



# Disease Modelling of Cognitive Outcomes and Biomarkers in the European Prevention of Alzheimer's Dementia Longitudinal Cohort

James Howlett<sup>1†</sup>, Steven M. Hill<sup>1‡</sup>, Craig W. Ritchie<sup>2</sup> and Brian D. M. Tom<sup>1\*</sup>

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge, United Kingdom, <sup>2</sup>Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom

## OPEN ACCESS

### Edited by:

Neil P. Oxtoby,  
University College London,  
United Kingdom

### Reviewed by:

Vikram Venkatraghavan,  
Erasmus Medical Center, Netherlands  
Cameron Shand,  
University College London,  
United Kingdom

### \*Correspondence:

Brian D. M. Tom  
brian.tom@mrc-bsu.cam.ac.uk

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### ‡Present Address:

Steven M. Hill,  
Cancer Research UK Manchester  
Institute Cancer Biomarker Centre,  
University of Manchester, Manchester,  
United Kingdom

### Specialty section:

This article was submitted to  
Medicine and Public Health,  
a section of the journal  
Frontiers in Big Data

Received: 04 March 2021

Accepted: 30 July 2021

Published: 20 August 2021

### Citation:

Howlett J, Hill SM, Ritchie CW and  
Tom BDM (2021) Disease Modelling of  
Cognitive Outcomes and Biomarkers  
in the European Prevention of  
Alzheimer's Dementia  
Longitudinal Cohort.  
Front. Big Data 4:676168.  
doi: 10.3389/fdata.2021.676168

A key challenge for the secondary prevention of Alzheimer's dementia is the need to identify individuals early on in the disease process through sensitive cognitive tests and biomarkers. The European Prevention of Alzheimer's Dementia (EPAD) consortium recruited participants into a longitudinal cohort study with the aim of building a readiness cohort for a proof-of-concept clinical trial and also to generate a rich longitudinal data-set for disease modelling. Data have been collected on a wide range of measurements including cognitive outcomes, neuroimaging, cerebrospinal fluid biomarkers, genetics and other clinical and environmental risk factors, and are available for 1,828 eligible participants at baseline, 1,567 at 6 months, 1,188 at one-year follow-up, 383 at 2 years, and 89 participants at three-year follow-up visit. We novelly apply state-of-the-art longitudinal modelling and risk stratification approaches to these data in order to characterise disease progression and biological heterogeneity within the cohort. Specifically, we use longitudinal class-specific mixed effects models to characterise the different clinical disease trajectories and a semi-supervised Bayesian clustering approach to explore whether participants can be stratified into homogeneous subgroups that have different patterns of cognitive functioning evolution, while also having subgroup-specific profiles in terms of baseline biomarkers and longitudinal rate of change in biomarkers.

**Keywords: Alzheimer's disease, biomarkers, cognitive functioning, disease modelling, European prevention of Alzheimer's dementia, latent class mixed models, precision medicine, Bayesian profile regression**

## 1 INTRODUCTION

Alzheimer's disease (AD), the leading cause of dementia globally (Livingston et al., 2017), is characterised by synaptic dysfunction and neurodegeneration (e.g., neuronal loss), triggered by sequential accumulation of amyloid plaques and neurofibrillary tangles (aggregates of hyperphosphorylated tau proteins) (Braak and Braak, 1991). The exact ordering of the pathological cascade of events, leading to clinical symptoms of cognitive deterioration and dementia, has been actively researched over the last decade. Jack and colleagues (Jack et al., 2010; Jack et al., 2013) hypothesised that there is an underlying disease process and that the temporal ordering of changes in key biomarkers and their dynamics characterise the full spectrum of the disease throughout the different successive stages of pre-clinical, prodromal and dementia.

On the whole there have been few treatment successes (and none of these are disease-modifying) despite substantive investment in pharmacological compounds for Alzheimer's disease in symptomatic populations and early promise shown in pre-clinical studies (Gauthier et al., 2016; Winblad et al., 2016; Anderson et al., 2017). There may be a number of possible explanations for the many failures including inadequate drug dosages, incorrect treatment targets and inappropriate trial populations where the disease process is too far along to be amenable to treatment (Raket, 2020; Shi et al., 2020; Yiannopoulou and Papageorgiou, 2020). There is a consensus that the genesis of AD pathology occurs decades before the onset of dementia symptoms (Braak and Braak, 1997; Hardy and Selkoe, 2002; Jack et al., 2010; Bateman et al., 2012; Braak and Del Tredici, 2012; Jack et al., 2013). This thus presents an opportunity for early disease course modification before dementia onset and even prior to clinical symptoms. As such there is great interest—from both academia and industry—in accurately identifying groups of individuals with higher likelihood of progressing to AD dementia for natural history studies, early phase treatment trials and for participation in secondary prevention trials where, for example, they may have evidence of AD pathology through relevant biomarker abnormalities but no clinical evidence of symptoms of dementia (Ritchie et al., 2016; Watts, 2018).

Current proposals for defining an individual's probability for developing AD dementia or for modelling cognitive deterioration based on biomarkers and/or clinical symptoms have been focused on the stage of AD close to dementia onset. Various disease progression and sub-type approaches have been proposed and developed. These include survival and multi-state models for investigating transitions between disease states (Hubbard and Zhou, 2011; Vos et al., 2013; van den Hout, 2016; Wei and Kryscio, 2016; Robitaille et al., 2018; Zhang et al., 2019); mixed effects models (linear, generalized, non-linear) that incorporate subject-specific random effects and can be extended to handle latent time shifts, random change points, latent factors, processes and classes, hidden states, and multiple outcomes (Hall et al., 2000; Jedynak et al., 2012; Liu et al., 2013; Proust-Lima et al., 2013; Donohue et al., 2014; Samtani et al., 2014; Lai et al., 2016; Zhang et al., 2016; Geifman et al., 2018; Li et al., 2018; Wang et al., 2018; Lorenzi et al., 2019; Proust-Lima et al., 2019; Villeneuve et al., 2019; Younes et al., 2019; Bachman et al., 2020; Kulason et al., 2020; Raket, 2020; Segalas et al., 2020; Williams et al., 2020) and can be combined with models for event-history data (Marioni et al., 2014; Blanche et al., 2015; Proust-Lima et al., 2016; Rouanet et al., 2016; Li et al., 2017; Iddi et al., 2019; Li and Luo, 2019; Wu et al., 2020); event-based models which attempt to model the pathological cascade of events occurring as the disease develops and progresses through disease stages (Fonteijn et al., 2012; Young et al., 2014; Chen et al., 2016; Goyal et al., 2018; Oxtoby et al., 2018); and various clustering approaches for discovering risk stratification/disease progression groups and endotypes. For example, those based on hierarchical, partitioning and model-based clustering algorithms/methods (Dong et al., 2016; Racine et al., 2016; Dong et al., 2017; ten Kate et al., 2018; Young et al., 2018). Moreover, various machine

learning and other statistical approaches have been proposed for both disease progression, prediction and subgroup identification in Alzheimer's disease (Fiot et al., 2014; Schmidt-Richberg et al., 2016; Cheng et al., 2017; Bhagwat et al., 2018; Khanna et al., 2018; de Jong et al., 2019; Martí-Juan et al., 2019; Brand et al., 2020; Golriz Khatami et al., 2020; Lei et al., 2020; Martí-Juan et al., 2020; Lin et al., 2021; Zhang et al., 2021).

However, in the earlier stages of disease, the development of disease models is far more challenging due to the relatively slow progression of the disease and clinical measures being insufficiently sensitive to detect such subtle changes. In order to develop disease models in the early stages when individuals do not have symptoms, or express only subjective complaints of cognitive decline or have only mild cognitive symptoms, it is necessary to undertake longitudinal follow-up of these individuals measuring reliable biomarkers of pathological changes alongside clinical outcomes. Ideally individuals would be followed-up over an extended period of time to ensure sufficient proportions make transitions through the various disease stages to dementia. Ultimately, these disease models would better inform patient selection into trials, improve understanding of AD progression in individuals and allow a more tailored approach to clinical management and targeting of disease modifying treatments to individuals (i.e., precision medicine) based on a range of biomarker modalities (e.g., neuroimaging, cerebrospinal fluid (CSF), blood), cognitive and clinical measures and risk factors.

Against this backdrop, the European Prevention of Alzheimer's Dementia (EPAD) consortium (Ritchie et al., 2016) was initiated as a large public-private partnership, and funded by the Innovative Medicines Initiative (IMI) Joint Undertaking. A total of 39 European organisations or "partners" were involved in the EPAD consortium. EPAD was developed as an interdisciplinary research initiative with an aim of improving the understanding of the early stages of Alzheimer's disease and delivering new preventative treatments.

The EPAD Longitudinal Cohort Study (LCS) was a prospective, multi-centre, pan-European study set up with the dual objectives of developing accurate longitudinal models over the entire course of Alzheimer's disease (AD) prior to the onset of dementia and creating a trial-ready cohort for potential recruitment into the EPAD Proof-of-Concept (PoC) Trial (Solomon et al., 2018). It was designed as a long-term observational study with recruitment from different types of existing parent cohorts (PCs) across Europe (e.g., population-based, memory clinics) and then, later on, more directly from clinical settings. It aimed to provide both a well-phenotyped population covering the full continuum of risk of subsequent AD dementia development and enough participants with particular profiles potentially eligible for an adaptive designed trial. This aim was achieved through monitoring of the evolving characteristics of the EPAD cohort and use of a flexible and dynamic approach to selection into the LCS that allowed over- and under-sampling by particular characteristics already available in the PCs. The other component of the EPAD programme, the EPAD PoC Trial, was designed to provide an environment for testing multiple interventions for the secondary prevention of AD dementia.

Using the data collected in the LCS on cognitive and clinical outcomes, biomarkers and risk factors, we aim to develop state-of-the-art models for disease progression and stratification which can be used 1) to inform selective recruitment and adaptation in clinical trials, 2) for longitudinal prediction and stratification, 3) for subgroup identification based on both baseline and longitudinal biomarker profiles and, ultimately, 4) to help improve treatment and clinical management decisions. We adopt a two-stage approach, where we first identify subpopulations/classes with different underlying, potentially AD-related cognitive/functional trajectory patterns (i.e., latent clinical phenotypes) over time after controlling for known exogenous risk factors (constitutional and genetic). These latent phenotypes are then jointly modelled with endogenous neuroimaging and CSF biomarkers to identify homogeneous subgroups/clusters based on biomarker profiles (i.e., neuropathological endotypes) that are linked to these trajectory patterns.

## 2 METHODS

### 2.1 Data

We performed all analyses on the V. IMI data release from the EPAD cohort (<http://ep-ad.org/open-access-data/access/>). Briefly, a total of 2,096 participants were screened and entered the cohort. Any participants who failed screening, had a baseline global clinical dementia rating (CDR)  $\geq 1$ , or had a diagnosis of Alzheimer's dementia at baseline were excluded, leaving 1,828 eligible participants. Participants were aged at least 50 years old, with either a CDR global score of 0 ( $n = 1,313$ ) or 0.5 ( $n = 498$ ) (The CDR global scores for seventeen participants were missing.) Recruitment occurred across 31 centres from 10 different European countries. Follow-up visits were designed to occur at 6 months, 1 year and yearly thereafter. Unfortunately, the LCS closed at the end of the IMI-funding period and therefore the maximum number of visits was five. Of the 1,828 participants with a baseline visit, 1,567 attended the 6-months visit, 1,188 attended the 1-year visit and 396 and 89 attended the 2-years and 3-years visits respectively. Two hundred and fifty four participants only had a baseline visit, 389 had two visits (including five who had a baseline and 1-year visit but not 6-months), 791 had three visits (including 2 who had baseline, 1-year and 2-years visits but not a 6-months visit; the remaining attended the first three visits), 307 had four visits (including 2 who had baseline, 6-months, 2-years and 3-years visits but not a 1-year visit; the remaining had all visits up to 2 years) and 87 had five visits. We restrict our study to the 1,574 participants who had more than one visit.

The variables used in the models can be considered to belong to four domains: 1) outcomes, 2) baseline risk factors, 3) baseline biomarkers, and 4) longitudinal biomarkers.

### Outcomes

The outcomes used were transformations of CDR sum of boxes (CDRSB) and Mini-Mental State Examination (MMSE) scores.

To deal with floor and ceiling effects of CDRSB, a logistic transformation was applied to CDRSB as defined in Eq. 1:

$$tCDRSB = -\log\left(\frac{(CDRSB + 0.1)}{(18 - CDRSB + 0.1)}\right) \quad (1)$$

A normalising transformation was applied to MMSE values, converting MMSE from a 0–30 scale to nMMSE on a 0–100 scale to deal with curvilinearity (Philipps et al., 2014). CDRSB was scheduled to be collected at all visits but MMSE was not designed to be collected at the 6-months visit.

### Baseline Risk Factors

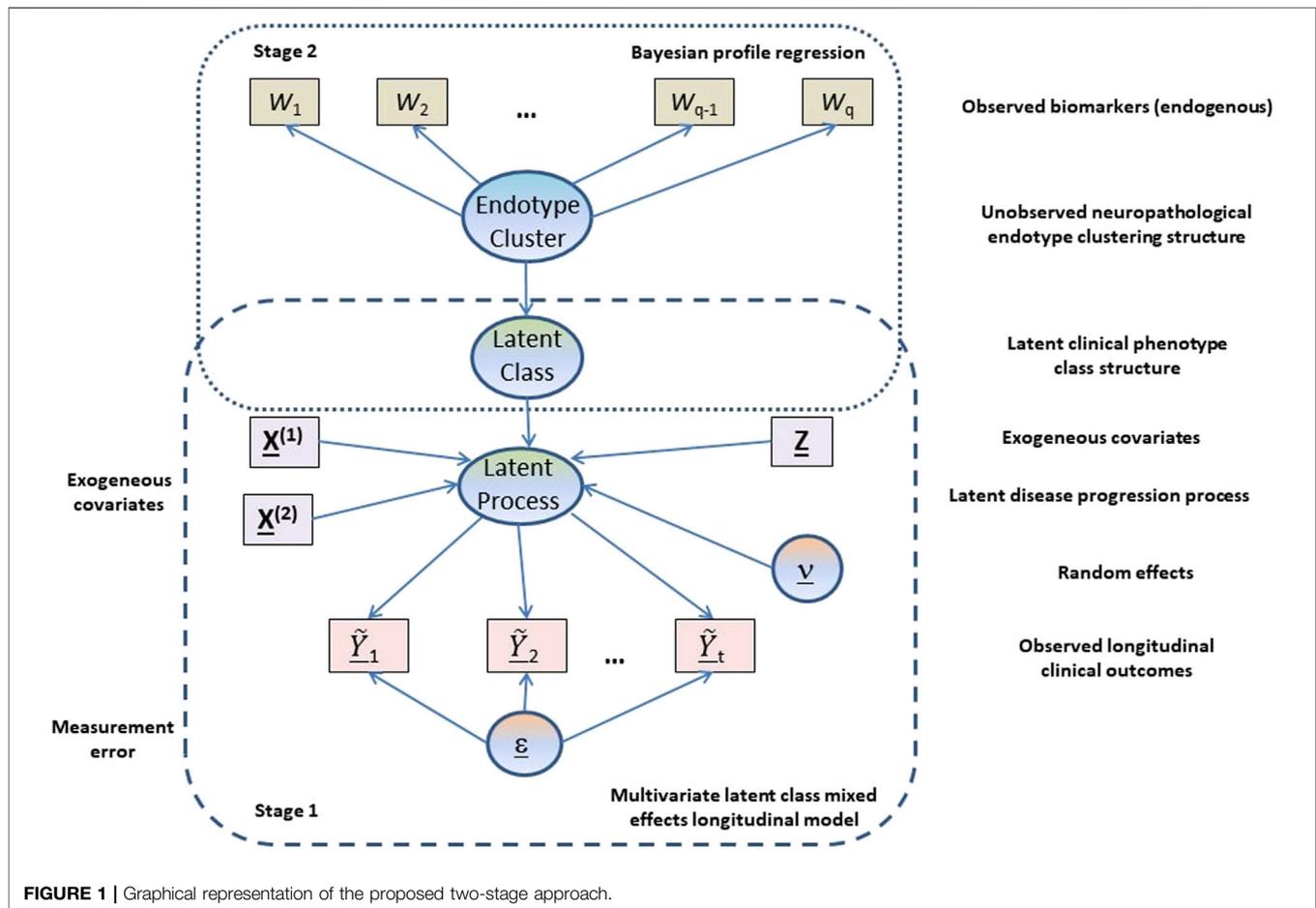
Baseline risk factors included age, sex, education, family history of AD (first-degree relatives), and APOE $\epsilon$ 4 carrier status. Age is treated as a continuous variable. Sex, family history, and APOE are binary. Education was recorded in the LCS as years of formal education. However, as the values have different interpretations for different countries, years of education was converted to a three-category highest educational attainment level variable labelled 1, 2, and 3 on a country-specific basis (European Commission/EACEA/Eurydice, 2018). Level 1 is defined as up to secondary education, level 2 as beyond secondary education up to undergraduate ordinary degree, and level 3 as postgraduate studies.

### Baseline Biomarkers

Baseline biomarkers included:

- the ratio of phosphorylated tau (pTau) to amyloid-beta 42 ( $A\beta$ ), derived from CSF samples using the fully automated Roche Elecsys System in a single laboratory;
- volumetric imaging variables of the total of the left and right hippocampi and of the total of the four ventricles adjusting for head size by dividing by the pseudo total intracranial factor (HV and VV), processed by IXICO using the learning embeddings for atlas propagation (LEAP) method (Wolz et al., 2010);
- neurological radiological reads variables obtained through central assessment of magnetic resonance (MR) images by IXICO raters following a standardised, compliant and efficient workflow (Ritchie et al., 2020; ten Kate et al., 2018):
  - average of left and right medial temporal lobe atrophy (MTA);
  - Fazekas scale deep (FSD) and Fazekas scale periventricular (FSPV); and
  - five regional age-related white matter change (ARWMC) variables.

For EPAD participants, values of pTau/ $A\beta$   $> 0.024$  are here defined as CSF “AD positive” based on the biomarker cut-offs derived by Roche for EPAD using the methodology in (Hansson et al., 2018; Schindler et al., 2018), and reflect either decreased concentrations of  $A\beta$  (a marker of amyloidosis) or increased levels of pTau (a marker of neurofibrillary tangles). All radiological reads biomarkers were converted to binary variables  $< 1$  and  $\geq 1$ , except for Fazekas scale deep which was dichotomised instead at 2. A score of 0 for all radiological reads



**FIGURE 1** | Graphical representation of the proposed two-stage approach.

variables indicates no pathology and scores  $\geq 1$  (and  $\geq 2$  for Fazekas scale deep) indicate some pathology. A score of 0.5 in the average of left and right MTA is assumed to provide inconclusive evidence of pathology. A combined ARWMC variable was created that counted the number of age-related regions with evidence of white matter lesion cerebrovascular pathology, and a count  $\geq 3$  indicated that the majority of regions had signs of pathology.

### Longitudinal Biomarkers

Longitudinal biomarkers considered were derived from the MR volumetric imaging variables of total hippocampal volume and total ventricular volume adjusting for head size. The processing of the longitudinal volumetric variables was also performed by IXICO using LEAP. The rates of change in the adjusted total hippocampal and ventricles volumes were calculated by dividing the difference between the last observed and baseline volumes by the time in study (in years) between the taking of the last and baseline volumes. These rate of change (i.e., annualised change) variables were used in our analyses to describe the longitudinal changes in biomarkers.

## 2.2 Statistical Methods

Our analysis is based on a two-stage approach (see **Figure 1**) where in the first stage a multivariate latent class linear mixed effects modelling approach is adopted to model the longitudinal

cognitive and clinical outcomes adjusting for constitutional and genetic risk factors purported to be important in AD disease progression or related to selection into the EPAD LCS. From the multivariate latent class linear mixed effects model, latent clinical phenotypes corresponding to the latent classes are extracted to characterise the various mean trajectory profiles which individuals may follow over time. These latent phenotypes result from a hard assignment of individuals to specific latent classes based on their posterior probabilities of class membership. In the second stage, a probabilistic outcome-guided clustering approach based on Dirichlet process mixture modelling called Bayesian profile regression is applied to the latent phenotypes alongside the CSF and neuroimaging biomarkers. This aims to identify homogeneous clusters of participants with particular neuropathological endotypes characterised by biomarker profiles linked to clinical disease progression. Note that the latent phenotypes and endotypes are not meant to represent a grouping orthogonal to disease severity or stage, but reflect and characterise potential underlying processes and features that give rise to or are associated with disease severity or stage.

The specific statistical formulation of this two-stage modelling approach for disease progression, trajectory stratification and subgroup identification are outlined in the next two subsections. Missing response data are assumed to be missing at random

(MAR) for both stages, which allows valid inference using likelihood approaches.

### 2.2.1 Multivariate Latent Class Linear Mixed Effects Model

We used a multivariate latent class linear mixed effects model (MLCMM) to identify  $G$  mean profiles of trajectories corresponding to  $G$  latent classes or sub-populations of individuals (Lai et al., 2016; Proust-Lima et al., 2017). This model assumes that a latent process  $\Lambda_i(t)$  generates the  $K$  longitudinal outcomes at time  $t$ , and this latent process is characterised by the mean trajectory profile corresponding to the latent class membership of individual  $i$ .  $Y_{ijk}$  is a measure of outcome  $k$  ( $k = 1, \dots, K$ ) for subject  $i$  ( $i = 1, \dots, N$ ) at measurement occasion  $j$  ( $j = 1, \dots, n_{ik}$ ), with associated time of outcome measurement from start of study  $t_{ijk}$ .  $Y_{ijk}$  is related to  $\Lambda_i(t_{ijk})$  via an outcome-specific link function. Here for the purposes of this paper, we assume a linear transformation link function (others are possible) for the outcomes with outcome-specific parameters. That is,  $\tilde{Y}_{ijk} = \frac{Y_{ijk} - \eta_{1k}}{\eta_{2k}}$ ,  $k = 1, \dots, K$ . These transformations allow the transformed outcomes to be interpreted as noisy measurements of the underlying latent process with outcome-specific measurement errors.

The general formulation of the linear mixed effects part of our model given membership to latent class  $g$  is

$$\tilde{Y}_{ijk|c_i=g} = \Lambda_i(t_{ijk})|_{c_i=g} + \epsilon_{ijk} \tag{2}$$

with

$$\Lambda_i(t_{ijk})|_{c_i=g} = X_{ijk}^{(1)T} \beta + X_{ijk}^{(2)T} \gamma_g + Z_{ijk}^T v_{ig}, \tag{3}$$

where  $c_i$  is the latent class variable,  $X_{ijk}^{(1)}$  are the covariates associated with the class-independent fixed effects  $\beta$  and  $X_{ijk}^{(2)}$  are the covariates associated with the class-specific fixed effects  $\gamma_g$ .  $Z_{ijk}$  are the covariates associated with the class-specific random effects  $v_{ig}$ , which are from a zero-mean multivariate normal with variance-covariance matrix  $\omega_g B$ , where  $B$  is left unspecified and  $\omega_g$  is a positive proportionality factor (with  $\omega_g = 1$  to ensure identifiability). The measurement errors  $\{\epsilon_{ijk}\}$  are assumed to be independent Gaussian random variables with mean 0 and outcome-specific variances  $\sigma_k^2$  ( $k = 1, \dots, K$ ).

The latent variable  $c_i$  equals  $g$  when subject  $i$  belongs to latent class  $g$ . To complete the specification of our multivariate latent class mixed model, the probability of individual  $i$  belonging to class  $g$  is described by the multinomial logistic submodel without covariates given by Eq. 4:

$$\pi_{ig} = P(c_i = g) = \frac{e^{\xi_{0g}}}{\sum_{l=1}^G e^{\xi_{0l}}}, \tag{4}$$

where  $\xi_{0g}$  is the intercept parameter for class  $g$ . Extension of this latent class membership submodel to include covariates is straightforward. The full MLCMM is fitted using maximum likelihood estimation within R (R Core Team, 2017) using the `multlcm` function in the `lcm` package (Proust-Lima et al., 2017).

In our application, we included the logistic transformed CDRSB, tCDRSB, and normalised MMSE, nMMSE, as outcomes ( $K = 2$ ) in our MLCMM formulation. For both these outcomes and the latent process, a higher value indicates

less cognitive/functional impairment (i.e., better cognitive functioning). We used time in study as the time scale and allowed class-specific fixed intercepts and slopes (time in study effects). As maximum follow-up in the EPAD study population was 3 years and 4 months and the majority of subjects had two or three visits, we considered only linear trends in an individual's underlying disease process. The baseline risk factors described in Section 2.1 were introduced into Eq. 3 with associated class-independent fixed effects. We included only class-specific random intercepts into the latent process model, which are introduced to induce correlation across the longitudinal observations of an outcome for an individual and to better align participants in terms of where they fall on the disease time scale. The variance of the random intercept for the reference class is not estimated by the model and is set to be 1. The best choice of the number of latent classes was made using the Bayesian Information Criterion (BIC) and the relative entropy.

All observations with either a recorded CDRSB or MMSE were considered for inclusion in the model provided that individuals had 2 or more visits. These corresponded to 4,795 visits on 1,574 participants. Of which, there were 3,228 visits with both CDRSB and MMSE present, 1,558 visits with only CDRSB present, and nine visits with only MMSE present. Of the 1,574 individuals, 86 had five observation-visits, 305, 789, 384 and 10 had 4, 3, 2 and 1 observation-visits with either CDRSB or MMSE or both present respectively. However, 31 individuals had missing APOE $\epsilon$ 4 carrier status information and were excluded. This thus resulted in 1,543 individuals be included in the MLCMM analysis.

### 2.2.2 Bayesian Profile Regression

Bayesian profile regression (Molitor et al., 2010) is a non-parametric outcome-guided clustering approach that links an outcome variable to covariates via cluster membership. Here, it was applied to identify  $G^*$  clusters of participants, with each cluster characterised by particular clinical disease progression phenotypes (latent classes from the MLCMM analysis) and a particular CSF/neuroimaging biomarker profile. These clusters can be interpreted as corresponding to different neuropathological endotypes.

Bayesian profile regression uses a Dirichlet process mixture model (DPMM), which can be regarded as the limit of a finite mixture model as the number of components goes to infinity. That is, for observed data  $D_i$  for subject  $i$ , we have the following DPMM likelihood:

$$p(D_i|\pi^*, \Theta) = \sum_{h=1}^{\infty} p(c_i^* = h|\pi^*)p(D_i|c_i^* = h, \Theta) \tag{5}$$

$$= \sum_{h=1}^{\infty} \pi_h^* f(D_i|\Theta_h), \tag{6}$$

where  $c_i^* \in \mathbb{Z}^+$  denotes latent cluster membership,  $\pi^* = (\pi_1^*, \pi_2^*, \dots)^T$  are mixture component (cluster) weights and  $\Theta^T = (\Theta_1^T, \Theta_2^T, \dots)$  are component-specific parameters for the mixture component densities, indexed by  $h \in \mathbb{Z}^+$ .

In addition to covariates  $W_i$  for subject  $i$ , Bayesian profile regression models an outcome  $Y_i^*$  that also informs the clustering and is assumed to be conditionally independent of the covariates given cluster assignment  $c_i^*$ . Furthermore, covariates can be a mix of discrete and continuous,  $W_i^* = (W_i^{(d)T}, W_i^{(c)T})$ , with discrete

covariates  $W_i^{(d)}$  and continuous covariates  $W_i^{(c)}$  also assumed to be conditionally independent given  $c_i^*$ . We therefore have observed data  $D_i^T = (Y_i^*, W_i^{(d)T}, W_i^{(c)T})$  for subject  $i$  and Eq. 6 becomes

$$p(Y_i^*, W_i^{(d)}, W_i^{(c)} | \pi^*, \Theta) = \sum_{h=1}^{\infty} \pi_h^* f(Y_i^* | \Theta_h^{(o)}) f(W_i^{(d)} | \Theta_h^{(d)}) f(W_i^{(c)} | \Theta_h^{(c)}), \tag{7}$$

where  $\Theta_h^T = (\Theta_h^{(o)T}, \Theta_h^{(d)T}, \Theta_h^{(c)T})$  are the component-specific parameters for the outcome, discrete covariate and continuous covariate densities respectively.

The stick-breaking construction of the DPMM (Sethuraman, 1994) is used within Bayesian profile regression which gives the following formulation for the prior on the mixture weights:  $\pi_1^* = V_1$  and  $\pi_h^* = V_h \prod_{l < h} (1 - V_l)$  for  $h \geq 2$  with  $V_h \stackrel{iid}{\sim} \text{Beta}(1, \alpha)$ . The concentration hyperparameter  $\alpha$ , which itself has a gamma prior distribution, affects the mixture weight distribution and implicitly informs the number of non-empty clusters. One of the key desirable properties of a DPMM approach to clustering is the removal of the need to pre-specify the number of clusters. Prior distributions are also placed on the component-specific parameters  $\Theta_h$  and Markov chain Monte Carlo (MCMC) is used to fit the resulting profile regression model (see (Liverani et al., 2015) for details of the prior distributions and for computational aspects of the MCMC).

In our application, the outcome variable for each subject is the latent class predicted from the MLCMM analysis, i.e.  $Y_i^* = \hat{c}_i$ . This is treated as a categorical variable with cluster-dependent parameters:  $Y_i^* | \Theta_h^{(o)} \sim \text{Cat}(\theta_{h,1}^{(o)}, \theta_{h,2}^{(o)}, \dots, \theta_{h,\hat{G}}^{(o)})$ , where  $\hat{G}$  is the estimated number of latent classes in the MLCMM. The covariates used in the model are the baseline and longitudinal biomarkers described in Section 2.1. In particular, we included five binary baseline covariates (pTau/Aβ, MTA, FSD, FSPV, ARWMC combined) for each subject, each independently taking a Bernoulli distribution given cluster assignment:  $W_{i,q}^{(d)} | \Theta_{h,q}^{(d)} \sim \text{Bern}(\theta_{h,q}^{(d)})$  for  $q = 1, \dots, 5$ . Additionally, four continuous covariates (standardised) were included—adjusted total hippocampal and ventricles volumes at baseline, HV and VV, and their corresponding longitudinal rate of changes, HV rate and VV rate—jointly taking a multivariate Gaussian distribution given cluster assignment:  $W_i^{(c)} | \Theta_h^{(c)} \sim \mathcal{N}_4(\mu_h, \Sigma_h)$ . This allows for the correlation between the continuous covariates to be taken into account.

Since the clustering assignments and number of clusters vary across the MCMC iterations, it is useful to obtain a “representative” clustering that summarises the MCMC output. Following (Molitor et al., 2010; Liverani et al., 2015), we find a “representative” clustering based on the  $N \times N$  posterior similarity matrix  $S$ , where element  $(i, j)$  of  $S$  is the proportion of MCMC iterations where subjects  $i$  and  $j$  are assigned to the same cluster. The partitioning around medoids (PAM) clustering algorithm (Kaufman and Rousseeuw, 1990) is applied to the posterior dissimilarity matrix  $1 - S$  to find a clustering of the subjects that is consistent with  $S$ , with the optimal number of clusters selected using the silhouette width method (Rousseeuw, 1987).

An advantage of the DPMM clustering framework is that it takes uncertainty in the clustering (including the number of

clusters) into account. This allows the uncertainty associated with the “representative” clustering to be investigated. If we let  $C_h^{(\text{rep})}$  denote the subset of subjects allocated to cluster  $h$  in the “representative” clustering, then at MCMC iteration  $r$  we can calculate the average value of mixture component parameters for subjects in  $C_h^{(\text{rep})}$ . For example, for the Bernoulli distribution parameter for binary covariate  $q$  we calculate

$$\bar{\theta}_{h,q}^{(d)}(r) = \frac{1}{n_h} \sum_{i \in C_h^{(\text{rep})}} \theta_{c_i^*(r),q}^{(d)}(r) \tag{8}$$

where  $n_h$  is the number of subjects in  $C_h^{(\text{rep})}$  and  $\theta_{c_i^*(r),q}^{(d)}(r)$  is the sampled Bernoulli parameter for the cluster  $c_i^*(r)$  that subject  $i$  is allocated to at MCMC iteration  $r$ . The distribution of  $\bar{\theta}_{h,q}^{(d)}(r)$  across the MCMC iterations (i.e., the posterior distribution) gives an insight into the uncertainty of cluster  $h$  in the “representative” clustering; narrower credible intervals indicates a more consistent clustering. These distributions can be computed for all of the “representative” clusters and for all of the mixture component parameters associated with the outcome variable and covariates.

Bayesian profile regression is implemented in the R package PReMiuM (Liverani et al., 2015) and this was used to fit the model and perform the post-processing analysis (PReMiuM package version 3.2.3; R version 3.6.3; default settings for hyperparameters used; run for 350,000 MCMC iterations with first 100,000 discarded as burn-in). Convergence of the MCMC procedure was investigated by checking agreement between the “representative” clusterings from six independent chains (quantified using the adjusted Rand index) and by inspection of posterior parameters (see (Liverani et al., 2015) for more details of convergence diagnostics). Consensus clustering of the consensus dissimilarity matrix, obtained through averaging of the dissimilarity matrices from the six independent chains and applying PAM to this matrix, resulted in the final representative clustering structure. The adjusted Rand indices assessing agreement between the final representative consensus clustering with the representative clusterings from the six independent chains are calculated and reported. Moreover, the lower triangular part of the individual posterior dissimilarity matrices from the six independent chains are compared to the lower triangular part of the consensus posterior dissimilarity matrix using Pearson’s correlation. Risk and covariate profiles are derived through pooling of MCMC iterations across the six chains and using the final representative consensus clustering. Additionally, Bayesian profile regression without the latent classes as outcome was performed to obtain a baseline/reference clustering structure based purely on the biomarkers. All 1,543 subjects included in the MLCMM analysis were included in the Bayesian profile regression analysis.

### 2.2.3 Validation

The final results of our multivariate latent class mixed model and Bayesian profile regression analysis on the full data-set were assessed for class and cluster validity through stability assessment under repeated sub-setting. We repeatedly (i.e., ten times) split the full data-set into two subsets, by first stratifying the full data-set by number of visits and then randomly allocating (with equal probability) within each strata a participant to belong to either

the first or second subset. Our proposed two-stage approach is then applied in turn to each subset to estimate a latent class structure followed by clustering structure as described in the previous two subsections. For assessing class validity for each split, we begin by using the multivariate latent class model trained on one subset to predict the class membership of participants in the other subset and vice versa. We next cross-tabulate the out-of-sample predictions of class memberships (based on the model trained on the subset that does not include the participant for whom a prediction is being made) with the in-sample class membership assignments (obtained from the model trained on the subset which includes the participant for whom a prediction is being made) to assess out-of-sample performances in the models trained on the two subsets and stability of class structure across the two subsets. Cohen's kappa statistic and the adjusted Rand index are used to measure the out-of-sample performances and across subset stability. Finally, the validity/stability of the class structure obtained from the full data-set is evaluated by a comparison of the in-sample class assignments based on the final multivariate latent class mixed model on the full data-set to the in-sample class assignments obtained from the multivariate latent class models for the subsets, using again Cohen's kappa and the adjusted Rand index.

For assessing clustering validity, we first apply Bayesian profile regression to the biomarker and latent class assignment data for each subset in turn and obtain the consensus results over six chains as described earlier. Next, the consensus dissimilarity matrices for the subsets are compared to their corresponding block diagonals of the consensus dissimilarity matrix from our Bayesian profile regression on the full data-set using Pearson's correlation. To assess cluster stability, the corresponding PAM consensus representative clustering structures from each subset are compared to the final representative clustering from the full data-set using the adjusted Rand index. Moreover, we make predictions for the held-out subsets that allow us to compare 1) their predicted dissimilarity matrices with the corresponding off-diagonal blocks of the final consensus dissimilarity matrix from the full data-set using Pearson's correlation, and 2) their predicted clustering structures with the PAM consensus representative clustering obtained using a model trained on the held-out subset (clustering predictions are obtained by using the predicted dissimilarity matrices to assign participants in the held-out subset to the PAM consensus representative cluster from the training subset that they are closest to).

External validation was not possible as we do not have access to data from studies on similar populations with the corresponding extensive baseline and longitudinal biomarker and phenotypic information to EPAD.

### 3 RESULTS

#### 3.1 Baseline Characteristics of the European Prevention of Alzheimer's Dementia Longitudinal Cohort Study Population

**Table 1** describes the group of 1,574 participants with two or more visits in the EPAD longitudinal cohort. The mean age of

these participants was 65.4 years with a standard deviation of 7.4 years. Around 56% were female and 63% had their highest educational attainment beyond secondary education—an indication of a highly educated cohort of participants recruited; reflecting the eligibility criterion on minimum years of formal education. The cohort was enriched for participants with a family history of AD (first degree relatives) and APOE $\epsilon$ 4 carriers, without diminished decision-making capacity. For the group, this enrichment corresponded to 65.5 and 37.5% with a known family history of AD and a known carrier for APOE $\epsilon$ 4 respectively. 78% of this group ( $n = 1,226$ ) had a global CDR of 0, while the remaining 22% had a score of 0.5 ( $n = 346$ ); two participants had unknown baseline CDR global. Around 82% of those with a family history of AD had a CDR global score of 0. Whereas 70% of those without a family history of AD had a CDR score of 0. Thus there was a clear association between CDR global score and family history of AD favouring the recruitment of participants with a family history who do not have any baseline cognitive impairment and for those without a family history enriching for early symptomatic ( $p < 0.0001$ ;  $\chi^2$ -test). No evidence for an association between CDR global score and APOE $\epsilon$ 4 carrier status was found ( $p = 0.10$ ), with 80% of non-carriers and 76% of carriers having CDR global equal 0.

**Table 1** also summarises the distributions of the EPAD cognitive and clinical outcomes and CSF and neuroimaging biomarkers at baseline. Ten percent of participants had an MMSE score below 27 and 12.5% had a CDRSB score of 1 or above; suggesting that the majority of participants had high levels of cognitive functioning at baseline. However, varying degrees of disease pathology at baseline were indicated on considering a range of biomarkers. AD positivity was estimated around 20% using the ratio of phosphorylated tau to amyloid-beta 42 in CSF. Convincing evidence for the widening of the choroid fissure to different degrees (average of left and right MTA  $\geq 1$ ) was found in about a quarter of the participants, whilst varying percentages of white matter lesion cerebrovascular pathology were seen ranging from 6 to 68% based on age-related regional white matter changes or based on an overall impression of the brain using the Fazekas scales (approximately 16 and 39%). Nearly a quarter of the participants (23.5%) had indications of cerebrovascular pathology in three or more of the five age-related white matter regions. The mean adjusted total hippocampal and ventricles volumes at baseline (with standard deviation) were 5,793mm<sup>3</sup> (703mm<sup>3</sup>) and 32,991mm<sup>3</sup> (17,669mm<sup>3</sup>).

#### 3.2 Disease Progression and Latent Phenotypes—Results From MLCMM

Our MLCMM was able to identify four distinct mean trajectories. **Figure 2** shows these four mean trajectory profiles on the latent process scale and on the original scales for CDRSB and MMSE. Latent clinical phenotype classes 0 to 3 had, respectively, 1,050 (68.0%), 97 (6.3%), 106 (6.9%), and 290 (18.8%) individuals hard assigned to them based on a posterior classification of participants' class membership through the selection of the participant's class with the highest posterior class-membership probability. Latent phenotype class 0, which had the majority of participants, is

**TABLE 1** | Baseline characteristics of the 1,574 participants with more than one visit.

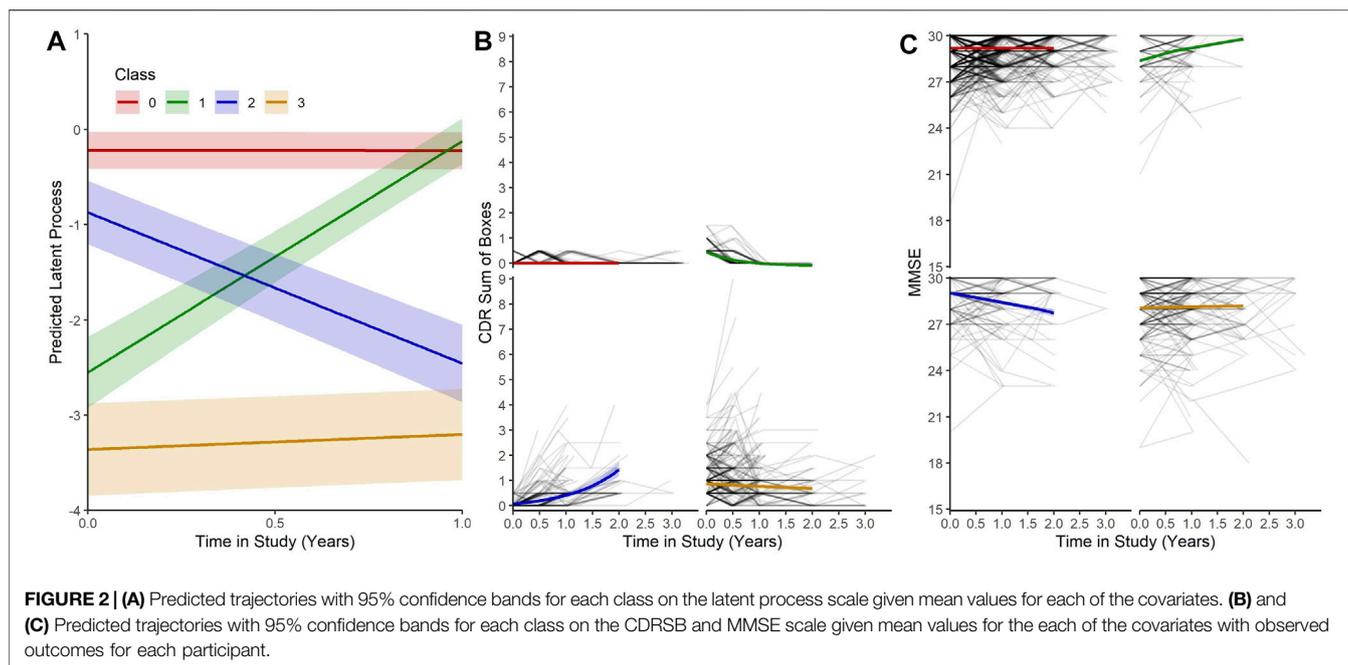
Variable		Mean (SD)	Frequency (%)	No. Unknown	
Risk Factors	Age, years	65.4 (7.4)		0	
	Sex	Female		888 (56.4)	0
		Male		686 (43.6)	
	Education	Level 1		587 (37.3)	0
		Level 2		393 (25.0)	
		Level 3		594 (37.7)	
	Family history of AD	No		543 (34.5)	0
		Yes		1,031 (65.5)	
APOE $\epsilon$ 4 carrier	No		965 (62.5)	31	
	Yes		578 (37.5)		
Outcomes	CDRSB	0		1,162 (73.9)	2
		0.5		214 (13.6)	
		$\geq 1$		196 (12.5)	
	MMSE	29–30		999 (63.5)	1
		27–28		417 (26.5)	
		$\leq 26$		157 (10.0)	
	Transformed CDRSB, tCDRSB	4.60 (1.04)			2
	Normalised MMSE, nMMSE	83.6 (14.6)			1
Biomarkers	pTau/A $\beta$	$\leq 0.024$		1,240 (80.5)	33
		$> 0.024$		301 (19.5)	
	MTA average	0		800 (51.2)	13
		0.5		375 (24.0)	
		$\geq 1$		386 (24.7)	
	Fazekas scale deep	$< 2$		1,317 (84.4)	13
		$\geq 2$		244 (15.6)	
	Fazekas scale periventricular	$< 1$		947 (60.7)	13
		$\geq 1$		614 (39.3)	
	ARWMC basal ganglia	$< 1$		1,379 (88.3)	13
		$\geq 1$		182 (11.7)	
	ARWMC frontal	$< 1$		506 (32.4)	13
		$\geq 1$		1,055 (67.6)	
	ARWMC infratentorial	$< 1$		1,465 (93.9)	13
		$\geq 1$		96 (6.1)	
	ARWMC parieto-occipital	$< 1$		786 (50.4)	13
		$\geq 1$		775 (49.6)	
	ARWMC temporal	$< 1$		1,268 (81.2)	13
		$\geq 1$		293 (18.8)	
	ARWMC combined	$< 3$		1,194 (76.5)	13
$\geq 3$			367 (23.5)		
Total hippocampal volume (adj), $mm^3$	5,793 (703)			62	
Total ventricular volume (adj), $mm^3$	32,991 (17,669)			168	

characterised by individuals having the highest levels of cognitive functioning with no signs of impairment at baseline and no decline throughout the course of the study. Class 1 contained individuals who showed some signs of cognitive/functional impairment at baseline but appeared to improve over time. Class 2 was characterised by individuals who appeared cognitively and functionally unimpaired at baseline (although cognitive functioning levels were not as high as those in class 0) but then declined on follow-up. Whereas class 3 contained individuals who showed the most evident signs of early cognitive/functional impairment at baseline and continued to show impairment on follow-up.

**Table 2** reports the results from our four-class MLCMM. A higher baseline age and a lower level of education are associated with higher levels of cognitive/functional impairment; consistent

with findings from the neurodegenerative and AD literature. Due to how individuals were recruited into the study (through use of a flexible and dynamic approach to selection), biased effects of family history of AD and APOE $\epsilon$ 4 carrier status were expected and therefore the corresponding estimates of these effects were not interpreted as they were notably affected by the selection mechanism. For example, both were found not to be statistically significantly associated with cognitive/functional impairment and the effect of family history of AD was in the opposite direction to that reported in the literature.

The measurement error variances for tCDRSB and nMMSE are 0.531 and 3.527 respectively indicating that tCDRSB has a stronger relationship to the underlying latent disease process. The estimated class-specific proportionality factors,  $\hat{\omega}_g$  ( $g = 0, 1, 2$ ),



**TABLE 2 |** Results of the 4-class MLCMM on the 1,543 participants.

		Coefficient (SE)	p-value
Class membership model	Intercept class 0	1.30 (0.07)	<0.0001
	Intercept class 1	-1.04 (0.12)	<0.0001
	Intercept class 2	-0.87 (0.13)	<0.0001
Fixed effects model	Intercept class 0	0 (not estimated)	—
	Intercept class 1	-2.33 (0.15)	<0.0001
	Intercept class 2	-0.65 (0.14)	<0.0001
	Intercept class 3	-3.14 (0.22)	<0.0001
	Time in study class 0	-0.0040 (0.014)	0.773
	Time in study class 1	2.43 (0.16)	<0.0001
	Time in study class 2	-1.58 (0.11)	<0.0001
	Time in study class 3	0.16 (0.04)	0.0001
	Age	-0.0033 (0.0014)	0.022
	Sex male	-0.015 (0.019)	0.443
	Education level 2	0.022 (0.024)	0.367
Education level 3	0.043 (0.022)	0.049	
Family history of AD	0.016 (0.021)	0.453	
APOE $\epsilon$ 4	-0.021 (0.019)	0.290	
Link function parameters	tCDRSB $\eta_1$	5.31 (0.07)	<0.0001
	tCDRSB $\eta_2$	0.73 (0.04)	<0.0001
	nMMSE $\eta_1$	87.96 (0.48)	<0.0001
	nMMSE $\eta_2$	3.80 (0.30)	<0.0001

which here correspond to the variances of the class-specific random intercepts are 0.0004 for class 0, 0.397 for class 1, and 0.957 for class 2. (The variance for the random intercept corresponding to class 3 was set to 1 for identifiability.) The log-likelihood of this model was -16,842.95, the BIC 33,869.44, and the (relative) entropy 0.947. By comparison, the equivalent three-class model had a log-likelihood of -17,097.26, a higher BIC of 34,348.70, and a lower entropy of 0.932. Thus our four-class model was preferred. It showed excellent ability to discriminate between latent trajectory classes.

Moreover assessment of validity of our four-class model through class stability under repeated sub-setting gave mean Cohen’s kappa and adjusted Rand index values (with standard deviations) of 0.987 (0.008) and 0.989 (0.006), respectively, across the twenty subset comparisons and 0.993 (0.004) and 0.995 (0.003) for the ten comparisons against the full data-set. These results indicate near perfect agreement with evidence for stability across subsets and validity of the class structure derived based on the full data-set. Across the ten splits, the number of discordant classifications seen when the in-sample latent class membership predictions for subsets are compared to the class memberships predicted by our four-class model on the full data-set ranged from 3 to 13 out of the 1,543 participants (0.19–0.84%). For the twenty subsets across the ten splits, four-class multivariate latent class mixed models were always found to provide a better fit (based on BIC) than the alternative three-class multivariate latent class mixed models, and these four-class models had similar class structure as our four-class model on the full data-set.

We further characterised these four latent phenotype classes by baseline and change variables and (marginally) compared these variables across classes using analysis of variance (ANOVA) tests for the continuous variables and  $\chi^2$  tests for binary and categorical variables. The results are shown in **Table 3**. We observe increasing trends in mean age and mean baseline ventricles volume across the latent classes from 0 to 3 and a decreasing trend in mean baseline hippocampal volume. Class 3 differed from the other three classes in having the highest proportions of males, lowest educational level attainments, those with AD positivity at baseline and with evidence on baseline MTA of widening of choroid fissure in varying degrees from widen to end stage atrophy. There was evidence found for differences amongst the

**TABLE 3** | Characterisation of the baseline and change variables by latent phenotype classes.

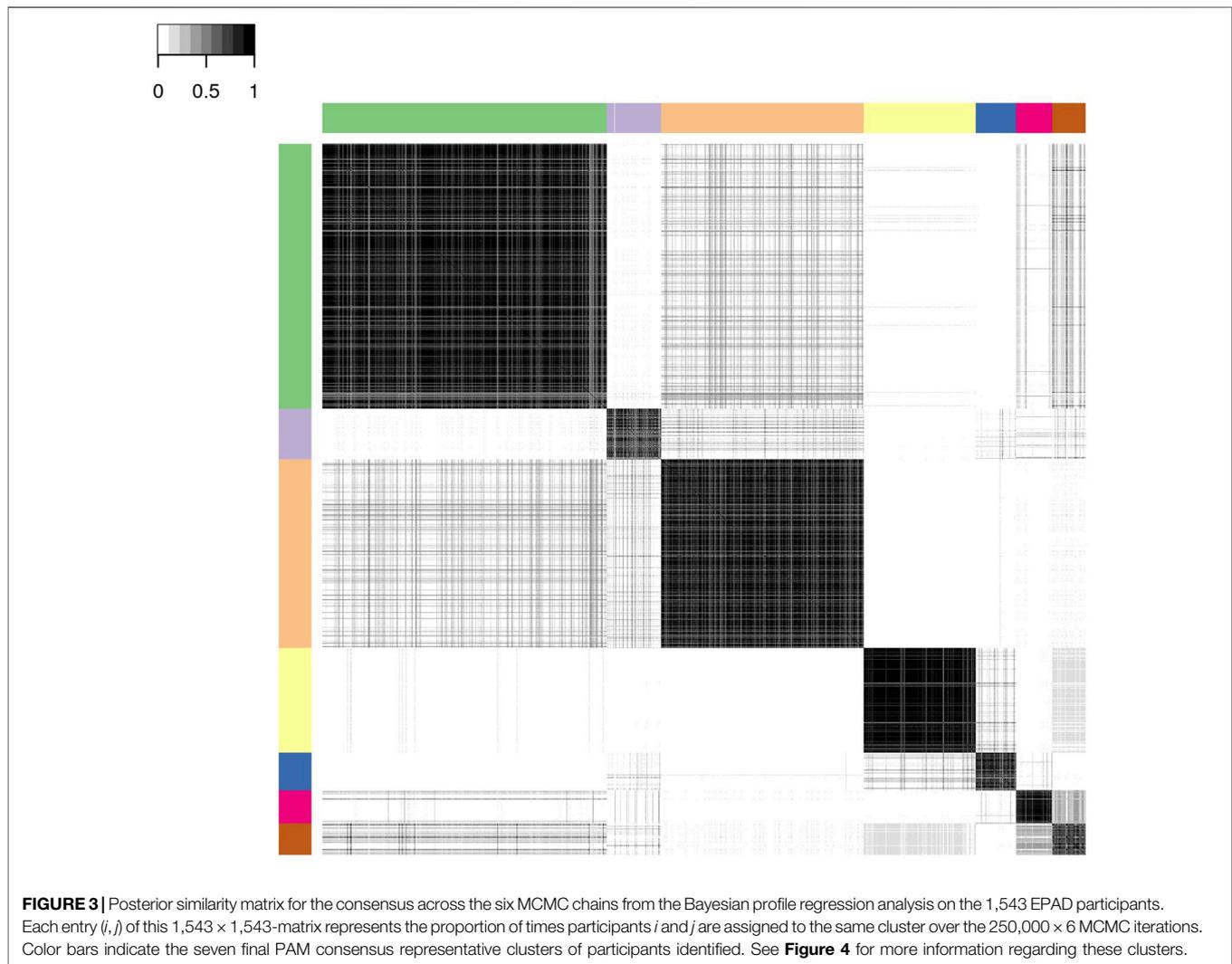
Variable	Mean (SD)				ANOVA <i>p</i> -value	
	Class 0	Class 1	Class 2	Class 3		
Age, years	63.9 (7.0)	65.6 (6.6)	68.1 (7.0)	69.5 (7.0)	<0.0001	
Total hippocampal volume (adj), <i>mm</i> <sup>3</sup>	5,911 (644)	5,814 (725)	5,609 (715)	5,429 (768)	<0.0001	
Total ventricular volume (adj), <i>mm</i> <sup>3</sup>	30,715 (16,348)	35,404 (19,823)	37,962 (18,838)	38,396 (19,405)	<0.0001	
Annual (adj) hippocampal volume change, <i>mm</i> <sup>3</sup> /yr	-9.4 (83.5)	-30.4 (61.9)	-40.2 (85.3)	-55.3 (99.1)	<0.0001	
Annual (adj) ventricular volume change, <i>mm</i> <sup>3</sup> /yr	988 (910)	1,430 (1,354)	1,958 (1,651)	1,688 (1,586)	<0.0001	
Variable		Frequency (%)				$\chi^2$ <i>p</i> -value
		Class 0	Class 1	Class 2	Class 3	
Sex	Female	614 (58.5)	54 (55.7)	65 (61.3)	137 (47.2)	0.005
	Male	436 (41.5)	43 (44.3)	41 (38.7)	153 (52.8)	—
Education	Level 1	368 (35.0)	35 (36.1)	34 (32.1)	138 (47.6)	0.009
	Level 2	267 (25.4)	24 (24.7)	30 (28.3)	63 (21.7)	—
	Level 3	415 (39.5)	38 (39.2)	42 (39.6)	89 (30.7)	—
Family history of AD	No	319 (30.4)	36 (37.1)	40 (37.7)	136 (46.9)	<0.0001
	Yes	731 (69.6)	61 (62.9)	66 (62.3)	154 (53.1)	—
APOE $\epsilon$ 4 carrier	No	656 (62.5)	68 (70.1)	73 (68.9)	168 (57.9)	0.078
	Yes	394 (37.5)	29 (29.9)	33 (31.1)	122 (42.1)	—
pTau/A $\beta$	≤0.024	906 (87.5)	74 (77.9)	71 (71.0)	167 (59.4)	<0.0001
	>0.024	130 (12.5)	21 (22.1)	29 (29.0)	114 (40.6)	—
MTA average	<1	856 (82.1)	63 (67.0)	75 (71.4)	158 (54.7)	<0.0001
	≥1	186 (17.9)	31 (33.0)	30 (28.6)	131 (45.3)	—
Fazekas scale deep (FSD)	<2	893 (85.7)	83 (88.3)	76 (72.4)	236 (81.7)	0.002
	≥2	149 (14.3)	11 (11.7)	29 (27.6)	53 (18.3)	—
Fazekas scale periventricular (FSPV)	<1	660 (63.3)	61 (64.9)	53 (50.5)	155 (53.6)	0.002
	≥1	382 (36.7)	33 (35.1)	52 (49.5)	134 (46.4)	—
ARWMC basal ganglia	<1	929 (89.2)	83 (88.3)	90 (85.7)	246 (85.1)	0.248
	≥1	113 (10.8)	11 (11.7)	15 (14.3)	43 (14.9)	—
ARWMC frontal	<1	346 (33.2)	36 (38.3)	27 (25.7)	86 (29.8)	0.182
	≥1	696 (66.8)	58 (61.7)	78 (74.3)	203 (70.2)	—
ARWMC infratentorial	<1	988 (94.8)	87 (92.6)	96 (91.4)	266 (92.0)	0.195
	≥1	54 (5.2)	7 (7.4)	9 (8.6)	23 (8.0)	—
ARWMC parieto-occipital	<1	549 (52.7)	48 (51.1)	45 (42.9)	127 (43.9)	0.025
	≥1	493 (47.3)	46 (48.9)	60 (57.1)	162 (56.1)	—
ARWMC temporal	<1	863 (82.8)	77 (81.9)	78 (74.3)	223 (77.2)	0.043
	≥1	179 (17.2)	17 (18.1)	27 (25.7)	66 (22.8)	—
ARWMC combined	<3	815 (78.2)	71 (75.5)	72 (68.6)	211 (73.0)	0.061
	≥3	227 (21.8)	23 (24.5)	33 (31.4)	78 (27.0)	—

classes in the presence of white matter hyperintensities in the entire brain as measured by the Fazekas scales, with latent class 2 having the highest proportion of participants with abnormal pathology. No evidence of any further differences between classes 3 and 2 (or between classes 0 and 1) was found with regard to age-related regional white matter changes. However evidence of differences between the lower two classes (0 and 1) compared to the upper two classes (2 and 3) was found for a number of these neuroimaging variables associated with white matter lesions (Table 3).

On examining possible associations of longitudinal changes in volumetric imaging measures and the latent phenotype classes, we observe an increasing annualised hippocampal shrinkage with increasing class, without a similar trend being seen between annualised ventricular enlargement and latent phenotype class. Notably, latent phenotype class 2 had the largest annualised increase in ventricles volume (Table 3).

### 3.3 Neuropathological Endotypes—Results From Profile Regression

For the profile regression analysis, which links CSF and neuroimaging biomarkers to the latent clinical trajectory phenotype, we ran six independent MCMC chains, and obtained a posterior similarity matrix and associated PAM “representative” clustering from the output of each chain (see Section 2.2.2 for details). Agreement between these six chains was high with mean pairwise Pearson’s correlation of 0.95 (standard deviation 0.03) between the dissimilarity matrices and mean pairwise adjusted Rand index of 0.90 (standard deviation 0.05) between the representative clusterings. This, together with inspection of posterior parameters for each chain, suggests that there is no strong evidence against convergence of the MCMC and there is a good level of robustness of the clustering structure. Three of the “representative” clusterings have six clusters, while the other three had seven.

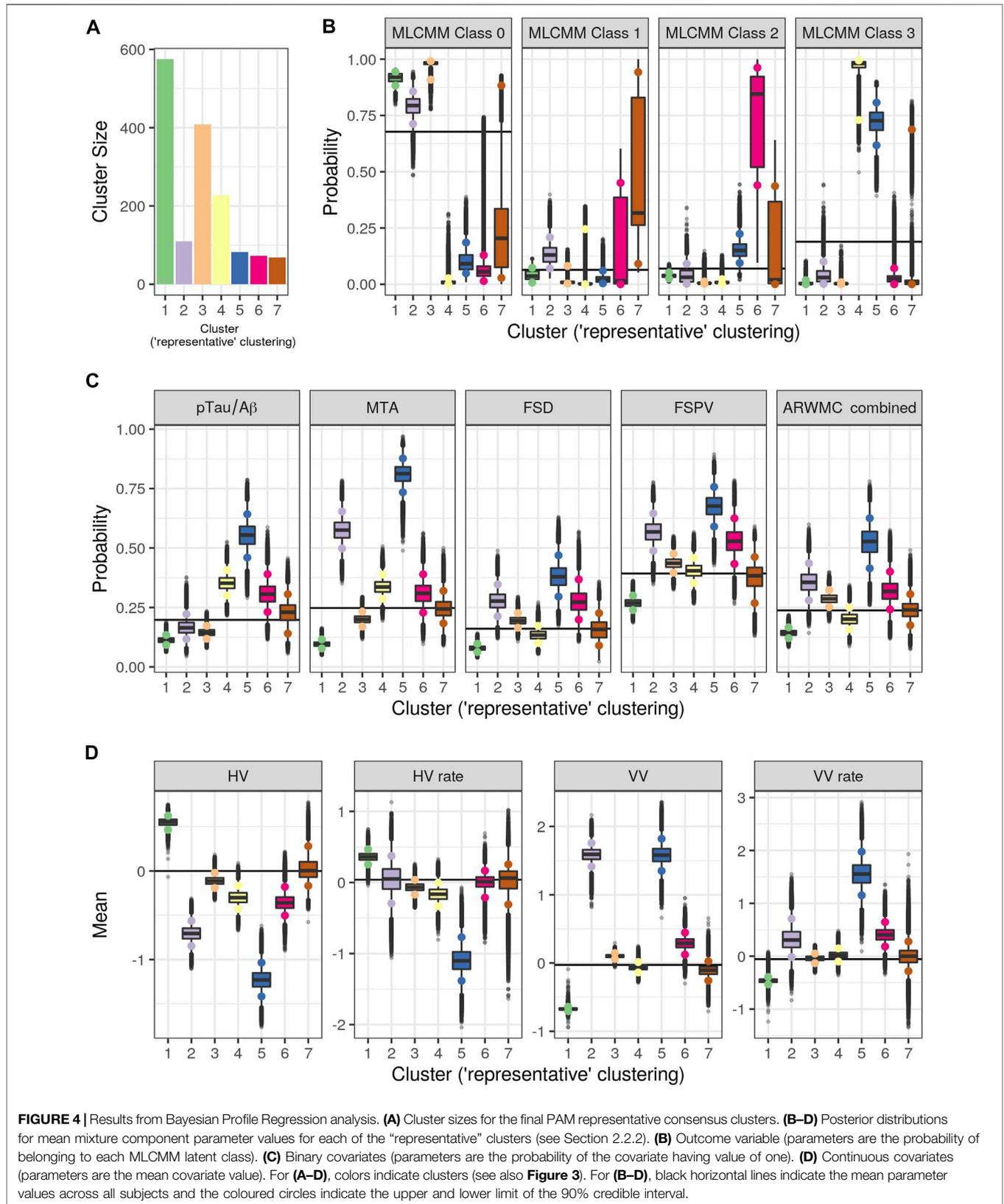


We present below the results of applying consensus clustering to aggregate the output from the six MCMC chains. The mean Pearson's correlation between the consensus dissimilarity matrix and the six independent chains' dissimilarity matrices was 0.98 (standard deviation 0.01). Similarly, the mean adjusted Rand index between the consensus representative clustering and the six representative clusterings from each chain was 0.93 (standard deviation 0.03). The final consensus representative clustering has seven clusters.

**Figure 3** shows the consensus posterior similarity matrix that summarises the output from across the six MCMC chains and the seven representative clusters that were identified from this matrix. **Figure 4** and **Table 4** describe these seven clusters and their distinct biomarker profiles (i.e., neuropathological endotypes). Cluster 1, which is the largest cluster (comprising of 575 out of 1,543 participants), estimated the posterior mean probability of belonging to latent phenotype class 0 to be 92% (in agreement with the empirical estimate of 94%). It was characterised by participants with lower than expected/average probabilities of

having abnormal pathology on the various biomarkers and above average healthy indicators of baseline and longitudinal volumetric measures for hippocampus and ventricles. We label this cluster as a “healthy brain” neuropathological endotype. It had on average the youngest participants, with a mean age (SD) of 61.4 (6.2) years.

Cluster 2, which is a mixture of participants from both latent phenotype classes 0 and 1 (85 and 15% respectively), had somewhat lower than average AD positivity risk (but within the margin of uncertainty of the overall mean) and had stable hippocampal volume over time, but otherwise had higher than expected risk of abnormal pathology on the other biomarkers, including medial temporal lobe atrophy (MTA) indicating hippocampal involvement, and 1.59 standard deviations (SDs) higher baseline ventricles volume and 0.32 SD faster annual rate of increase in ventricles volume above their average, which is being tolerated so far. This cluster appears to be a non-AD driven cluster with “kindling” cerebrovascular disease. We label it as an “at-risk-of-vascular dementia” neuropathological endotype. The



**TABLE 4** | Results from the Bayesian profile regression analysis on the 1,543 participants.

Clusters	N (%)	Posterior means												
		Probability of abnormal pathology					SD distance from overall mean				Class membership probability			
		pTau/ A $\beta$	MTA	FSD	FSPV	ARWMC combined	Mean HV	Mean HV rate	Mean VV	Mean VV rate	Class 0	Class 1	Class 2	Class 3
1	575 (37.3)	0.113	0.096	0.079	0.269	0.143	0.549	0.362	-0.674	-0.463	0.917	0.040	0.038	0.005
2	110 (7.1)	0.166	0.575	0.278	0.567	0.357	-0.706	0.047	1.590	0.321	0.791	0.133	0.038	0.039
3	409 (26.5)	0.145	0.200	0.195	0.436	0.287	-0.111	-0.066	0.101	-0.042	0.976	0.015	0.006	0.003
4	227 (14.7)	0.353	0.337	0.135	0.405	0.202	-0.300	-0.166	-0.063	0.020	0.010	0.040	0.009	0.941
5	82 (5.3)	0.553	0.810	0.380	0.675	0.524	-1.229	-1.093	1.583	1.558	0.101	0.024	0.154	0.721
6	72 (4.7)	0.308	0.309	0.276	0.529	0.319	-0.354	0.000	0.287	0.407	0.064	0.177	0.731	0.028
7	68 (4.4)	0.228	0.247	0.157	0.376	0.240	0.025	0.026	-0.109	-0.004	0.248	0.464	0.173	0.115
Overall empirical mean		0.194	0.247	0.158	0.393	0.236	5,793	-23.8	32,997	1,274	0.680	0.063	0.069	0.188
SD							705	88.7	17,687	1,264				

mean age (SD) of participants was 69.1 (7.0) years and percentage APOE $\epsilon$ 4 positive was 28%, the lowest amongst the clusters.

Cluster 3, which is the second largest in size (409 participants; all from latent phenotype class 0), is characterised by lower than expected risk of both AD positivity and MTA abnormal pathology and no clinically meaningful pathological indications on volumetric neuroimaging; but with evidence of white matter lesion pathology. We describe this cluster as a “healthy ageing” endotype especially as participants are on average older than those in cluster 1, with a mean age (SD) of 66.1 (6.5) years, and they appear to be able to compensate for some cerebrovascular disease. Moreover, none of the differences from overall average for any of the biomarkers were particularly large.

Cluster 4 has 15% of the participants—with average age (SD) of 68.1 (6.8) years—all of whom belong to latent phenotype class 3 (questionably cognitively impaired class). It is characterised by clinically meaningful increased risk of AD positivity and MTA abnormal pathology, early pathological indications on hippocampal volume markers and slightly increased proportion of APOE $\epsilon$ 4 carriers relative to the overall average (0.4 versus 0.375) and may represent a subgroup of “AD high risk” participants.

Cluster 5 represents the 5% of the cohort who have the highest risk, worst baseline levels and fastest rate of worsening on markers. It comprises of a mixture of participants from latent phenotype classes 0 (6%), 2 (17%) and 3 (77%). We consider this to be an “AD-related cluster”. Moreover, it has the highest mean age of 74.2 years (SD 5.6 years) amongst the seven clusters and, notably, the highest proportion of APOE $\epsilon$ 4 carriers (0.46) despite the EPAD selection mechanism.

Finally clusters 6 and 7, which are the most uncertain ones (i.e., empirical class membership proportions of 100% in class 2 for cluster 6 and class 1 for cluster 7 do not match with the

corresponding mean posterior probabilities for these classes of 73 and 46% respectively), correspond to clusters where there are, respectively, evidence of increased abnormal pathology on all markers (except hippocampal atrophy and MTA) and no particular overall evidence of increased abnormal pathology beyond expected on any particular biomarker. Cluster 6 may be another AD-related cluster, but one, possibly, in an earlier stage of progression (cf cluster 5) as they are on average 5.6 years younger, with a mean age (SD) of 68.6 (6.3) years. Cluster 7 appears to have individuals with both unclear biomarker profiles and unclear cognitive trajectories, and therefore we describe it as an “ambiguous” cluster. The mean age (SD) here is 66.0 (6.5) years.

We assessed clustering validity through stability under repeated sub-setting (10 splits, totalling 20 subsets of the data). Out of the twenty consensus representative clustering structures obtained from applying Bayesian profile regression to the twenty subsets, 8 and 10 of these clustering structures consisted of four and five clusters respectively, while the other two comprised three and six clusters. Agreement between the consensus clusterings and the clusterings from the corresponding six independent MCMC chains across the twenty subsets were again high with mean adjusted Rand index of 0.93 (standard deviation of 0.10). The reduced number of clusters relative to the seven clusters found using the full data-set is likely due to the 50% reduction in sample size for the subsets. A comparison of the consensus representative clustering obtained using the subsets of data with the consensus representative clustering obtained using the full data-set (restricted to those individuals in each subset for the comparisons) resulted in a mean adjusted Rand index of 0.69 (standard deviation of 0.09). Furthermore, comparing the 20 consensus posterior dissimilarity matrices obtained from the subsets against those obtained using the corresponding

submatrices of the full data-set resulted in a mean Pearson's correlation of 0.860 (standard deviation of 0.049). These comparisons indicate good agreement between the results obtained on the subsets and those obtained using the full data-set, giving evidence for stability of our results.

Additionally, the held-out prediction analyses (training on one subset and predicting for the other in each split) resulted in a mean Pearson's correlation of 0.674 (standard deviation of 0.045) between the predicted posterior dissimilarity matrices and the corresponding submatrices obtained using the full data-set. Comparing each of the estimated consensus representative clustering obtained from one subset of the split with the predicted clustering for this subset (predictions obtained using a model trained on the other subset of the data split only) resulted in a mean adjusted Rand index of 0.464 (standard deviation of 0.076) over the 20 comparisons. This performance on challenging held-out prediction tasks gives further support for the validity and stability of our clustering results.

The consensus representative clustering structure obtained using Bayesian profile regression without the MLCMM class outcome (i.e., only using biomarker covariates) had an adjusted Rand index of 0.48 with the clustering that did include the outcome, indicating that the outcome is playing an influential role in the clustering analysis and is facilitating interpretation of the clusters in terms of linking them to latent clinical phenotypes.

## 4 DISCUSSION

In this paper, we demonstrate the usefulness of our two-stage approach in, firstly, characterising the evolution of correlated cognitive and clinical outcomes for LCS participants via an underlying latent process in which its trajectory depends on one of four latent clinical phenotypes, and then in providing biological insight through the identification of subgroups based on distinct biomarker profiles (i.e., neuropathological endotypes) linked to the latent phenotypes. Our approach recognises that the longitudinal cognitive and clinical outcomes are the downstream clinical manifestations/consequences of earlier endogenous biological changes occurring within the brain whether they be due to normal brain ageing or pathological due to a specific underlying disease process. It however does not attempt to assess the exact ordering of the pathological cascade of events.

Our intention here was not to provide a comprehensive clinical and biological investigation of the EPAD LCS data but to demonstrate the utility of our two-stage strategy in uncovering meaningful clinical and biological structure within this heterogeneous population. Therefore we chose to use a reduced set of coarser, but still relevant, ATN (amyloid-beta deposition (A), pathologic tau (T), and neurodegeneration (N)) and cerebrovascular biomarkers to demonstrate our two-stage approach. If interest lies in a more thorough investigation, then our approach can be extended to incorporate a larger set of biomarkers, providing more granular information (e.g., both left and right MTAs and hippocampal volumes and all five ARWMC regions could be considered instead of the average, total or majority as was done in this paper; with additional markers such as the Koedam score, which measures parietal atrophy,

included), and additional correlated cognitive or clinical outcomes (e.g., specific cognitive domains). However, with more biomarkers being considered, this could result in increased uncertainty and instability in clustering structure obtained through use of Bayesian profile regression. Therefore we would recommend the incorporation of a variable selection component into the Bayesian profile regression analysis in order to identify the actual drivers of the clustering structure. Related issues may arise regarding both the number and relevance of latent classes arrived at when additional outcomes are added to the multivariate latent class mixed effects analysis, especially when weakly informative or conflicting outcomes are included.

The latent process arising from the multivariate latent class mixed modelling (MLCMM) approach appeared to be more highly correlated with the observed transformed CDR sum of boxes score than to the normalised Mini-Mental State Examination score, possibly reflecting the former being more sensitive to underlying changes than the latter early on. Nevertheless both CDRSB and MMSE produced concurring patterns with each other across the four latent phenotype classes (see **Figure 2**). These four trajectories correspond to a normal cognitive functioning class throughout, a reversion class, a declining class and a (questionable) cognitively impaired class. They are consistent with what has been reported previously in the literature, although the reversion class probably reflects measurement error. Interestingly, with our Bayesian profile regression analysis, we were able to find endotypes covering the full spectrum from "healthy brain" to "AD-related" within the EPAD cohort; reflecting one of the aims of the EPAD LCS to provide a well-phenotyped population covering the full continuum of risk of subsequent AD dementia development.

We note that the diminishing numbers at each visit reflect both the staggered opening of the 31 recruitment centres across Europe and the LCS concluding at the end of the IMI funding period. Attempts to further fund the cohort as a whole across Europe were not successful, in large part due to the COVID-19 pandemic. Attempts are ongoing to follow-up these participants in a series of studies across Europe to provide longer term clinical and biological outcomes.

The second objective of the EPAD LCS was to create a trial-ready cohort for potential recruitment into the EPAD PoC Trial. Unfortunately, this trial was not realised. However, our approach can still be used to demonstrate trial-readiness with respect to both minimising screen failures and identifying participants with particular biomarker profiles eligible for recruitment. For example, participants identified/pre-screened as belonging to the "healthy brain" or "healthy ageing" clusters would not be considered for inclusion into trials thereby reducing screen-failure rates currently seen in AD-related trials due to the low prevalence of AD pathology in individuals without dementia, especially among cognitively unimpaired. Whereas, for example, individuals in clusters 4, 5 or 6 may be specifically targeted for phase II trials in which volumetric neuroimaging biomarkers are used as "surrogate" endpoints. While secondary prevention trials in pre-clinical populations with no baseline cognitive impairment may be more inclined to focus recruitment on participants from cluster 6 (or class 2) when the primary endpoint is a cognitive one.

The novelty of our approach is not only in characterising the longitudinal cognitive and clinical outcomes into latent

phenotype trajectories and in identifying neuropathological endotypes, but going beyond identifying substructures to also being able to do future longitudinal clinical prediction in individuals. Briefly, we would combine the posterior predictive probabilities of class membership obtained from both the Bayesian profile regression and the MLCMM, based on the observed relevant biomarker, cognitive and risk factor data, to update the individual's mixture component probabilities in the MLCMM. We would then use these as weights to average over the linear mixed effects submodels corresponding to the four classes in order to predict future transformed CDRSB and normalised MMSE. Currently, the uncertainty attached to the latent trajectory classes is not taken account of in the Bayesian profile regression analysis in our two-stage approach, although this can be rectified by using Markov melding (Goudie et al., 2019). However, we expect this to have little impact on our findings.

In conclusion, we have introduced a two-stage approach for the modelling of longitudinal cognitive and clinical outcomes, biomarkers (baseline and longitudinal) and risk factors to analyse the data from the EPAD Longitudinal Cohort Study and shown its clinical and biological utility in the areas of trajectory stratification, subgroup identification and prediction. In the long term we envisage this approach to be applicable more widely to precision medicine and secondary prevention in Alzheimer's dementia research and practice.

## DATA AVAILABILITY STATEMENT

A publicly available dataset was analyzed in this study. This dataset can be found at: <http://ep-ad.org/erap/>, doi: 10.34688/epadlcs\_v.imi\_20.10.30.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Independent Ethics Committee or other relevant

ethical review board for written approval as required by local laws and regulations. A copy of approval is required by the University of Edinburgh as Sponsor before the study commences at each site. The study is designed and conducted in accordance with the guidelines for Good Clinical Practice (GCP), and with the ethical principles as proclaimed in the Declaration of Helsinki. All participants are required to provide written informed consent prior to participation in any research activities laid out in the EPAD LCS protocol. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

JH, SH, and BT developed the disease progression and probabilistic clustering models, analyzed the data, and drafted the manuscript. CR and BT conceptualised the approach for Alzheimer's disease. CR is the chief investigator of the EPAD Longitudinal Cohort Study and provided clinical input, reviewed and gave critical feedback on the manuscript. The authors read and approved the final manuscript.

## FUNDING

This work was funded through the EU/EFPIA Innovative Medicines Initiative Joint Undertaking EPAD grant agreement 115736. BT is supported through the United Kingdom Medical Research Council programme grant No. (MC\_UU\_00002/2) and supported by the NIHR the Cambridge Biomedical Research Centre.

## ACKNOWLEDGMENTS

We thank all the EPAD LCS participants for their enthusiastic involvement in this study.

## REFERENCES

- Anderson, R. M., Hadjichrysanthou, C., Evans, S., and Wong, M. M. (2017). Why Do So many Clinical Trials of Therapies for Alzheimer's Disease Fail?. *The Lancet* 390, 2327–2329. doi:10.1016/s0140-6736(17)32399-1
- Bachman, A. H., and Ardekani, B. A.; for the Alzheimer's Disease Neuroimaging Initiative (2020). Change point Analyses in Prodromal Alzheimer's Disease. *Biomarkers in Neuropsychiatry* 3, 100028. doi:10.1016/j.bionps.2020.100028
- Bateman, R. J., Xiong, C., Benzinger, T. L. S., Fagan, A. M., Goate, A., Fox, N. C., et al. (2012). Clinical and Biomarker Changes in Dominantly Inherited Alzheimer's Disease. *N. Engl. J. Med.* 367, 795–804. doi:10.1056/nejmoa1202753
- Bhagwat, N., Viviano, J. D., Voineskos, A. N., and Chakravarty, M. M.; Alzheimer's Disease Neuroimaging Initiative (2018). Modeling and Prediction of Clinical Symptom Trajectories in Alzheimer's Disease Using Longitudinal Data. *Plos Comput. Biol.* 14, e1006376. doi:10.1371/journal.pcbi.1006376
- Blanche, P., Proust-Lima, C., Loubère, L., Berr, C., Dartigues, J.-F., and Jacqmin-Gadda, H. (2015). Quantifying and Comparing Dynamic Predictive Accuracy of Joint Models for Longitudinal Marker and Time-To-Event in Presence of Censoring and Competing Risks. *Biom* 71, 102–113. doi:10.1111/biom.12232
- Braak, H., and Braak, E. (1997). Frequency of Stages of Alzheimer-Related Lesions in Different Age Categories. *Neurobiol. Aging* 18, 351–357. doi:10.1016/s0197-4580(97)00056-0
- Braak, H., and Braak, E. (1991). Neuropathological Staging of Alzheimer-Related Changes. *Acta Neuropathol.* 82, 239–259. doi:10.1007/bf00308809
- Braak, H., and Del Tredici, K. (2012). Alzheimer's Disease: Pathogenesis and Prevention. *Alzheimer's Dement.* 8, 227–233. doi:10.1016/j.jalz.2012.01.011
- Brand, L., Nichols, K., Wang, H., Shen, L., and Huang, H. (2020). Joint Multi-Modal Longitudinal Regression and Classification for Alzheimer's Disease Prediction. *IEEE Trans. Med. Imaging* 39, 1845–1855. doi:10.1109/tmi.2019.2958943
- Chen, G., Shu, H., Chen, G., Ward, B. D., Antuono, P. G., Zhang, Z., et al. (2016). Staging Alzheimer's Disease Risk by Sequencing Brain Function and Structure, Cerebrospinal Fluid, and Cognition Biomarkers. *Jad* 54, 983–993. doi:10.3233/jad-160537
- Cheng, B., Liu, M., Liu, M., Shen, D., Li, Z., and Zhang, D. (2017). Multi-Domain Transfer Learning for Early Diagnosis of Alzheimer's Disease. *Neuroinform* 15, 115–132. doi:10.1007/s12021-016-9318-5

- de Jong, J., Emon, M. A., Wu, P., Karki, R., Sood, M., Godard, P., et al. (2019). Deep Learning for Clustering of Multivariate Clinical Patient Trajectories with Missing Values. *GigaScience* 8, giz134. doi:10.1093/gigascience/giz134
- Dong, A., Toledo, J. B., Honnorat, N., Doshi, J., Varol, E., Sotiras, A., et al. (2017). Heterogeneity of Neuroanatomical Patterns in Prodromal Alzheimer's Disease: Links to Cognition, Progression and Biomarkers. *Brain* 140, 735–747. doi:10.1093/brain/aww319
- Dong, A., Honnorat, N., Gaonkar, B., and Davatzikos, C. (2016). CHIMERA: Clustering of Heterogeneous Disease Effects via Distribution Matching of Imaging Patterns. *IEEE Trans. Med. Imaging* 35, 612–621. doi:10.1109/tmi.2015.2487423
- Donohue, M. C., Jacqmin-Gadda, H., Le Goff, M., Thomas, R. G., Raman, R., Gamst, A. C., et al. (2014). Estimating Long-Term Multivariate Progression from Short-Term Data. *Alzheimer's Dement.* 10, S400–S410. doi:10.1016/j.jalz.2013.10.003
- European Commission/EACEA/Eurydice (2018). *The Structure of the European Education Systems 2018/19: Schematic Diagrams, Eurydice Facts and Figures*. Luxembourg: Publications Office of the European Union Luxembourg.
- Fiot, J.-B., Raguet, H., Risser, L., Cohen, L. D., Fripp, J., Vialard, F.-X., et al. (2014). Longitudinal Deformation Models, Spatial Regularizations and Learning Strategies to Quantify Alzheimer's Disease Progression. *NeuroImage: Clin.* 4, 718–729. doi:10.1016/j.nicl.2014.02.002
- Fontejn, H. M., Modat, M., Clarkson, M. J., Barnes, J., Lehmann, M., Hobbs, N. Z., et al. (2012). An Event-Based Model for Disease Progression and its Application in Familial Alzheimer's Disease and Huntington's Disease. *NeuroImage* 60, 1880–1889. doi:10.1016/j.neuroimage.2012.01.062
- Gauthier, S., Albert, M., Fox, N., Goedert, M., Kivipelto, M., Mestre-Ferrandiz, J., et al. (2016). Why Has Therapy Development for Dementia Failed in the Last Two Decades?. *Alzheimer's Dement.* 12, 60–64. doi:10.1016/j.jalz.2015.12.003
- Geifman, N., Kennedy, R. E., Schneider, L. S., Buchan, I., and Brinton, R. D. (2018). Data-driven Identification of Endophenotypes of Alzheimer's Disease Progression: Implications for Clinical Trials and Therapeutic Interventions. *Alzheimers Res. Ther.* 10, 1–7. doi:10.1186/s13195-017-0332-0
- Golriz Khatami, S., Robinson, C., Birkenbihl, C., Domingo-Fernández, D., Hoyt, C. T., and Hofmann-Apitius, M. (2020). Challenges of Integrative Disease Modeling in Alzheimer's Disease. *Front. Mol. Biosci.* 6, 158. doi:10.3389/fmolb.2019.00158
- Goudie, R. J. B., Presanis, A. M., Lunn, D., De Angelis, D., and Wernisch, L. (2019). Joining and Splitting Models with Markov Melding. *Bayesian Anal.* 14, 81–109. doi:10.1214/18-BA1104
- Goyal, D., Tjandra, D., Migrino, R. Q., Giordani, B., Syed, Z., Wiens, J., et al. (2018). Characterizing Heterogeneity in the Progression of Alzheimer's Disease Using Longitudinal Clinical and Neuroimaging Biomarkers. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* 10, 629–637. doi:10.1016/j.dadm.2018.06.007
- Hall, C. B., Lipton, R. B., Sliwinski, M., and Stewart, W. F. (2000). A Change point Model for Estimating the Onset of Cognitive Decline in Preclinical Alzheimer's Disease. *Statist. Med.* 19, 1555–1566. doi:10.1002/(sici)1097-0258(20000615/30)19:11<1555::aid-sim445>3.0.co;2-3
- Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J. Q., Bittner, T., et al. (2018). CSF Biomarkers of Alzheimer's Disease concord with Amyloid- $\beta$  PET and Predict Clinical Progression: A Study of Fully Automated Immunoassays in BioFINDER and ADNI Cohorts. *Alzheimer's Dement.* 14, 1470–1481. doi:10.1016/j.jalz.2018.01.010
- Hardy, J., and Selkoe, D. J. (2002). The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science* 297, 353–356. doi:10.1126/science.1072994
- Hubbard, R. A., and Zhou, X. H. (2011). A Comparison of Non-homogeneous Markov Regression Models with Application to Alzheimer's Disease Progression. *J. Appl. Stat.* 38, 2313–2326. doi:10.1080/02664763.2010.547567
- Iddi, S., Li, D., Aisen, P. S., Rafii, M. S., Thompson, W. K., and Donohue, M. C. (2019). Predicting the Course of Alzheimer's Progression. *Brain Inform.* 6, 1–18. doi:10.1186/s40708-019-0099-0
- Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Tracking Pathophysiological Processes in Alzheimer's Disease: an Updated Hypothetical Model of Dynamic Biomarkers. *Lancet Neurol.* 12, 207–216. doi:10.1016/s1474-4422(12)70291-0
- Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., et al. (2010). Hypothetical Model of Dynamic Biomarkers of the Alzheimer's Pathological cascade. *Lancet Neurol.* 9, 119–128. doi:10.1016/s1474-4422(09)70299-6
- Jedynak, B. M., Lang, A., Liu, B., Katz, E., Zhang, Y., Wyman, B. T., et al. (2012). A Computational Neurodegenerative Disease Progression Score: Method and Results with the Alzheimer's Disease Neuroimaging Initiative Cohort. *NeuroImage* 63, 1478–1486. doi:10.1016/j.neuroimage.2012.07.059
- Kaufman, L., and Rousseeuw, P. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons.
- Khanna, S., Domingo-Fernández, D., Iyappan, A., Emon, M. A., Hofmann-Apitius, M., and Fröhlich, H. (2018). Using Multi-Scale Genetic, Neuroimaging and Clinical Data for Predicting Alzheimer's Disease and Reconstruction of Relevant Biological Mechanisms. *Sci. Rep.* 8, 1–13. doi:10.1038/s41598-018-29433-3
- Kulason, S., Xu, E., Tward, D. J., Bakker, A., Albert, M., Younes, L., et al. (2020). Entorhinal and Transentorhinal Atrophy in Preclinical Alzheimer's Disease. *Front. Neurosci.* 14, 804. doi:10.3389/fnins.2020.00804
- Lai, D., Xu, H., Koller, D., Foroud, T., and Gao, S. (2016). A Multivariate Finite Mixture Latent Trajectory Model with Application to Dementia Studies. *J. Appl. Stat.* 43, 2503–2523. doi:10.1080/02664763.2016.1141181
- Lei, B., Yang, M., Yang, P., Zhou, F., Hou, W., Zou, W., et al. (2020). Deep and Joint Learning of Longitudinal Data for Alzheimer's Disease Prediction. *Pattern Recognition* 102, 107247. doi:10.1016/j.patcog.2020.107247
- Li, D., Iddi, S., Thompson, W. K., Rafii, M. S., Aisen, P. S., Donohue, M. C., et al. (2018). Bayesian Latent Time Joint Mixed-effects Model of Progression in the Alzheimer's Disease Neuroimaging Initiative. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* 10, 657–668. doi:10.1016/j.dadm.2018.07.008
- Li, K., Chan, W., Doody, R. S., Quinn, J., and Luo, S. (2017). Prediction of Conversion to Alzheimer's Disease with Longitudinal Measures and Time-To-Event Data. *Jad* 58, 361–371. doi:10.3233/jad-161201
- Li, K., and Luo, S. (2019). Dynamic Predictions in Bayesian Functional Joint Models for Longitudinal and Time-To-Event Data: An Application to Alzheimer's Disease. *Stat. Methods Med. Res.* 28, 327–342. doi:10.1177/0962280217722177
- Lin, J., Li, K., and Luo, S. (2021). Functional Survival Forests for Multivariate Longitudinal Outcomes: Dynamic Prediction of Alzheimer's Disease Progression. *Stat. Methods Med. Res.* 30, 99–111. doi:10.1177/0962280220941532
- Liu, W., Zhang, B., Zhang, Z., and Zhou, X.-H. (2013). Joint Modeling of Transitional Patterns of Alzheimer's Disease. *PloS ONE* 8, e75487. doi:10.1371/journal.pone.0075487
- Liverani, S., Hastie, D. I., Azizi, L., Papatomas, M., and Richardson, S. (2015). PREMIUM: An R Package for Profile Regression Mixture Models Using Dirichlet Processes. *J. Stat. Softw.* 64, 1–30. doi:10.18637/jss.v064.i07
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., et al. (2017). Dementia Prevention, Intervention, and Care. *The Lancet* 390, 2673–2734. doi:10.1016/s0140-6736(17)31363-6
- Lorenzi, M., Filippone, M., Frisoni, G. B., Alexander, D. C., and Ourselin, S.; Alzheimer's Disease Neuroimaging Initiative (2019). Probabilistic Disease Progression Modeling to Characterize Diagnostic Uncertainty: Application to Staging and Prediction in Alzheimer's Disease. *NeuroImage* 190, 56–68. doi:10.1016/j.neuroimage.2017.08.059
- Marioni, R. E., Proust-Lima, C., Amieva, H., Brayne, C., Matthews, F. E., Dartigues, J.-F., et al. (2014). Cognitive Lifestyle Jointly Predicts Longitudinal Cognitive Decline and Mortality Risk. *Eur. J. Epidemiol.* 29, 211–219. doi:10.1007/s10654-014-9881-8
- Martí-Juan, G., Sanroma, G., and Piella, G.; Alzheimer's Disease Neuroimaging Initiative and Alzheimer's Disease Metabolomics Consortium (2019). Revealing Heterogeneity of Brain Imaging Phenotypes in Alzheimer's Disease Based on Unsupervised Clustering of Blood Marker Profiles. *PloS ONE* 14, e0211121. doi:10.1371/journal.pone.0211121
- Martí-Juan, G., Sanroma-Guell, G., and Piella, G. (2020). A Survey on Machine and Statistical Learning for Longitudinal Analysis of Neuroimaging Data in Alzheimer's Disease. *Comp. Methods Programs Biomed.* 189, 105348. doi:10.1016/j.cmpb.2020.105348
- Molitor, J., Papatomas, M., Jerrett, M., and Richardson, S. (2010). Bayesian Profile Regression with an Application to the National Survey of Children's Health. *Biostatistics* 11, 484–498. doi:10.1093/biostatistics/kxq013

- Oxtoby, N. P., Young, A. L., Cash, D. M., Benzinger, T. L. S., Fagan, A. M., Morris, J. C., et al. (2018). Data-driven Models of Dominantly-Inherited Alzheimer's Disease Progression. *Brain* 141, 1529–1544. doi:10.1093/brain/awy050
- Philipps, V., Amieva, H., Andrieu, S., Dufouil, C., Berr, C., Dartigues, J.-F., et al. (2014). Normalized Mini-Mental State Examination for Assessing Cognitive Change in Population-Based Brain Aging Studies. *Neuroepidemiology* 43, 15–25. doi:10.1159/000365637
- Proust-Lima, C., Amieva, H., and Jacqmin-Gadda, H. (2013). Analysis of Multivariate Mixed Longitudinal Data: a Flexible Latent Process Approach. *Br. J. Math. Stat. Psychol.* 66, 470–487. doi:10.1111/bmsp.12000
- Proust-Lima, C., Philipps, V., and Dartigues, J. F. (2019). A Joint Model for Multiple Dynamic Processes and Clinical Endpoints: Application to Alzheimer's Disease. *Stat. Med.* 38, 4702–4717. doi:10.1002/sim.8328
- Proust-Lima, C., Philipps, V., and Liqueur, B. (2017). Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcmsm. *J. Stat. Softw.* 78, 1–56. doi:10.18637/jss.v078.i02
- Proust-Lima, C., Dartigues, J.-F., and Jacqmin-Gadda, H. (2016). Joint Modeling of Repeated Multivariate Cognitive Measures and Competing Risks of Dementia and Death: a Latent Process and Latent Class Approach. *Statist. Med.* 35, 382–398. doi:10.1002/sim.6731
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Racine, A. M., Kosciak, R. L., Berman, S. E., Nicholas, C. R., Clark, L. R., Okonkwo, O. C., et al. (2016). Biomarker Clusters Are Differentially Associated with Longitudinal Cognitive Decline in Late Midlife. *Brain* 139, 2261–2274. doi:10.1093/brain/aww142
- Raket, L. L. (2020). Statistical Disease Progression Modeling in Alzheimer Disease. *Front. Big Data* 3, 24. doi:10.3389/fdata.2020.00024
- Ritchie, C. W., Muniz-Terrera, G., Kivipelto, M., Solomon, A., Tom, B., and Molinuevo, J. L. (2020). The European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study: Baseline Data Release V500.0. *J. Prev. Alzheimers Dis.* 7, 8–13. doi:10.14283/jpad.2019.46
- Ritchie, C. W., Molinuevo, J. L., Truyen, L., Satlin, A., Van der Geyten, S., and Lovestone, S. (2016). Development of Interventions for the Secondary Prevention of Alzheimer's Dementia: the European Prevention of Alzheimer's Dementia (EPAD) Project. *The Lancet Psychiatry* 3, 179–186. doi:10.1016/s2215-0366(15)00454-x
- Robitaille, A., van den Hout, A., Machado, R. J. M., Bennett, D. A., Čukić, I., Deary, I. J., et al. (2018). Transitions across Cognitive States and Death Among Older Adults in Relation to Education: A Multistate Survival Model Using Data from Six Longitudinal Studies. *Alzheimer's Dement.* 14, 462–472. doi:10.1016/j.jalz.2017.10.003
- Rouanet, A., Joly, P., Dartigues, J. F., Proust-Lima, C., and Jacqmin-Gadda, H. (2016). Joint Latent Class Model for Longitudinal Data and Interval-censored Semi-competing Events: Application to Dementia. *Biom* 72, 1123–1135. doi:10.1111/biom.12530
- Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* 20, 53–65. doi:10.1016/0377-0427(87)90125-7
- Samtani, M., Raghavan, N., Novak, G., Nandy, P., and Narayan, V. A. (2014). Disease Progression Model for Clinical Dementia Rating–Sum of Boxes in Mild Cognitive Impairment and Alzheimer's Subjects from the Alzheimer's Disease Neuroimaging Initiative. *Ndt* 10, 929. doi:10.2147/ndt.s62323
- Schindler, S. E., Gray, J. D., Gordon, B. A., Xiong, C., Batrla-Utermann, R., Quan, M., et al. (2018). Cerebrospinal Fluid Biomarkers Measured by Elecsys Assays Compared to Amyloid Imaging. *Alzheimer's Dement.* 14, 1460–1469. doi:10.1016/j.jalz.2018.01.013
- Schmidt-Richberg, A., Ledig, C., Guerrero, R., Molina-Abril, H., Frangi, A., Rueckert, D., et al. (2016). Learning Biomarker Models for Progression Estimation of Alzheimer's Disease. *PLoS ONE* 11, e0153040. doi:10.1371/journal.pone.0153040
- Segalas, C., Helmer, C., and Jacqmin-Gadda, H. (2020). A Curvilinear Bivariate Random Change-point Model to Assess Temporal Order of Markers. *Stat. Methods Med. Res.* 29, 2481–2492. doi:10.1177/0962280219898719
- Sethuraman, J. (1994). A Constructive Definition of Dirichlet Priors. *Stat. Sinica* 4, 639–650.
- Shi, J., Sabbagh, M. N., and Vellas, B. (2020). Alzheimer's Disease beyond Amyloid: Strategies for Future Therapeutic Interventions. *BMJ* 371, m3684. doi:10.1136/bmj.m3684
- Solomon, A., Kivipelto, M., Molinuevo, J. L., Tom, B., and Ritchie, C. W. (2018). European Prevention of Alzheimer's Dementia Longitudinal Cohort Study (EPAD LCS): Study Protocol. *BMJ Open* 8, e021017. doi:10.1136/bmjopen-2017-021017
- ten Kate, M., Ingala, S., Schwarz, A. J., Fox, N. C., Chételat, G., van Berckel, B. N. M., et al. (2018). Secondary Prevention of Alzheimer's Dementia: Neuroimaging Contributions. *Alzheimers Res. Ther.* 10, 1–21. doi:10.1186/s13195-018-0438-z
- ten Kate, M., Dicks, E., Visser, P. J., van der Flier, W. M., Teunissen, C. E., Barkhof, F., et al. (2018). Atrophy Subtypes in Prodromal Alzheimer's Disease Are Associated with Cognitive Decline. *Brain* 141, 3443–3456. doi:10.1093/brain/awy264
- van den Hout, A. (2016). *Multi-state Survival Models for Interval-Censored Data*. New York: CRC Press.
- Villeneuve, S. C., Houot, M., Cacciamani, F., Verrijp, M., Dubois, B., Sikkes, S., et al. (2019). Latent Class Analysis Identifies Functional Decline with Amsterdam IADL in Preclinical Alzheimer's Disease. *Alzheimer's Dement. Translational Res. Clin. Interventions* 5, 553–562. doi:10.1016/j.trci.2019.08.009
- Vos, S. J., Xiong, C., Visser, P. J., Jasielec, M. S., Hassenstab, J., Grant, E. A., et al. (2013). Preclinical Alzheimer's Disease and its Outcome: a Longitudinal Cohort Study. *Lancet Neurol.* 12, 957–965. doi:10.1016/s1474-4422(13)70194-7
- Wang, G., Xiong, C., McDade, E. M., Hassenstab, J., Aschenbrenner, A. J., Fagan, A. M., et al. (2018). Simultaneously Evaluating the Effect of Baseline Levels and Longitudinal Changes in Disease Biomarkers on Cognition in Dominantly Inherited Alzheimer's Disease. *Alzheimer's Dement. Translational Res. Clin. Interventions* 4, 669–676. doi:10.1016/j.trci.2018.10.009
- Watts, G. (2018). Prospects for Dementia Research. *The Lancet* 391, 416. doi:10.1016/s0140-6736(18)30190-9
- Wei, S., and Kryscio, R. J. (2016). Semi-Markov Models for Interval Censored Transient Cognitive States with Back Transitions and a Competing Risk. *Stat. Methods Med. Res.* 25, 2909–2924. doi:10.1177/0962280214534412
- Williams, O. A., An, Y., Armstrong, N. M., Kitner-Triolo, M., Ferrucci, L., and Resnick, S. M. (2020). Profiles of Cognitive Change in Preclinical and Prodromal Alzheimer's Disease Using Change-Point Analysis. *Jad* 75, 1169–1180. doi:10.3233/jad-191268
- Winblad, B., Amouyel, P., Andrieu, S., Ballard, C., Brayne, C., Brodaty, H., et al. (2016). Defeating Alzheimer's Disease and Other Dementias: a Priority for European Science and Society. *Lancet Neurol.* 15, 455–532. doi:10.1016/s1474-4422(16)00062-4
- Wolz, R., Aljabar, P., Hajnal, J. V., Hammers, A., and Rueckert, D. (2010). LEAP: Learning Embeddings for Atlas Propagation. *NeuroImage* 49, 1316–1325. doi:10.1016/j.neuroimage.2009.09.069
- Wu, Y., Zhang, X., He, Y., Cui, J., Ge, X., Han, H., et al. (2020). Predicting Alzheimer's Disease Based on Survival Data and Longitudinally Measured Performance on Cognitive and Functional Scales. *Psychiatry Res.* 291, 113201. doi:10.1016/j.psychres.2020.113201
- Yiannopoulou, K. G., and Papageorgiou, S. G. (2020). Current and Future Treatments in Alzheimer Disease: An Update. *J. Cent. Nervous Syst. Dis.* 12, 1–12. doi:10.1177/1179573520907397
- Younes, L., Albert, M., Moghekar, A., Soldan, A., Pettigrew, C., and Miller, M. I. (2019). Identifying Change-points in Biomarkers during the Preclinical Phase of Alzheimer's Disease. *Front. Aging Neurosci.* 11, 74. doi:10.3389/fnagi.2019.00074
- Young, A. L., Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., et al. (2018). Uncovering the Heterogeneity and Temporal Complexity of Neurodegenerative Diseases with Subtype and Stage Inference. *Nat. Commun.* 9, 1–16. doi:10.1038/s41467-018-05892-0
- Young, A. L., Oxtoby, N. P., Daga, P., Cash, D. M., Fox, N. C., Ourselin, S., et al. (2014). A Data-Driven Model of Biomarker Changes in Sporadic Alzheimer's Disease. *Brain* 137, 2564–2577. doi:10.1093/brain/awu176
- Zhang, L., Lim, C. Y., Maiti, T., Li, Y., Choi, J., Bozoki, A., et al. (2019). Analysis of Conversion of Alzheimer's Disease Using a Multi-State Markov Model. *Stat. Methods Med. Res.* 28, 2801–2819. doi:10.1177/0962280218786525

- Zhang, X., Mormino, E. C., Sun, N., Sperling, R. A., Sabuncu, M. R., Yeo, B. T. T., et al. (2016). Bayesian Model Reveals Latent Atrophy Factors with Dissociable Cognitive Trajectories in Alzheimer's Disease. *Proc. Natl. Acad. Sci. USA* 113, E6535–E6544. doi:10.1073/pnas.1611073113
- Zhang, X., Yang, Y., Li, T., Zhang, Y., Wang, H., and Fujita, H. (2021). CMC: A Consensus Multi-View Clustering Model for Predicting Alzheimer's Disease Progression. *Comp. Methods Programs Biomed.* 199, 105895. doi:10.1016/j.cmpb.2020.105895

**Author Disclaimer:** The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Howlett, Hill, Ritchie and Tom. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*