Analyzing the effect of APOE on Alzheimer's disease progression using an event-based model for stratified populations

Alzheimer's Disease Neuroimaging Initiative

DOI
10.1016/j.neuroimage.2020.117646

Publication date
2020

Document Version
Final published version

Published in
NeuroImage

NeuroImage 227 (2021) 117646

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Analyzing the effect of *APOE* on Alzheimer's disease progression using an event-based model for stratified populations

Vikram Venkatraghavan[a,*], Stefan Klein[a], Lana Fani[c], Leontine S. Ham[a], Henri Vrooman[a], M. Kamran Ikram[c,d], Wiro J. Niessen[a,b], Esther E. Bron[a], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a] *Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear Medicine, Erasmus MC, University Medical Center Rotterdam, the Netherlands*
[b] *Quantitative Imaging Group, Dept. of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, the Netherlands*
[c] *Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, the Netherlands*
[d] *Department of Neurology, Erasmus MC, University Medical Center Rotterdam, the Netherlands*

ARTICLE INFO

Keywords:
Disease Progression Modeling
Event-Based Model
Alzheimer's Disease
APOE

ABSTRACT

Alzheimer's disease (AD) is the most common form of dementia and is phenotypically heterogeneous. *APOE* is a triallelic gene which correlates with phenotypic heterogeneity in AD. In this work, we determined the effect of *APOE* alleles on the disease progression timeline of AD using a discriminative event-based model (DEBM). Since DEBM is a data-driven model, stratification into smaller disease subgroups would lead to more inaccurate models as compared to fitting the model on the entire dataset. Hence our secondary aim is to propose and evaluate novel approaches in which we split the different steps of DEBM into group-aspecific and group-specific parts, where the entire dataset is used to train the group-aspecific parts and only the data from a specific group is used to train the group-specific parts of the DEBM. We performed simulation experiments to benchmark the accuracy of the proposed approaches and to select the optimal approach. Subsequently, the chosen approach was applied to the baseline data of 417 cognitively normal, 235 mild cognitively impaired who convert to AD within 3 years, and 342 AD patients from the Alzheimers Disease Neuroimaging Initiative (ADNI) dataset to gain new insights into the effect of *APOE* carriership on the disease progression timeline of AD. In the $\epsilon 4$ carrier group, the model predicted with high confidence that CSF Amyloidβ_{42} and the cognitive score of Alzheimer's Disease Assessment Scale (ADAS) are early biomarkers. Hippocampus was the earliest volumetric biomarker to become abnormal, closely followed by the CSF Phosphorylated Tau_{181} (PTAU) biomarker. In the homozygous $\epsilon 3$ carrier group, the model predicted a similar ordering among CSF biomarkers. However, the volume of the fusiform gyrus was identified as one of the earliest volumetric biomarker. While the findings in the $\epsilon 4$ carrier and the homozygous $\epsilon 3$ carrier groups fit the current understanding of progression of AD, the finding in the $\epsilon 2$ carrier group did not. The model predicted, with relatively low confidence, CSF Neurogranin as one of the earliest biomarkers along with cognitive score of Mini-Mental State Examination (MMSE). Amyloid β_{42} was found to become abnormal after PTAU. The presented models could aid understanding of the disease, and in selecting homogeneous group of presymptomatic subjects at-risk of developing symptoms for clinical trials.

1. Introduction

Dementia affects roughly 5% of the world's elderly population of whom 60 – 70% are affected by Alzheimer's Disease (AD), which is the most common form of dementia (Organization, 2017). There are several neurobiological subtypes of AD (Ferreira et al., 2020) and each subtype potentially needs a different strategy to prevent or slow the progression of AD. Understanding the pathophysiological processes in AD is thus crucial for selecting novel preventive or therapeutic targets for clinical trials of disease modifying treatments, identifying target groups for such trials and tracking the disease progression in patients.

While several studies have looked into the pathophysiology of AD (Bloom, 2014; Jack Jr. et al., 2013; Weigand et al., 2019), it is still not completely understood. Although it has been observed that AD

* Corresponding author.
E-mail address: v.venkatraghavan@erasmusmc.nl (V. Venkatraghavan).

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

https://doi.org/10.1016/j.neuroimage.2020.117646
Received 15 September 2020; Received in revised form 12 November 2020; Accepted 10 December 2020
Available online 16 December 2020

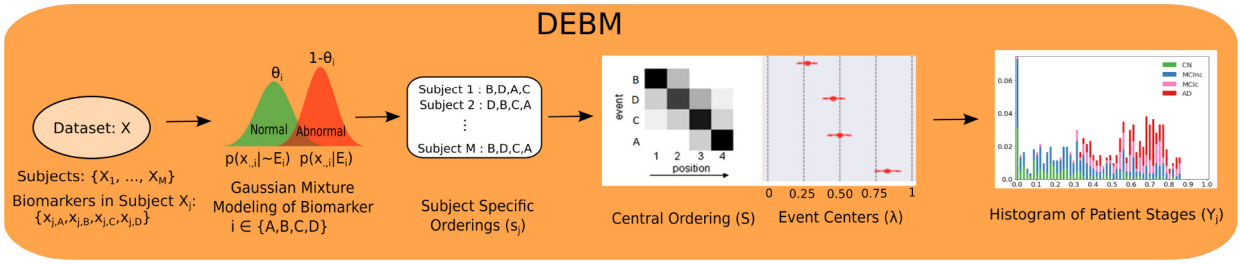


Fig. 1. Overview of the steps involved in DEBM. Input for the DEBM model is a cross-sectional dataset X with M subjects and various biomarkers (A, B, C and D) representing different aspects of neuro-degeneration. Using Gaussian mixture modeling (GMM), mixing parameters (θ_i) and probability density functions of normal ($p(x_{.,i}|\neg E_i)$) and abnormal ($p(x_{.,i}|E_i)$) levels are estimated for each biomarker. This is followed by the estimation of subject-specific orderings (s_j), for each subject in the dataset. Disease progression timeline consisting of central ordering (S) and event-centers (λ) are estimated based on these subject-specific orderings. Based on the constructed disease progression timeline, patient stages (Υ_j) of subjects in an independent test-set can be estimated.

is phenotypically heterogeneous (Au et al., 2015; Murray et al., 2011; Patterson, 2018) with potentially different pathways for disease progression, these pathways remain unclear. There is hence a need to understand the phenotypic heterogeneity in AD while leveraging neuroimaging, fluid and cognitive biomarkers.

APOE is a triallelic gene in which the $\epsilon 2$ allele reduces the risk of AD (van der Lee et al., 2018), the $\epsilon 3$ allele acts as a reference allele and the $\epsilon 4$ allele is a major genetic risk factor of AD (Genin et al., 2011; Kim et al., 2009; Saunders et al., 1993). *APOE* has been shown to correlate with phenotypic heterogeneity in AD (Weintraub et al., 2019). Hence we hypothesize that the pathophysiology of AD can be better understood when considering the effect of *APOE* carriership on biomarker changes.

In the context of data-driven methods for understanding AD pathophysiology, disease progression models have been used to study the trajectories of individual biomarkers (Jedynak et al., 2012; Lorenzi et al., 2019; Schiratti et al., 2015) as well as their progression with respect to each other (Fonteijn et al., 2012; Huang and Alexander, 2012; Venkatraghavan et al., 2017; Young et al., 2014). Unlike typical machine learning approaches, these models are interpretable by design and provide insight for understanding the mechanisms of disease progression. Event-based models (EBMs) are a class of such interpretable disease progression models that estimate the timeline of neuropathologic change during AD progression using cross-sectional data (Fonteijn et al., 2012; Venkatraghavan et al., 2019a).

Our primary aim is to use the discriminative event-based model (DEBM), which was shown to be more accurate than previously proposed EBMs (Venkatraghavan et al., 2019a), to understand the effect of different *APOE* alleles on the disease timeline of AD. To shed light on different aspects of neurodegeneration and identify the earliest brain regions affected, we included commonly studied cerebrospinal fluid (CSF) biomarkers, cognitive scores, and volumetric biomarkers from neuroimaging.

The default approach for estimating the disease progression timeline would be to stratify the population based on their *APOE* $\epsilon 2-4$ carrier status and independently train the DEBM model on the stratified populations (Young et al., 2014). However, since DEBM is a data-driven model, stratification into smaller groups would lead to less accurate models than those obtained by the original method on the entire dataset. Hence our secondary aim is to propose and evaluate a novel approach in which we split the different steps of DEBM into group-aspecific and group-specific parts, where the entire dataset is used to train the group-aspecific parts and only the data from a specific group is used to train the group-specific parts of the DEBM. We present two different variations of this approach and we hypothesize that the optimal split of the DEBM steps into the group-aspecific and group-specific parts would result in better accuracy of the estimated disease progression timeline. Since the ground-truth timelines are unknown in a clinical setting, we evaluate the accuracy of the proposed variations using simulation experiments and we select the optimal method for the analysis on the effect of *APOE* on the AD progression timeline on patient data.

To summarize, our contributions in this paper include proposing and evaluating a novel approach for using DEBM in stratified populations and estimating a comprehensive timeline of AD progression, in terms of biomarker changes, in the presence of different *APOE* alleles.

2. Methods

An introduction to the DEBM model (Venkatraghavan et al., 2019a) is provided in Section 2.1. In Section 2.2 we propose our novel approach for using DEBM in stratified populations with its two variations.

2.1. Discriminative event-based modeling

In a cross-sectional dataset (X) of M subjects, including cognitively normal individuals (CN), subjects with mild cognitive impairment (MCI) and patients with AD, let X_j denote a measurement of biomarkers for subject $j \in [1, M]$, consisting of scalar biomarker values $x_{j,i}$ for $i \in [1, N]$. $x_{.,i}$ denotes the ith biomarker for any unspecified j. DEBM estimates the posterior probabilities of individual biomarkers being abnormal. These posterior probabilities are used to estimate the ordering of biomarker changes for each subject independently. The central ordering and disease progression timeline for the entire dataset are estimated based on these subject-specific orderings. The resulting disease progression timeline is used for assessing the severity of disease in an individual based on his/her biomarker values. Figure 1 shows the different steps involved in DEBM.

Step 1 - Mixture Modeling: As AD is characterized by a cascade of neuropathological changes that occurs over several years, presymptomatic CN subjects can have some abnormal biomarker values. On the other hand, in some clinically diagnosed AD subjects, a proportion of biomarkers may still have normal values, as they might not have an underlying AD pathology or could have atypical AD. Hence clinical labels cannot directly be propagated to individual biomarkers to label normal and abnormal biomarker values. We shall refer to this as biomarker label noise in the rest of the paper. In order to estimate the posterior probabilities of individual biomarkers being abnormal, DEBM, similar to previously proposed EBMs (Fonteijn et al., 2012; Huang and Alexander, 2012; Young et al., 2014), fits a Gaussian mixture model (GMM) to construct the normal / pre-event probability density function (PDF), $p(x_{.,i}|\neg E_i)$, and abnormal / post-event PDF, $p(x_{.,i}|E_i)$. Event E_i in this notation is used to denote the corresponding biomarker becoming abnormal and $\neg E_i$ denotes the corresponnding biomarer being normal. The aforementioned PDFs can be expressed as:

$$p(x_{.,i}|\neg E_i) = \mathcal{N}(\mu_{i,\neg E}; \sigma_{i,\neg E}) \tag{1}$$

$$p(x_{.,i}|E_i) = \mathcal{N}(\mu_{i,E}; \sigma_{i,E}) \tag{2}$$

Where, $\mathcal{N}(\not{f}, \int)$ is the normal distribution with mean μ and standard deviation σ.

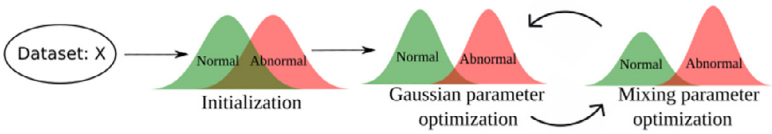


Fig. 2. Overview of GMM optimization in DEBM.

For estimating these parameters robustly in the presence of biomarker label noise, the normal and abnormal PDF estimates are first initialized using the mean and standard deviations after truncating the overlapping tails of the observed distributions in CN and AD subjects. This can be observed in Fig. 2, where the initialization is performed only based on the non-overlapping parts of green and red curves, while the overlapping part is left out to account for biomarker label noise. At this stage of GMM initialization, MCI subjects are left out as well, because it is unsure *a priori* whether their biomarkers are normal or abnormal. The resulting initialized PDFs are denoted as $\widehat{p}(x_{\cdot,i}|\neg E_i))$ and $\widehat{p}(x_{\cdot,i}|E_i)$.

This is followed by an alternating GMM maximum likelihood optimization scheme until both the Gaussian parameters as well as the mixing parameters converge. All the subjects, including MCI, are used for GMM optimization. After convergence, these Gaussians are used to represent the PDFs $p(x_{\cdot,i}|\neg E_i)$ and $p(x_{\cdot,i}|E_i)$. The mixing parameters (θ_i) are used as prior probabilities to convert these PDFs to posterior probabilities $p(\neg E_i|x_{\cdot,i})$ and $p(E_i|x_{\cdot,i})$. Fig. 2 shows an overview of this optimization scheme.

Step 2 - Subject-specific Orderings: $p(E_i|x_{j,i})\forall i$ are used to estimate the subject-specific orderings s_j. s_j is established such that:

$$s_j \ni p\left(E_{s_j(1)}\middle|x_{j,s_j(1)}\right) > \ldots > p\left(E_{s_j(N)}\middle|x_{j,s_j(N)}\right) \tag{3}$$

Step 3 - Central Ordering: DEBM computes the central event ordering S from the subject-specific estimates s_j. To describe the distribution of s_j, a generalized Mallows model is used (Fligner and Verducci, 1988). The central ordering is defined as the ordering that minimizes the sum of distances to all subject-specific orderings s_j, with probabilistic Kendall's Tau being the distance measure (Venkatraghavan et al., 2019a). While S denotes the sequence of biomarker events, the relative position of these events (event-centers) in a normalized scale of $[0, 1]$ is denoted by the vector λ. The pair $\{S, \lambda\}$ together forms a disease progression timeline.

Step 4 - Patient Staging: Once the disease progression timeline is created, subjects in an independent test set (T) can be placed on this timeline to estimate disease severity. This is achieved by converting the biomarker values of the test subjects to posterior probabilities $p(E_i|x_{j,i})$, $\forall j \in T$. These can be used to estimate disease severities in test subjects by first estimating the conditional distribution $p(i|S, X_j)$, which estimates the probability that the first i events of S have occurred for a test-subject and the rest are yet to occur.

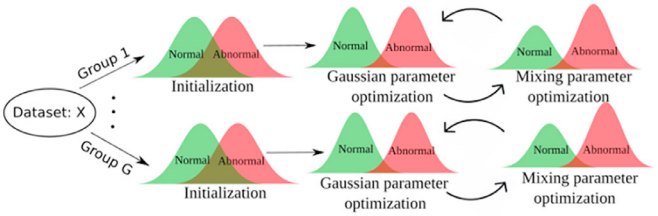
$$p(i|S, X_j) \propto \prod_{l=1}^{i} p\left(E_{S(l)}\middle|x_{j,S(l)}\right) \times \prod_{l=i+1}^{N} p\left(\neg E_{S(l)}\middle|x_{j,S(l)}\right) \tag{4}$$

The patient stage of a test subject (Υ_j) is defined as the expectation of $\lambda(i)$ with respect to the conditional distribution $p(i|S, X_j)$.

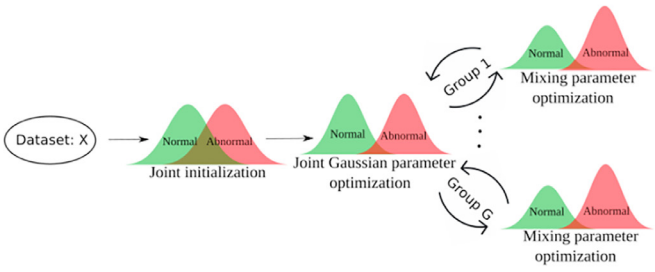
$$\Upsilon_j = \frac{\sum_{i=1}^{N} \lambda(i) p(i|S, X_j)}{\sum_{i=1}^{N} p(i|S, X_j)} \tag{5}$$

2.2. Group-specific and group-aspecific parts of DEBM

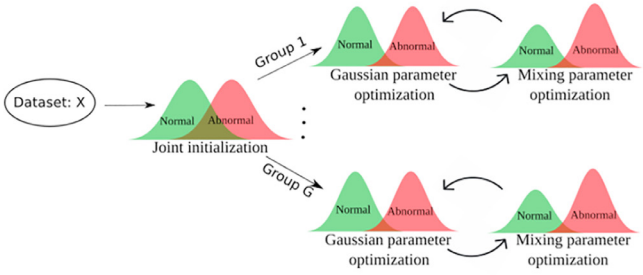
We propose extensions of DEBM for stratified populations, i.e., when the dataset X can be subdivided in groups $g \in [1, G]$, based on, e.g., genotype or phenotype of the subjects. Since DEBM is a data-driven model, data stratification into smaller groups would lead to more inaccurate models (Venkatraghavan et al., 2019a). To obtain better DEBM accuracies in such scenario, we propose to co-train DEBM for estimating disease timelines $\forall g$ by splitting DEBM into group-aspecific and group-specific parts. The group-aspecific parts of DEBM are estimated using the entire dataset and group-specific parts are estimated for each group independently.



(a) GMM in independent DEBM



(b) GMM in Coupled DEBM



(c) GMM in Co-init DEBM

Fig. 3. Overview of GMM optimization strategies in the different approaches for DEBM analysis in stratified populations. (a) The default approach in which GMM in each group is trained independently. (b) GMM in coupled DEBM, where the different groups share the Gaussian parameters, but the mixing parameters are estimated independently. (c) GMM in co-init DEBM in which the different groups are jointly initialized before the GMM optimization, but the optimization is done independently for each group.

We first discuss the default way of independently training DEBM in the different groups and then propose two different approaches for splitting DEBM into group-aspecific and group-specific parts.

Approach 1: Independent DEBM

In this default approach, each group is considered as an independent dataset and the disease progression timeline in each group is estimated independently. GMM in such a scenario is illustrated in Fig. 3a.

Approach 2: Coupled DEBM

$$\text{DEBM} \rightarrow \begin{cases} p(x_{\cdot,i}|\neg E_i), p(x_{\cdot,i}|E_i) & \text{group-aspecific} \\ \theta_{i,g}, \{S_g, \lambda_g\}, & \text{group-specific} \end{cases} \tag{6}$$

In this approach, we assume that the different groups share the normal and abnormal PDFs, but the ordering in which these biomarkers become abnormal are different. The mixing parameters ($\theta_{i,g}$) are considered as group-specific part of the DEBM algorithm because the proportion of subjects with normal and abnormal biomarker values in each group g is correlated with the position of the biomarker along the ordering S_g, which we expect to be different in each group.

Hence, in our approach, we modify the alternating GMM optimization scheme to jointly optimize the GMM parameters of multiple groups. First, the GMM algorithm is initialized without considering the groups, as explained in Section 2.1. Secondly, as with the default DEBM, Gaussian parameters and mixing parameters are alternately optimized. In contrast in coupled DEBM, the Gaussian parameters are estimated jointly for all groups, while mixing parameters are estimated separately for each group. This has been illustrated in Figure 3b.

Once the GMM optimization has been performed, S_g and λ_g are estimated in each group. Patient staging (Υ_j) of the test-subjects in group g are computed based on the disease progression timeline $\{S_g, \lambda_g\}$.

Approach 3: Co-init DEBM

$$\text{DEBM} \rightarrow \begin{cases} \widehat{p}(x_{.,i}|\neg E_i), \widehat{p}(x_{.,i}|E_i) & \text{group-aspecific} \\ p_g(x_{.,i}|\neg E_i), p_g(x_{.,i}|E_i) & \text{group-specific} \\ \theta_{i,g}, \{S_g, \lambda_g\} & \text{group-specific} \end{cases} \tag{7}$$

In this approach, we assume that the different groups do not share the normal and abnormal PDFs, but that they are close to each other. Hence, in co-init DEBM, we relax the constraint on $p(x_{.,i}|\neg E_i)$ and $p(x_{.,i}|E_i)$ and instead consider the initialized values of normal and abnormal PDFs ($\widehat{p}(x_{.,i}|\neg E_i)$ and $\widehat{p}(x_{.,i}|E_i)$) to be group-aspecific part of DEBM. We estimate $p_g(x_{.,i}|\neg E_i)$ and $p_g(x_{.,i}|E_i)$ independently for each group. This is illustrated in Fig. 3c.

As with the previous approach, S_g, λ_g and the patient staging of the test-subjects in group g are computed independently for each group.

3. Experiments

Section 3.1 describes the experiments to evaluate the proposed DEBM approaches on a stratified population. Since ground-truth orderings are unknown in real clinical data, we use simulated datasets for evaluating the methods. After evaluating the proposed approaches, we select the best approach for analyzing the effect of *APOE* on AD progression using subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Section 3.2 descibes the details of these experiments.

3.1. Simulation experiments

We used the framework developed by Young et al. (2015) for simulating cross-sectional data consisting of scalar biomarker values for CN, MCI and AD subjects in two groups. In this framework, disease progression in a subject is modeled by a series of biomarker changes representing the temporal cascade of biomarker abnormality as estimated by an EBM. Individual biomarker trajectories are represented by sigmoids varying from the biomarker's normal value to its abnormal value. To account for inter-subject variability, the normal and abnormal values for different subjects are drawn randomly from Gaussian distributions.

The simulation dataset used in our experiments are based on a set of seven biomarkers as described in the simulation experiments of Venkatraghavan et al. (2019a). The simulated datasets were stratified into two groups, with each group having its own distinct disease progression patterns. There are two ways in which the progression of disease in the groups can differ: 1. difference in ground-truth orderings S_1 and S_2; 2. difference in the abnormal biomarker PDFs in the two groups *i.e.* $p_1(x_{.,i}|E_i)$ and $p_2(x_{.,i}|E_i)$. Each of these differences could affect the accuracy of the proposed approaches. Hence, we evaluated the proposed approaches in the presence of each of these differences. Normalized Kendall's Tau distance between the estimated ordering (S) and the ground-truth ordering (S_{gt}) was used as an evaluation measure in these experiments:

$$\epsilon_S = \frac{K(S, S_{gt})}{\binom{N}{2}} \tag{8}$$

where $K(A, B)$ is the number of swaps required to obtain ordering B from ordering A.

The normalization ensures that ϵ_S falls in the range $[0, 1]$, with 0 as the distance when the two orderings are the same, and 1 as the distance when the two orderings are the reverse of each other.

Experiment 1: The first simulation experiment studied the effect of the difference in ordering between the two groups. The ordering in the first group (Group 1) was fixed and the ordering in the second group (Group 2) was selected randomly such that the normalized Kendall's Tau distance between the two groups was a fixed number, say ϵ_O. ϵ_O was varied from 0 to 1 in steps of 0.2. The number of subjects in Group 2 was kept constant at 900. The number of subjects in Group 1 was varied from 100 to 900 in steps of 200, to study how the different approaches perform in small as well as large groups. The normal and abnormal biomarkers levels in the two groups were sampled from the same Gaussian distribution for this experiment. We generated 50 random repetitions of the simulated datasets, and reported mean and standard deviation of ϵ_S for independent DEBM, coupled DEBM, and co-init DEBM in groups 1 and 2.

Experiment 2: This experiment studied the performance of the proposed approaches with the $\mu_{g,i,E}$ parameter of the $p_g(x_{.,i}|E_i)$ distribution being different in the two groups. $\mu_{1,i,E}$ was fixed, and $\mu_{2,i,E}$ was varied such that the difference $\mu_{2,i,E} - \mu_{1,i,E}$ (ϵ_G) was one of $\{-0.2d, 0, +0.2d\}$ where $d = \mu_{1,i,E} - \mu_{1,i,\neg E}$. 0 is considered the reference level, where the abnormal Gaussians are the same in the two groups. $\mu_{g,i,\neg E}$ were kept the same in the two groups. Hence, when $\epsilon_G = -0.2d$, the abnormal biomarker levels are closer to the normal biomarker levels in Group 2 than in Group 1. This results in Group 2 biomarkers being weaker than their Group 1 counterparts when $\epsilon_G = -0.2d$ and stronger when $\epsilon_G = +0.2d$. The number of subjects in Group 2 was kept a constant at 900, while the subjects in Group 1 increased from 100 to 900. ϵ_O between the two groups was fixed at 0.4. We again generated 50 random repetitions of the simulated datasets, and reported mean and standard deviation of ϵ_S for coupled DEBM, co-init DEBM and DEBM.

These experiments were used to evaluate the different approaches mentioned in Section 2 and select the best method for analyzing the effect of *APOE* alleles in AD progression.

3.2. Studying the effect of APOE

We considered the baseline measurements from 417 CN, 235 MCI converters and 342 AD subjects in ADNI1, ADNIGO and ADNI2 studies.[2] The MCI converters are subjects who had MCI at baseline but converted to AD within 3 years of baseline measurement. We excluded subjects with significant memory concerns (without a diagnosis of AD or MCI) and MCI non-converters in our experiments to select a more phenotypically homogeneous group of subjects with prevalent or incident AD. In each of the experiments, the dataset was divided into three groups ($\varepsilon 2$ carriers, homozygous $\varepsilon 3$ carriers, and $\varepsilon 4$ carriers) based on the subject's *APOE* carriership (van der Lee et al., 2018). Subjects with *APOE* $\varepsilon 2, 4$ (n=34) were not included in either group because of the presence of both $\varepsilon 2$ and $\varepsilon 4$ alleles

[2] The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimers disease (AD). For up-to-date information, see www.adni-info.org.

Table 1
Demographics for the used population. 2⋆ represents the subjects with *APOE* alleles $\epsilon 2, 2$ and $\epsilon 2, 3$. 33 represents the subjects with reference *APOE* allele $\epsilon 3, 3$. ⋆4 represents the subjects with *APOE* alleles $\epsilon 3, 4$ and $\epsilon 4, 4$. Subjects with both $\epsilon 2$ and $\epsilon 4$ alleles were excluded from this study (n=34). Edu. is an abbreviation used for Education.

Demographics			
Diagnosis	CN	MCIc	AD
n	417	235	342
APOE 2⋆/33/⋆4	57/244/110	6/66/156	12/101/219
Sex M/F	209/208	145/90	189/153
Age [yrs.] ($\mu \pm \sigma$)	74.8 ± 5.7	73.7 ± 7.0	75.0 ± 7.8
Edu [yrs.] ($\mu \pm \sigma$)	16.3 ± 2.7	15.9 ± 2.7	15.2 ± 3.0

Table 2
Biomarker availability in number of subjects in the *APOE* based groups of $\epsilon 2$ carriers, homozygous $\epsilon 3$ carriers, and $\epsilon 4$ carriers.

Biomarker availability			
Biomarker	$\epsilon 2$ carriers ($N = 75$)	Homozygous $\epsilon 3$ carriers ($N = 411$)	$\epsilon 4$ carriers ($N = 485$)
Imaging	74	408	481
ABETA	57	301	357
PTAU	57	301	357
TAU	57	299	348
NG	21	113	131
NFL	23	118	137
MMSE	75	411	485
ADAS	74	410	477

Subject demographics and their *APOE* carrierships are summarized in Table 1. The modalities considered were structural imaging biomarkers, biomarkers extracted from cerebrospinal fluid (CSF), and cognitive biomarkers. Structural imaging biomarkers were obtained from T1-weighted MRI acquired at 1.5T or 3T. Details of the MRI acquisition protocols of ADNI can be found in Jack Jr. et al., 2008, 2015.

Imaging biomarkers were estimated from T1-weighted MRI scans analysed with FreeSurfer software v6.0 cross-sectional stream and outputs were visually checked. We assumed a symmetric pattern of atrophy in AD and averaged imaging biomarkers between the left and right hemisphere.

Experiment 3: For this experiment, the selected imaging biomarkers were: hippocampal volume, volume of the entorhinal cortex, fusiform gyrus volume, middle-temporal gyrus volume, precuneus volume, together with whole brain volume and volume of the ventricles (Archetti et al., 2019; Frisoni et al., 2010; Vemuri and Jack, 2010). The selected CSF based biomarkers were: CSF concentrations of Amyloid-β_{42} (ABETA), total Tau (TAU) and phosphorylated Tau_{181} (PTAU) proteins (Blennow and Hampel, 2003; Blennow et al., 2010), Neurogranin (Thorsell et al., 2010) and Neurofilament light chain (Jin et al., 2019; de Wolf et al., 2020). Mini mental state examination (MMSE) and Alzheimer's Disease Assessment Scale - Cognitive (13 items) (ADAS13) were used as cognitive biomarkers. The availability of these multimodal biomarkers in the ADNI database is summarized in Table 2.

We downloaded the CSF measurements from the ADNI database. The measurements of ABETA, TAU and PTAU had been made using the microbead-based multiplex immunoassay, the INNO-BIA AlzBio3 RUO (Olsson et al., 2005). The measurement of NFL had been made with enzyme-linked immunosorbent assay NF-light ELISA kit (Mattsson et al., 2017). NG had been measured by electrochemiluminescence technology (Meso Scale Discovery) using a monoclonal antibody specific for NG (Ng7) for coating together with a detector antibody polyclonal neurogranin anti-rabbit (ab 23570, Upstate) (Portelius et al., 2015). As described previously in Venkatraghavan et al. (2019a), the TAU and PTAU measurements were transformed to logarithmic scales to make the distributions less skewed and more suitable for DEBM analysis.

The volumes of the selected regions were regressed with age, sex and intra-cranial volume (ICV) and the effects of these factors were subsequently corrected for, before being used as biomarkers. The effects of age and sex were regressed out of CSF features, whereas effects of age, sex and education were regressed out of cognitive scores.

For the 12 selected biomarkers, we estimated the disease timelines in the three aforementioned groups using the method selected after simulation experiments. We studied the positional variance of the estimated orderings by creating 100 bootstrapped samples of the data. In order to evaluate if the estimated orderings in the three groups were significantly different from one another, we used permutation testing and estimated the distribution of the Kendall's Tau distance under the null hypothesis. To compute this distribution, we generated 10,000 random permutations of the three groups. We then computed the one-sided *p*-values for the actual Kendall's Tau distances between the orderings of the three groups, calculated as the proportion of sampled permutations where the distance was greater than or equal to the actual distance, and using Bonferroni correction to account for multiple testing.

Experiment 4: In this experiment, we validated the disease stage (Υ_j) by computing its correlation with the subjects' MMSE and ADAS13 values. We used a 10-fold cross validation, where the training set was used to estimate the disease timeline in the aforementioned groups and the test subjects' disease stage was evaluated by placing them on this disease timeline. We used the volume-based and CSF-based biomarkers from Experiment 3, but excluded MMSE and ADAS13 scores from the model.

4. Results

4.1. Simulations

Experiment 1: Fig. 4 (a) and (b) show the ordering errors (ϵ_S) in Group 1 of the simulation datasets for DEBM, coupled DEBM and co-init DEBM as a function of number of subjects in Group 1, when ϵ_O between the two groups changes from 0 to 1. Fig. 4 (c)–(e) show ϵ_S in Group 2 of the simulation datasets for the aforementioned methods, as a function of number of subjects in Group 1. In our experiments, Group 1 dataset remains the same while Group 2 dataset changes as ϵ_O increase. Hence DEBM results do not change with change in ϵ_O in Fig. 4 (a) and (b), whereas in Fig. 4 (c), DEBM results do not change with increase in number of subjects in Group 1.

It can be seen that both coupled-training methods (i.e., co-init DEBM and coupled DEBM) outperform the default method of independently training DEBM models. It can also be observed that in both co-init DEBM and coupled DEBM the ordering errors decrease as ϵ_O increases and that co-init DEBM outperforms coupled DEBM for lower values of ϵ_O, whereas the performance is on par with coupled DEBM for higher values of ϵ_O.

Experiment 2: Fig. 5 (a) and (b) show ϵ_S in Group 1 and Fig. 5 (c)–(e) show the same in Group 2, when varying ϵ_G. Even with $\epsilon_G \neq 0$, coupled training (i.e., co-init DEBM and coupled DEBM) outperformed the default method of independently training DEBM models. Co-init DEBM showed negligible change in the errors when $\epsilon_G \neq 0$. The performance of coupled DEBM in Group 1 worsened for $\epsilon_G = +0.2d$ (Fig. 5 (a)) and in Group 2 for $\epsilon_G = -0.2d$ (Fig. 5 (d)).

4.2. Studying the effect of APOE

The results in Experiments 1 and 2 show that the performance of co-init DEBM is more accurate and robust than coupled DEBM in most scenarios. We hence analyzed Experiments 3 and 4 using co-init DEBM.

Experiment 3: Fig. 6 shows orderings of CSF, global cognition and volumetric biomarkers in the *APOE* based groups of $\epsilon 2$ carriers, homozygous $\epsilon 3$ carriers, and $\epsilon 4$ carriers along with their uncertainty estimates. It can be seen that the uncertainty of the ordering in the $\epsilon 2$ carriers group was high. Despite this uncertainty, some biomarkers (i.e. MMSE,

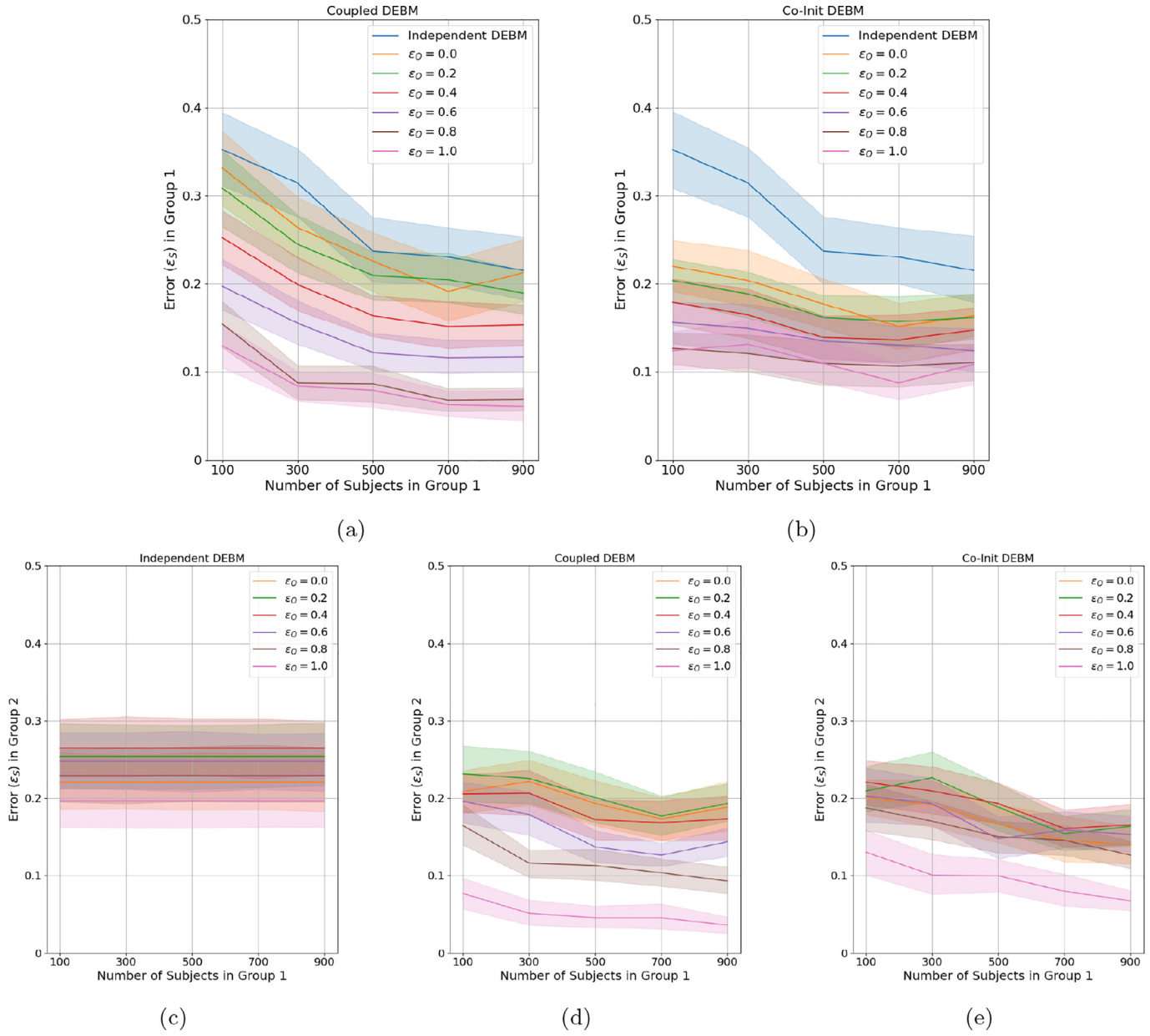


Fig. 4. Experiment 1: The effect of ϵ_O (the difference in groundtruth event orderings in the two groups) on the performance of the proposed methods. The shaded region in these plots represents standard deviation of the error in estimation of the proposed methods in 50 random iterations of simulations. The plots in (a) and (b) show the ordering errors in Group 1 using Coupled DEBM and Co-init DEBM with independent DEBM shown in both (a) and (b), as a function of number of subjects in Group 1. The plots in (c), (d) and (e) show the ordering errors in Group 2 using independent DEBM, Coupled DEBM and Co-init DEBM respectively as a function of number of subjects in Group 1.

NG and PTAU) seem to occur earlier than the other biomarkers in this group.

In the homozygous $\epsilon 3$ carrier group, ABETA was very prominently the earliest biomarker, followed by cognitive scores of MMSE and ADAS13. Among the CSF biomarkers, PTAU followed immediately after ABETA, which was inturn followed by TAU. NFL and NG were late biomarkers. Among the structural biomarkers, volumes of fusiform and middle-temporal gyri were the first to become abnormal, followed by ventricular volume and wholebrain volume. Hippocampus, precuneus and entorhinal volumes were late biomarkers in this group.

In the $\epsilon 4$ carrier group, the CSF biomarkers followed a pattern that was similar to that of the homozygous $\epsilon 3$ carrier group. The cognitive biomarkers were early biomarkers in this group as well. However the ordering in structural biomarkers was very different from that in the homozygous $\epsilon 3$ carrier group. Hippocampus and entorhinal volumes were early biomarkers in this group, followed by middle-temporal and fusiform gyri volumes. Wholebrain, ventricular and precuneus volumes were late biomarkers.

The ordering of the $\epsilon 2$ carrier group was significantly different from that of the homozygous $\epsilon 3$ carrier group ($p = 0.0156$, after Bonferroni correction for multiple testing). Similarly, the orderings for the other two groups were significant as well: $p = 0.0147$ for the difference between $\epsilon 2$ carrier group and $\epsilon 4$ carrier group and $p = 0.0003$ for the difference between the homozygous $\epsilon 3$ carrier group and $\epsilon 4$ carrier group.

Experiment 4: The variation of MMSE and ADAS13 scores with respect to the estimated disease stages has been plotted in Fig. 7, for all three groups. The patient stages showed a significant correlation with both MMSE and ADAS13 scores. The correlation coefficients were also comparable in the three groups.

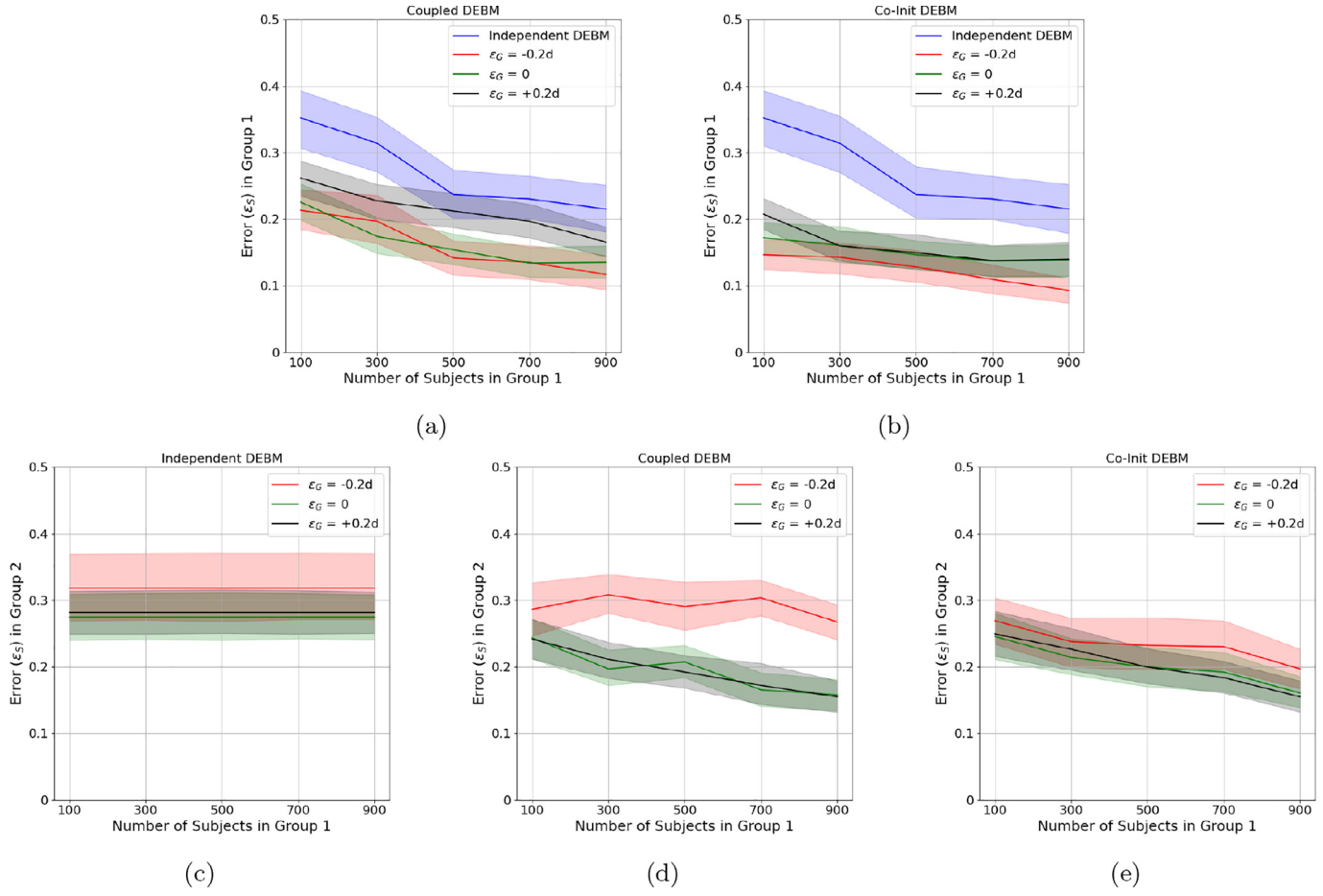


Fig. 5. Experiment 2: The effect of ϵ_G (difference in abnormal biomarker levels in the two groups), on the performance of the proposed methods. The shaded region represents standard deviation of the error in 50 random iterations. The plots in (a) and (b) show the ordering errors in Group 1 using Coupled DEBM and Co-init DEBM with independent DEBM shown in both (a) and (b), as a function of number of subjects in Group 1. The plots in (c), (d) and (e) show the ordering errors in Group 2 using independent DEBM, Coupled DEBM and Co-init DEBM respectively as a function of number of subjects in Group 1.

5. Discussion

DEBM models have been shown to be effective in determining the temporal cascade of biomarker abnormality as AD progresses, from cross-sectional data. In this work, we introduced a novel concept of splitting the different steps of DEBM into group-specific and group-aspecific parts for coupled training in stratified population. We considered two novel variations to split the steps of DEBM in this manner and through thorough experimentation in simulation datasets we observed that co-init DEBM helps in obtaining more accurate orderings in a stratified population. Using this method, we estimated the biomarker cascades in AD progression with $\epsilon 2$ alleles, homozygous $\epsilon 3$ alleles, and $\epsilon 4$ alleles of *APOE*, based on cross-sectional ADNI data. While the findings in the homozygous $\epsilon 3$ carrier and $\epsilon 4$ carrier groups fit the current understanding of progression of AD with high-confidence, the finding in the $\epsilon 2$ carrier group shows evidence for an alternative pathway (with relatively low confidence). In this section, we discuss the insights provided by the simulation experiments (Section 5.1) used for method selection as well as the insights into the AD progression pathways provided by our experiments on the ADNI dataset (Section 5.2).

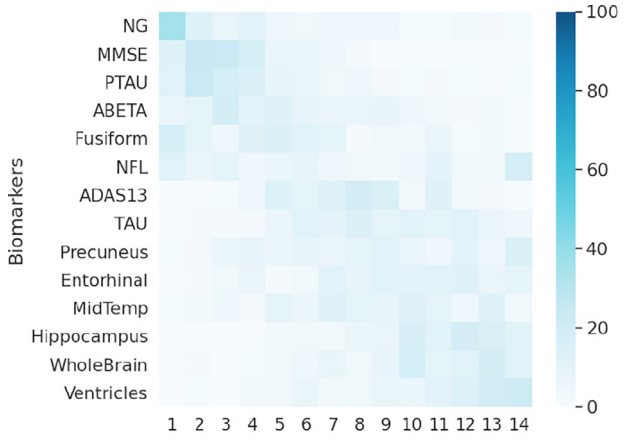
5.1. Choice of the method

Coupled DEBM and co-init DEBM both split DEBM into group-specific and group-aspecific steps for coupled training of an EBM in stratified populations. Experiment 1 and 2 showed that coupled training of the group-aspecific parts of DEBM and independently training the group-specific parts of DEBM results in more accurate orderings in the groups better than the default approach of independently training a DEBM model in each group.

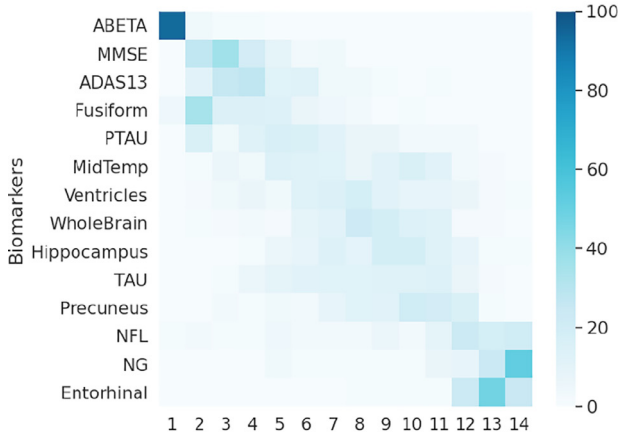
While splitting DEBM into group-specific and group-aspecific parts, we started with the assumption that the latent true normal and abnormal biomarker distributions in the groups are either same or similar. The difference between co-init DEBM and coupled DEBM is that, co-init DEBM accounts for slight differences in the underlying biomarker distributions between the groups whereas coupled DEBM does not.

The simulation dataset generated in Experiment 1 had the same true normal and abnormal biomarker distributions in the different groups, from which the simulated subjects were randomly sampled, aligning well with the assumption of coupled DEBM. However, this did not result in overall better accuracies for coupled DEBM than that of co-init DEBM. Co-init DEBM was also more robust than coupled DEBM as its accuracy was less dependent on ϵ_O, the distance between the ground-truth orderings in the two groups.

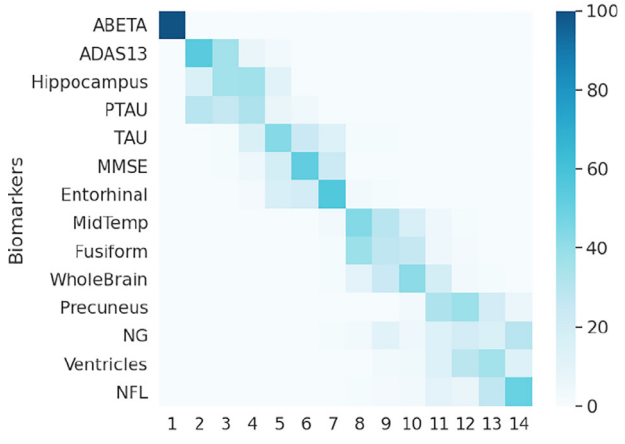
Another observation in Experiment 1, which was rather counter-intuitive, was that the errors made by the co-init and coupled DEBM models decreased as the distance between the ground-truth orderings in the two groups increased. When the orderings are further apart, the combined biomarker distributions in CN and AD groups have a larger overlap. The non-overlapping initialization (before the GMM optimization) thus results in the normal and abnormal distributions to be further apart. We hypothesize that this results in a better estimation of



(a) $\varepsilon 2$ carriers



(b) Homozygous $\varepsilon 3$ carriers



(c) $\varepsilon 4$ carriers

Fig. 6. Experiment 3: Orderings of CSF, global cognition and volumetric biomarkers in the *APOE* based groups of $\varepsilon 2$ carriers, homozygous $\varepsilon 3$ carriers, and $\varepsilon 4$ carriers along with their uncertainty estimates. Uncertainty in the estimation of the ordering was measured by 100 repetitions of bootstrapping, in the three *APOE* based groups. The color-map is based on the number of times a biomarker is at a position in 100 repetitions of bootstrapping. The number of subjects in the three groups were 75, 411 and 485 respectively. The orderings were obtained using Co-init DEBM.

the mixing parameters during GMM optimization and in-turn resulted in more accurate orderings, as mixing-parameters are dependent on the biomarker's position in the ordering.

In Experiment 2, we checked the performance of our approaches when the assumption (true normal and abnormal biomarker distributions being same across groups) is violated in the dataset. This experiment showed that the orderings obtained using co-init DEBM are more robust to differences between the abnormal Gaussians across groups than those obtained with coupled DEBM. With coupled DEBM, the error increased in the group with weaker biomarkers i.e., Group 1 in the case of $\epsilon_G = +0.2d$ and Group 2 in the case of $\epsilon_G = -0.2d$. This shows that coupled DEBM introduces a systematic bias in the estimation of ordering that is detrimental to the group with weaker biomarkers. Co-init DEBM also showed a similar bias, but to a much lesser extent.

We hence selected co-init DEBM as the preferred approach for splitting and performed our analysis on ADNI dataset using this approach. We expect that this idea of splitting DEBM into group-specific and group-aspecific parts can be easily extended to the EBM introduced by Fonteijn et al. (2012).

5.2. *Cascade of biomarker changes in the* APOE *based groups*

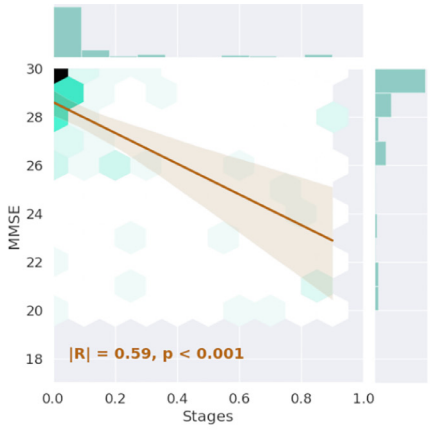
Dividing the total population into groups based on *APOE* carriership enabled us to create more phenotypically homogeneous groups (Weintraub et al., 2019), each with potentially specific disease progression timeline. In this section, we discuss our results in these *APOE* carriership based groups.

Our findings show that the three *APOE*-carriership based groups have significantly different temporal cascades of disease progression. This suggests that the underlying pathways of progression are different for the three genotypes. Among the CSF biomarkers in the homozygous $\varepsilon 3$ carrier and the $\varepsilon 4$ carrier groups, ABETA abnormality is the earliest biomarker event followed by PTAU. This fits current understanding of AD progression (Bloom, 2014). It also confirms the need for preventing the accumulation of ABETA in high-risk patients. NFL and NG are late biomarkers in the homozygous $\varepsilon 3$ carrier and $\varepsilon 4$ carrier groups, which suggests that axonal (Ashton et al., 2019) and synaptic (Thorsell et al., 2010) degeneration do not occur until very late in the disease process in these groups. NG being abnormal after PTAU and TAU in the homozygous $\varepsilon 3$ carrier and $\varepsilon 4$ carrier groups is also consistent with the previous findings that Tau mediates synaptic damage in AD (Jadhav et al., 2015).

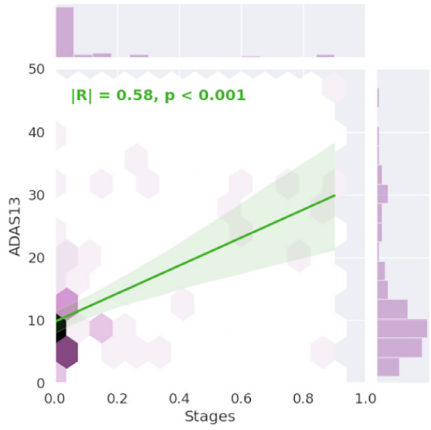
In the $\varepsilon 2$ carrier group, we found that the abnormal NG and PTAU are the earliest CSF events, even before ABETA becomes abnormal. This could hint at the existence of an alternative pathway for the formation of tau tangles in the brain before ABETA accumulation, as suggested in Weigand et al. (2019), but needs more extensive validation.

Among the volumetric biomarkers, Entorhinal cortex is one of the early biomarkers in the $\varepsilon 4$ carrier group which is supported by the findings in Huijbers et al. (2014), but is one of the last biomarkers to become abnormal in the homozygous $\varepsilon 3$ carrier group. Ventricular volume is a late biomarker in the $\varepsilon 4$ carrier group but it becomes abnormal quite early in the homozygous $\varepsilon 3$ carrier group as also observed by Nestor et al. (2008). Hippocampus volume is the earliest biomarker in the $\varepsilon 4$ carrier group, but is a relatively late biomarker in the homozygous $\varepsilon 3$ carrier and $\varepsilon 2$ carrier groups. This suggests that incidence of hippocampal sparing AD (Ferreira et al., 2017) could correlate with *APOE* carriership.

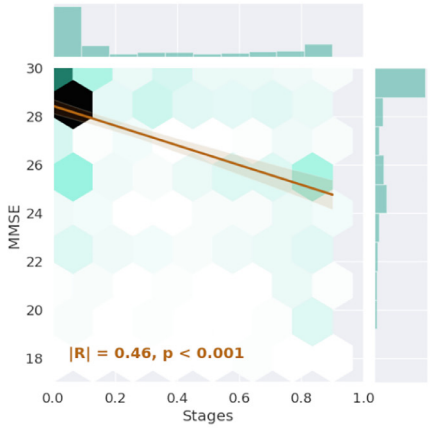
The findings related to these orderings of biomarker events were validated by correlating the patient stages derived from these orderings with MMSE and ADAS13 scores. Patient stages of subjects in all three groups, when used as test-subjects in a cross-validated manner, showed a significant correlation ($p < 0.001$) with these scores. These correlations validate our findings and suggest that these genotype-specific disease progression timelines could be used for patient monitoring.



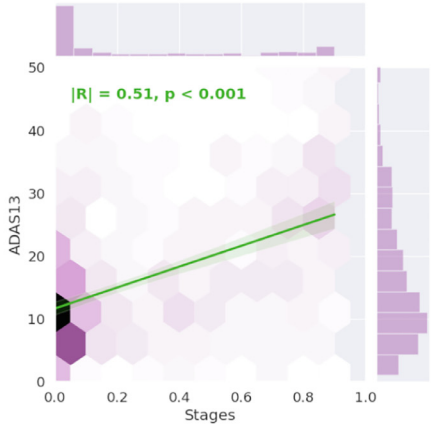
(a) $\varepsilon 2$: Stage vs MMSE



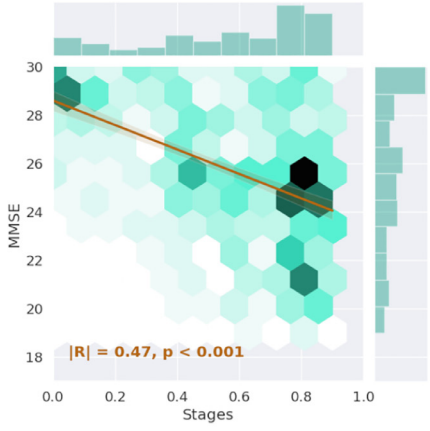
(b) $\varepsilon 2$: Stage vs ADAS13



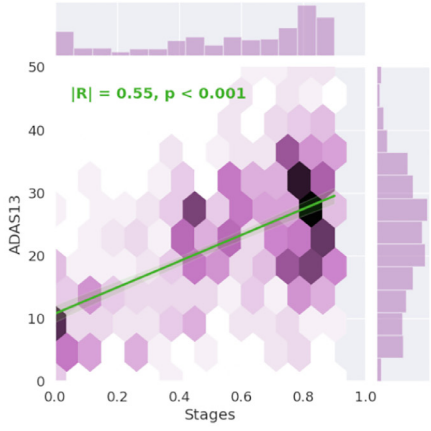
(c) Homozygous $\varepsilon 3$: Stage vs MMSE



(d) Homozygous $\varepsilon 3$: Stage vs ADAS13



(e) $\varepsilon 4$: Stage vs MMSE



(f) $\varepsilon 4$: Stage vs ADAS13

Fig. 7. Experiment 4: Correlation of estimated disease stages with MMSE and ADAS scores in the *APOE* based groups of $\varepsilon 2$ carriers, homozygous $\varepsilon 3$ carriers, and $\varepsilon 4$ carriers. The plot on top of each subfigure shows the probability density function of the disease stages, and the plot on the right of each subfigure shows the probability density function of the cognitive score in the subfigure. The 2D plot in each subfigure shows the joint density function of the two axes. The line in each subfigure shows the linear regression of MMSE / ADAS scores with the estimated disease stage and the shaded area around the line shows its 95% confidence interval. Figures (a),(c) and (e) depict correlation between MMSE score and obtained disease stages in the three *APOE* based groups. Figures (b), (d) and (f) depict correlation between ADAS13 score and the obtained disease stages in the three *APOE* based groups.

6. Conclusion and future work

We conclude that co-init DEBM provides the best accuracy and robustness when estimating orderings in stratified populations. Future work on co-init DEBM can focus on extending the approach for high-dimensional imaging biomarkers (Venkatraghavan et al., 2019b). This work also provides groundwork for extending the method towards hypothesis-free, data-driven stratification of phenotypes.

We gained new insights into the disease progression timeline of AD in the *APOE* based groups of $\varepsilon 2$ carriers, homozygous $\varepsilon 3$ carriers, and $\varepsilon 4$ carriers. While we observed that the estimated disease progression timelines in the $\varepsilon 4$ carrier and the homozygous $\varepsilon 3$ carrier groups fit

the current understanding of AD progression with high confidence, the estimated timelines in the $\epsilon 2$ carrier group may suggest an alternative pathway for the formation of tau tangles in the brain before amyloid β accumulation, albeit with relatively low condence. We expect that these genotype-specific disease progression timelines will benefit patient monitoring in the future, and may help optimize selection of eligible subjects for clinical trials.

Acknowledgement

This work is part of the EuroPOND initiative, which is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 666992. E.E. Bron acknowledges support from the Dutch Heart Foundation (PPP Allowance, 2018B011), Medical Delta Diagnostics 3.0: Dementia and Stroke, and the Netherlands Cardio-Vascular Research Initiative (Heart-Brain Connection: CVON2012-06, CVON2018-28). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimers Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICOLtd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

References

Archetti, D., Ingala, S., Venkatraghavan, V., Wottschel, V., Young, A.L., Bellio, M., Bron, E.E., Klein, S., Barkhof, F., Alexander, D.C., Oxtoby, N.P., Frisoni, G.B., Redolfi, A., 2019. Multi-study validation of data-driven disease progression models to characterize evolution of biomarkers in Alzheimer's disease. NeuroImage Clinical 24, 101954.

Ashton, N.J., Leuzy, A., Lim, Y.M., Troakes, C., Hortobgyi, T., Hglund, K., Aarsland, D., Lovestone, S., Schll, M., Blennow, K., Zetterberg, H., Hye, A., 2019. Increased plasma neurofilament light v ctchain concentration correlates with severity of post-mortem neurofibrillary tangle pathology and neurodegeneration. Acta Neuropathol. Commun. 7, 5.

Au, R., Piers, R.J., Lancashire, L., 2015. Back to the future: Alzheimer's disease heterogeneity revisited. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1, 368–370.

Blennow, K., Hampel, H., 2003. Csf markers for incipient Alzheimer's disease. Lancet Neurol. 2, 605–613.

Blennow, K., Hampel, H., Weiner, M., Zetterberg, H., 2010. Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. Nat. Rev. Neurol. 6, 131–144.

Bloom, G.S., 2014. Amyloid and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol. 71, 505–508.

Ferreira, D., Nordberg, A., Westman, E., 2020. Biological subtypes of alzheimer disease. Neurology 94, 436–448.

Ferreira, D., Verhagen, C., Hernndez-cabrera, J.A., Cavallin, L., Guo, C.j., Ekman, U., Muehlboeck, J., Simmons, A., Barroso, J., Wahlund, L.o., Westman, E., 2017. Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: longitudinal trajectories and clinical applications. Sci. Rep. 7, 46263.

Fligner, M.A., Verducci, J.S., 1988. Multistage ranking models. J. Am. Stat. Assoc. 83, 892–901.

Fonteijn, H.M., Modat, M., Clarkson, M.J., Barnes, J., Lehmann, M., Hobbs, N.Z., Scahill, R.I., Tabrizi, S.J., Ourselin, S., Fox, N.C., Alexander, D.C., 2012. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. NeuroImage 60, 1880–1889.

Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6, 67–77.

Genin, E., Hannequin, D., Wallon, D., Sleegers, K., Hiltunen, M., Combarros, O., Bullido, M., Engelborghs, S., Paul, D., Berr, C., Pasquier, F., Dubois, B., Tognoni, G., Fivet, N., Brouwers, N., Bettens, K., Arosio, B., Coto, E., Zompo, M., Campion, D., 2011. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol. Psychiatry 16, 903–907.

Huang, J., Alexander, D., 2012. Probabilistic event cascades for Alzheimer's disease. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 3095–3103.

Huijbers, W., Mormino, E.C., Wigman, S.E., Ward, A.M., Vannini, P., McLaren, D.G., Becker, J.A., Schultz, A.P., Hedden, T., Johnson, K.A., Sperling, R.A., 2014. Amyloid deposition is linked to aberrant entorhinal activity among cognitively normal older adults. J. Neurosci. 34, 52005210.

Jack Jr., C.R., Barnes, J., Bernstein, M.A., Borowski, B.J., Brewer, J., Clegg, S., Dale, A.M., Carmichael, O., Ching, C., DeCarli, C., Desikan, R.S., Fennema-Notestine, C., Fjell, A.M., Fletcher, E., Fox, N.C., Gunter, J., Gutman, B.A., Holland, D., Hua, X., Insel, P., Kantarci, K., Killiany, R.J., Krueger, G., Leung, K.K., Mackin, S., Maillard, P., Malone, I.B., Mattsson, N., McEvoy, L., Modat, M., Mueller, S., Nosheny, R., Ourselin, S., Schuff, N., Senjem, M.L., Simonson, A., Thompson, P.M., Rettmann, D., Vemuri, P., Walhovd, K., Zhao, Y., Zuk, S., Weiner, M., 2015. Magnetic resonance imaging in Alzheimer's disease neuroimaging initiative 2. Alzheimer's & Dementia 11, 740–756.

Jack Jr., C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., L. Whitwell, J., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.

Jack Jr., C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Pankratz, V.S., Donohue, M.C., Trojanowski, J.Q., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol. 12, 207–216.

Jadhav, S., Cubinkova, V., Zimova, I., Brezovakova, V., Madari, A., Cigankova, V., Zilka, N., 2015. Tau-mediated synaptic damage in Alzheimer's disease. Transl. Neurosci. 6, 214226.

Jedynak, B.M., Lang, A., Liu, B., Katz, E., Zhang, Y., Wyman, B.T., Raunig, D., Jedynak, C.P., Caffo, B., Prince, J.L., 2012. A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease neuroimaging initiative cohort. NeuroImage 63, 1478–1486.

Jin, M., Cao, L., Dai, Y.p., 2019. Role of neurofilament light chain as a potential biomarker for Alzheimer's disease: a correlative meta-analysis. Front. Aging Neurosci. 11, 254.

Kim, J., Basak, J.M., Holtzman, D.M., 2009. The role of apolipoprotein e in Alzheimer's disease. Neuron 63, 287–303.

van der Lee, S.J., Wolters, F.J., Ikram, M.K., Hofman, A., Ikram, M.A., Amin, N., van Duijn, C.M., 2018. The effect of APOE and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. Lancet Neurol. 17, 434–444.

Lorenzi, M., Filippone, M., Frisoni, G.B., Alexander, D.C., Ourselin, S., 2019. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. NeuroImage 190, 56–68.

Mattsson, N., Andreasson, U., Zetterberg, H., Blennow, K., Initiative, f. t.A.D.N., 2017. Association of plasma neurofilament light with neurodegeneration in patients with alzheimer disease. JAMA Neurol. 74, 557–566.

Murray, M.E., Graff-Radford, N.R., Ross, O.A., Petersen, R.C., Duara, R., Dickson, D.W., 2011. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. Lancet Neurol. 10, 785–796.

Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R., Initiative, A.D.N., 2008. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain 131, 24432454.

Olsson, A., Vanderstichele, H., Andreasen, N., De Meyer, G., Wallin, A., Holmberg, B., Rosengren, L., Vanmechelen, E., Blennow, K., 2005. Simultaneous measurement of β-amyloid(142), total tau, and phosphorylated tau (thr181) in cerebrospinal fluid by the xMAP technology. Clin. Chem. 51, 336–345.

Patterson, C., 2018. World Alzheimer Report 2018. Alzheimer's Disease International.

Portelius, E., Zetterberg, H., Skillbck, T., Trnqvist, U., Andreasson, U., Trojanowski, J.Q., Weiner, M.W., Shaw, L.M., Mattsson, N., Blennow Kaj, f. t. A.D.N.I., 2015. Cerebrospinal fluid neurogranin: relation to cognition and neurodegeneration in Alzheimer's disease. Brain 138, 3373–3385.

Saunders, A., Strittmatter, W., Schmechel, D., St. George-Hyslop, P., Pericak-Vance, M., Joo, S., Rosi, B., Gusella, J., Crapper-Mac Lachlan, D., Alberts, M., Hulette, C., Crain, B., Goldgaber, D., Roses, A., 1993. Association of apolipoprotein e allele 4 with late-onset familial and sporadic Alzheimers disease. Neurology 43, 1467–1472.

Schiratti, J.B., Allassonnière, S., Routier, A., Colliot, O., Durrleman, S., 2015. A Mixed-effects Model With Time Reparametrization for Longitudinal Univariate Manifold-valued Data. Springer International Publishing, Cham, pp. 564–575.

Thorsell, A., Bjerke, M., Gobom, J., Brunhage, E., Vanmechelen, E., Andreasen, N., Hansson, O., Minthon, L., Zetterberg, H., Blennow, K., 2010. Neurogranin in cerebrospinal fluid as a marker of synaptic degeneration in alzheimer's disease. Brain Research 1362, 13–22.

Vemuri, P., Jack, C.R., 2010. Role of structural MRI in Alzheimer's disease. Alzheimer's Res. Ther. 2, 23.

Venkatraghavan, V., Bron, E.E., Niessen, W.J., Klein, S., 2017. A discriminative event based model for Alzheimer's disease progression modeling. In: Niethammer, M., Styner, M., Aylward, S., Zhu, H., Oguz, I., Yap, P.T., Shen, D. (Eds.), Information Processing in Medical Imaging. Springer International Publishing, Cham, pp. 121–133.

Venkatraghavan, V., Bron, E.E., Niessen, W.J., Klein, S., 2019. Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. NeuroImage 186, 518–532.

Venkatraghavan, V., Dubost, F., Bron, E., Niessen, W., de Bruijne, M., Klein, S., 2019. Event-based modeling with high-dimensional imaging biomarkers for estimating spatial progression of dementia. In: Chung, A., Gee, J., Yushkevich, P., Bao, S. (Eds.), Information Processing in Medical Imaging - 26th International Conference, IPMI 2019, Proceedings. Springer, pp. 169–180.

Weigand, A.J., Bangen, K.J., Thomas, K.R., Delano-Wood, L., Gilbert, P.E., Brickman, A.M., Bondi, M.W., Initiative, A.D.N., 2019. Is tau in the absence of amyloid on the alzheimers continuum?: A study of discordant PET positivity. Brain Commun. 2.

Weintraub, S., Teylan, M., Rader, B., Chan, K.C., Bollenbeck, M., Kukull, W.A., Coventry, C., Rogalski, E., Bigio, E., Mesulam, M.M., 2019. APOE is a correlate of phenotypic heterogeneity in alzheimer disease in a national cohort. Neurology.

World Health Organization., 2017. Global action plan on the public health response to dementia 2017–2025.

de Wolf, F., Ghanbari, M., Licher, S., McRae-McKee, K., Gras, L., Weverling, G.J., Wermeling, P., Sedaghat, S., Ikram, M.K., Waziry, R., Koudstaal, W., Klap, J., Kostense, S., Hofman, A., Anderson, R., Goudsmit, J., Ikram, M.A., 2020. Plasma tau, neurofilament light chain and amyloid levels and risk of dementia; a population-based cohort study. Brain 143, 1220–1232.

Young, A.L., Oxtoby, N.P., Daga, P., Cash, D.M., Fox, N.C., Ourselin, S., Schott, J.M., Alexander, D.C., 2014. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 137, 2564–2577.

Young, A.L., Oxtoby, N.P., Ourselin, S., Schott, J.M., Alexander, D.C., 2015. A simulation system for biomarker evolution in neurodegenerative disease. Med. Image Anal. 26, 47–56.