Prediction of amyloid β accumulation from multiple biomarkers using a hierarchical Bayesian model

Yuichiro Yada[1,**], Honda Naoki[1,2,3,4,*]

Affiliations:

1) Laboratory of Data-driven Biology, Graduate School of Integrated Sciences for Life, Hiroshima University, Kagamiyama, Higashi-hiroshima, Hiroshima, 739-8526, Japan
2) Kansei-Brain Informatics Group, Center for Brain, Mind and Kansei Sciences Research (BMK Center), Hiroshima University, Kasumi, Minami-ku, Hiroshima 734-8551, Japan
3) Laboratory of Theoretical Biology, Graduate School of Biostudies, Kyoto University, Yoshidakonoecho, Sakyo, Kyoto, 606-8315, Japan
4) Theoretical Biology Research Group, Exploratory Research Center on Life and Living Systems (ExCELLS), National Institutes of Natural Sciences, Okazaki, Aichi, 444-8787, Japan

*Corresponding author: Honda Naoki
Address: Graduate School of Integrated Sciences for Life, 1-3-1 Kagamiyama, Higashi-Hiroshima City, Hiroshima, 739-8526, Japan
Tel.: +81-82-424-7336
E-mail: nhonda@hiroshima-u.ac.jp

**Co-corresponding author: Yuichiro Yada
Address: Graduate School of Integrated Sciences for Life, 1-3-1 Kagamiyama, Higashi-Hiroshima City, Hiroshima, 739-8526, Japan
Tel.: +81-82-424-7836
E-mail: yuyada@hiroshima-u.ac.jp

Abstract:

Accumulation of amyloid-beta (Aβ) in the brain is associated with neurodegeneration in Alzheimer's disease and can be an indicator of early disease progression. Thus, the non-invasively and inexpensively observable features related to Aβ accumulation are promising biomarkers. However, in the experimental discovery of biomarkers in preclinical models, Aβ and biomarker candidates are usually not observed in identical sample populations. This study established a hierarchical Bayesian model that predicts Aβ accumulation level solely from biomarker candidates by integrating incomplete information. The model was applied to 5xFAD mouse behavioral experimental data. The predicted Aβ accumulation level obeyed the observed amount of Aβ when multiple features were used for learning and prediction. Based on the evaluation of predictability, the results suggest that the proposed model can contribute to discovering novel biomarkers, that is, multivariate biomarkers relevant to the accumulation state of abnormal proteins.

Introduction:

Alzheimer's disease (AD) is the most common cause of dementia[1,2]. AD is a neurodegenerative disease in which neurons in the brain gradually die, causing progressive cognitive decline characterized by memory loss, impaired judgment and reasoning skills, communication difficulties, and changes in personality and behavior. However, most cases of AD progress sporadically, and many patients have no family history of AD[3]. Thus, the genetic background of sporadic AD has been extensively investigated, revealing the complex genetic architecture of late-onset neurodegenerative diseases[4–7].

The diagnosis or risk foresight of AD before initiating the decisive progression of neuronal loss may enable the potential treatment of the disease or the administration of appropriate symptomatic medication. The rapidly growing population of affected people raises the urgent demand for developing novel conveniently observed biomarkers and prediction methods for AD before this time point[8–10]. In AD, the gradual accumulation of abnormal proteins, amyloid beta (Aβ), and tau precede irreversible neurodegeneration[1,2,11]. Aβ accumulates earlier, followed by tau. Abnormal proteins may propagate from the brain to other regions[12–14]. Thus, the accumulation level of abnormal protein forms may represent the progression stage of neurodegenerative diseases in the early phase[15–17]. The imaging of Aβ in the brain using positron emission tomography (PET) scan and its assessment in cerebrospinal fluid (CSF) potentially facilitate AD onset prediction[1,2,18–20]. Yet, the demand for more inexpensive and non-invasive

methods to predict the risk of onset has led to further investigation of novel biomarker candidates in humans and preclinical models[21–23].

To discover new AD biomarkers using animal models, researchers usually conduct statistical testing to detect differing features between the AD model and healthy wildtype samples at the same age. Progression timing and speed of pathology, however, may be heterogeneous even among AD model animals with the same genetic background. Such heterogeneity could mask the differences researchers aimed to detect (**Fig. 1A**). The accumulation level of Aβ in the brains of individual animals, which may reflect the progression state of AD pathology, may have been overlooked in past biomarker discovery in model animals. Ideally, time-series paired data of abnormal protein levels and multidimensional biomarkers are required to identify novel biomarkers regarding the relevance of Aβ accumulation in AD (**Fig. 1B**). However, longitudinal measurements of individuals are impossible due to animal sacrifices; data are obtained as snapshots (**Fig. 1C**). Another problem is that most studies have only measured either aberrant proteins or multidimensional biomarkers; few studies have measured both from enough samples (**Fig. 1C**).

In the present study, to find Aβ-accumulation-relevant biomarkers overcoming the problem of limited data, we proposed a hierarchical Bayesian model that describes the Aβ accumulation process and observation of biomarker candidates. In the model, an abnormal form of Aβ deposits over time according to the logistic function, whose parameters are unique to each sample. Owing to the Bayesian probabilistic formulation, the model can naturally integrate unpaired and partially incomplete data, increasing the accumulation level's predictability from the observed biomarkers. By applying the model to the behavioral data sets of public 5xFAD mice[24], we predicted the accumulation level of Aβ solely from behavioral data, supporting the concept of biomarker discovery based on predictability.

Results:

The hierarchical Bayesian model of Aß accumulation

To represent the pathogenesis of AD in animal models, we developed a mathematical model describing the accumulation process of Aβ in the brain and the observation process of Aβ and biomarker candidates. In the model (**Fig. 1D**), Aβ accumulates over time according to the logistic function:

$$z_n = \frac{\alpha_n}{1 + exp\{-\beta_n(t_n - \tau_n)\}},$$

where $\theta_n = \{\alpha_n, \beta_n, \tau_n\}$ is a set of parameters depending on the individual animals; α_n, β_n, and τ_n denote the maximum level of Aβ accumulation, steepness of Aβ accumulation, and

critical period to reach half of the maximum Aβ, respectively. To express heterogeneity in the Aβ accumulation process among animals, different values of $\theta_n = \{\alpha_n, \beta_n, \tau_n\}$ are assigned to the individual animals, following the distribution shared within the same type of animals s_n, that is, wildtype or AD model animals.

The observed amount of Aβ, y_n, is subject to noise as $y_n = z_n + \sigma_y \xi$, where σ_y and ξ indicate the noise strength and Gaussian noise with zero mean and unit variance, respectively. The observed data of biomarker candidates $\mathbf{x}_n \in \mathbb{R}^L$ were assumed to reflect z_n and its temporal derivative z'_n, i.e., dz_n/dt, as

$$\mathbf{x}_n = W\mathbf{z}_n + S_x\boldsymbol{\varepsilon},$$

where $\mathbf{z}_n = (z_n, Cz'_n, 1)^{\mathrm{T}}$, $W \in \mathbb{R}^{L\times 3}$ indicates weight matrix, $S_x = (\sigma_{x_1}, \sigma_{x_2}, \dots, \sigma_{x_L})$ and $\boldsymbol{\varepsilon} \in \mathbb{R}^L$ indicate Gaussian independent noises with zero mean and unit variance (**Fig. 1D**). C is a scaling factor common across samples. Here, we assumed that biomarker candidate data $\mathbf{x}_n$ was generated (**Fig. 1E**) via the same biological process among animals. W is thus the shared parameter among the animals sampled from the same distribution. This model was formulated based on a hierarchical Bayesian model (see Methods).

Semi-supervised procedure to predict Aß accumulation

Based on this model, we aimed to estimate the latent Aβ accumulation z_n from the observed data of the biomarker candidate $\mathbf{x}_n$. To this end, we also needed to train the model by estimating the parameters from the data. Here, we made three assumptions, following the actual situation of available data: first, animals are sampled once as snapshots, not as time series, due to the requirement of euthanization. Second, most of the samples only had either Aβ accumulation or biomarker candidates, meaning that these were observed in different animal populations. Third, only a subset of the samples had paired data for the biomarker candidate and Aβ accumulation at the same age. Therefore, there were three kinds of data: unpaired data of Aβ accumulation, unpaired data of biomarker candidates, and paired data containing both (**Fig. 1C**).

Given the above assumption, we proposed a Bayesian probabilistic treatment to train the model by integrating the paired and incomplete unpaired data in a semi-supervised manner (see Methods). In the first step, using only the unpaired dataset of Aβ accumulation, we pre-trained the model by estimating the distribution of the logistic function parameter θ for each of the wildtype and AD model animals (**Fig. 2**). Next, using the estimated distributions in the first step as prior knowledge, we inferred the distribution of all model parameters from the remaining datasets; the unpaired dataset of biomarker candidates was used for unsupervised learning, whereas the paired dataset was used for supervised learning.

After learning the parameters, the model could predict the accumulated level of Aβ z_n from the observed data of biomarker candidate x_n (**Fig. 2**) (see Methods).

Experiment data of AD model mice

As the application of our estimation method, we adopted real-world experimental data from the AD model and wildtype (WT) mice[24]. The AD model mice were 5xFAD transgenic mice with five human familial AD mutations in APP and PSEN1 on the background of a C57BL/6J strain, showing robust Aβ pathological accumulation and neuronal cell death. For AD model mice, most of the data contained unpaired data either of the Aβ accumulation (24 samples) or behavioral features (82 samples). In contrast, some paired data existed (18 samples only for 8- and 12-month-old mice) (**Fig. 3A**). Aβ accumulation was evaluated by the insoluble fraction level of Aβ40 and Aβ42 in the hippocampus of 42 AD model mice at 4, 8, 12, and 18 months of age (**Fig. 3D**). For WT mice, the observation of Aβ is unavailable in the dataset. We assumed that the insoluble Aβ in WT mice remained undetectable during the mouse's lifetime based on the reports that WT mice do not develop Aβ plaque during their normal life span[25]. Each AD model and WT mouse were behaviorally evaluated; we selected three kinds of behavioral experiments where 11 features were obtained at 4, 8, or 12 months of age (Supplementary Table S1) (**Fig. 3B**). These 11-dimensional data were addressed as biomarker candidate data, and visualized using principal component analysis (PCA) (**Fig. 3C**).

Prediction of Aß accumulation in AD model mice

Using our method, we predicted Aβ accumulation in the hippocampus using behavioral features as biomarker candidates. First, we pre-trained the model to learn the distributions of the parameters of the logistic function from the insoluble fraction level of Aβ. The pre-trained model generated logistic time courses, representing the observed insoluble fraction level of Aβ40 and Aβ42 (**Fig. 3D**). Next, we trained the model using an unpaired dataset of behavioral features and a paired dataset.

Using the trained model, we predicted the accumulation level of Aβ from the behavioral features of the paired data using leave-one-out cross-validation (**Fig. 3E**). The prediction errors were almost the same for the two types of Aβ (mean squared error (MSE)= 0.060 for Aβ40 and MSE = 0.111 for Aβ42). The majority of the predictions followed the observed amount of Aβ both in Aβ40 and Aβ42. However, several data were unpredictable with large errors (asterisks in **Fig. 3E**). Most of the unpredictable data were from 8-month-old mice and shared between Aβ40 and Aβ42. Moreover, 8-month-old mice showed larger prediction errors than 12-month-old mice (MSE of Aβ40 = 0.083 (8 months), 0.046 (12 months); MSE of Aβ42 = 0.214 (8 months), 0.046 (12 months)). If the training was implemented unsupervised, the prediction failed (**Fig. 3F**), indicating the importance of supervised information for prediction. The predictive performance for Aβ accumulation in the cortex was not as good as that in the hippocampus (Supplementary

Fig. S1).

Selection of biomarkers for prediction of Aß accumulation

To assess the importance of each behavioral feature in predicting Aβ accumulation, we removed each behavioral feature and evaluated the prediction error for each (**Fig. 4A**). We then found that predictive performance was considerably decreased by removing "time in the center" in the open field experiment and "time spent in the open arm" in the elevated plus maze experiment, which exhibited significant differences in WT and 5xFAD mice[24]. Next, we ranked the features by their impact on the prediction error and assessed the prediction performance by including features individually from the top of the ranks (**Fig. 4B**). The prediction error decreased considerably until the top five features of the ranks were recruited, which were composed of features from three different experiments. The results show that multivariate features from different experiments could be potential AD biomarkers.

We also evaluated the number of features required for the proposed model to determine its predictability. Here, we randomly selected several behavioral features by varying their number, trained the model, and made a prediction (**Fig. 5A**). Statistically significant differences were detected between the predictions using 1–7 and 11 features (Mann–Whitney U-test, $p<0.05$). In contrast, there was no significant difference between the predictions using 10 and 11 features (Fig. 5B), suggesting that as many diverse features as possible are preferable to achieve better prediction performance.

Prediction of Aß accumulation by usual machine learning methods

The predictive performance of the proposed model was compared with that of standard machine learning techniques, namely ordinary linear and random forest regressions. We then demonstrated that our proposed model outperformed standard machine learning techniques both in prediction of Aβ40 and Aβ42 in the hippocampus (**Fig. 6A–D**). Indeed, the proposed model showed a smaller median absolute prediction error than that of the standard models (**Table 1**).

Application to synthetic data

In the 5xFAD experimental data used in this study, the paired data with Aβ accumulation and behavioral features were limited only to the phase when the accumulation process of Aβ has vastly progressed, that is, 8- and 12-month-old mice. Thus, the predictive performance of the proposed model for the earlier phase of Aβ accumulation remains unknown. To evaluate this, the model was applied to synthetic data that contained earlier phase samples.

To this end, we prepared synthetic unpaired and paired data at various ages, including the early phases, by simulating the model. The synthetic data were composed of 20 samples of Aβ

accumulation alone, 50 samples of biomarker candidates alone, and 50 samples paired for the AD model and WT (**Fig. 7A**). First, we pretrained the model using the synthesized unpaired data of Aβ accumulation, representing the variation in the observed accumulation level of Aβ (**Fig. 7B**). We then trained the model using unpaired data of biomarker candidates and paired data using semi-supervised learning. We confirmed that the estimated parameters followed the ground truth used for the synthesized data (**Fig. 7C**), indicating that our estimation method was efficient. Using the trained model, we predicted z^*, the Aβ accumulation level of unknown samples, from the observed biomarker features X^* (**Fig. 7D**, top right). We evaluated the prediction accuracy by changing the ratio of supervised paired samples for training. The MSE changed slightly with varying ratios, suggesting that not many supervised samples were required for prediction (right and bottom left in **Fig. 7D**).

Finally, we evaluated the predictive performance dependency of the samples on age. The discrimination of sample types, such as WT and AD model mice, were relatively accurate in the older samples compared to that in the younger ones (**Fig. 7E**). The predictive performance of Aβ accumulation in the AD model samples did not show a significant change according to age (**Fig. 7F**). The results suggest that the model can predict Aβ accumulation levels in samples in the early phase of the accumulation process, while the types of samples are sometimes mispredicted.

Discussion:

This study demonstrates the proof-of-concept of Aβ-predictability-based multivariate AD biomarker discovery. We proposed a hierarchical Bayesian model that describes how biomarker candidates are generated in response to the accumulation of Aβ in the hippocampus. Using this model, we predicted Aβ accumulation from the behavioral features obtained in the 5xFAD behavioral experiments. Based on the effect of each biomarker candidate on predictability, we revealed that multiple behavioral features from three different experiments could be important biomarkers for predicting Aβ accumulation level.

The proposed model can naturally integrate the information from paired and unpaired data. The inference of the distribution of the logistic function parameters from Aβ-observed unpaired data constrains the dynamic range of Aβ accumulation level. The information from unpaired data that lack an observed amount of Aβ provides the generation process of biomarker candidates and their variability. Learning information from the paired supervised data further helps to adjust the generation process. Notably, as the simultaneous observation of biomarker candidates and Aβ from the same sample is labor-intensive, expensive, and technically challenging, most samples usually lack information on either biomarker candidates or Aβ. In this situation, the proposed

model makes it possible to make such incomplete datasets available for Aβ predictability-based biomarker discovery.

The predictive performance of the proposed model was found to be inferior in young mice than in older mice. The nature of the analyzed biomarker candidates and the proposed model limitations would have regulated the results. The principal component space of the 5xFAD mouse behavioral test data showed little difference between WT and AD modelo mice in the younger samples. The observed data of biomarker candidates may be generated as a non-linear transformation from Aβ accumulation, in contrast to the linearity assumed in the proposed model. We confirmed whether the model with nonlinear transformation yields better prediction performance by adopting a 3-layer neural network at the generation from z to X. However, the prediction performance only improved slightly (data not shown).

In humans, the accumulation of Aβ may initiate 10–20 years before the recognized cognitive decline[1,2]. However, in the model animals, behaviors and cognitive abilities were altered before the saturation of Aβ accumulation. For example, 5xFAD mice had a decline in memory function before 6 months of age, when Aβ was still in the accumulation process[24,26,27]. These facts suggest that a probabilistic model sensitive to an earlier stage of Aβ accumulation is preferable to discover biomarkers in model animals. Furthermore, the gradient term of accumulated Aβ at the generation of biomarker data in the model might bestow such specificity to the proposed model when analyzing data that contain phase-specific biomarkers.

Whether Aβ acts upstream in the cascade leading to cell death, as in the amyloid hypothesis, is controversial[28,29]. However, Aβ indeed accumulates in the early stages of AD. Even if there is no causal relationship between the accumulation of Aβ and AD progression, Aβ may be a useful precursor for predicting AD. Notably, studies have reported alterations in the phenotypes of AD model mice that appear earlier than the onset of Aβ accumulation[30]. Such lesions may allow for an earlier definition of the latent state of AD progression.

Machine learning approaches for identifying the latent progression states of AD have recently attracted attention. Probabilistic models, such as mixed effect[31–35] and hidden Markov models[36–38] found the latent trajectories of the disease progression from longitudinal data of the clinical cohorts in unsupervised learning. Our approach assumed the accumulation level of Aβ as “the latent progression state” and estimated the state using semi-supervised learning. The molecular biological observation of Aβ is challenging in humans. However, the proposed framework is potentially beneficial for discovering non-invasive convenient biomarkers[8,9,39,40] that are relevant to the amount of PET-detected Aβ or CSF Aβ in human data.

In other neurodegenerative diseases, such as Parkinson’s disease, Lewy body dementia, multiple system atrophy, Huntington’s disease, amyloid lateral sclerosis, and frontotemporal lobar degeneration, abnormal proteins accumulate in specific brain regions, leading to neuronal

death[41–43]. The modeling approach presented in this paper is also potentially applicable to such neurodegenerative diseases. The risk of neurodegenerative diseases is a growing concern in an aging society. Predictability-based biomarker discovery by the proposed model may contribute to identifying biomarkers that make available predictions and potential interventions for diseases.

Methods

The generative model of Aß and biomarker candidates

In the model, Aß is assumed monotonically accumulated as the logistic function:

$$z_n = \frac{\alpha_n}{1 + exp(-\beta_n t_n + \gamma_n)},$$

where $\theta_n = \{\alpha_n, \beta_n, \gamma_n\}$ is a set of parameters, and n is the animal index. This equation can be rewritten from Equation (1), where $\gamma_n = \beta_n \tau_n$. Accordingly, the temporal derivative of the Aβ accumulation is

$$z'_n = \beta_n z_n (1 - \frac{z_n}{\alpha_n}).$$

The parameters θ_n depend on the individual, following distributions as

$$P(\alpha_n) = \prod_{k\in\{1,2\}} \mathcal{N}_+\left(\alpha_n | \mu_{\alpha,k}, \sigma^2_{\alpha,k}\right)^{s_{n,k}},$$

$$P(\beta_n) = \prod_{k\in\{1,2\}} \mathcal{N}_+\left(\beta_n | \mu_{\beta,k}, \sigma^2_{\beta,k}\right)^{s_{n,k}},$$

$$P(\gamma_n) = \prod_{k\in\{1,2\}} \mathcal{N}\left(\gamma_n | \mu_{\gamma,k}, \sigma^2_{\gamma,k}\right)^{s_{n,k}},$$

where $\mathcal{N}(x|\mu, \sigma^2)$ and $\mathcal{N}_+(x|\mu, \sigma^2)$ indicate normal distribution with mean μ and variance σ^2 and truncated normal distribution with range of $x > 0$, respectively; $\mathbf{s}_n$ is one-hot vector representing the type of the samples, i.e., wildtype as $(1,0)^{\mathrm{T}}$ or AD model mouse as $(0,1)^{\mathrm{T}}$; $\mu_{\phi,k}$ and ${\sigma^2}_{\phi,k}$ ($\phi \in \{\alpha, \beta, \gamma\}$) indicates parameters of wildtype ($k = 1$) or AD model mouse ($k = 2$).

The observed amount of Aß, y_n, is generated from z_n as

$$P(y_n | z_n) = \mathcal{N}\left(y_n | z_n, \sigma^2_y\right).$$

The observed data of L-dimensional biomarker candidates $\mathbf{x}_n \in \mathbb{R}^L$ was generated from z_n and z'_n as

$$P(\mathbf{x}_n | \mathbf{z}_n) = \mathcal{N}(\mathbf{x}_n | W\mathbf{z}_n, \Sigma_x),$$

where $\mathbf{z}_n = (z_n,\ Cz'_n, 1)^{\mathrm{T}}$, $W \in \mathbb{R}^{L\times 3}$ indicates weight matrix, and $\Sigma_x = diag(\sigma^2_{x_1}, \sigma^2_{x_2}, \ldots, \sigma^2_{x_L})$. C indicates a scaling factor common among samples that adjusts the range of z'.

The prior distribution of parameters

For parameter estimation in a Bayesian manner, we introduced the prior distributions of parameters. The hyper-parameters of $P(\alpha_n)$, $P(\beta_n)$ and $P(\gamma_n)$ are sampled from the following distributions:

$$P(\mu_{\phi,k}) = \mathcal{N}\left(\mu_{\phi,k} \middle| m_{\phi,k}, v_{\phi,k}^2\right),$$
$$P\left(\sigma_{\phi,k}^{-2}\right) = \mathrm{Gamma}\left(\sigma_{\phi,k}^{-2} \middle| a_{\phi,k}, b_{\phi,k}\right),$$

where $\phi \in \{\alpha, \beta, \tau\}$, $\mathrm{Gamma}(x|a,b)$ indicates a Gamma distribution with shape parameter a and rate parameter b.

The prior distributions of W was

$$P(\mathbf{w}_l) = \mathcal{N}\left(\mathbf{w}_l | \mathbf{0}, \sigma_{w_l}^2 \mathbf{I}\right),$$

where $\mathbf{w}_l$ indicates the l-row of the weight matrix W. Prior of its hyper-parameter $\sigma_{w_l}^2$ was hierarchically introduced as

$$P\left(\sigma_{W_l}^{-2}\right) = \mathrm{Gamma}\left(\sigma_{W_l}^{-2} \middle| a_w, b_w\right).$$

The prior distributions of $\sigma_{x_l}^2$ was

$$P\left(\sigma_{x_l}^{-2}\right) = \mathrm{Gamma}\left(\sigma_{x_l}^{-2} \middle| a_x, b_x\right).$$

In the MCMC, we used $v_{\alpha,k}^2 = 0.1$, $a_{\alpha,k} = 25$, $b_{\alpha,k} = 0.5$, $v_{\beta,k}^2 = 0.1$, $a_{\beta,k} = 100$, $b_{\beta,k} = 1$, $v_{\gamma,k}^2 = 1$, $a_{\gamma,k} = 10$, $b_{\gamma,k} = 1$ $v_{\alpha,k} = 0.1$, $a_{\alpha,k} = 25$, $b_{\alpha,k} = 0.5$. $a_w = 100$ $b_w = 1000$ $a_x = 0.5$ and $b_x = 1$ and $\sigma_y^2 = 0.05$.

When the model learned the distributions of the hyper-parameters from the Aß observation (step 1), we set $m_{\alpha,k}$, $m_{\beta,k}$ and $m_{\gamma,k}$ as the values estimated by the least square method.

Bayesian inference of parameters

The model parameters were learned in two steps. In the first step (step 1), we inferred the posterior distributions of the parameters of a logistic function and those of hyper-parameters, given the data of Aβ accumulation as

$$P\left(\theta_{1:N_y}, \mu_\theta, \sigma_\theta^2 \middle| y_{1:N_y}, t_{1:N_y}, \mathbf{s}_{1:N_y}\right)$$
$$\propto \prod_{n \in \mathcal{S}_y} P\left(y_n \middle| \theta_n, t_n, \sigma_y^2\right) \prod_{k \in \{1,2\}} P\left(\theta_n \middle| \mu_{\theta,k}, \sigma_{\theta,k}^2, \mathbf{s}_n\right)^{s_{n,k}} P(\mu_{\theta,k}) P(\sigma_{\theta,k}^2),$$

where $\mu_\theta = \{\mu_{\theta,1}, \mu_{\theta,2}\}$, $\mu_{\theta,k} = \{\mu_{\alpha,k}, \mu_{\beta,k}, \mu_{\gamma,k}\}$, $\sigma_\theta^2 = \{\sigma_{\theta,1}^2, \sigma_{\theta,2}^2\}$, $\sigma_{\theta,k}^2 = \{\sigma_{\alpha,k}^2, \sigma_{\beta,k}^2, \sigma_{\gamma,k}^2\}$, $\mathcal{S}_y$ is the set of samples with Aβ accumulation in the training data, and N_y is the number of samples in $\mathcal{S}_y$. We assumed that the accumulation level of Aβ $z_n = 0$ at time $t_n = 0$ in all samples. This posterior distribution was estimated using the MCMC sampling algorithms. No-U-Turn samplers (NUTS) were used in this step because it is difficult to derive the closed form of the posterior distribution. From the posterior samples, we estimated the parameters of the Gaussian and Gamma distributions for $\mu_{\theta,k}$ and $\sigma_{\theta,k}^2$, respectively, and these distributions with the estimated parameters were used as priors for the hyper-parameters. In the first step, MCMC

sampling was performed across three independent chains, where 3,000 samples were drawn for burn-in, and another 3,000 were drawn to estimate the distribution.

In the second step (step 2), we inferred the posterior distributions of all parameters in the model, given the unpaired data of biomarker candidates and the paired data of Aβ accumulation and biomarker candidates. We used the NUTS-within-Gibbs approach for the inference. The weight matrix W, the variance of the weight matrix $S_W = \{\sigma_{W_1}^2, \sigma_{W_2}^2, \ldots, \sigma_{W_L}^2\}$, the variance of the observation noise σ_x were sampled using Gibbs sampling as follows:

The Gibbs sampler for weight matrix W:

$$P\left(W \middle| \mathbf{x}_{1:N_x}, y_{1:N_x}, t_{1:N_x}, \mathbf{s}_{1:N_x}, \Sigma_x, S_W, \theta_{1:N_x}, \mu_\theta, \sigma_\theta^2\right) \propto \prod_{n \in \mathcal{S}_X} P(\mathbf{x}_n | W, \mathbf{z}_n, \Sigma_x) P(W | S_W)$$
$$= \prod_{l=1} \mathcal{N}\left(\mathbf{w}_l \middle| \tilde{\mu}_{w_l}, \tilde{\sigma}_{w_l}^2 I\right),$$

where $\mathcal{S}_X$ is a set of samples at least having biomarker candidates in the training data, N_X is the number of samples in $\mathcal{S}_X$, and

$$\tilde{\mu}_{w_l} = \tilde{\sigma}_{w_l}^2 \sum_{n \in \mathcal{S}_X} \frac{\mathbf{z}_n x_{n,l}}{\sigma_{x_l}^2},$$

$$\frac{1}{\tilde{\sigma}_{w_l}^2} = \sum_{n \in \mathcal{S}_X} \frac{\mathbf{z}_n \mathbf{z}_n^{\mathrm{T}}}{\sigma_{x_l}^2} + \frac{1}{\sigma_{w_l}^2}.$$

The Gibbs sampler for observation noise of biomarker candidates S_x:

$$P\left(S_x^{-2} \middle| \mathbf{x}_{1:N_x}, y_{1:N_x}, t_{1:N_x}, \mathbf{s}_{1:N_x}, W, S_W, \theta_{1:N_x}, \mu_\theta, \sigma_\theta^2\right) \propto \prod_{n \in \mathcal{S}_X} P(\mathbf{x}_n | W, \mathbf{z}_n, \Sigma_x) \prod_{l=1} P\left(\sigma_{x_l}^{-2}\right)$$
$$= \prod_{l=1} Gamma\left(\sigma_{x_l}^{-2} \middle| \tilde{a}_{x_l}, \tilde{b}_{x_l}\right),$$

where

$$\tilde{a}_{x_l} = a_{x_l} + \frac{2}{N_X},$$

$$\tilde{b}_{x_l} = b_{x_l} + \sum_{n \in \mathcal{S}_X} \frac{1}{2}\left(x_{n,l} - \mathbf{w}_l^{\mathrm{T}} \mathbf{z}_n\right)^2.$$

The Gibbs sampler for variance of the coefficient matrix S_W:

$$P\left(S_W^{-1} \middle| \mathbf{x}_{1:N_x}, y_{1:N_x}, t_{1:N_x}, \mathbf{s}_{1:N_x}, W, \Sigma_x, \theta_{1:N_x}, \mu_\theta, \sigma_\theta^2\right) \propto p(W | S_w) \prod_{l=1} P\left(\sigma_{w_l}^{-2}\right)$$
$$= \prod_{l=1} Gamma\left(\sigma_{W_l}^{-2} \middle| \tilde{a}_{w_l}, \tilde{b}_{w_l}\right),$$

where

$$\tilde{a}_{w_l} = a_{w_l} + \frac{3}{2},$$

$$\tilde{b}_{w_l} = b_{w_l} + \frac{1}{2}\mathbf{w}_l^T\mathbf{w}_l.$$

The posterior distribution of other parameters was sampled by MCMC using NUTS as

$$P\left(\theta_{1:N_{pair}}, \mu_\theta, \sigma_\theta^2 \,\middle|\, \mathbf{x}_{1:N_x}, z^o_{1:N_x}, t_{1:N_x}, \mathbf{s}_{1:N_x}, W, \Sigma_x, S_W\right)$$

$$= P(\mu_\theta)P(\sigma_\theta^2)\left[\prod_{n\in\mathcal{S}_{XY}} P(\mathbf{x}_n|\mathbf{z}_n, W, \Sigma_x)P(y_n|\mathbf{z}_n, \sigma_z^2)P\left(\theta_n\middle|\mu_\theta, \sigma_\theta^2, \mathbf{s}_n\right)\right]\times$$

$$\left[\prod_{n'\in\mathcal{S}_X} P(\mathbf{x}_{n'}|\mathbf{z}_{n'}, W, \Sigma_x)P\left(\theta_{n'}\middle|\mu_\theta, \sigma_\theta^2, \mathbf{s}_{n'}\right)\right]$$

where $\mathcal{S}_{XY}$ and $\mathcal{S}_X$ are sets of samples with paired data of Aβ accumulation and biomarker candidates and the unpaired data of biomarker candidates, respectively. We adopt the distributions estimated in Step 1 as $P(\mu_\theta)$ and $P(\sigma_\theta^2)$. In the second step, MCMC sampling was performed across three independent chains, where 5,000 samples were drawn for burn-in, and another 5,000 were drawn to estimate the distribution. The inference program was implemented with Python using the NumPyro framework.

Prediction of Aß accumulation from biomarkers

In the prediction of Aß accumulation in the test data (step 3), we computed a conditional posterior predictive distribution of y^* given $\mathbf{x}^*$ as the following equation:

$$P(y^*|\mathbf{x}^*, D) \propto P(\mathbf{x}^*, y^*|\, D)$$

$$\propto \frac{1}{2N_t I}\sum_{i=1}\sum_{t}\sum_{s} P\left(\mathbf{x}^*\middle|t, s, W^{(i)}, \Sigma_x^{(i)}, S_W^{(i)}, \theta^{(i)}\right)P\left(y^*\middle|t, s, \theta^{(i)}\right),$$

where D is the learned training data, $W^{(i)}$, $\Sigma_x^{(i)}$, $S_W^{(i)}$, and $\theta^{(i)}$ are the posterior samples of the parameters, I is the number of the posterior samples and N_t is the number of time points considered in the prediction $t = \{2, 3, \ldots, 18\}$.

Behavioral experiments with 5xFAD mice

We used the dataset from Forner et al.[24], freely available in the public repository. Eleven features from three experiments were analyzed using our proposed model. In the open-field experiment, the velocity and time ratio in the center (the time in the center divided by the time in the arena)

was used in the analysis. In the elevated plus maze experiment, the amount of time a mouse spent cumulatively in the open arm, closed arm, and center area of the maze was used in the analysis. In the contextual fear conditioning experiment, the activity level, inactive freezing frequency, and cumulative duration of inactive freezing were monitored for each mouse during a 2-min habituation and exploration in a chamber. Subsequently, electrical shock was applied to the mouse. After 24 h, the same behavioral features were monitored for 5 min in the chamber.

Preprocessing

Preprocessing was performed to analyze 5xFAD mouse behavioral data. Behavioral features were standardized such that the mean of each feature was zero and the standard deviation of each feature was 1.0. The observed amount of Aβ was scaled so that the maximum observed value from 12 months of age mice equals 1.0. Based on the assumption that insoluble Aβ in the brain of WT mice remains undetectable throughout their lives, we virtually generated WT samples at 8, 12, and 18 months of age, where the observed amount of Aβ at each time sample was 0.0. A 5xFAD mouse “individual ID = 572” was excluded from the paired-data samples because the measurement of insoluble Aβ in the sample may have failed.

Acknowledgments

This work is supported in part by Moonshot R&D–MILLENNIA Program [grant number JPMJMS2024-9] by JST.

Author Contributions

Y.Y. and H.N. conceived the project and developed the method. Y.Y. implemented the software and analyzed data. Y.Y. and H.N. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Data availability

Datasets for the current study are available from Forner et al., *Scientific Data*, 2021 (https://doi.org/10.1038/s41597-021-01054-y).

Figures

Fig. 1

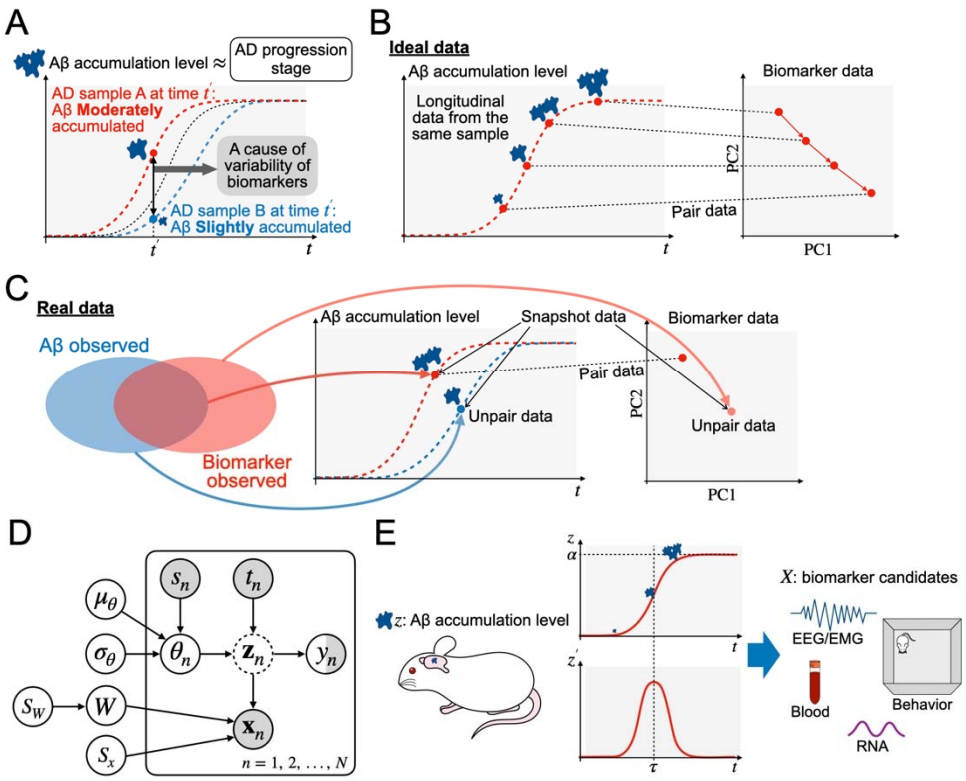

Fig. 1: Concept and graphical model of the proposed model.

(A) Alzheimer's disease (AD) model samples (A and B) at the same age may have unignorably different amounts of accumulated amyloid beta (Aβ), which may correlate with the progression stage of early AD. (B) Ideally, Aβ-accumulation-relevant biomarkers are estimated from the longitudinal paired data of Aβ observation and biomarker candidate observation simultaneously. (C) In a real experiment, Aβ and biomarker candidates are usually observed primarily in distinct groups of samples (unpaired data). The limited number of samples may have paired data. Each sample has data from a single time point (snapshot). (D) The graphical model representation of our proposed model. (E) The concept of our model. Aβ accumulates in a particular brain region according to a logistic function of time. Observed data of biomarker candidates are generated depending on the accumulation level and the instantaneous accumulation speed of Aβ in the brain.

Fig. 2

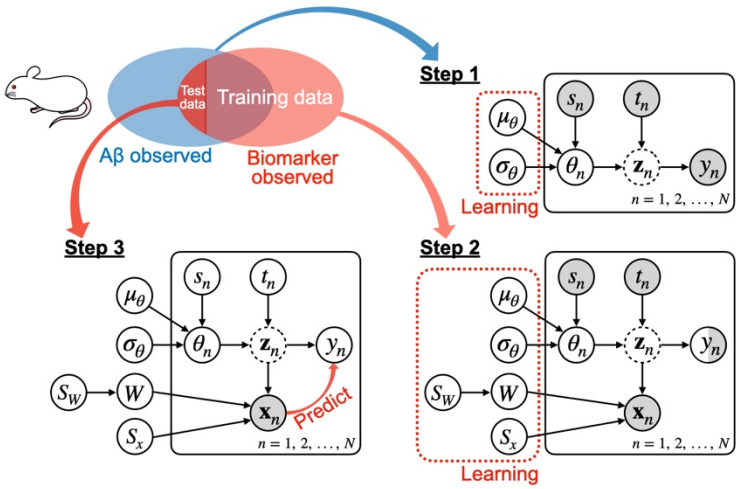

Fig. 2: Learning and prediction steps of the proposed model.

The model learns the parameters that integrate the incomplete data and predicts the Aβ accumulation from the observed data. In the first step, the distributions of the hyperparameters are inferred from the observed Aβ value. In the second step, the distributions of all parameters and hyperparameters were updated using the observed biomarker candidate data. Third, the predictability of the accumulated level of Aβ from the observed biomarker candidates was evaluated.

Fig. 3

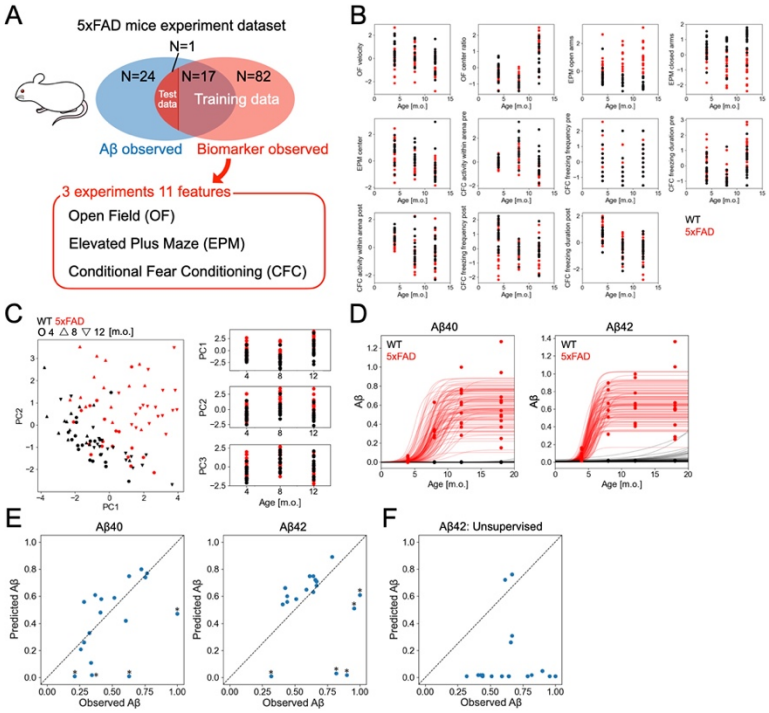

Fig. 3: Application of the proposed model to 5xFAD mice data

(A) Components of the experimental data. The accumulation levels of Aβ were observed in 42 samples. The biomarker candidates, 11 features from 3 different behavioral experiments, were observed in 100 samples. The two groups partially overlapped: 18 samples have Aβ and biomarker candidate data. The predictability is evaluated by the leave-one-out cross-validation. (B) Time course of the behavioral features. The black points are from WT mice, and the red points are from 5xFAD mice. (C) PCA of the behavioral features obtained from the experiment. The red points correspond to 5xFAD samples, and the black points correspond to WT samples. In the scatter plot (left), the points' shape represents the samples' months of age (circle: 4, triangle: 8, inverted triangle: 12). (D) Inference of hyper-parameters of a logistic function. The data points are scaled insoluble fractions of $A\beta_{40}$ and $A\beta_{42}$ at the hippocampus. The red points are of 5xFAD mice, and the black points are of WT mice. The lines represent the example logistic functions with hyper-parameters randomly sampled from the learned distribution (red: 5xFAD, black: WT). (E) Accumulation levels of $A\beta_{40}$ and $A\beta_{42}$ at the hippocampus predicted by the trained model from the behavioral features. (F) The accumulation level of Aβ42 at the hippocampus predicted by the unsupervised-trained model.

Fig. 4

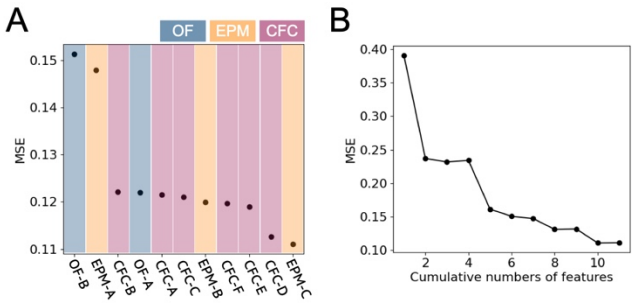

Fig. 4: Evaluation of the behavioral features based on the predictability of the Aβ accumulation level.

(A) Prediction error in the case of excluding a particular behavioral feature from the observation. (B) Variation of the prediction performance when the number of the observed features increases, where the features are sequentially included from the top of the list in Fig. 4A.

Fig. 5

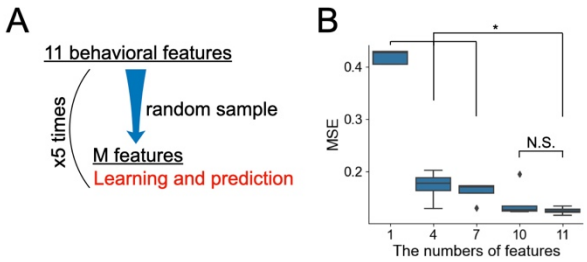

Fig. 5: Learnings and predictions in randomly sampled features.

(A) M features are randomly sampled from 11 behavioral features; the model learns the data using the M features, and the prediction performance is evaluated. (B) Mean squared errors between the observed amount of Aβ ($A\beta_{42}$ at the hippocampus) and the predicted amount using $M = \{1, 4, 7, 10, 11\}$ features for the learning and prediction. M features were sampled five times. (*p<0.05. Mann–Whitney U-test).

Fig. 6

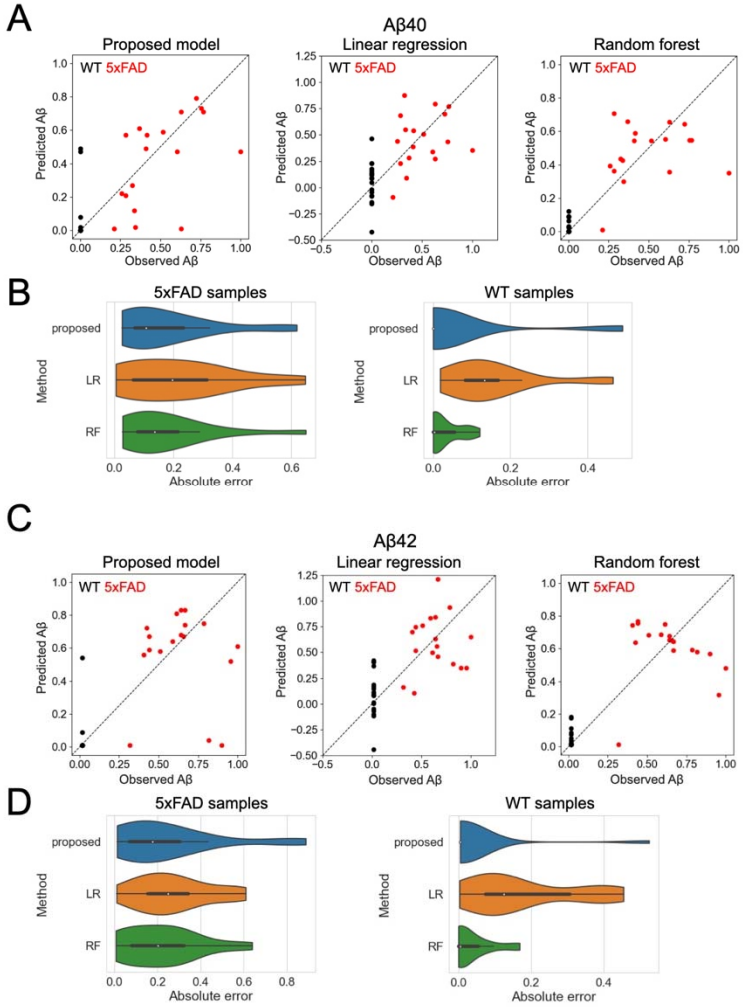

Fig. 6: Comparison with standard machine learning methods.

(A, C) Predictive performances against (A) Aβ40 and (C) Aβ42 of the proposed model, linear regression, and random forest regression. Eighteen virtually prepared supervised WT samples were used as test samples; the Aβ accumulation levels of such samples were assumed to be minimum quantity in the observation of Aβ. Three 5xFAD mice and three WT mice were used as test samples and evaluated with 6-fold cross-validation. (B, D) Distributions of absolute prediction errors when the accumulation levels of (B) Aβ40 and (D) Aβ42 were predicted for 5xFAD samples (left) and WT samples (right).

Fig. 7

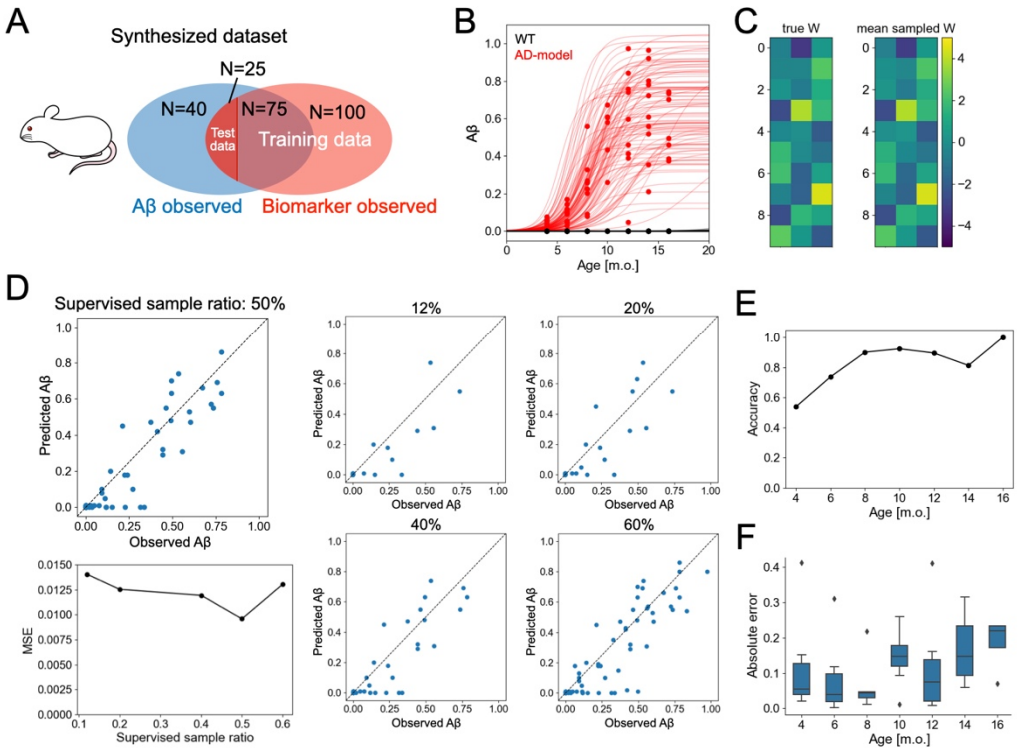

Fig. 7: Learning and prediction performance of the proposed model is evaluated by synthesized data.

(A) Components of the synthesized data. The accumulation levels of Aβ were observed in 140 samples. The biomarker candidates were observed in 200 samples. The two groups partially overlap: 100 samples have Aβ and biomarker candidate data. The ages of the samples are randomly sampled from $t_n = \{4, 6, \ldots, 16\}$ and WT and AD model samples are in the same proportion. The predictability is evaluated by 4-fold cross-validation. (B) Inference of hyper-parameters of a logistic function from the observed amount of Aβ. The red points are the observed value of AD model samples, and the black points are those of WT samples. The example logistic functions with hyper-parameters are randomly sampled from the learned distribution are represented with red (AD model) and black (WT) lines. (C) Representative result of the inference of the model parameters from synthetically generated biomarker data. The left is true generated W, and the right is the mean of the posterior samples of W. (D) Prediction of the accumulation level of Aβ from the observed biomarker data by the learned model. The mean prediction errors according to the ratio of supervised samples (bottom left) and the prediction results at the respective conditions (right). (E, F) Accuracy of predicting the type of the samples (E) and absolute prediction errors in the AD model samples (F) according to the age of the samples.

Table 1: Comparison of predictive performance with standard machine learning techniques.

Method	Aß	Sample type	Median absolute prediction error
Proposed	$Aß_{40}$	5xFAD	0.108
		WT	0.0
	$Aß_{42}$	5xFAD	0.177
		WT	0.005
Linear regression (LR)	$Aß_{40}$	5xFAD	0.197
		WT	0.133
	$Aß_{42}$	5xFAD	0.249
		WT	0.125
Random forest (RF)	$Aß_{40}$	5xFAD	0.136
		WT	0.003
	$Aß_{42}$	5xFAD	0.208
		WT	0.005

References

1. Long, J. M. & Holtzman, D. M. Alzheimer disease: An update on pathobiology and treatment strategies. *Cell* **179**, 312–339 (2019).
2. Bateman, R. J. *et al.* Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* **367**, 795–804 (2012).
3. Bekris, L. M., Yu, C.-E., Bird, T. D. & Tsuang, D. W. Genetics of Alzheimer disease. *J. Geriatr. Psychiatry Neurol.* **23**, 213–227 (2010).
4. Lambert, J.-C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
5. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
6. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019).
7. Kondo, T. *et al.* Dissection of the polygenic architecture of neuronal Aβ production using a large sample of individual iPSC lines derived from Alzheimer's disease patients. *Nature Aging* **2**, 125–139 (2022).
8. Karikari, T. K. *et al.* Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. *Lancet Neurol.* **19**, 422–433 (2020).
9. Schindler, S. E. *et al.* High-precision plasma β-amyloid 42/40 predicts current and future brain amyloidosis. *Neurology* **93**, e1647–e1659 (2019).
10. Eyigoz, E., Mathur, S., Santamaria, M., Cecchi, G. & Naylor, M. Linguistic markers predict onset of Alzheimer's disease. *EClinicalMedicine* **28**, 100583 (2020).
11. Jack, C. R., Jr *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* **12**, 207–216 (2013).
12. Jaunmuktane, Z. *et al.* Evidence for human transmission of amyloid-β pathology and cerebral amyloid angiopathy. *Nature* **525**, 247–250 (2015).
13. Lam, S. *et al.* Transmission of amyloid-beta and tau pathologies is associated with cognitive impairments in a primate. *Acta Neuropathol Commun* **9**, 165 (2021).
14. Vogel, J. W. *et al.* Spread of pathological tau proteins through communicating neurons in human Alzheimer's disease. *Nat. Commun.* **11**, 1–15 (2020).
15. Collij, L. E. *et al.* Spatial-Temporal Patterns of β-Amyloid Accumulation. *Neurology* **98**, e1692–e1703 (2022).

16. Collij, L. E. *et al.* Multitracer model for staging cortical amyloid deposition using PET imaging. *Neurology* **95**, e1538–e1553 (2020).
17. Grothe, M. J. *et al.* In vivo staging of regional amyloid deposition. *Neurology* **89**, 2031–2038 (2017).
18. Palmqvist, S. *et al.* Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. *Neurology* **85**, 1240–1249 (2015).
19. Caminiti, S. P. *et al.* FDG-PET and CSF biomarker accuracy in prediction of conversion to different dementias in a large multicentre MCI cohort. *NeuroImage: Clinical* **18**, 167–177 (2018).
20. Therriault, J. *et al.* Biomarker modeling of Alzheimer's disease using PET-based Braak staging. *Nature Aging* 1–10 (2022) doi:10.1038/s43587-022-00204-0.
21. Vitek, M. P. *et al.* Translational animal models for Alzheimer's disease: An Alzheimer's Association Business Consortium Think Tank. *Alzheimers. Dement.* **6**, e12114 (2020).
22. Sabbagh, J. J., Kinney, J. W. & Cummings, J. L. Alzheimer's disease biomarkers: Correspondence between human studies and animal models. *Neurobiol. Dis.* **56**, 116–130 (2013).
23. Sasaguri, H. *et al.* APP mouse models for Alzheimer's disease preclinical studies. *EMBO J.* **36**, 2473–2487 (2017).
24. Forner, S. *et al.* Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. *Sci. Data* **8**, 270 (2021).
25. Xu, G. *et al.* Murine Aβ over-production produces diffuse and compact Alzheimer-type amyloid deposits. *Acta Neuropathol Commun* **3**, 72 (2015).
26. Jawhar, S., Trawicka, A., Jenneckens, C., Bayer, T. A. & Wirths, O. Motor deficits, neuron loss, and reduced anxiety coinciding with axonal degeneration and intraneuronal Aβ aggregation in the 5XFAD mouse model of Alzheimer's disease. *Neurobiol. Aging* **33**, 196.e29–40 (2012).
27. Richard, B. C. *et al.* Gene Dosage Dependent Aggravation of the Neurological Phenotype in the 5XFAD Mouse Model of Alzheimer's Disease. *J. Alzheimers. Dis.* **45**, 1223–1236 (2015).
28. Holmes, C. *et al.* Long-term effects of $A\beta_{42}$ immunisation in Alzheimer's disease: follow-up of a randomised, placebo-controlled phase I trial. *Lancet* **372**, 216–223 (2008).
29. Dickson, D. W. *et al.* Correlations of synaptic and pathological markers with cognition of the elderly. *Neurobiol. Aging* **16**, 285–98; discussion 298-304 (1995).
30. Jacobsen, J. S. *et al.* Early-onset behavioral and synaptic deficits in a mouse model of Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5161–5166 (2006).
31. Wang, Z. *et al.* AD risk score for the early phases of disease based on unsupervised

machine learning. *Alzheimers. Dement.* **16**, 1524–1533 (2020).

32. Raket, L. L. Statistical Disease Progression Modeling in Alzheimer Disease. *Front Big Data* **3**, 24 (2020).
33. Platero, C. Categorical predictive and disease progression modeling in the early stage of Alzheimer's disease. *J. Neurosci. Methods* **374**, 109581 (2022).
34. Bilgel, M., Prince, J. L., Wong, D. F., Resnick, S. M. & Jedynak, B. M. A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. *Neuroimage* **134**, 658–670 (2016).
35. Ishida, T. *et al.* A Novel Method to Estimate Long-Term Chronological Changes From Fragmented Observations in Disease Progression. *Clin. Pharmacol. Ther.* **105**, 436–447 (2019).
36. Sukkar, R., Katz, E., Zhang, Y., Raunig, D. & Wyman, B. T. Disease progression modeling using Hidden Markov Models. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2012**, 2845–2848 (2012).
37. Williams, J. P., Storlie, C. B., Therneau, T. M., Jr, C. R. J. & Hannig, J. A Bayesian Approach to Multistate Hidden Markov Models: Application to Dementia Progression. *J. Am. Stat. Assoc.* **115**, 16–31 (2020).
38. Liu, Y.-Y., Li, S., Li, F., Song, L. & Rehg, J. M. Efficient Learning of Continuous-Time Hidden Markov Models for Disease Progression. *Adv. Neural Inf. Process. Syst.* **28**, 3599–3607 (2015).
39. Hampel, H. *et al.* Blood-based biomarkers for Alzheimer disease: mapping the road to the clinic. *Nat. Rev. Neurol.* **14**, 639–652 (2018).
40. Leuzy, A. *et al.* Blood-based biomarkers for Alzheimer's disease. *EMBO Mol. Med.* **14**, e14408 (2022).
41. Soto, C. & Estrada, L. D. Protein Misfolding and Neurodegeneration. *Arch. Neurol.* **65**, 184–189 (2008).
42. Shastry, B. S. Neurodegenerative disorders of protein aggregation. *Neurochem. Int.* **43**, 1–7 (2003).
43. Spires-Jones, T. L., Attems, J. & Thal, D. R. Interactions of pathological proteins in neurodegenerative diseases. *Acta Neuropathol.* **134**, 187–205 (2017).