, and –4,102. The residual error parameterization that gave the lowest –2LL value was the logit-normal distribution. The three-parameter logistic model with logit-normal residuals was considered as the base model. However, before proceeding to covariate model building, a sensitivity analysis was performed for the selection of a δ value that would allow inclusion of boundary data.

Boundary data handling: sensitivity analysis

The results of the sensitivity analysis for handling of boundary data using logit-normal residuals with the three-parameter logistic model are reported in Table S1. The reference set of parameter estimates is based on the model estimation, without rescaling data and excluding data at the boundaries (a total of 30 observations). For the sensitivity analysis, subsequent model estimations were performed with data sets where the observations have been rescaled by a small amount δ, where δ varied from 0.001 to 0.33. Sets of parameter estimates for each tested value of δ were compared with the reference parameter set. The choice of the optimal δ was based on two criteria: a) the set of parameter estimates for a given value of δ should be as close as possible to the reference set; b) the set of parameter estimates for a given value of δ should be estimated with reasonable precision ie, the percent coefficient of variation (CV) for imprecision should be close to the reference set. Both these criteria were

met for a δ value of 0.01. This is also the δ value recommended by Verkuilen and Smithson (2012).[31] Thus, for all further testing of the model, all the data were used after rescaling with a δ value of 0.01.

Biomarker mixture modeling and base reference disease progression model

The estimate of the IIV for the slope parameter in the base model was close to 100% (Table S1). To understand the basis of this heterogeneity, the baseline CSF biomarker values were subjected to mixture modeling. The aim was to assess whether the large variability in the progression rate could be explained by the dichotomy in the biomarker distributions, as all baseline CSF and volumetric biomarkers exhibited bimodality. The parameter estimates of the mixture models for the various biomarkers are displayed in Table S2.

As a vignette, the results of the mixture model with one of the CSF biomarkers (log p-tau$_{181p}$/Aβ$_{1–42}$) is shown in Figure 1. Figure 1A shows the baseline distribution of CSF p-tau$_{181p}$/Aβ$_{1–42}$ pooled from CN, LMCI, and mild AD subjects on a log scale as a histogram that exhibits a clear bimodal pattern. Overlaid on the observed distribution are the density curves based on the parameter estimates of the mixture model and the estimated threshold of 0.147. The relevance of this threshold to all three patient types (CN vs LMCI vs AD) is shown in Figure 1B. The observed density (kernel density estimate) ie, a kind of moving average of the distribution of CN, LMCI, and AD subjects, is overlaid on top of the estimated threshold for log p-tau$_{181p}$/Aβ$_{1–42}$. The CN and AD distributions are quite distinct and are characterized by low and high CSF p-tau$_{181p}$/Aβ$_{1–42}$ values, respectively. In contrast, the LMCI subjects represent a mixture, some with AD pathology and some without. This is reflected by the impact of the disease label on the mixing fraction, which increases from 20% to 68% to 91% for CN, LMCI, and AD subjects, respectively, for CSF p-tau$_{181p}$/Aβ$_{1–42}$ (Table S2). The impact of the estimated CSF

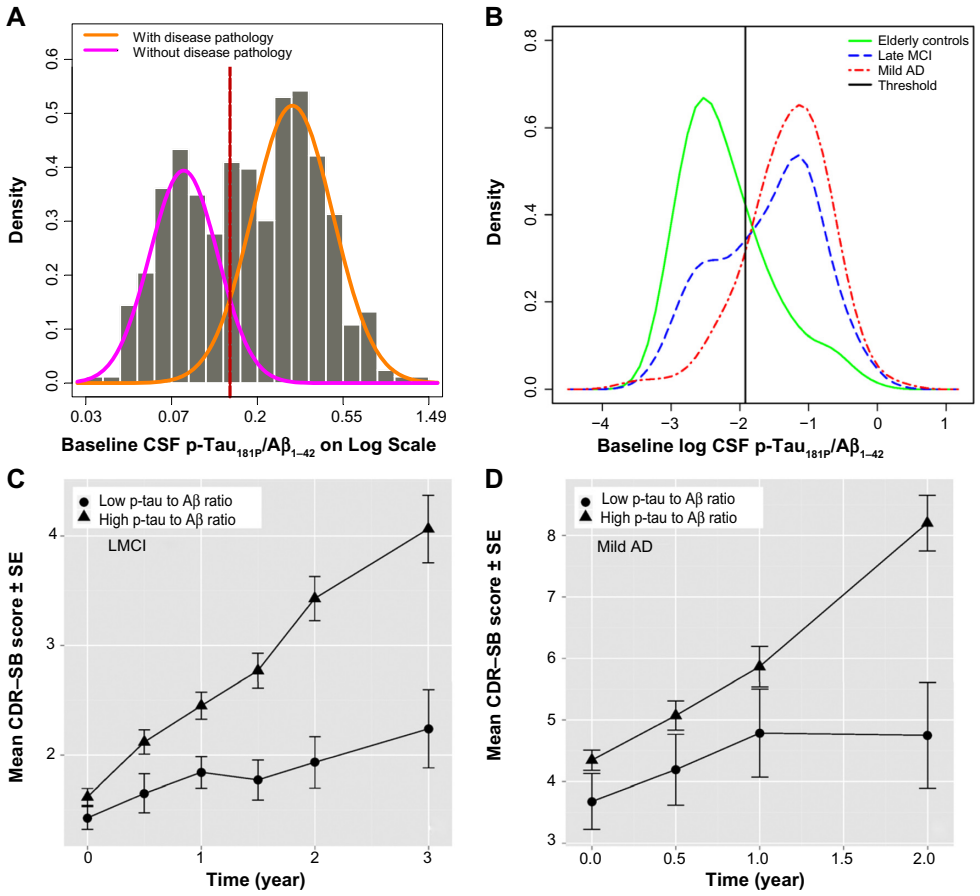

Figure 1 (**A**) Distribution of CSF p-tau$_{181P}$/Aβ$_{1–42}$ on log scale in CN, LMCI, and mild AD subjects. The curves and the vertical line are the density of the subpopulations and threshold of 0.147 (untransformed value), based on the CSF mixture model. (**B**) The observed density for log CSF p-tau$_{181P}$/Aβ$_{1–42}$ in CN, LMCI, and AD subjects is overlaid on top of the estimated threshold of –1.92 (natural log transformed value). (**C**) Mean CDR–SB profiles by biomarker status, in LMCI subjects. (**D**) Mean CDR–SB profiles by biomarker status, in mild AD subjects.

Note: The CSF p-tau$_{181P}$/Aβ$_{1–42}$ threshold of 0.147 separating the two subgroups was used to dichotomize the population for Panels (**C**) and (**D**).

Abbreviations: Aβ$_{1–42}$, 42 amino acid isoform of amyloid beta peptide; AD, Alzheimer's disease; CN, cognitively normal; CSF, cerebrospinal fluid; LMCI, late mild cognitive impairment; p-tau$_{181P}$, tau protein phosphorylated at position threonine 181; SE, standard error.

p-tau$_{181p}$/A$\beta_{1–42}$ threshold on mean CDR–SB progression curves in LMCI (Figure 1C) and mild AD (Figure 1D) is quite marked. Subjects with baseline CSF p-tau$_{181p}$/A$\beta_{1–42}$ below the 0.147 threshold exhibit a much slower rate of progression than those subjects above the threshold.

All six biomarkers were dichotomized as categorical covariates based on the threshold from the mixture models. They were introduced back into the CDR–SB disease progression model as covariates on bCDRSB and progression rate parameters and were tested to check which biomarker gave the best improvement in model performance. The AIC values for the CDR–SB progression model, in improving order of model performance, for the six biomarkers were: a) CSF t-tau =–3,879.0; b) CSF p-tau$_{181P}$ =–3,879.7; c) hippocampal volume =–3,902.1; d) CSF t-tau/A$\beta_{1–42}$ ratio =–3,910.8; e) CSF A$\beta_{1–42}$ =–3,919.6; and f) CSF p-tau$_{181P}$/A$\beta_{1–42}$ =–3,921.1. Since CSF p-tau$_{181P}$/A$\beta_{1–42}$ gave the lowest AIC value, it was also tested as a continuous covariate, and the various continuous parameterizations worsened the AIC by 8–35 points, suggesting that the biomarker threshold may function as a cutoff above which disease progression becomes discernible.

Two additional assumptions were tested. The first assumption was whether the progressors with the slower progression rate (CSF p-tau$_{181P}$/A$\beta_{1–42}$ ≤0.147) represented nonprogressors. To test this assumption, the typical progression rate parameter for these subjects was fixed to zero. This led to an increase in the minimum value of the OFV by only 1.7 point (P=0.19), suggesting that simplification does not affect model performance. Secondly, a separate residual error parameter was tested for progressors vs nonprogressors, and it was found that the residual error for nonprogressors was greater than that for progressors (the OFV decreased by 34 points, which is highly significant). This suggests that studying MCI patients with pathologic biomarkers has the potential to reduce intraindividual residual variability. In conclusion, a comparison between the base model (no covariates) and a model with CSF p-tau$_{181P}$/A$\beta_{1–42}$ as a covariate involved the addition of two parameters that decreased the OFV by 106 points (P=9.6 · 10^{-24}), a huge improvement that was guided by prior knowledge about the disease biology. Thus, this model with the optimal CSF p-tau$_{181P}$/A$\beta_{1–42}$ biomarker was accepted as the base reference model.

Influence of additional covariates on CDR–SB disease progression

Further covariate model building was performed via a forward addition/backward elimination procedure in NONMEM®. PsN was used to optimize and finalize the covariate model. The forward addition/backward elimination procedure employed the likelihood ratio test. A total 37 covariates were tested across 477 models, using the automated process in PsN. Covariates were identified both on the disease progression rate parameter and baseline disease score. The statistically significant covariates on the disease progression rate parameter were delayed logical memory (a measure of episodic memory) and Trail-Making Test part A (Trails A) (a measure of cognitive processing speed). The statistically significant covariates for bCDRSB score were baseline disease status (mild AD vs LMCI), hippocampal volume, and AD symptomatic comedication use. The impact of the statistically significant covariates on observed bCDRSB scores is summarized in Table S3. The bCDRSB was more than twofold higher in mild AD subjects as compared with LMCI subjects. Higher bCDRSB scores were associated with increase in the use of AD comedications (from zero to one to two drugs). Hippocampal volume has been previously identified as an influential covariate on baseline severity[24,25] and was significant in the current analysis as well, ie, smaller hippocampal volume was associated with higher bCDRSB score.

To allow visualization of important covariate effects on progression rate, some simple diagnostic plots were created (Figure 2). For these plots, the important covariates were dichotomized (> median and ≤ median) to create roughly equal groups in the AD and LMCI cohorts, and the mean CDR–SB was plotted vs time as a function of this newly created categorical variable for subjects with pathologic CSF. The impact of delayed logical memory, which is a measure of episodic memory, on progression rate is clearly visible in the LMCI cohort in Figure 2A, where low baseline delayed logical memory score (poor episodic memory) is associated with faster progression rate. The impact of delayed logical memory in mild AD is difficult to discern due to a floor effect in episodic memory in this population (Figure 2B). A total of 56% of the mild AD subjects (vs 22% for LMCI) had zero delayed logical memory score at baseline. A longer completion time on the Trails A test was associated with faster progression for AD and LMCI subjects with pathologic CSF (Figure 2C and 2D), which indicates that patients with poor cognitive processing speed progress rapidly. However, the impact of Trails A test in the LMCI cohort becomes less apparent between year 2 and 3 (Figure 2C), probably due to missing data, which is described later.

Model refinement and disease progression parameter estimates

The model obtained after the covariate search indicated that the impact of CSF biomarker status on baseline disease

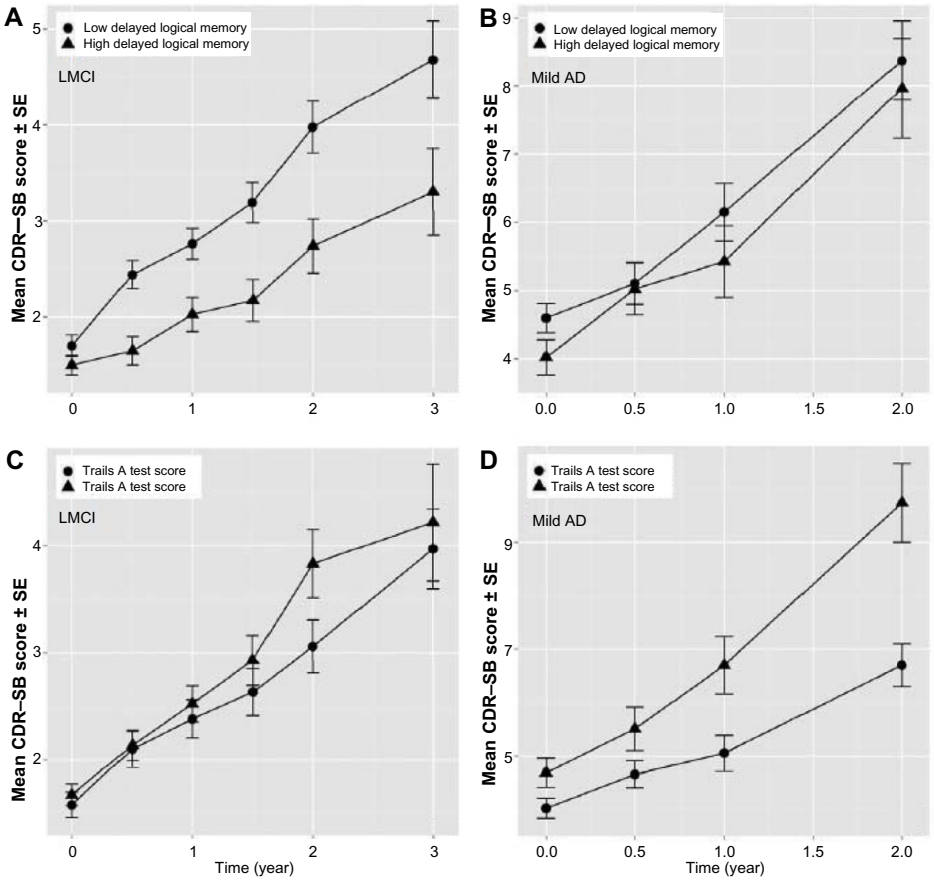

Figure 2 Baseline covariates influencing CDR–SB disease progression rate in subjects with CSF p-tau$_{181P}$/A$\beta_{1–42}$>0.147.

Notes: (**A**) delayed logical memory in LMCI; (**B**) delayed logical memory in mild AD; (**C**) Trails A test in LMCI; and (**D**) Trails A test in mild AD. Covariates were dichotomized (>median and ≤median) to create roughly two equal groups in each panel. Median delayed logical memory scores for the ADNI-1 mild AD and LMCI biomarker-positive population at baseline were 0 and 3, respectively. Median Trails A test scores for the ADNI-1 mild AD and LMCI biomarker-positive population at baseline were 57 and 43 seconds, respectively.

Abbreviations: A$\beta_{1–42}$, 42 amino acid isoform of amyloid beta peptide; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; CSF, cerebrospinal fluid; LMCI, late mild cognitive impairment; p-tau$_{181P}$, tau protein phosphorylated at position threonine 181; SE, standard error; Trails A test, Trail-Making Test part A.

severity was no longer statistically significant. Subjects who were biomarker-positive (CSF p-tau$_{181p}$/A$\beta_{1–42}$ ≥0.147) had a baseline score that was only 0.18% higher than those who were biomarker negative, and this parameter was estimated with high imprecision (3,706%CV). Thus the model was reduced for predictive purposes and for reasons of parsimony. Removal of this covariate effect could be justified because it did not impact coefficients for other remaining effects and because the increase in OFV was 0.002 (P=0.96, df=1). The parameter estimates of the final disease progression model are reported in Table 2. Based on Equation 10, parameter estimates in Table 2, and the baseline scores reported in Table S3, it was estimated that the CDR–SB disease progression rate in a typical patient with late MCI and mild AD was approximately 0.5 and 1.4 points/year, respectively. The estimated inflection point, where the progression rate is fastest, was found to be at a CDR–SB score of 10 (Table 2). The estimated η shrinkage for the progression rate and baseline parameters was 23.5% and 14.1% in the final model. The condition number was 127, suggesting that the final model was fairly stable.

Missing data mechanism

As the first step in the missing data analysis, an exploratory analysis was performed, where the average CDR–SB observed scores as a function of different study discontinuation times were plotted (Figure S3). These results suggested that subjects with higher CDR–SB scores tended to discontinue (have missing data) compared with subjects with lower CDR–SB scores. The results of this exploratory analysis were formally assessed in the missing data analysis, where the likelihood of a subject's data being missing was related to the CDR–SB score prior to the event. Subject's baseline age was also assessed as a covariate in the missing

Table 2 Final CDR–SB disease progression model parameters

Parameter[a]	Estimate	90% CI[b]
CDR–SB baseline score	1.53	1.40 to 1.66
LMCI population (θRo)		
Hippocampal volume (θRo_HVOL)	–0.916	–1.18 to –0.65
Mild AD (θRo_DIS)	1.22	1.02 to 1.42
Zero comedication (θRo_COMED_0)	–0.117	–0.22 to –0.02
Two comedications (θRo_COMED_2)	0.169	0.05 to 0.28
Progression rate parameter pathologic CSF population (θα)	0.306	0.23 to 0.39
Delayed logical memory (θα_DLM)	–0.057	–0.09 to –0.02
Trails A test (θα_TAT)	0.411	0.24 to 0.58
Shape parameter β[c]	1.22	1.02 to 1.42
IIV, progression rate parameter (%CV)[d]	62.2	39 to 86
IIV, baseline CDR–SB (%CV)[d]	39.0	35 to 43
Residual variability, SD (pathologic CSF population)	0.422	0.38 to 0.46
Residual variability, SD (nonpathologic CSF population)	0.551	0.48 to 0.62

Notes: [a]Relationships between covariates and the TV of final parameter estimates: TVRo=θRo·(HVOL/2954)^θRo_HVOL·(1+θRo_DIS)^DIS·(1+θRo_COMED_0)^C_0·(1+θRo_COMED_2)^C_2 and TV α = θα·CSFFLAG·(TAT/44)^θα_TAT·exp(θα_DLM·(DLM-2)), where; DIS is a 0/1 flag for LMCI/mild AD respectively; C_0 is a 0/1 flag for subjects taking vs not taking comedications respectively; C_2 is a 0/1 flag for subjects taking less than two comedications – vs those taking two comedications respectively; and CSFFLAG is a 0/1 flag for subjects without/with pathologic CSF respectively (CSF p-tau$_{181P}$/Aβ_{1-42} ratio>0.147 is considered pathologic). [b]Confidence interval represents precision of parameter estimates. [c]Estimated inflection based on the formula β·18/(1+β) is 10. [d]Between the base model and final covariate model, the IIV estimates improved from 99.0% and 65.5% to 62.2% and 39.0%CV, respectively.

Abbreviations: Aβ_{1-42}, 42 amino acid isoform of amyloid beta peptide; AD, Alzheimer's disease; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; CI, confidence interval; COMED, comedication use; CSF, cerebrospinal fluid; CV, percent coefficient of variation; DIS, disease state; DLM, delayed logical memory; HVOL, age and head size–adjusted hippocampal volume; IIV, interindividual variability; LMCI, late mild cognitive impairment; p-tau$_{181P}$, tau protein phosphorylated at position threonine 181; SD, standard deviation; TAT, Trails A test; TV, typical value; vs, versus.

Table 3 Parameter estimates from the dropout model

	Estimate	90% confidence interval[a]
Baseline hazard (α parameters)		
Period 1 in years [0,0.5)[b]	–4.020	–4.58 to –3.46
Period 2 in years [0.5,1)	–3.602	–4.08 to –3.13
Period 3 in years [1,1.5)	–3.349	–3.77 to –2.93
Period 4 in years [1.5,2)	–3.036	–3.49 to –2.58
Period 5 in years [2,3)	–1.352	–1.63 to –1.08
Coefficients (β parameters)[c]		
CDR–SB score prior to drop out	0.189*	0.15 to 0.23

Notes: *$P<0.0001$. [a]Confidence interval represents precision of parameter estimates. [b][lower boundary, upper boundary) indicates that the range includes the lower boundary but not the upper boundary. [c]Hazard ratio can be obtained by exponentiating the parameter estimate, indicating there is approximately a 21% increase in the hazard of dropping out (ie, data being missing) with 1-point increase in CDR–SB.

Abbreviation: CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes.

data model, but it did not reach statistical significance (P=0.4). This formal analysis confirmed that the likelihood of missing data was dependent on the CDR–SB score prior to the event ($P<0.0001$) (Table 3) and time. Hazard ratios indicated that there was a 21% increase in the hazard of dropping out (ie, data being missing) with a 1-point increase in CDR–SB score. This finding confirmed that "missing completely at random" was not the missing data mechanism and that MAR might be a more reasonable assumption for the current analysis data set.

Model qualification using diagnostic plots and visual predictive checks

Figure S4 shows the goodness-of-fit plots for the final model, based on the ADNI-1 LMCI and mild AD CDR–SB data set. For the residual plots, a random distribution of residuals against time was observed. Furthermore, there was also a random distribution of observed versus individual predicted values across the identity line. However, a mismatch between data at the lower boundary of the CDR–SB scale and population predictions was observed. Only 28 of the 1,602 (<2%) observations in the data set were at the lower boundary of the CDR–SB scale, and a majority (20/28) of these observations came from subjects that were CSF biomarker-negative. The mismatch between the model predictions and observations at the lower boundary of the scale is probably not clinically relevant because CSF biomarker-negative subjects will most likely be excluded from future MCI clinical trials (see the "Discussion" section).

To examine the predictive performance of the model, a VPC was performed. Data sets were simulated based on the fixed and random effect estimates of the final model. The simulated data sets had study design features similar to the ADNI-1 study. In one scenario, simulations were performed without the dropout model from the realized design. For the other scenario, simulations were performed with the combined disease progression plus dropout model, in which the simulated scores replaced the observed CDR–SB score in the dropout model. The results in Figure S5 indicate that the VPCs improve when simulating from the combined disease progression plus dropout model, which allowed simulated scores to be discarded following a simulated dropout event (especially at later time points, where scores are high and the probability of dropout is higher). Thus, the remaining stratified VPCs were based on the disease progression model that accounted for the missing data based on the dropout model described in the previous section. Figure 3 provides the percentiles of the observed scores and the VPC consisting of the model-based intervals superimposed on the observed percentiles. To assess the robustness of the model, the VPC was stratified by disease status (mild AD

versus LMCI) and CSF-positive/negative biomarker status (CSF p-tau$_{181p}$/Aβ_{1-42} >0.147 vs CSF p-tau$_{181p}$/Aβ_{1-42} ≤0.147). The VPC suggests that the final model describes the longitudinal progression of CDR–SB scores reasonably well in both the LMCI and mild AD subjects. It further suggests a lack of disease progression in the biomarker-negative subgroup of both LMCI and mild AD subjects.

The results of the VPC and diagnostic plots substantiated the inference that the disease progression model was able to describe the temporal profile and the IIV because the observed percentiles fell within the model predicted intervals. The VPC, and therefore the model, can be considered adequate for describing the disease progression profile.

Discussion

The choice of the CDR–SB end point for disease progression analysis, in a patient population proximal to the onset of dementia, with biomarker information was guided by a recent FDA draft guidance on drug development for early AD.[4] This model-based analysis involved a thorough investigation of CDR–SB disease progression, including an investigation of a) structural models; b) residual error structure, and c) covariate effects. Disease progression has commonly been described using linear, exponential, and logistic structural models in the AD literature. In this analysis, we formally compared those structural models. The Verhulst logistic structural model was further enhanced by including a shape parameter. The linear and exponential models have the potential to predict scores outside the range of the CDR–SB scale of 0 to 18. The Verhulst logistic model helped overcome this limitation and also tested the assumption that disease progression is nonlinear. The shape parameter in the final disease progression model remedied the inherent problem with the Verhulst logistic model that required the inflection point to be at the midpoint of the scale.[23]

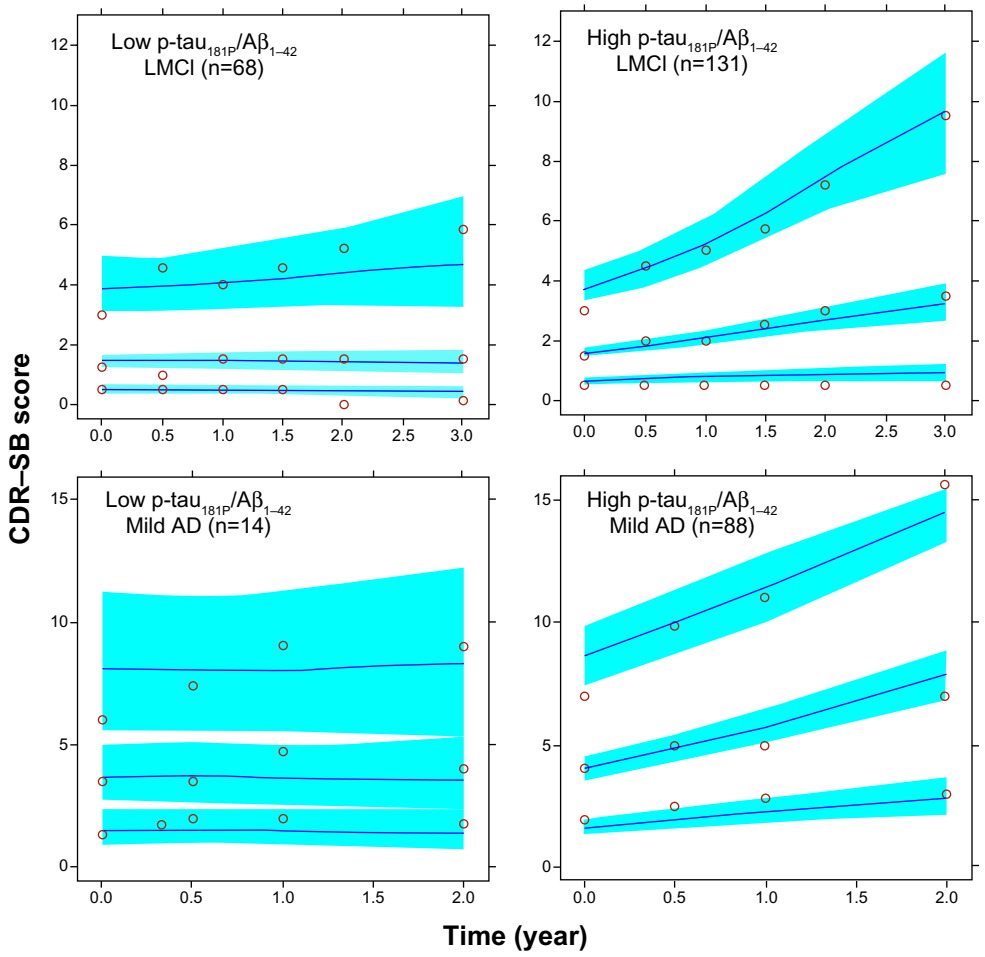

Figure 3 Stratified visual predictive check.

Notes: The upper, middle, and lower profiles indicated by the open circles represent the 95th, 50th, and fifth percentiles of the observed data, respectively. The upper, middle, and lower curves indicated by the lines are the median model-based predictions for the 95th, 50th, and fifth percentiles, respectively, and these predictions account for missing data. The shaded areas are the 90% confidence intervals of the corresponding percentiles of the simulations based on the model.

Abbreviations: Aβ_{1-42}, 42 amino acid isoform of amyloid beta peptide; AD, Alzheimer's disease; LMCI, late mild cognitive impairment; p-tau$_{181P}$, tau protein phosphorylated at position threonine 181; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes.

The results show that scores increase exponentially during the early phase of the disease and as AD worsens the scores theoretically level off as the CDR–SB approaches 18 (flooring of cognitive and functional abilities). This nonlinear process was mathematically described using the three-parameter logistic function in the final model, which agrees with the expectation that pharmacodynamic systems typically follow saturable relationships.[37] Interestingly, we have found that this structural model also describes disease progression for another AD end point.[24,25] The three-parameter logistic model also performed slightly better than the Richard's function. This is probably related to the fact that the three-parameter logistic model is a more flexible function at low scores,[25] where most of the data reside for LMCI and mild AD subjects. The three-parameter logistic function that describes an S-shaped disease progression curve appears to embody various end points in AD that follow a nonlinear and saturable trajectory.

The final model for the CDR–SB score also had certain salient features. The IIV on the baseline scores was assumed to follow a log-normal distribution. This, unlike the normal distribution, it does not predict negative baseline scores at the individual level. The random effect on the progression rate parameters was assumed to follow a normal distribution. This allowed subjects to improve, deteriorate, or stay unchanged in terms of disease status, thus allowing greater flexibility to accommodate the wide range of progression rates (seen in Figure S2). Various types of residual error structures were also tested in the current analysis (normal, log-normal, logit-normal, and beta distribution), and it was found that the logit-normal distribution captured the behavior of the residual error appropriately. This residual error structure had two advantages. First, the model predictions, even after accounting for residual error, stayed within the boundaries of the CDR–SB scale. Second, logit-normal residuals accounted for the interdependence between the mean and variance of the data. The variance was generally lower as scores approached the boundaries of the scale, while variance was greater around the middle portion of the scale, and logit-normal residuals were able to accommodate this complexity.

The model also took into account effect of previously known[25] and highly influential CSF biomarker covariate for disease progression rate. The analysis found that CSF p-tau$_{181p}$/A$\beta_{1–42}$, parameterized as a categorical covariate, was able to distinguish CDR–SB progressors from nonprogressors in both LMCI and mild AD subjects. In our recent analysis with LMCI ADNI-1 subjects, this covariate, parameterized as a threshold determinant, worked in an identical manner for describing ADAS-cog/11 progression rate.[25] The previous LMCI ADNI-1 analysis[25] identified a threshold of 0.156 for CSF p-tau$_{181p}$/A$\beta_{1–42}$, which is almost identical to the threshold of 0.147 derived from pooling CN, LMCI, and mild AD biomarker data. This adds credence to the derived threshold and suggests that this threshold is similar across disease states. The bimodal distribution in values in cross-sectional data is compatible with the notion that there are two subpopulations. Thus, subjects with biomarkers below the critical threshold have a low likelihood of disease progression within a 2–3 year time frame of a clinical trial and should potentially be excluded from studies aimed at showing disease modification with drug treatment. This potential application of CSF biomarkers for trial enrichment has received regulatory attention from the EMA,[12] and there are at least three ongoing or completed MCI clinical trials that have used this technique for population enrichment.[38–40]

The finding that CSF biomarkers do not influence CDR–SB baseline severity is not unusual because the other covariates influencing baseline disease score (hippocampal volume, disease state, and comedication use) are correlated with CSF biomarker status. A total of 86% of AD subjects were CSF biomarker-positive vs 66% of LMCI subjects. Further, 78% of subjects taking an AD comedication were CSF biomarker-positive vs 63% of those not taking comedications. Finally, 81% of subjects with smaller hippocampal volume ($\leq$3,233 mm^3 [cutoff based on the mixture model in Table S2]) were CSF biomarker-positive vs 52% whose hippocampal volume was greater than the cutoff. With these three covariates influencing baseline severity, there was no additional impact of CSF biomarker status on baseline severity. However, as pointed out previously, baseline CSF biomarker status has a large impact on progression rate, where it influences the discrimination of progressors from nonprogressors. Thus, CSF biomarkers are appropriate indicators of progression rate, while hippocampal volume, disease state, and comedication use are likely predictors of baseline severity.

An expanded list of covariates was assessed in our analysis given the richness of the ADNI-1 database, and the analysis found several noteworthy predictor variables. Higher bCDRSB scores reflected a clinical impression (which can be formed in the absence of knowledge of biomarkers or detailed neuropsychological testing) that the subject ought to be treated and therefore, the use of AD comedications was associated with higher baseline scores. However, AD comedication-use did not influence progression rate, which is in agreement with the mechanism of action for these drugs

ie, they are symptomatic in their drug effects and do not alter the disease progression rate. As expected, AD subjects had a higher baseline score compared with LMCI subjects. The structural model accounted for the nonlinear relationship between disease severity and progression rate. Once this nonlinearity was incorporated, the disease label (LMCI vs mild AD) had no further impact on the progression rate parameter. This suggests that LMCI and mild AD subjects have a single progression rate parameter and that the various stages of the disease represent a continuum.

In a univariate analysis, baseline hippocampal volume had a highly significant effect (Table S2) on both progression rate and baseline severity as a categorical covariate ($P=3.6 \cdot 10^{-19}$). However, based on the results in Table S2, the CSF biomarkers performed better than hippocampal volume in this univariate analysis. Interestingly, once the CSF biomarkers were incorporated as a covariate on progression rate, hippocampal volume did not contribute further to the progression rate (however, it still had an impact on baseline severity). An explanation for this finding could be that volumetric measurements themselves exhibit progressive hippocampal atrophy. Structural MRI changes start to manifest at an earlier stage before changes in clinical measures become clearly evident.[41] Thus, baseline hippocampal volume could be a weak predictor of future cognitive decline because cognitive changes occur at a different pace at a later stage in the disease.[42] It seems that hippocampal volume is a good indicator of baseline severity, while the CSF biomarkers are likely predictors of progression rate. The lack of hippocampal volume effect on progression rate in a model with CSF biomarkers is not unexpected since hippocampal volume and CSF biomarkers are correlated variables, as discussed before. More noteworthy is the fact that the correlation between hippocampal volume and CSF biomarker status was much stronger in ApoE4 carriers vs noncarriers. In ApoE4 carriers, 82% of subjects (145/178) had smaller hippocampi (ie, $\leq$3,233 mm^3), and 94% of these subjects (136/145) were CSF biomarker-positive. In contrast, for ApoE4 noncarriers, 57% of subjects (70/123) had smaller hippocampi, and 54% of these subjects (38/70) were CSF biomarker-positive. One implication of these findings is that future drug trials could be enriched by including patients with ApoE4, who are more likely to have smaller hippocampi and are also more likely to be CSF biomarker-positive (and therefore likely to progress within the time frame of a clinical trial).

Two additional covariates were identified that influenced progression rate. These were the baseline delayed logical memory score and the baseline Trails A test score. The delayed logical memory score is a reflection of the episodic memory deficit, and episodic memory deficit is a well-recognized predictor of progression rate.[43] The other covariate associated with progression rate was the Trails A test score, which is a measure of cognitive processing speed. However, the Trails A test score is also highly correlated with the Trail-Making Test part B (Trails B) test score (Pearson's $r=0.6$; $P<0.001$). The Trails B test score is a measure of executive functioning, and in two previous progression analyses,[24,25] this test was associated with progression rate. The Trails A test may perform somewhat better over Trails B test in the current analysis because only 3% of baseline Trails A test scores reside at the upper boundary of the scale compared with 16% of the baseline Trails B scores that do so, suggesting about 16% of the subjects would be unable to complete the Trails B test within the allotted 300 seconds (recorded in the covariate database as 300 seconds for 16% of subjects). The Trails A test is not associated with this data limitation and may therefore perform slightly better in the covariate analysis than the Trails B test. Thus the influence of Trails A test on progression rate may be indicative of the influence of executive functioning on progression rate, through its correlation with the Trails B test. It also appears that delayed logical memory is more predictive of progression in LMCI than in AD (Figure 2A and B), while the reverse appears to be true for Trails A. This suggests that episodic memory is predictive of progression earlier in the disease, while executive functioning is a better predictor later in the disease.

We have previously shown[24] that younger age is associated with faster progression rate in AD subjects. This earlier analysis did not assess CSF biomarkers. There is a correlation between log CSF p-tau_{181p}/$A\beta_{1-42}$ and age (Pearson's $r=-0.32$; $P=0.001$); suggesting subjects with AD at a younger age tend to be biomarker-positive. This agrees well with another recent report,[25] where it was shown that the CSF biomarkers were positive in ApoE4 carriers and subjects with high cholesterol, who also tended to exhibit faster disease progression. Thus, previously reported covariate effects, such as the influence of age, ApoE4-carrier status, and cholesterol, on progression rate can all be explained by the presence of CSF biomarkers. Once the influence of the primary covariate (CSF biomarkers) on progression rate is factored into the model, the other covariates correlated with CSF biomarkers (age, ApoE4, serum cholesterol, and hippocampal volume) are no longer necessary.

There is one key limitation of the current model. This model of disease progression is built on CDR–SB, a well-known scale, with certain limitations for the early stages of

MCI. This limitation arises from the fact that three out of the six CDR–SB subscales cover various aspects of daily functioning that are not impaired earlier in the disease. Scales with a broader dynamic range and better sensitivity than CDR–SB (such as a sensitive cognitive measure) at a lower spectrum of the disease may be more suitable to explore in the future, for early MCI. An additional limitation is that the CDR–SB evaluation, which is based on a semistructured interview with the patient and an informant, is a lengthy instrument to administer.[6] This requires the raters to be well trained, and the test relies on the accuracy of the information provided by the subject's informant. However, it does not include numerically scored psychometric tests and therefore, does not exhibit learning effects, which are generally described as a "placebo" effects in disease progression models.[6] In addition, it is a continuous outcome measure that is capable of tracking the progression of the disease on the cognitive and functional scale with no floor and ceiling effects (minimal data at the boundaries) for the LMCI and mild AD population. Moreover, disease progression as assessed by the CDR–SB correlates well with other independent cognitive and functional tests used in LMCI and mild AD, which gives it additional face validity.[6] Finally, CDR–SB is able to track the progression of cognitive and functional decline since it exhibits reasonable change over 2- to 3-year periods.[6,7] Thus the CDR–SB, used for >20 years in clinical trials, represents a validated, well-described, and reliable measure of disease progression.

Conclusion

This work provides a model-based approach for describing CDR–SB disease progression, in LMCI and mild AD subjects. CDR–SB disease progression is nonlinear and follows a sigmoidal shape. CDR–SB exhibited negligible floor and ceiling effects (very small proportion of data at the boundaries of the scale) in LMCI and mild AD subjects. The dropout process for this population was informative in nature, and as the CDR–SB scores increased, the probability of dropout increased. Covariates associated with baseline disease severity were baseline hippocampal volume, mild AD vs LMCI status, and AD comedication use. Similarly, factors that influence disease progression rate were current CDR–SB score, baseline CSF biomarkers, baseline delayed logical memory and Trails A score. The model allowed the identification of a subpopulation, based on CSF biomarkers, that have a low likelihood of disease progression and could be excluded from future clinical trials. The model could serve as a useful simulation tool for the efficient design of clinical trials.

Acknowledgments

Data used in preparation of this article were obtained from the ADNI database. 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.

The authors would like to thank the subjects, investigators, and the staff who participated in the ADNI-1 trial. The authors thank Dr An Vermeulen (Janssen Research and Development) for providing scientific review of the manuscript. We are also thankful to Jenny Yang and Haoli Chai for formatting the NONMEM-ready data files for pharmacometric analysis.

Data collection and sharing for this project was funded by the ADNI (National Institutes of Health grant number U01 AG024904) and Department of Defense 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: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; F Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, NV; IXICO Ltd; Janssen Alzheimer Immunotherapy Research and Development, LLC; Johnson and Johnson Pharmaceutical Research and Development LLC; Medpace, Inc.; Merck and Co, Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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 (http://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 of Neuro Imaging at the University of Southern California.

Disclosure

The authors of this manuscript are employees of Janssen Research and Development, LLC and own Johnson and

Johnson stocks. All authors met the International Committee of Medical Journal Editors criteria. The authors report no other conflicts of interest in this work.

References

1. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*. 2002;297(5580):353–356.
2. Trinh NH, Hoblyn J, Mohanty S, Yaffe K. Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: a meta-analysis. *JAMA*. 2003;289(2):210–216.
3. Jicha GA, Parisi JE, Dickson DW, et al. Neuropathologic outcome of mild cognitive impairment following progression to clinical dementia. *Arch Neurol*. 2006;63(5):674–681.
4. Food and Drug Administration. *Draft Guidance for Industry: Alzheimer's Disease: Developing Drugs for the Treatment of Early Stage Disease.* Washington, DC: Center for Drug Evaluation and Research; 2013.
5. Food and Drug Administration. *Draft Guidance for Industry: Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products.* Washington, DC: Center for Drug Evaluation and Research; 2013.
6. Cedarbaum JM, Jaros M, Hernandez C, et al; Alzheimer's Disease Neuroimaging Initiative. Rationale for use of the Clinical Dementia Rating Sum of Boxes as a primary outcome measure for Alzheimer's disease clinical trials. *Alzheimers Dement*. 2013;9(Suppl 1):S45–S55.
7. Williams MM, Storandt M, Roe CM, Morris JC. Progression of Alzheimer's disease as measured by Clinical Dementia Rating Sum of Boxes scores. *Alzheimers Dement*. 2013;9(Suppl 1):S39–S44.
8. European Medicines Agency. *Qualification Opinion of a Novel Data Driven Model of Disease Progression and Trial Evaluation in Mild and Moderate Alzheimer's Disease.* London: European Medicines Agency; 2013.
9. Romero K, Stephenson D, Rogers J, et al. Advancement of a drug-disease trial model for Alzheimer's disease through a regulatory science path: The CAMD experience. *Alzheimers Dement.* 2012;8(4):P587–P588.
10. Rogers JA, Polhamus D, Gillespie WR, et al. Combining patient-level and summary-level data for Alzheimer's disease modeling and simulation: a β regression meta-analysis. *J Pharmacokinet Pharmacodyn*. 2012;39(5):479–498.
11. European Medicines Agency. *Qualification Opinion òf Low Hippocampal Volume (Atrophy) by MRI for Use in Clinical Trials for Regulatory Purpose – in Pre-Dementia Stage Of Alzheimer's Disease*. London: European Medicines Agency, 2011. Available from http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2011/12/WC500118737.pdf. Accessed April 4, 2014.
12. European Medicines Agency. *Qualification Opinion of Alzheimer's Disease Novel Methodologies/Biomarkers for BMS-708163*. London: European Medicines Agency, 2011.
13. European Medicines Agency. *Qualification Opinion of Alzheimer's Disease Novel Methodologies/Biomarkers for PET Amyloid Imaging (Positive/Negative) as a Biomarker for Enrichment for Use – in Predementia AD Clinical Trials.* London: European Medicines Agency, 2011.
14. Jagust WJ, Landau SM, Shaw LM, et al; Alzheimer's Disease Neuroimaging Initiative. Relationships between biomarkers in aging and dementia. *Neurology*. 2009;73(15):1193–1199.
15. Liu Y, Mattila J, Ruiz MÁ, et al; Alzheimer's Disease Neuroimaging Initiative. Predicting AD conversion: comparison between prodromal AD guidelines and computer assisted Predict AD tool. *PLoS One*. 2013;8(2):e55246.
16. Alzheimer's Disease Cooperative Study (ADCS). ADNI: Alzheimer's Disease Neuroimaging Initiative. In: ClinicalTrials.gov [website on the Internet]. Bethesda, MD: US National Library of Medicine; 2013 [updated May 15, 2013]. Available from http://clinicaltrials.gov/ct2/show/NCT00106899. NLM identifier: NCT00106899. Accessed March 26, 2014.
17. ADNI data training part 2. ADNI biostatistics core team. August 1, 2013. Available from http://adni.loni.usc.edu/wp-content/uploads/2012/08/slide_data_training_part2_reduced-size.pdf. Accessed April 4, 2014.
18. Beal SL, Sheiner LB, Boeckmann AJ, Bauer RJ. *NONMEM 7.2.0 Users Guide*. Ellicott City, MD: Icon Development Solutions; 1989–2011.
19. Lindbom L, Pihlgren P, Jonsson EN, Jonsson N. PsN-Toolkit – a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Programs Biomed*. 2005;79(3):241–257.
20. Ito K, Corrigan B, Zhao Q, et al; Alzheimer's Disease Neuroimaging Initiative. Disease progression model for cognitive deterioration from Alzheimer's Disease Neuroimaging Initiative database. *Alzheimers Dement*. 2011;7(2):151–160.
21. Ito K, Hutmacher MM, Corrigan BW. Modeling of Functional Assessment Questionnaire (FAQ) as continuous bounded data from the ADNI database. *J Pharmacokinet Pharmacodyn*. 2012;39(6):601–618.
22. William-Faltaos D, Chen Y, Wang Y, Gobburu J, Zhu H. Quantification of disease progression and dropout for Alzheimer's disease. *Int J Clin Pharmacol Ther*. 2013;51(2):120–131.
23. Tsoularis A, Wallace J. Analysis of logistic growth models. *Math Biosci*. 2002;179(1):21–55.
24. Samtani MN, Farnum M, Lobanov V, et al; Alzheimer's Disease Neuroimaging Initiative. An improved model for disease progression in patients from the Alzheimer's disease neuroimaging initiative. *J Clin Pharmacol*. 2012;52(5):629–644.
25. Samtani MN, Raghavan N, Shi Y, et al; Alzheimer's Disease Neuroimaging Initiative. Disease progression model in subjects with mild cognitive impairment from the Alzheimer's disease neuroimaging initiative: CSF biomarkers predict population subtypes. *Br J Clin Pharmacol*. 2013;75(1):146–161.
26. Ashford JW, Schmitt FA. Modeling the time-course of Alzheimer dementia. *Curr Psychiatry Rep*. 2001;3(1):20–28.
27. Xu XS, Samtani MN, Dunne A, Nandy P, Vermeulen A, De Ridder F; Alzheimer's Disease Neuroimaging Initiative. Mixed-effects beta regression for modeling continuous bounded outcome scores using NONMEM when data are not on the boundaries. *J Pharmacokinet Pharmacodyn*. 2013;40(4):537–544.
28. Burnham KP, Anderson DR. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach.* 2nd ed. New York, NY: Springer; 2002.
29. Frederic P, Lad F. Two moments of the logitnormal distribution. *Commun Stat Simul Comput*. 2008;37(7):1263–1269.
30. Smithson M, Verkuilen J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol Methods*. 2006;11(1):54–71.
31. Verkuilen J, Smithson M. Mixed and mixture regression models for continuous bounded responses using the beta distribution. *J Educ Behav Stat*. 2012;37(1):82–113.
32. Jack CR, Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. *Radiology*. 1989;172(2):549–554.
33. Siddiqui O, Hung HM, O'Neill R. MMRM vs LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. *J Biopharm Stat*. 2009;19(2):227–246.
34. Kleinbaum DG, Klein M. *Survival Analysis. A Self-Learning Text*. 2nd ed. New York, NY: Springer; 2005.
35. Friberg LE, de Greef R, Kerbusch T, Karlsson MO. Modeling and simulation of the time course of asenapine exposure response and dropout patterns in acute schizophrenia. *Clin Pharmacol Ther*. 2009;86(1):84–91.
36. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201–209.
37. Mager DE, Wyska E, Jusko WJ. Diversity of mechanism-based pharmacodynamic models. *Drug Metab Dispos*. 2003;31(5):510–518.

38. Janssen Research and Development, LLC. A study to evaluate the effects of JNJ-54861911 on amyloid beta processing in cerebrospinal fluid and plasma in patients with prodromal Alzheimer's disease. Available from: http://clinicaltrials.gov/ct2/show/NCT01978548?term=CSF Aβ1-42&rank=2. NLM identifier: NCT01978548. Accessed December 1, 2013.
39. Merck Sharp and Dohme Corp. Efficacy and safety trial of MK-8931 in participants with prodromal Alzheimer's disease (MK-8931-019) (APECS). Available from: http://clinicaltrials.gov/ct2/show/NCT01953601?term=abetaMK&rank=3. NLM identifier: NCT01953601. Accessed December 1, 2013.
40. Bristol-Myers Squibb. A multicenter, double blind, placebo-controlled, safety and tolerability study of BMS-708163 in patients with prodromal Alzheimer's disease. Available from: http://clinicaltrials.gov/ct2/show/NCT00890890?term=BMS-708163&rank=8. NLM identifier: NCT00890890. Accessed December 1, 2013.
41. Tondelli M, Wilcock GK, Nichelli P, De Jager CA, Jenkinson M, Zamboni G. Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. *Neurobiol Aging*. 2012;33(4):825.e25–e825.e36.
42. Jack CR, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12(2):207–216.
43. Landau SM, Harvey D, Madison CM, et al; Alzheimer's Disease Neuroimaging Initiative. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology*. 2010;75(3):230–238.

Supplementary materials

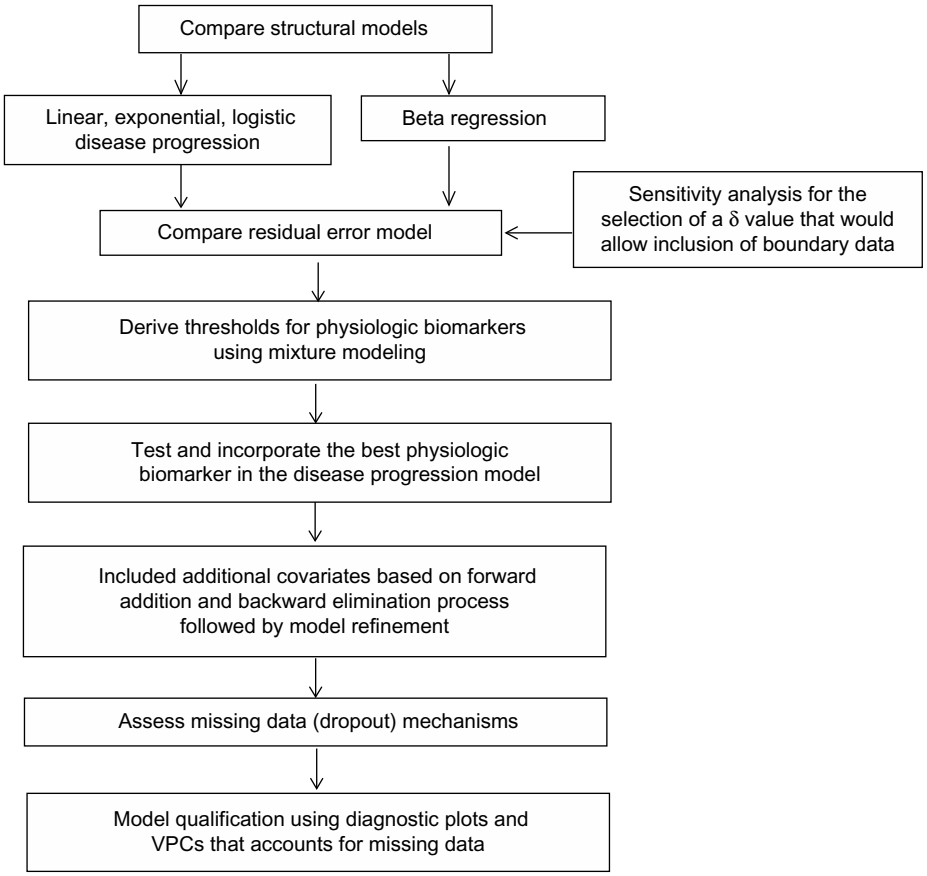

Figure S1 Overview of the model building process.
Note: δ is a small noise used for rescaling all the data to move the boundary observations slightly within the edges.
Abbreviation: VPC, visual predictive check.

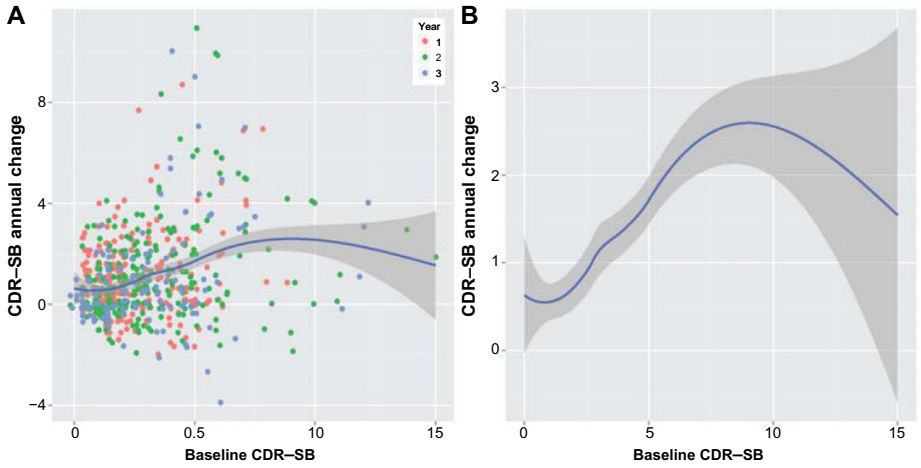

Figure S2 (**A**) 1-year change scores vs disease severity on CDR–SB scale for mild AD and LMCI subjects. The line and gray areas represent the loess smoother and the 95% confidence bands. (**B**) The loess smoother and confidence bands are shown alone to allow illustration of the inverted U-shaped relationship between progression rate and disease severity.
Notes: The calculations are based on the 12-month interval method. The difference in scores between a pair of points separated by 12 months was computed (later score minus the earlier baseline score) and plotted against the baseline score for that pair. If subjects contributed more than one pair of scores separated by 12 months, those subjects generated more than one data point for the plot (the legend in Panel A indicates the year of the study during which the pair of scores were generated). Due to the statistical issues associated with subjects contributing different numbers of observations and the correlations among data points from one subject, these graphs are used only for diagnostic/exploratory purposes.
Abbreviations: AD, Alzheimer's disease; LCMI, late mild cognitive impairment; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; vs, versus.

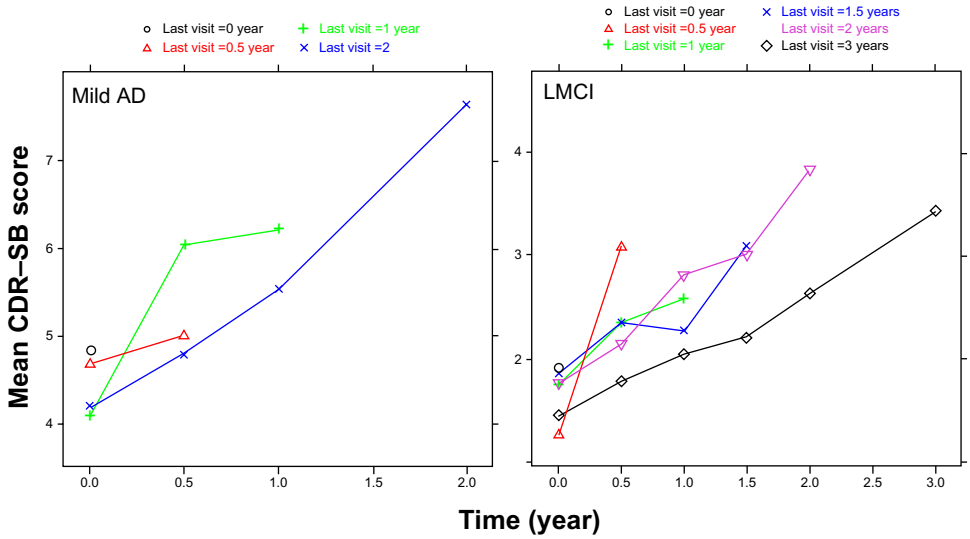

Figure S3 Results of the exploratory analysis for missing data: mean observed scores as a function of different study discontinuation times.
Abbreviations: AD, Alzheimer's disease; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; LMCI, late mild cognitive impairment.

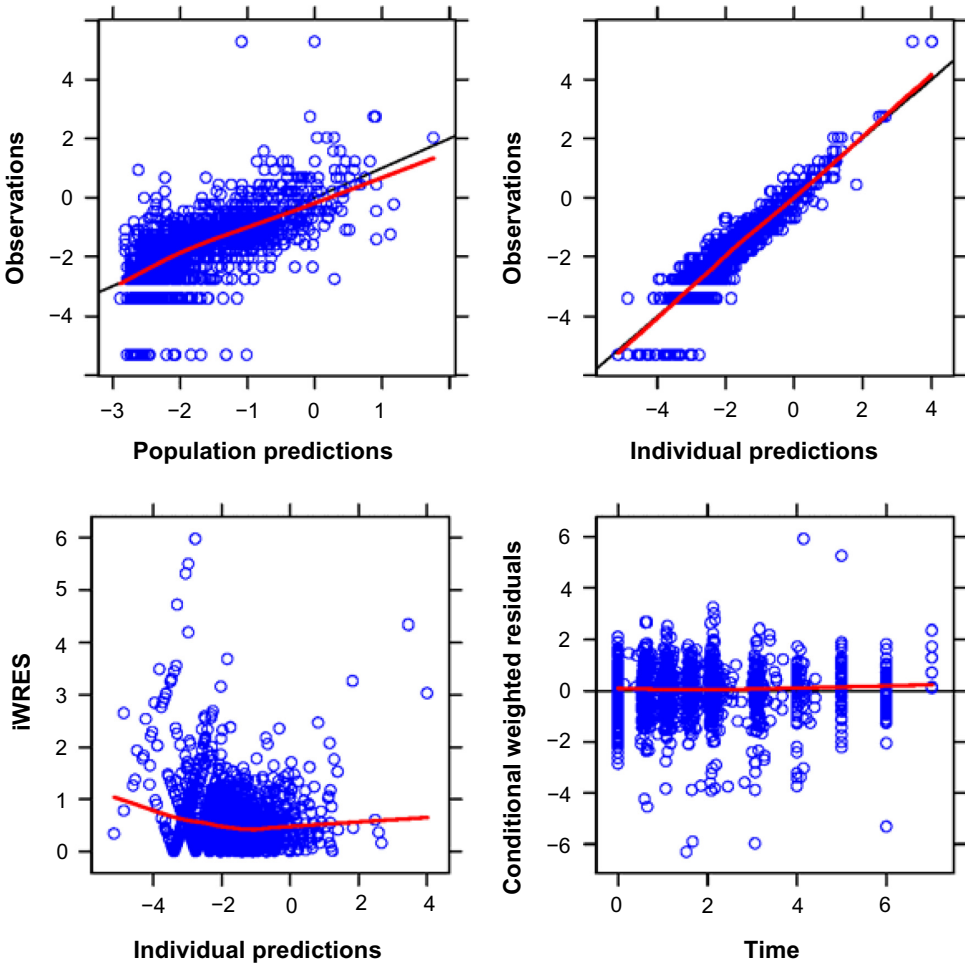

Figure S4 Goodness-of-fit plots.
Notes: The observed data versus the population and individual predictions are shown in the top panels. The bottom left panel shows the absolute values of the individually weighted residuals, and the bottom right panel shows the population conditional weighted residuals. Individual data points are indicated by blue circles; the red lines are loess smoothers, the black diagonal lines (top panels) represent the lines of identity, and the black horizontal line (bottom right) indicates the ordinate value of 0.
Abbreviation: iWRES, the individually weighted residuals.

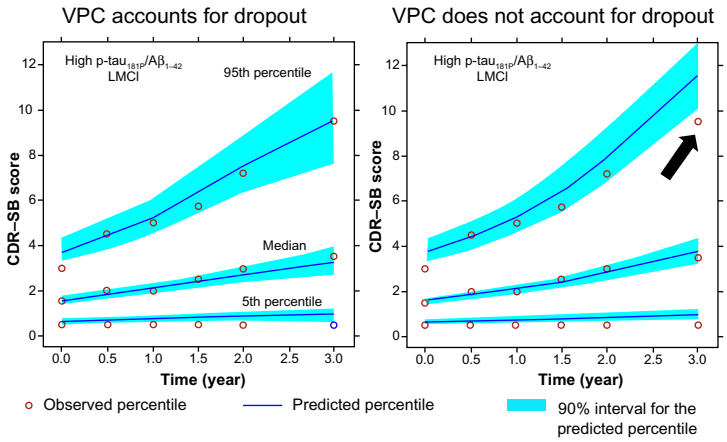

Figure S5 VPC with and without dropout.
Notes: In the left panel, simulations were performed with the combined disease progression plus dropout model. In the right panel the simulations were performed without the dropout model. The black arrow in the right panel indicates that predictive performance is hampered at later times (where scores are high and probability of dropout higher) when dropouts are not accounted.
Abbreviations: $A\beta_{1-42}$, 42 amino acid isoform of amyloid beta peptide; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; LMCI, late mild cognitive impairment; p-tau_{181P}, tau protein phosphorylated at position threonine 181; VPC, visual predictive check.

Table S1 Boundary data handling: results of sensitivity analysis*

δ	R_0	α	β	RUV SD	R_0 IIV%CV	α IIV%CV
Reference parameters						
N/A[a]	1.94 (4.1%)	0.259 (11.2%)	1.25 (6.4%)	0.385 (3.2%)	68.3% (7%)	74% (27.4%)
Comparable parameter estimates[b] based on the sensitivity analysis						
0.001	1.87 (4.4%)	0.336 (13.6%)	1 (10.9%)	0.672 (8%)	69.5% (10.3%)	115% (27.7%)
0.0018	1.89 (4.3%)	0.316 (13.9%)	1.05 (10.6%)	0.614 (7.4%)	69.1% (9.6%)	112% (29.3%)
0.003	1.91 (4.2%)	0.298 (14%)	1.09 (10.2%)	0.566 (6.7%)	68.5% (9.1%)	108% (31.1%)
0.005	1.94 (4.1%)	0.278 (14%)	1.14 (9.6%)	0.52 (6%)	67.6% (8.6%)	105% (32.2%)
0.0075	1.97 (4%)	0.262 (13.9%)	1.18 (9.2%)	0.485 (5.3%)	66.6% (8.3%)	102% (33.1%)
0.01	2.00 (4%)	0.25 (12.5%)	1.21 (8.4%)	0.46 (4.9%)	65.5% (8.1%)	99% (32.2%)
0.02	2.11 (3.8%)	0.219 (13.1%)	1.3 (7.5%)	0.404 (3.9%)	62% (7.6%)	93% (32.6%)
0.025	2.16 (3.6%)	0.207 (13.2%)	1.33 (7.3%)	0.386 (3.7%)	60.5% (7.5%)	91% (32.9%)
0.03	2.2 (3.5%)	0.197 (12.5%)	1.35 (6.7%)	0.372 (3.4%)	59% (7.5%)	89% (31.4%)
0.04	2.3 (3.4%)	0.18 (12.8%)	1.4 (6.5%)	0.348 (3.2%)	56.3% (7.4%)	87% (31.9%)
0.05	2.39 (3.2%)	0.165 (12.9%)	1.44 (6.2%)	0.33 (3.1%)	53.9% (7.3%)	85% (31.8%)
0.1	2.8 (2.7%)	0.112 (13.9%)	1.6 (5.5%)	0.27 (3%)	44.5% (7.3%)	80% (32.3%)
0.33	4.5 (1.3%)	0.019 (27.4%)	2.26 (6.5%)	0.152 (3.5%)	22.4% (8.6%)	79% (57.8%)

Notes: [a]δ was not used, ie, Model estimation without rescaling data and excluding data on boundaries; [b]sets of parameter estimates (parameter imprecision, expressed as percent coefficient of variation) are reported for each tested value of δ. δ is a small noise used for rescaling all the data to move the boundary observations slightly within the edges. *Parameter estimates (with parameter precision) reported for each value of δ that was tested.
Abbreviations: α, progression rate parameter; β, shape parameter in the logistic structural model; CV, coefficient of variation; IIV, interindividual variability; N/A, not applicable; R_0, baseline CDR–SB score; RUV SD, standard deviation of the residual unexplained variability.

Table S2 Parameters of the mixture models fitted to the baseline CSF biomarker data and hippocampal volume in the ADNI-1 CN, LMCI, and mild AD subjects

Baseline biomarker	Mean of subpopulation with pathologic biomarker*	SD of subpopulation with pathologic biomarker*	Mean of subpopulation without pathologic biomarker*	SD of subpopulation without pathologic biomarker*	Threshold between the two subpopulations	Mixing proportion: pathologic* group in CN	Mixing proportion: pathologic* group in LMCI	Mixing proportion: pathologic* group in mild AD	AIC for CDR–SB disease progression model***
log t-tau	4.79	0.43	4.15	0.37	4.5‡	4%	55%	83%	–3,879.0
log p-tau$_{181P}$	3.69	0.39	2.89	0.29	3.2§	25%	69%	91%	–3,879.7
Hippocampal volume (mm^3)	2,810	397	3,570	369	3,233	0%	64%	97%	–3,902.1
log t-tau/ A$\beta_{1–42}$ ratio	–0.14	0.47	–1.29	0.39	–0.78**	15%	66%	89%	–3,910.8
A$\beta_{1–42}$ (pg/mL)	136	27	241	30	192	34%	74%	92%	–3,919.6
log p-tau$_{181P}$/ A$\beta_{1–42}$ ratio	–1.19	0.47	–2.46	0.39	–1.92†	20%	68%	91%	–3,921.1

Notes: The LMCI and AD populations were dichotomized based on the thresholds reported below, and AIC values are reported in the last column for the CDR–SB disease progression model, where the biomarker was introduced as a dichotomized covariate.*Pathologic biomarker (low CSF A$\beta_{1–42}$, low hippocampal volume, high p-tau$_{181P}$, high p-tau$_{181P}$/A$\beta_{1–42}$, high t-tau, or high t-tau/A$\beta_{1–42}$). ‡This corresponds to a cutoff value of 90 pg/mL on the untransformed scale for baseline CSF t-tau. §This corresponds to a cutoff value of 24.5 pg/mL on the untransformed scale for baseline CSF p-tau$_{181P}$. **This corresponds to a cutoff value of 0.458 on the untransformed scale for baseline CSF t-tau/A$\beta_{1–42}$ ratio. †This corresponds to a cutoff value of 0.147 on the untransformed scale for baseline CSF p-tau$_{181P}$/A$\beta_{1–42}$ ratio. ***As a comparison, the base model for CDR–SB disease progression, which included data at the boundaries, had an AIC value of –3,819.1.

Abbreviations: A$\beta_{1–42}$, 42 amino acid isoform of amyloid beta peptide; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; AIC, Akaike information criterion; CDR–SB, Clinical Dementia rating Scale–Sum of Boxes; CN, cognitively normal; LMCI, late mild cognitive impairment; p-tau$_{181P}$, tau protein phosphorylated at position threonine 181; SD, standard deviation; t-tau, total tau protein; CSF, cerebrospinal fluid.

Table S3 Mean baseline CDR–SB for the subgroups, based on the statistically significant covariates for the baseline parameter in the disease progression model

Disease status	Stratification variable	N	Mean baseline CDR–SB ± SD
Disease group			
LMCI	N/A	199	1.55±0.9
Mild AD	N/A	102	4.25±1.6
Comedication status			
LMCI	0 comedication	90	1.39±0.7
LMCI	1 comedication	86	1.64±0.9
LMCI	2 comedications	23	1.85±1.1
Mild AD	0 comedication	7	3.50±1.1
Mild AD	1 comedication	54	4.08±1.6
Mild AD	2 comedications	41	4.61±1.6
Hippocampal volume			
LMCI	Normal hippocampi	66	1.28±0.6
LMCI	Shrunken hippocampi	133	1.69±1.0
Mild AD	Normal hippocampi	20	3.15±1.1
Mild AD	Shrunken hippocampi	82	4.52±1.6

Notes: Normal hippocampi (hippocampal volume >3,233 mm^3); Shrunken hippocampi (hippocampal volume ≤3,233 mm^3). The hippocampal volume threshold of 3,233 mm^3 was based on the biomarker mixture model (Table S2). The threshold of 3,233 mm^3 is also very close to the threshold of 3,226 mm^3, which is 1 SD below the mean hippocampal volume for the elderly normal controls.

Abbreviations: AD, Alzheimer's disease; CDR–SB, Clinical Dementia Rating Scale–Sum of Boxes; LMCI, late mild cognitive impairment; N/A, not applicable; SD, standard deviation.