



## OPEN Toward precision medicine using a “digital twin” approach: modeling the onset of disease-specific brain atrophy in individuals with multiple sclerosis

Steven Cen<sup>1✉</sup>, Mulugeta Gebregziabher<sup>2</sup>, Saeed Moazami<sup>3</sup>, Christina J. Azevedo<sup>4</sup> & Daniel Pelletier<sup>4</sup>

Digital Twin (DT) is a novel concept that may bring a paradigm shift for precision medicine. In this study we demonstrate a DT application for estimating the age of onset of disease-specific brain atrophy in individuals with multiple sclerosis (MS) using brain MRI. We first augmented longitudinal data from a well-fitted spline model derived from a large cross-sectional normal aging data. Then we compared different mixed spline models through both simulated and real-life data and identified the mixed spline model with the best fit. Using the appropriate covariate structure selected from 52 different candidate structures, we augmented the thalamic atrophy trajectory over the lifespan for each individual MS patient and a corresponding hypothetical twin with normal aging. Theoretically, the age at which the brain atrophy trajectory of an MS patient deviates from the trajectory of their hypothetical healthy twin can be considered as the onset of progressive brain tissue loss. With a tenfold cross validation procedure through 1000 bootstrapping samples, we found the onset age of progressive brain tissue loss was, on average, 5–6 years prior to clinical symptom onset. Our novel approach also discovered two clear patterns of patient clusters: earlier onset versus simultaneous onset of brain atrophy.

The Digital Twin (DT) concept was first introduced in 2002 as a Fourth Industrial Revolution (Industry4.0) solution to manufacturing intelligence<sup>1</sup>. It was later brought into the medical field as a potential solution for precision medicine, the so-called Health Digital Twin (HDT). In the context of precision medicine, the HDT approach can be defined as “virtual mirror of ourselves that allows us to simulate our personal medical history and state of health using data-driven analytical algorithms and theory-driven physical knowledge”<sup>2</sup>. Recently, HDT has been applied in multiple disease areas, such as oncology<sup>3–5</sup>, geriatrics<sup>6</sup>, cardiology<sup>7–12</sup>, infectious disease<sup>13</sup>, genomic medicine<sup>14</sup>, neurodegenerative diseases<sup>15,16</sup>, vascular medicine<sup>17</sup>, and mental health<sup>18</sup>. The applications of HDT include patient safety, wellbeing management, and health care decision support<sup>19</sup>. Most of the published literature is oriented to system design or concept illustration. There are no standardized methodologies for HDT to date<sup>20,21</sup>. The implementation of HDT includes several major components such as personal devices, AI algorithms, right data to train the AI, and Internet of Things (IoT) for rapid data synchronization while providing real time decision-making<sup>22</sup>.

Towards this AI based solution, identifying appropriate statistical models for specific data structures is critical. In this study we demonstrate an application for modeling neurodegeneration data, specifically brain atrophy data measured by magnetic resonance imaging (MRI) in multiple sclerosis (MS) individuals. For a chronic disease like MS, we must realize that neurodegeneration is a long process occurring over decades<sup>15</sup>. The decision-making for patient care such as drug selection should be based on the entire disease course and not only over a short period of time<sup>23,24</sup>. Moreover, the response to treatment may take years to observe. Therefore, the HDT for a patient with neurodegenerative disease must cover the entire disease-span or even lifespan of patients.

<sup>1</sup>Department of Radiology/Neurology, University of Southern California, Los Angeles, USA. <sup>2</sup>Department of Public Health Sciences, Medical University of South Carolina, Charleston, USA. <sup>3</sup>Department of Aerospace and Mechanical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, USA. <sup>4</sup>Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, USA. ✉email: cen@usc.edu

MS is a chronic, immune-mediated, inflammatory and neurodegenerative disorder of the central nervous system and the most common cause of nontraumatic neurologic disability in young adults, affecting over 1,000,000 people in the U.S. and 2.8 million worldwide<sup>25</sup>. Clinically, the diagnosis is defined by the presence of typical neurological symptoms and demyelinating-appearing white matter lesions MRI<sup>26</sup>. Brain atrophy is considered a fundamental aspect of MS, occurring about 3 × faster in MS patients than in healthy controls and likely representing the net accumulation of tissue damage due to the disease<sup>27</sup>. As a major relay nucleus, the thalamus is particularly susceptible to neurodegeneration and has been shown to be one of the earliest regions impacted by atrophy in MS<sup>28,29</sup>. Interestingly, MS plaques and thalamic atrophy can be observed on MRI several years before the onset of first clinical symptoms<sup>28,30</sup>, suggesting that the biological onset of the disease may precede the clinical onset by several years. As such, the ‘true’ biological onset of MS remains unknown. This represents a major barrier to understanding the earliest events in the MS pathophysiology and even the natural history of MS, which is typically based on clinical disease duration.

Using brain atrophy (and more specifically, thalamic atrophy) is an appealing application of the HDT concept in MS. Brain atrophy occurs as part of normal aging, which has been studied extensively in healthy individuals<sup>31</sup>. In the absence of disease, brain volume trajectories are relatively predictable; in fact, recent work has presented normative brain growth charts across the human lifespan<sup>32</sup>. In principle, healthy brain trajectories can be leveraged to create HDTs for patients with neurologic and psychiatric diseases. In the case of MS, one could estimate when an individual MS patient’s thalamic atrophy trajectory deviates from that of a healthy individual. The onset of progressive brain tissue loss should be closer to the true biological onset of the disease, and may further our fundamental understanding of MS.

Using normalized thalamic volumes (thalamic volume/intracranial volume\*1000) from brain MRI images, our main objective was to develop statistical learning models to estimate when the thalamic atrophy trajectory of an MS patient deviated from their expected thalamic atrophy trajectory based on their corresponding HDT. The age when MS atrophy trajectory departed from normal aging was defined as the onset of progressive brain tissue loss. We hypothesized that the age of progressive brain tissue loss would be statistically earlier than the age of clinical onset determined by clinicians.

## Challenges

The first challenge is to identify a large longitudinal normal aging dataset with subjects imaged across several decades using brain MRI scans to create the HDT as a normal aging reference for a given MS patient. However, such a longitudinal dataset rarely exists. Most of the datasets collected to study normal aging are cross-sectional. As such, it is difficult to gather real-life datasets to generate reliable trajectories over the entire lifespan.

Even if we had repeated scans for both MS and normal aging, the best statistical method to fit an accurate trajectory curve has not yet been identified. Moreover, it has been shown that the aging brain trajectory is not linear<sup>32–39</sup>. The conventional statistical approach for longitudinal data is a mixed model using year(s) at study entry as the time unit to fit a linear or quadratic trend. However, linear or quadratic approaches may not be the most effective method for representing the complexity of aging data<sup>33,40</sup>, as they can result in biased estimates and low power in statistical tests<sup>40</sup>. On the other hand, as a nonparametric method, a spline model is recommended for its flexibility and robustness to accurately model the age trajectories of neuroimaging markers<sup>39</sup>. There remain several unanswered questions, including whether it is possible to augment longitudinal data from cross-sectional data, whether lifespan data can be augmented from only a few longitudinal data points, how to choose an appropriate spline setting from many mixed spline candidates, and how to select covariates having two-way or three-way interactions with the spline slope.

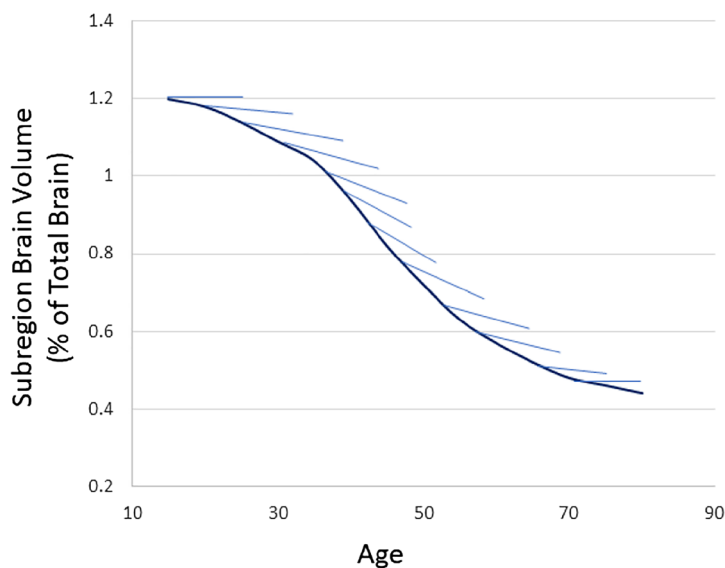
## Study design

### Overall study design and concept

Most longitudinal MRI datasets only cover a few years of an individual’s lifespan. For such a short period, when using years of follow-up as the time variable, a linear trend may be the best fit to the data, even though true brain atrophy over lifespan is non-linear. However, when using the actual age in years as the time variable, the model will look very different. For an entire sample, age has a wide coverage for the lifespan, but for each individual, age only covers a small fragment of the lifespan. This data structure can be conceptualized as a “fish bone” (Fig. 1), where the constructed spline curve can be considered the “back bone” and the straight lines (representing observed longitudinal data) can be considered “rib bones”. By using cross-sectional data or the intercept from a longitudinal model with age as the time variable, we should be able to construct the “back bone” of the spline. Adding the “rib bones” from large number of individuals in different age categories can enhance the shape of the spline.

Generally speaking, to obtain the lifespan trajectory, we should observe the “rib bones” in our typical longitudinal datasets and then attempt to construct the “back bone.” In this work, we attempt the converse; using the “back bone”, we attempt to grow the “rib bones”. In other words, our approach is to first fit an accurate spline model from cross-sectional data to model the non-linear trajectory across the lifespan. We then use the age slope to augment the longitudinal data for each normal aging subject, given that a linear model can suffice to model brain atrophy over a short (5-year) period of time.

The study design includes the following steps: (1) identify a well-fitted spline model using cross-sectional data from normal aging populations; (2) augment longitudinal data from this well-fitted spline model using a linear slope at a given age point; (3) compare 12 different mixed spline models through simulated data and identify the mixed spline model that fits the “fish bone” data structure; (4) combine augmented longitudinal normal aging data with longitudinal MS data to fit mixed spline model and compare across 12 mixed spline models; (5) use a manual forward then backward model building strategy to select the covariates from 52 covariate structures;



**The straight lines (light blue) represent the “rib bones” of the fish bone and are based on observed MRI trajectories in available longitudinal MRI datasets, which generally display a linear atrophy trend because of the short follow-up period of the study. The spline curve (black) can be conceptualized as the “back bone” of the fish bone and can be fit using cross-sectional data or the intercept from a longitudinal model.**

**Figure 1.** Concept diagram for “fish bone” data structure.

(6) identify the individual age of onset of progressive brain tissue loss with associated 95% confidence interval using a tenfold cross validation procedure through 1000 bootstrapping samples.

### Study sample

Our dataset was assembled from the following three sources (Table 1): (1) The Human Connectome Project (HCP: <http://www.humanconnectome.org>), (2) Alzheimer’s Disease Neuroimaging Initiative (ADNI: <http://www.adni-info.org>) and (3) a single-center, prospective cohort MS study conducted from January 2005 through December 2010<sup>29</sup>. Normal aging samples were from HCP, ADNI and 89 healthy control cases from the single-center study. Age at scan date and sex were extracted from each of the data sources. Healthy control subjects (N = 2053) had an overall mean age of  $44 \pm 21$  years (Q1: 27, Q3: 62) with 56% female, while MS subjects (N = 519) had a mean age of  $43 \pm 10$  years (Q1: 36, Q3: 50) with 70% female and an average of  $4 \pm 1.5$  annual scans per subject. Most of the normal aging sample only had one MRI scan, but 228 of them had repeated measures ( $2.9 \pm 1$  scans in  $2.5 \pm 1.4$  years). Subjects with age > 90 or < 16 (to avoid brain growth confounding) were excluded.

3 T/3D T1-weighted volumetric gradient echo images were processed with FreeSurfer v6.0 to extract thalamic and intracranial volumes. Thalamic volumes were normalized by total intracranial volume and multiplied by 1000. Healthy control subjects had normalized thalamic volumes of  $9.7 \pm 1$  (Q1: 9.1, Q3: 10.4) at study entry, while MS subjects had normalized thalamic volumes at study entry of  $9.3 \pm 1$  (Q1: 8.7, Q3: 9.9) (Table 2).

### Longitudinal normal aging data augmentation from large cross-sectional data

Multivariate Adaptive Regression Splines (MARS) was used to fit a cross-sectional spline so that we could augment the longitudinal data. MARS was chosen because of its robustness to outliers and its ability to auto-search non-linear associations with high dimensional interactions<sup>41</sup>. Age at scan, intracranial volume (ICV), and sex were used, with three-way interactions among them, as predictors of thalamic volume (percent of total brain volume). ICV and sex were treated as constant for each individual subject when augmenting the longitudinal data. Longitudinal thalamic volumes were augmented at  $\pm 2$  years from age at scan. We reserved 433 repeated measurements from 229 individuals as independent testing. ICC two-way mixed with absolute agreement and repeated measure correlation were used to assess the agreement/correlation between MARS model-augmented longitudinal data versus observed testing longitudinal data. SAS9.4 ADAPTIVEREG was used to fit the MARS model.

Dataset	# of Subjects	# of MRI time points	Age (Mean $\pm$ SD)	Age range	% Female
HCP-D	178	178	$19 \pm 2$	16–22	55.1
HCP	865	865	$29 \pm 4$	22–37	56.3
HCP-A	676	676	$58 \pm 14$	36–90	56.8
Single Center	87	152	$40 \pm 11$	22–65	67.8
ADNI	247	614	$75 \pm 7$	56–89	52.2
Total	2053	2485			

**Table 1.** Demographic distribution of health control cohorts.

Age at Study Entry (years)		42.7 ± 9.7, 42.2 (35.5 to 50.4)
Sex	Female	364 (70.13%)
	Male	155 (29.87%)
Age of Clinical Onset (years)		33.5 ± 9.1, 33 (27 to 40)
Disease Type	CIS	90 (17.34%)
	RRMS	391 (75.34%)
	SPMS	38 (7.32%)
History of DMT Use at Study Entry	Never Used	212 (40.85%)
	Ever Used	307 (59.15%)
DMT Exposure (years)		2.5 ± 2.5, 1.8 (0.6 to 3.5)
T2 White Matter Lesion Volume (mm <sup>3</sup> )		4636.9 ± 5775.2, 2589.4 (991.5 to 6133.2)
Normalized Thalamic Volume at Study Entry (thalamic volume/ICV*1000)		9.3 ± 1, 9.3 (8.7 to 9.9)

**Table 2.** Demographic distribution and disease characteristics of MS cohort (n = 519). CIS clinically isolated syndrome, DMT disease-modifying therapy, ICV intracranial volume, RRMS relapsing remitting multiple sclerosis, SPMS secondary progressive multiple sclerosis, T2 T2-weighted MRI. \*Mean ± SD, Median (Q1, Q3).

### Mixed spline model of thalamic atrophy trajectory

After data augmentation, we fit the mixed spline model. Let  $n$  be the number of subjects. For the  $i$ th participant, denote  $t_i$  as the age, denote  $Y_{ij}(t)$  as the thalamus volume at the  $j$ th measurement for subject  $i$ , and denote  $X_{ij}$  as other predictors such as sex. To model the age effects accurately and efficiently, we use a semiparametric model of the form given below:

$$Y_{ij}(t) = \mu_{ij}(t) + X_{ij}\beta + v_i(t) + \epsilon_{ij}(t), i = 1, \dots, n, j = 1, \dots, k$$

where  $\mu_{ij}(t)$  is the unspecified aging trajectory for subject  $i$  at the  $j$ th time evaluated at age  $t$ , and  $\beta$  are the regression coefficients of the other predictors at the  $t_j$  time.  $v_i$  is the random effect of each subject. The measurement errors  $\epsilon_{ij}$  are assumed to follow a normal distribution  $N(0, R)$ , where  $R$  is the covariance matrix. This semiparametric regression model is a parsimonious way to both capture the potential nonlinear age trajectory and investigate the effects of other predictors. The simplest special case of this model is the linear mixed model where  $\mu_{ij}(t) = \beta_0 + \beta_1 t_{ij}$ . Regression splines are a broader class of models and could be fitted under this framework, which can be based on truncated power function (TPF) basis, B-spline basis or natural spline basis. These models vary by the choice of the spline basis and tuning parameters (the number of knots and the knot positions) that have an impact on the estimated shape of a spline function. Parameter-function estimation contains two major steps: (i) approximation using basis functions (e.g., TPF, B-Spline) which allows to fit lower-order polynomials within very small interval partitions (based on knots) and (ii) smoothing the approximation via penalty (e.g., random SPLINE coefficients, TOEPLITZ G-side matrix, RSMOOTH G-side matrix). The smoothing could be done via generalized cross-validation (GCV)<sup>42</sup> or mixed effects approaches<sup>43,44</sup>, which are known to facilitate the choice of the knot positions in spline modelling<sup>45</sup>. They also allow a *penalty* to be applied directly to the model coefficients (P-spline penalty penalizes the squared differences between adjacent model coefficients, which in turn penalizes wiggles).

We then compared penalized splines (P-spline) with B-spline basis and truncated power function (TPF) basis with different random effect structures such as P-SPLINE and RSMOOTH (radial smoothing). For the P-spline, the unspecified function  $\mu_{ij}(t)$  is approximated with a cubic B-Spline or TPF basis. Following Ruppert, Wand and Carroll<sup>46</sup>, the cubic spline can be represented as:

$$\mu_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \sum_{j=1}^K \beta_{3+j} (x_i - t_j)^3$$

$$(x - t) = \begin{cases} x - t & x > t \\ 0 & 0 \end{cases}$$

The first part of the formula  $\mu_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$  represents the polynomial function. The second part of the formula  $\sum_{j=1}^K \beta_{3+j} (x_i - t_j)^3$  represents the truncated power function (TPF). The  $t_j$  represents the potential cut points for the knots. Estimation of parameters is made by minimizing the penalized log-likelihood function using proc GLIMMIX in SAS 9.4 with smoothing implemented using P-SPLINE smoothing (Random x/type = pspline) or radial smoothing (Random x/type = rsmooth). This mixed model formulation of spline smoothing has the advantage that the smoothing parameter is selected automatically<sup>46</sup> and is shown to be more robust with misspecification of error dependence structure, compared to GCV-based approach<sup>47</sup>.

For the 12 spline structures described above, model comparison was made using four criteria: (i) Akaike information criterion (AIC) and Bayesian information criterion (BIC) criterion, with lower values indicating better fit; (ii) repeated measure correlation coefficient<sup>48,49</sup> and intraclass correlation two-way mixed for longitudinal data between model-predicted versus observed data from reserved 10% testing dataset; (iii) visual inspection of the expected shape for projected lifespan spline (the normal curve must inherit the shape of the spline based on cross-sectional normal aging, and the MS curve must followed the shape of observed spaghetti plots

as in Fig. 3A & Supplemental Fig. 10); and iv) both MS and normal aging trajectory curves must have narrow predicted interval along age points.

### Simulation study design

The purpose of the simulation study is to compare spline models to choose the most appropriate spline model for the “fish bone” data structure. The simulated data mimic the fish bone data structure by combining 10 sets of data from 10 different age blocks ( $k = 1$  to 10) with age range from 30 to 80 by 5-year intervals (e.g., 30–34, 35–49). Each simulated data set was based on the covariance parameters estimated from a linear mixed model (with random intercept and slope). Block-specific weights ( $W_k, V_k$ ) were added to the fixed effects of intercept and slope respectively for block  $k$ .  $W_k$  and  $V_k$  were altered to mimic a spline shape (“back bone” as shown in Fig. 1). The final mixed effects model was as follows:

$$Y_{kij} = W_k \beta_{k00} + V_k \beta_{k10} (Year)_{kij} + \beta_{k01} MS + \beta_{k11} (MS) * (Year)_{kij} + b_{k0j} + b_{k1j} (Year)_{kij} + \epsilon_{kij},$$

$$\epsilon_{kij} \sim iidN(0, \sigma^2)$$

The training sample was a combination of 10 datasets with 50 MS subjects each (age span 30 to 80 years). Each MS subject had 5 longitudinal MRI data points within each block, simulated using the linear mixed model above. Therefore, we simulated 500 subjects total in the training data.  $W_k$  and  $V_k$  started with small values in younger age, e.g. 1% decrease from the previous age block, but larger in middle age, e.g. 5% decrease, then became smaller again in older age, e.g. 1% decrease. The testing data followed the same simulation procedure, but we used the same subject ID across the 10 blocks; thus, the testing data contained 50 MS subjects, and each subject had 50 simulated age points. Because our ultimate goal is to predict the thalamic volume at an age that is younger than the observed age, the testing data included 4 more younger age points: 26, 27, 28, and 29, in addition to the 50 age points.

We considered twelve different models with three G-side covariance types (TOEPLIZ, P-SPLINE and radial smoothing) and four basis functions (Cubic-B-Spline, Cubic -TPF, Natural-TPF, Natural-B-Spline). To estimate the prediction accuracy from the spline model, we made comparisons using AIC/BIC with 500 iterations. The testing data were scored through each of the 12 spline models. We then obtained the estimated thalamic volume with associated 95% confidence interval at each age point. We took the average of the 500 replicates for the model-predicted thalamic volume and associated 95% CIs, then graphed the spline plots to visually inspect the overlap between true spline curve and the model-estimated spline curves, as well as the width of 95% confidence band.

### Real-life data application

We applied 12 different scenarios of spline models (listed in Table 3 to a real-life dataset with 519 MS subjects and 2053 normal aging subjects. For the normal subjects, we used augmented longitudinal data with 5 follow-up years (actual scan year at the middle). Among the 519 MS subjects, we randomly selected and reserved 52 MS subjects as the independent testing data. We repeated this iteration 10 times with 10 mutually exclusive independent testing data. For the first iteration, we selected the optimal spline structure and covariates using the criteria defined in 3.3. For covariate selection, we used a forward then backward strategy. Age spline, MS status, and age spline  $\times$  MS status interaction were mandatory for each model. Other covariates included sex, baseline thalamic volume ( $Thalamus_0$ ), baseline ICV ( $ICV_0$ ), age of clinical onset (set as 0 for normal control), and cumulative years of exposure to MS disease modifying therapies at the first scan ( $DMT_0$ ; set as 0 for normal control). Each covariate entered the model first as the main effect, then as interaction terms with MS status or/and age spline. We categorized the covariate structure as the following: (a) only the main effect from each covariate; (b) two-way interactions with MS status, each covariate interaction term one by one, then multiple interaction terms together; (c) three-way interactions for each covariate one by one with MS status and age spline except  $DMT_0$  and age of clinical onset; (d) select any two covariates with the three-way interactions; (e) select any three covariates with the three-way interactions; (f) select any four covariates with the three-way interactions; (g) backwards selections with the terms showing model improvement from a to f. The same criteria defined in 3.3 were used for model selection. Once the final model was selected, we used the same model structure in the other 9 training datasets to obtain estimated coefficients, then applied them to independent testing data at each fold of iteration.

Cubic Spline Basis Function		G-side covariance type		
		TOEPLIZ	PSPLINE	RSMOOTH
Cubic-B-Spline	AIC	354.45 $\pm$ 74.55 <sup>†</sup>	5843.49 $\pm$ 250.76	797.04 $\pm$ 80.81
Cubic -TPF	AIC	448.76 $\pm$ 117.7	5840.56 $\pm$ 250.33	809.23 $\pm$ 83.55
Natural-TPF	AIC	673.43 $\pm$ 76.93	5921.87 $\pm$ 250.77	906.61 $\pm$ 80.24
Natural-B-Spline	AIC	673.43 $\pm$ 76.93	5921.87 $\pm$ 250.77	906.61 $\pm$ 80.24
Cubic-B-Spline	BIC	421.88 $\pm$ 74.57	5910.93 $\pm$ 250.75	864.46 $\pm$ 80.76
Cubic -TPF	BIC	516.25 $\pm$ 117.71	5907.99 $\pm$ 250.33	876.67 $\pm$ 83.52
Natural-TPF	BIC	707.14 $\pm$ 76.9	5955.57 $\pm$ 250.75	940.34 $\pm$ 80.23
Natural-B-Spline	BIC	707.14 $\pm$ 76.9	5955.57 $\pm$ 250.75	940.34 $\pm$ 80.23

**Table 3.** AIC and BIC from different spline structures based on simulation data. <sup>†</sup>Mean  $\pm$  Std from 500 iterations.

For each independent testing MS patient, we constructed the hypothetical individualized normal aging trajectory curve (Health Digital Twin) and MS lifespan trajectory curve using the patient-specific covariates, which included sex,  $\text{Thalamus}_0$ ,  $\text{ICV}_0$ , age of clinical onset (set as 0 for normal control),  $\text{DMT}_0$  (set as 0 for normal control) and sequential age points from 15 to 75 with an interval of 1. The age at which the MS trajectory curve began to depart from the HDT trajectory curve (or when both curves crossed in young age) was defined as the age of onset of progressive brain tissue loss. A bootstrapping procedure with 1000 iterations was used to determine the 95% confidence interval of this brain atrophy-defined age of onset. The bootstrapping procedure was conducted at the patient level, i.e., once a patient ID had been selected, all longitudinal scans associated with this patient were selected as a completed block. We repeated this procedure 10 times with 10 mutually exclusive independent validation data (tenfold cross validation procedure with 10% testing data each fold). Thus, each patient had an age of onset of progressive brain tissue loss (PBTL) with 95% CI estimated from their exclusive training dataset. In the end, we identified two groups of MS patients: earlier onset (the upper 95% CI limit of PBTL onset is younger than clinical onset age; in other words, the age of onset of PBTL was statistically significantly earlier than the age of clinical onset); and simultaneous onset (clinical onset did not differ statistically from the age onset of PBTL). We examined the different patterns of the onset age gap (age of clinical onset minus the age of onset of PBTL) between the earlier onset and simultaneous onset groups used Bland–Altman plots.

## Results

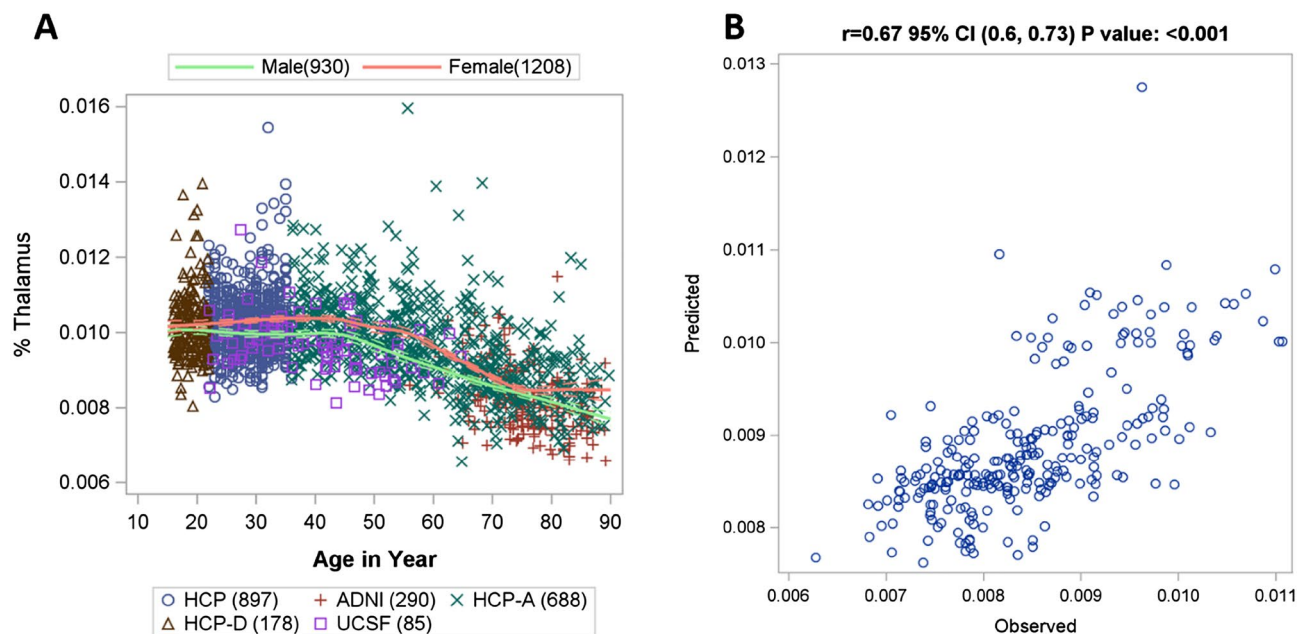
### Accuracy of augmented longitudinal normal aging data

Figure 2A shows the smooth data cloud across age for healthy controls. The spline constructed by MARS represents the trend in the data. After data augmentation, independent validation was conducted by comparing the MARS model-predicted longitudinal data vs. the 433 observed longitudinal data. The model-predicted data had good agreement with the observed longitudinal data with ICC 0.62 95% CI (0.56, 0.68), using two-way mixed with absolute agreement. The predicted value explains 45% of total variance in the observed value ( $r=0.67$ , 95% CI (0.6, 0.73),  $p < 0.001$ ;  $R=0.45$ ) based on repeated measure correlation (Fig. 2B).

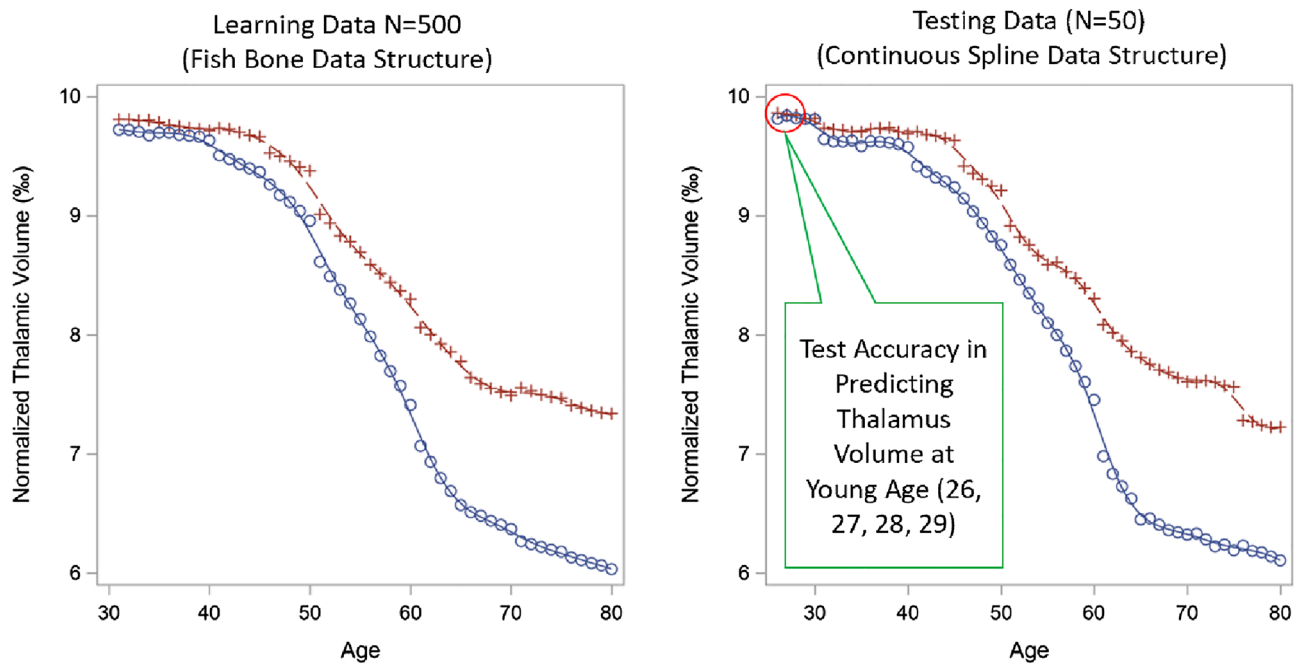
### Results of simulation study

Figure 3 shows the spline shapes for both simulated training and testing data from one of the 500 iterations, with red line as the normal aging trajectory and blue line as MS trajectory. The curve from the training data (left panel) shows the expected value of simulated ‘fish-bone’ structure which contains 10 datasets with 50 MS subjects each (age span 30–80 years). The curve from the testing data (right panel) shows the expected value of simulated ‘continuous spline’ data structure which contains 50 MS subjects with 50 simulated age points (referred as continuous age points across lifespan) plus 4 additional earlier age points (age 26–29). This represents the ‘ground truth’ spline shape that should be found by the mixed spline model.

Table 3 shows the mean and standard deviation of AIC and BIC comparing mixed-spline models and their corresponding G-side covariance from 500 iterations based on each of the 12 spline modeling scenarios. In general, unrestricted B-Spline had the smallest AIC or BIC showing the best fitting index, followed by unrestricted



**Figure 2.** (A) Distribution of normalized thalamic volume (%Thalamus) in healthy controls across age from cross-sectional data (estimated using MARS). (B) Correlation between MARS model-predicted (i.e. augmented) and observed longitudinal data.



**Figure 3.** Simulated data curves for normal aging (red line) and MS (blue line). Description of simulated spline shape based on piecewise linear mixed models with age varying slopes. Red line represents normal aging and blue line represents MS. This illustrates the ‘ground truth’ spline shape that should be found by the mixed spline model.

TPF. The restricted basis functions, both natural-B-Spline and natural-TPF, performed poorly. For G-side matrix, the TOEPLIZ had the best performance, followed by radial smoothing. P-SPLINE had the worst performance.

In addition to model fitting, we also assessed the prediction accuracy of each of the 12 spline models based on the prediction accuracy of the testing data. We visually inspected the overlap of the observed spline and the model-estimated spline curves, as well as the width of 95% confidence band. The visual inspection matched the AIC/BIC finding, with the best performance from B-Spline with a TOEPLIZ covariance model. The visual illustration of selected spline curves from TOEPLIZ is presented in Fig. 4.

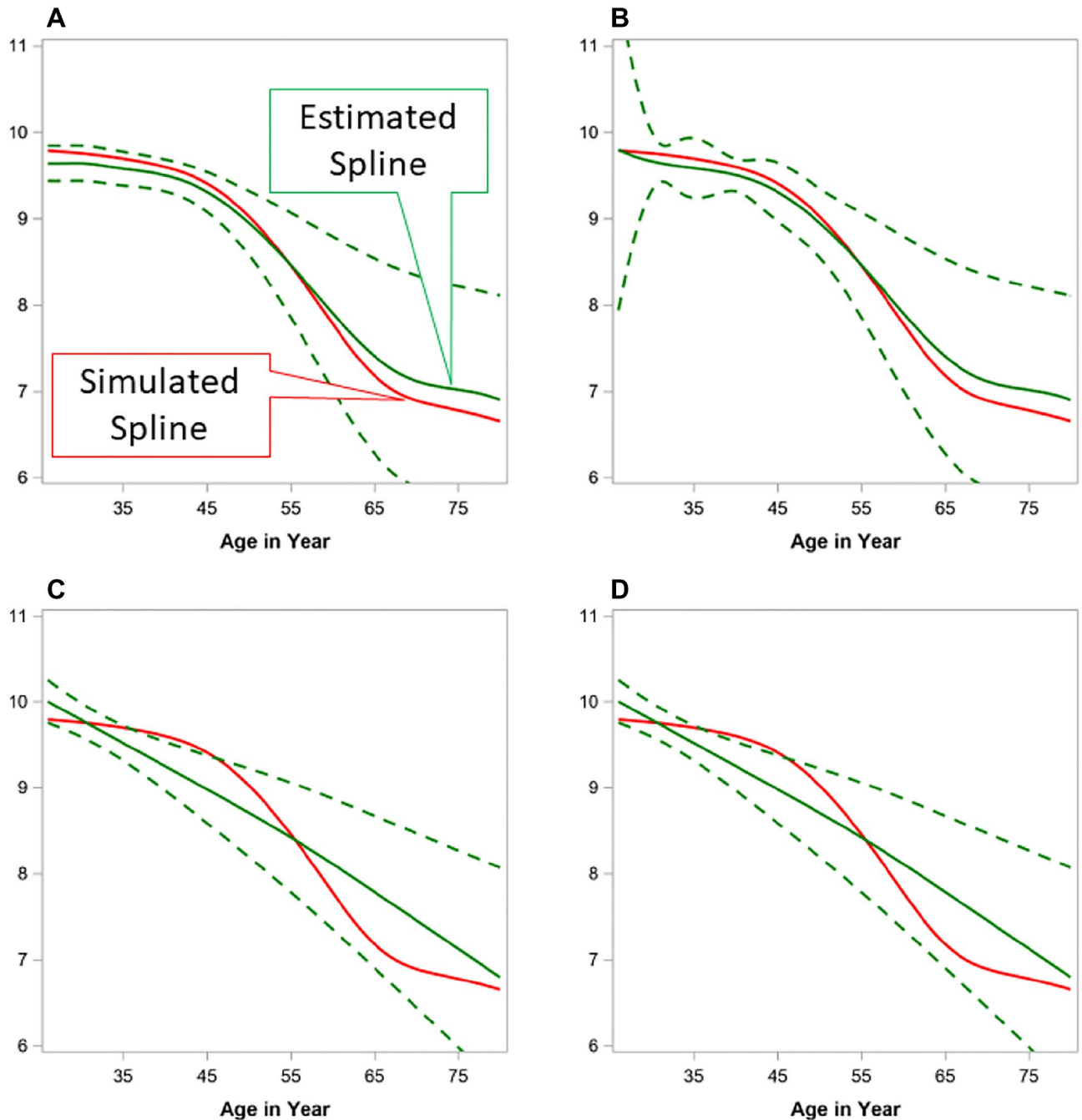
Figure 4 shows the patterns of the smoothed spline from simulated value (ground truth) vs. the predicted value constructed using the mean of predicted values with 95% confidence band over 500 iterations from the testing data. In Fig. 4A, the predicted spline from TOEPLIZ with Cubic-B-Spline overlapped well with the ground truth spline, and the predicted 95% confidence band is narrow for the younger age. When using TPF as basis function (Fig. 4B), the 95% confidence band became very wide at early ages. The restricted splines (Fig. 4C,D) fitted a line as straight as a linear line, which is largely deviated from the simulated spline. Because our overall objective is to model the disease onset, we are most interested in modeling accuracy around the younger ages.

In summary, we have demonstrated through simulation that a mixed spline model can be used in a “fish bone” data structure when longitudinal MRI datasets only cover a few years of an individual’s lifespan. Figure 3 shows the ‘ground truth’ simulated spline shape based on piecewise linear mixed models with age varying slopes, and Fig. 4 shows good overlap between the ‘ground truth’ simulated spline and the estimated trajectory from the mixed spline model.

### Results of real-life data analysis

Supplemental Figs. 1–7 illustrate the model fitting indices of 12 spline structures and 52 covariate structures. As described in 2.3, the modeling fitting criteria included (i) AIC; (ii) repeated measurement correlation from 10% independent testing data between observed and predicted longitudinal values; (iii) shape of trajectory curves; and (iv) predicted 95% confidence band for trajectory curves. The real-life data application results concurred with the simulation study, with the best fitting model being the B-Spline with TOEPLIZ. The covariate structure included: age-spline, MS-status, ICV<sub>0</sub>, sex, sex\*MS-status, Thalamus<sub>0</sub>, sex\*age-spline, age at study entry, age of clinical onset, and DMT<sub>0</sub>. The final model reached a repeated measure correlation coefficient of 0.88 based on tenfold cross validation. Trajectory curves and scatter plots (Fig. 5, Supplemental Fig. 8) demonstrate that the spline curve ran through the observed data points in most of the cases (illustrated as black diamonds in Fig. 5). The AIC and BIC from the final model was also the smallest among all model structures.

Using this final model, we were able to construct the lifespan thalamic atrophy trajectory curve for an MS individual and their corresponding normal aging digital twin. Figure 5 illustrates two example cases. In each, the age of onset from progressive brain tissue loss (green dot) was younger than the age of clinical onset (red dot). The 95% confidence interval was derived from a bootstrapping procedure with 1000 iterations. Figure 5A shows that the upper limit of the 95% CI of the age of PBTL onset was younger than the age of clinical onset



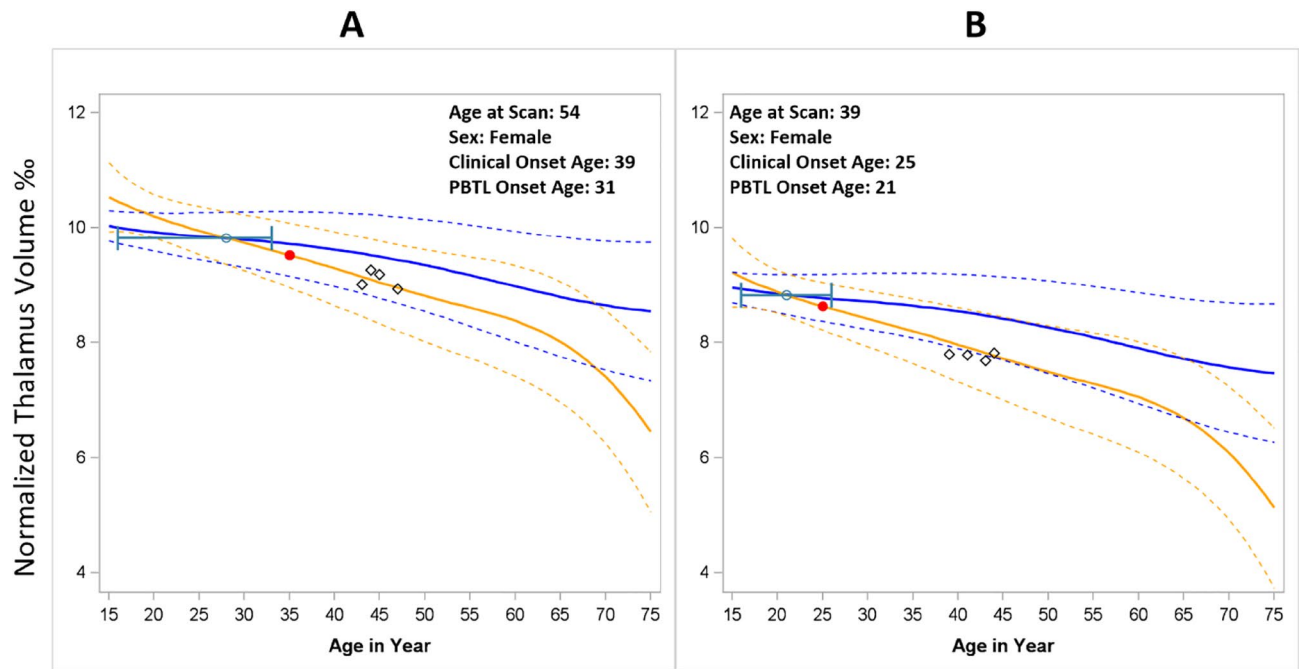
**Figure 4.** Illustration of simulated spline vs. predicted spline from simulation study. (A) cubic spline with bspline basis and random toep1; (B) cubic spline with TPF basis and random toep1; (C) restricted cubic spline with TPF basis and random toep1; (D) restricted cubic bspline with pspline basis and random toep1.

(earlier onset). In contrast, Fig. 5B shows that the 95% CI for the age of PBTL onset overlaps with the age of clinical onset (simultaneous onset).

Figure 6 shows the age of onset of progressive brain tissue loss based on 1000 bootstrapping samples of 519 MS patients (Fig. 6A). The x-axis was centered by the age of clinical onset; it was calculated as age of PBTL onset minus age of clinical onset. Therefore, the center vertical line (red) is the zero reference line. If the upper limit of 95% CI is left of the center line, it suggests that the onset of PBTL was significantly earlier than the clinical onset (“earlier” onset). Otherwise, if the 95% CI includes the center line (0 gap between age of clinical onset and age of onset of progressive brain tissue loss), it suggests simultaneous onset. Overall, the age of onset of progressive brain tissue loss was younger than the age of clinical onset, with a mean difference of  $5.1 \pm 3.8$  years and a median difference of 6 years (IQR 3.1–8.1).

Using our definitions, 55.4% of patients could be classified as earlier onset, while 44.6% of patient could be classified as simultaneous onset. In Fig. 6A, there are 26 patients with extremely high upper confidence limit of





**Figure 5.** Trajectory curves for earlier onset (A) versus simultaneous onset (B). Estimation of Progressive Brain Tissue Loss (PBTL) in two individuals with MS (Panels A and B). The solid yellow line represents the individual MS subjects' thalamic volume trajectories (95% CI = dashed yellow lines). The solid blue line represents the thalamic volume trajectory for a hypothetical normal aging individual with same demographic characteristics and baseline thalamic volume as the MS patient (their digital twin). 95% CI = dashed blue lines. The point at which the MS subjects' trajectory deviates from their digital twin is defined as the onset of PBTL, denoted with the green dot (95% CI from bootstrapping in green). The red dot denotes the age of clinical onset. Black diamonds represent MS subjects' observed thalamic volumes over time.

PBTL age. These patients (Supplemental Table 1) all had young clinical onset age and their mean projected PBTL onset age is close to the clinical onset age. Model fit was adequate in these patients and the wide upper confidence limits are mainly driven by outliers in the bootstrapping procedure. As such, they were correctly classified as "simultaneous onset." Wilcoxon rank sum tests showed that earlier onset patients had statistically significantly older onset age compared to simultaneous onset, for both age of clinical (Median 36, IQR 33–41 vs. Median 22, IQR 26–30,  $p < 0.01$ ) and progressive brain tissue loss onset (Median 29, IQR 26–32 vs. Median 24, IQR 23–26,  $p < 0.01$ ). Bland–Altman plots (Fig. 6B,C) showed age of onset of progressive brain tissue loss is much younger compared to age of clinical onset for patients in earlier onset group. Such difference is more scattered in the simultaneous onset group around the 0-reference line.

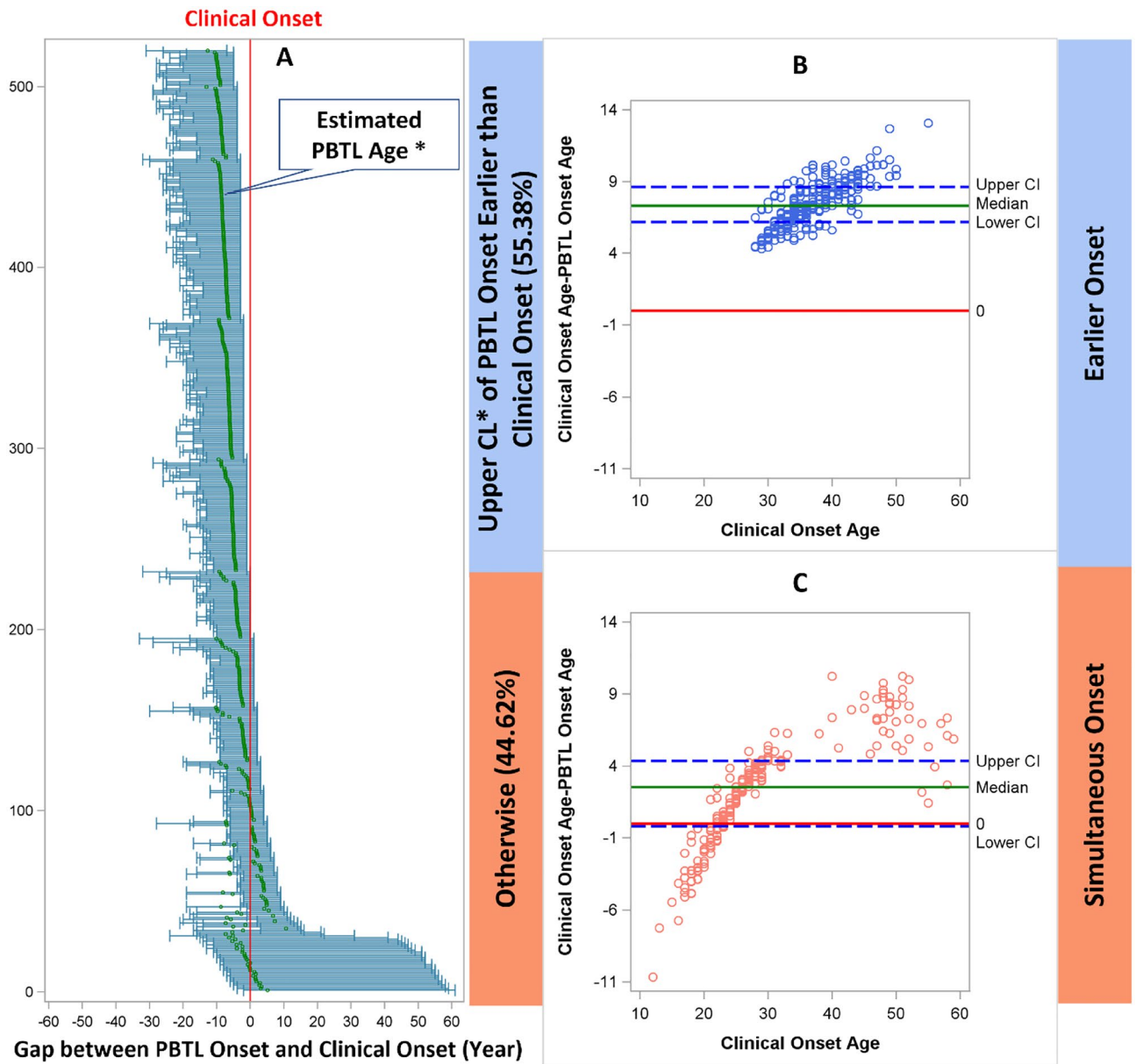
Demographics and disease characteristics between earlier and simultaneous onset patients are reported in Table 4. In general, patients with simultaneous onset were significantly younger at clinical disease onset and at study, were more likely to be male, had higher T2 lesion volumes, and had smaller thalamic volumes.

## Discussion

Health digital twin is a novel and promising concept to further advance precision medicine. It includes many major components such as personal devices, AI algorithms, right data to train the AI, and Internet of Things (IoT) for rapid data synchronization while providing real time decision-making<sup>22</sup>. The AI component can be considered the heart of the health digital twin approach. In aging-related fields such as neurodegenerative diseases, such AI algorithms must be developed from the scope of lifespan.

In this study, we apply the health digital twin conceptual framework to build an AI algorithm in addressing a fundamental clinical problem in multiple sclerosis, which is to identify the disease-related onset of brain atrophy. By estimating the deviation of the thalamic atrophy trajectory curve of an individual MS patient from their corresponding hypothetical normal aging health digital twin, our major finding is that progressive brain tissue loss precedes clinical disease onset in MS by a mean of  $5.1 \pm 3.8$  years and a median of 6 years (IQR 3.1–8.1). Although the onset of progressive brain tissue loss measured by MRI is not synonymous with the true biological disease onset, our results suggest a major improvement in estimating MS disease duration compared to the standard practice of defining the disease onset as the time of first clinical symptom. This may have significant implications for MS clinicians, researchers, and patients, and could lead to a fundamental shift in our disease understanding and, one day, determining its cause.

Our novel approach in developing this AI algorithm towards health digital twin has several innovations. The first innovation is the development of a novel statistical application of mixed splines to overcome the challenge of lacking lifespan longitudinal data. Longitudinal studies usually have limited sample sizes and short follow-up



\* Estimation and 95% Confidence Interval Based on Bootstrapping with 1000 Iterations

**Figure 6.** Panel A: Forest plot showing PBTL onset age minus clinical onset age (in years, x axis) for each of the 519 MS patients based on 1000 bootstrapping procedure. The center vertical red line represents the 0-reference line (clinical onset age). The tick marks in each horizontal line include the mean and 95% confidence interval. If the upper limit of 95% CI on the left of the center line, it suggests earlier onset. Otherwise, if the 95% CI includes the center line, it suggests simultaneous onset. Panels B and C: Bland–Altman plots. Each dot represents the difference between the age of onset of PBTL and the age of clinical onset. If the lower bound of 95% CI for the difference goes below the 0-reference line (red), it indicates weak evidence that the age of onset of progressive brain tissue loss is younger than the age of clinical onset. None of the dots are below the reference line in Panel B but they are in Panel C, consistent with our classification of “early” vs “simultaneous” onset.

periods. Generally, high-quality longitudinal MRI datasets, such as those that use the same pulse sequences and scanner, only contain 3–5 years of follow-up. To overcome this challenge, we describe the concept of the “fish bone” data structure. Structuring the data this way can have two benefits: (1) augment longitudinal data from large cross-sectional data; (2) augment individual lifespan trajectory based on small fragments of follow-up periods over a widespread age range.

To demonstrate these benefits, we successfully fit a spline model from a large cross-sectional dataset across a wide age range that reflects the non-linear trajectory of brain atrophy across the lifespan, and then used this to augment the longitudinal data with a good repeated measures correlation with the observed data ( $r=0.67$ ,  $p<0.001$ ). Moreover, rather than using follow-up time since study entry per the conventional approach, we used

	Category	Earlier onset (n = 288)	Simultaneous onset (n = 231)	p value
Age at study entry (years)		44.5 ± 7.6, 43.9 (39 to 50)*	40.4 ± 11.6, 38.9 (30.7 to 50.8)	< 0.001‡
Sex	Female	226 (78.5%)	138 (59.7%)	< 0.001†
	Male	62 (21.5%)	93 (40.3%)	
Age of clinical onset (years)		36.9 ± 5.2, 36 (33 to 41)	29.3 ± 11, 26 (22 to 30)	< 0.001‡
disease type	CIS	59 (20.5%)	31 (13.4%)	0.11†
	RRMS	209 (72.6%)	182 (78.8%)	
	SPMS	20 (6.9%)	18 (7.8%)	
History of DMT use at study entry	Never used	125 (43.4%)	87 (37.7%)	0.19†
	Ever used	163 (56.6%)	144 (62.3%)	
DMT exposure (years)		n = 163, 2 ± 1.8, 1.6 (0.6 to 3.1)	n = 144, 3.1 ± 3, 2.1 (0.8 to 4.8)	0.013#
T2 white matter Lesion volume (mm <sup>3</sup> )		3776.4 ± 4916.8, 2106.3 (874.7 to 4546.2)	5719.3 ± 6553.2, 3456.5 (1191.1 to 7743.8)	< 0.001‡
Normalized thalamic volume at study entry (thalamic volume/ICV*1000)		9.4 ± 1, 9.4 (8.8 to 10)	9.1 ± 1, 9.2 (8.5 to 9.7)	0.012‡

**Table 4.** Demographics and disease characteristics in patients with earlier onset versus simultaneous onset. CIS clinically isolated syndrome, DMT disease-modifying therapy, ICV intracranial volume, RRMS relapsing remitting multiple sclerosis, SPMS secondary progressive multiple sclerosis, T2 T2-weighted MRI. \*Mean ± SD, Median (Q1, Q3), #Wilcoxon Rank Sum Test, ‡Independent T-Test, †Chi Square Test.

age at follow-up as the time variable for longitudinal data. This approach has been used in a recent study to predict mild cognitive impairment in Alzheimer's disease based on the brain atrophy trajectory pattern by different periods<sup>50</sup>. However, the authors only explored a quadratic term as the non-linear effect. We have advanced the modeling strategy to a mixed spline model.

When fitting the longitudinal spline (mixed spline) model with an uneven time scale and short follow-up period, it was not trivial to determine the best fit spline structure from 12 different candidate spline structures. In our study, we used both simulation and real-life data to reach the conclusion that the best fitting model is B-Spline with TOEPLIZ as G-side matrix. This may be due to the simplicity of the mixed spline structure when applied to this special data structure. Determining the appropriate covariates was also not trivial, as those that interact with the spline slope will alter the slope, while non-interaction terms will affect the elevation of the spline and parallel distance between the MS curve and the normal aging trajectory curve. Interactions can be two-way, 3-way, or higher dimensional interactions. Given this, the covariate selection cannot follow the conventional forward, backward, or LASSO approach<sup>51</sup>. The interaction term must be intact with the marginal effect as a bundle. When adding higher dimensional interactions, each piece of lower-level interaction terms must be intact, too. Therefore, we used a manual forward then backward covariate selection strategy.

Given these complexities, determining the final model with the best fitting spline structure and covariates is challenging. The conventional approach is to use AIC and BIC; more stringent criteria require cross-validation. We initially used AIC/BIC plus a tenfold cross validation using the repeated measure correlation between model-predicted and observed values. However, we observed that this approach can be misleading. When constructing the 12 × 52 lifespan trajectory curves (Supplemental Figs. 11–7) for a given individual, we observed scenarios where fitting indices were strong (small AIC/BIC, large r), but the trajectory curve was wild. For example, Supplemental Fig. 5A row E\_01, column 3, a model with 3 three-way interaction terms, is likely overfit, and the spline shape did not inherit the shape we observed in large cross-sectional data (Fig. 3A). This phenomenon could be due to model tracing for few observed data points in the middle of the lifespan while sacrificing the fit of both far right (older) or far left (younger) ends. Since both AIC/BIC and repeated measure correlation were driven by the observed data, these fitting indices misled the lifespan trajectory. We added two additional criteria for model selection: (i) visual inspection of the shape of projected lifespan spline (normal aging curve must inherit the shape of the spline from large population based cross-sectional normal aging data and individual MS curve must follow the shape from group estimates and spaghetti plots); (ii) both MS and health digital twin trajectory curves must have narrow predicted bands along the age span. We were able to identify the optimal model based on this approach.

Our normal aging data were combined from four different studies. As such, one potential issue could be data heterogeneity due to slight differences in MRI protocols and scanner settings. A common practice in neuroimaging is to use a statistical model such as neuro-ComBat to harmonize the data before conducting further analysis<sup>52</sup>. Since each individual dataset of normal aging represents a subset of the age category, age must be added to neuro-ComBat as a covariate. Supplemental Fig. 9 shows the distribution of both the original (A) and ComBat harmonized (B) percent thalamic volumes along with age. The data distribution did not change from the original to ComBat harmonized thalamic volumes. In fact, the original data had a very smooth cloud across age. The robustness of our normalized thalamus volumes from different scanners and settings may be due to the use of relative values (thalamus as a percentage of ICV) instead of using absolute thalamic volumes. In a phantom study, we found that using relative values was robust to different scanner settings and protocols<sup>53</sup>. Forcing age as a covariate in neuro-ComBat removes the spline effect, which contradicts the non-linear brain trajectory reported from other major studies<sup>32</sup>. Therefore, we used and retained the original percent thalamic volumes throughout the study.

Although our overall model fit was adequate, we observed an interesting pattern in some patients with a young clinical onset age (Fig. 6, Supplemental Table 1). Within our dataset of 519 MS patients, 26 had extremely high upper confidence limit of their PBTL age. All of these patients had a young clinical onset age and a mean projected PBTL onset age that was close to their clinical onset age. Because they had young onset age, their trajectory curve will be close to and parallel with the normal aging trajectory curve. Thus, the age corresponding to the minimal distance between the two curves (when they cross) will be difficult to determine. In rare bootstrapping samples, the model identified an extreme older age as the PBTL onset age, and thus the wide upper confidence limits are mainly driven by outliers in the bootstrapping procedure. Those outliers will not affect our estimated PBTL onset age since we used the mean age from the bootstrapping sample, nor will they affect our classification of earlier vs. simultaneous onset. These patients very clearly belong to simultaneous onset group. However, this illustrates an issue that may arise in estimating PBTL age in patients with young clinical onset age and deserves future study.

Herein, we describe a novel statistical modeling strategy to overcome limitations in real-world longitudinal neuroimaging datasets and provide an application of the Health Digital Twin framework to address a fundamental clinical conundrum in MS. We found that the MS thalamic atrophy trajectory deviated from the corresponding hypothetical normal aging trajectory 5–6 years prior to clinical symptom onset. While there is no ground truth to validate our findings, this is consistent with clinical observations that white matter lesions and thalamic atrophy are already present prior to first clinical symptoms in MS<sup>28,30</sup>. Further work should replicate the feasibility of using a mixed spline model for this type of data structure. We performed an initial comparison of disease characteristics and demographics in the two groups of MS patients we identified (early vs. simultaneous onset). While the focus of this paper is to describe the statistical methods, further investigations are underway to examine the clinical impact of these findings, which is outside the scope of the current work.

The notion that the biological disease onset may precede the first clinical symptoms is common across many neurologic and psychiatric diseases<sup>54</sup>. Because the trajectories of cortical, subcortical, grey matter, white matter, and whole brain volumes have been described over the lifespan<sup>Bethlemén</sup>, our method of leveraging these known trajectories can be applied to other neurodegenerative diseases according to their respective brain region of interest. Many neuropsychiatric disease states may benefit from this Digital Twin approach to accelerate the clinical use of precision medicine.

## Informed consent

Informed consent was obtained from all MS and single-site healthy control participants and was approved by the Committee on Human Research at UCSF. The local Institutional Review Board Committee at USC approved the use of additional anonymized brain MRI scans from healthy control participants. All methods were performed in accordance with the relevant guidelines and regulations.

## Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to the condition and constraint from original sources. The data for the Human Connectome Project (HCP: <http://www.humanconnectome.org>), and Alzheimer's Disease Neuroimaging Initiative (ADNI: <http://www.adni-info.org>) should be requested directly through the study website. The data from the single-center, prospective case–control cohort MS study was sponsored by Biogen and GSK. Authorization is needed for using the raw imaging data.

Received: 18 April 2023; Accepted: 26 September 2023

Published online: 28 September 2023

## References

- Grieves, M. & V. J., *Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems* 85–113 (Springer, Cham, 2017).
- Alber, M. *et al.* Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ Digit. Med.* **2**, 115. <https://doi.org/10.1038/s41746-019-0193-y> (2019).
- Filippo, M. D. *et al.* Single-cell digital twins for cancer preclinical investigation. *Methods Mol. Biol. (Clifton, N.J.)* **2088**, 331–343. [https://doi.org/10.1007/978-1-0716-0159-4\\_15](https://doi.org/10.1007/978-1-0716-0159-4_15) (2020).
- Ardila, D. *et al.* End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **25**, 954–961. <https://doi.org/10.1038/s41591-019-0447-x> (2019).
- Rahmim, A. *et al.* Theranostic digital twins for personalized radiopharmaceutical therapies: Reimagining theranostics via computational nuclear oncology. *Front. Oncol.* **12**, 1062592. <https://doi.org/10.3389/fonc.2022.1062592> (2022).
- Calderita, L. V., Vega, A., Barroso-Ramírez, S., Bustos, P. & Núñez, P. Designing a cyber-physical system for ambient assisted living: A use-case analysis for social robot navigation in caregiving centers. *Sensors (Basel, Switzerland)* <https://doi.org/10.3390/s20144005> (2020).
- Corral-Acero, J. *et al.* The “digital twin” to enable the vision of precision cardiology. *Eur. Heart J.* **41**, 4556–4564. <https://doi.org/10.1093/eurheartj/ehaa159> (2020).
- Hirschvogel, M., Jagschies, L., Maier, A., Wildhirt, S. M. & Gee, M. W. An in silico twin for epicardial augmentation of the failing heart. *Int. J. Numer. Methods Biomed. Eng.* **35**, e3233. <https://doi.org/10.1002/cnm.3233> (2019).
- Hose, D. R. *et al.* Cardiovascular models for personalised medicine: Where now and where next?. *Med. Eng. Phys.* **72**, 38–48. <https://doi.org/10.1016/j.medengphy.2019.08.007> (2019).
- Mazumder, O., Roy, D., Bhattacharya, S., Sinha, A. & Pal, A. Synthetic PPG generation from haemodynamic model with baroreflex autoregulation: A Digital twin of cardiovascular system. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2019* 5024–5029. <https://doi.org/10.1109/embc.2019.8856691> (2019).
- Niederer, S. A. *et al.* Creation and application of virtual patient cohorts of heart models. *Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **378**, 20190558. <https://doi.org/10.1098/rsta.2019.0558> (2020).

12. Sharma, P., Suehling, M., Flohr, T. & Comaniciu, D. Artificial intelligence in diagnostic imaging: Status quo, challenges, and future opportunities. *J. Thorac. Imaging* **35**(Suppl 1), S11–s16. <https://doi.org/10.1097/rti.0000000000000499> (2020).
13. Ivanov, D. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *Transp. Res. Part E Logist. Transp. Rev.* **136**, 101922. <https://doi.org/10.1016/j.tre.2020.101922> (2020).
14. Tellechea-Luzardo, J. *et al.* Linking engineered cells to their digital twins: A version control system for strain engineering. *ACS Synth. Biol.* **9**, 536–545. <https://doi.org/10.1021/acssynbio.9b00400> (2020).
15. Voigt, I. *et al.* Digital twins for multiple sclerosis. *Front. Immunol.* **12**, 669811. <https://doi.org/10.3389/fimmu.2021.669811> (2021).
16. Wickramasinghe, N. *et al.* Digital twins to enable better precision and personalized dementia care. *JAMIA Open* **5**, oaac072. <https://doi.org/10.1093/jamiaopen/ooac072> (2022).
17. Lareyre, F., Adam, C., Carrier, M. & Raffort, J. Using digital twins for precision medicine in vascular surgery. *Ann. Vasc. Surg.* **67**, e577–e578. <https://doi.org/10.1016/j.avsg.2020.04.042> (2020).
18. Berger, K. *et al.* Multi-sensor spectral synergies for crop stress detection and monitoring in the optical domain: A review. *Remote Sens. Environ.* **280**, 113198. <https://doi.org/10.1016/j.rse.2022.113198> (2022).
19. Elkefi, S. & Asan, O. Digital twins for managing health care systems: Rapid literature review. *J. Med. Internet Res.* **24**, e37641. <https://doi.org/10.2196/37641> (2022).
20. Rasheed, A., San, O. & Kvamsdal, T. Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access* **8**, 32. <https://doi.org/10.1109/ACCESS.2020.2970143> (2020).
21. Tao, F. & Qi, Q. Make more digital twins. *Nature* **573**, 490–491. <https://doi.org/10.1038/d41586-019-02849-1> (2019).
22. Venkatesh, K. P., Raza, M. M. & Kvedar, J. C. Health digital twins as tools for precision medicine: Considerations for computation, implementation, and regulation. *NPJ Digit. Med.* **5**, 150. <https://doi.org/10.1038/s41746-022-00694-7> (2022).
23. Brown, J. W. L. *et al.* Association of initial disease-modifying therapy with later conversion to secondary progressive multiple sclerosis. *Jama* **321**, 175–187. <https://doi.org/10.1001/jama.2018.20588> (2019).
24. Cerqueira, J. J. *et al.* Time matters in multiple sclerosis: can early treatment and long-term follow-up ensure everyone benefits from the latest advances in multiple sclerosis?. *J. Neurol. Neurosurg. Psychiatry* **89**, 844–850. <https://doi.org/10.1136/jnnp-2017-317509> (2018).
25. Walton, C. *et al.* Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS. *Mult. Scler.* **26**, 1816–1821. <https://doi.org/10.1177/1352458520970841> (2020).
26. Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173. [https://doi.org/10.1016/S1474-4422\(17\)30470-2](https://doi.org/10.1016/S1474-4422(17)30470-2) (2018).
27. De Stefano, N. *et al.* Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs* **28**, 147–156. <https://doi.org/10.1007/s40263-014-0140-z> (2014).
28. Azevedo, C. J. *et al.* Early CNS neurodegeneration in radiologically isolated syndrome. *Neurol. Neuroimmunol. Neuroinflamm.* **2**, e102. <https://doi.org/10.1212/NXI.000000000000102> (2015).
29. Azevedo, C. J. *et al.* Thalamic atrophy in multiple sclerosis: A magnetic resonance imaging marker of neurodegeneration throughout disease. *Ann. Neurol.* **83**, 223–234. <https://doi.org/10.1002/ana.25150> (2018).
30. Okuda, D. T. *et al.* Incidental MRI anomalies suggestive of multiple sclerosis: The radiologically isolated syndrome. *Neurology* **72**, 800–805. <https://doi.org/10.1212/01.wnl.0000335764.14513.1a> (2009).
31. Scahill, R. I. *et al.* A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol.* **60**, 989–994. <https://doi.org/10.1001/archneur.60.7.989> (2003).
32. Bethlehem, R. A. I. *et al.* Brain charts for the human lifespan. *Nature* **604**, 525–533. <https://doi.org/10.1038/s41586-022-04554-y> (2022).
33. Fjell, A. M. & Walhovd, K. B. Structural brain changes in aging: Courses, causes and cognitive consequences. *Rev. Neurosci.* **21**, 187–221. <https://doi.org/10.1515/revneuro.2010.21.3.187> (2010).
34. Fjell, A. M. *et al.* One-year brain atrophy evident in healthy aging. *J. Neurosci.* **29**, 15223–15231. <https://doi.org/10.1523/JNEUROSCI.3252-09.2009> (2009).
35. Walhovd, K. B. *et al.* Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiol. Aging* **26**, 1261–1270; discussion 1275–1268. <https://doi.org/10.1016/j.neurobiolaging.2005.05.020> (2005).
36. Hedman, A. M., van Haren, N. E., Schnack, H. G., Kahn, R. S. & Hulshoff Pol, H. E. Human brain changes across the life span: A review of 56 longitudinal magnetic resonance imaging studies. *Hum. Brain Mapp.* **33**, 1987–2002. <https://doi.org/10.1002/hbm.21334> (2012).
37. Fjell, A. M. *et al.* Minute effects of sex on the aging brain: A multisample magnetic resonance imaging study of healthy aging and Alzheimer's disease. *J. Neurosci.* **29**, 8774–8783. <https://doi.org/10.1523/JNEUROSCI.0115-09.2009> (2009).
38. Fjell, A. M. *et al.* When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. *Neuroimage* **50**, 1376–1383. <https://doi.org/10.1016/j.neuroimage.2010.01.061> (2010).
39. Schippling, S. *et al.* Global and regional annual brain volume loss rates in physiological aging. *J. Neurol.* **264**, 520–528. <https://doi.org/10.1007/s00415-016-8374-y> (2017).
40. Chen, H. *et al.* Statistical approaches for the study of cognitive and brain aging. *Front. Aging Neurosci.* **8**, 176. <https://doi.org/10.3389/fnagi.2016.00176> (2016).
41. Hastie, T., Friedman, J. & Tibshirani, R. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* 313 (Springer, 2018).
42. Wahba, G. *Spline Models for Observational Data* (Society for Industrial and Applied Mathematics, 1990).
43. Wood, S. N. Thin plate regression splines. *J. R. Stat. Soc. B* **65**, 95–114. <https://doi.org/10.1111/1467-9868.00374> (2003).
44. Wood, S. N. *Generalized Additive Models: An Introduction with R* (CRC Press, 2017).
45. Eilers, P. H. C. & Marx, B. D. Flexible smoothing with B-splines and penalties. *Stat. Sci.* **11**, 89–102. <https://doi.org/10.1214/ss/1038425655> (1996).
46. Ruppert, D. W. M. & Carroll, R. J. *Semiparametric regression* 186–193 (Cambridge University Press, New York, 2003).
47. Krivobokova, T. & Kauermann, G. A note on penalized spline smoothing with correlated errors. *J. Am. Stat. Assoc.* **102**, 1328–1337. <https://doi.org/10.1198/016214507000000978> (2007).
48. Roy, A. Estimating correlation coefficient between two variables with repeated observations using mixed effects model. *Biom. J.* **48**, 286–301. <https://doi.org/10.1002/bimj.200510192> (2006).
49. Irimata, K., Paul, K. & Li, X. *Estimation of Correlation Coefficient in Data with Repeated Measures* Paper 2424–2018. (SAS Global, Washington, DC, USA, 2018)
50. Mofrad, S. A., Lundervold, A. J., Vik, A. & Lundervold, A. S. Cognitive and MRI trajectories for prediction of Alzheimer's disease. *Sci. Rep.* **11**, 2122. <https://doi.org/10.1038/s41598-020-78095-7> (2021).
51. Jain, R. & Xu, W. HDSI: High dimensional selection with interactions algorithm on feature selection and testing. *PLoS One* **16**, e0246159. <https://doi.org/10.1371/journal.pone.0246159> (2021).
52. Fortin, J. P. *et al.* Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage* **167**, 104–120. <https://doi.org/10.1016/j.neuroimage.2017.11.024> (2018).
53. Varghese, B. A. *et al.* Identification of robust and reproducible CT-texture metrics using a customized 3D-printed texture phantom. *J. Appl. Clin. Med. Phys.* **22**, 98–107. <https://doi.org/10.1002/acm2.13162> (2021).

54. Cacciaguerra, L. *et al.* Dynamic volumetric changes of hippocampal subfields in clinically isolated syndrome patients: A 2-year MRI study. *Mult. Scler.* **25**, 1232–1242. <https://doi.org/10.1177/1352458518787347> (2019).

## Acknowledgements

The authors thank the study participants, MS clinicians at the UCSF MS Center who referred patients to the EPIC study, and the research coordinators at UCSF for data collection. The authors also gratefully acknowledge the funding sources for this study, including the NIH (NINDS R01NS062885 to D.P.) and the National Multiple Sclerosis Society (RG-1802-30140 to C.J.A.) for MRI data acquisition, processing, and analysis, and Biogen and Glaxo-Smith-Klein for MRI data acquisition of MS subjects. The authors would also like to acknowledge The Human Connectome Project at <http://www.humanconnectome.org> and the Alzheimer's Disease Neuroimaging Initiative at <http://www.adni-info.org> for brain MRI data availability.

## Author contributions

Dr. Cen, Dr. M.G., Dr. C.J.A. and Dr. Pelletier contributed to the original research concept and design. Dr. D.P., and Dr. Gebregziabher contributed to methodology development and data analysis. Mr. Moazami contributed MRI imaging processing. Dr. M.G.C. wrote the main manuscript text. All authors reviewed and edited the manuscript.

## Competing interests

In the past 3 years, Dr. Azevedo has received grant support from the National Multiple Sclerosis Society and the NIH. She has received consulting fees from Horizon Therapeutics, Genentech, Sanofi Genzyme, TG Therapeutics, and EMD Serono. She has received honoraria for serving on grant review committees for the Department of Defense and the NIH and for participation in unbranded CME activities from the American Academy of Neurology, Efficient LLC, Spire Learning, and Catamount Medical Education. Dr. Cen, Dr. Pelletier, Dr. Gebregziabher and Mr. Moazami have no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-43618-5>.

**Correspondence** and requests for materials should be addressed to S.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023