doi:10.1093/brain/awy093 BRAIN 2018: 141; 1871–1883 | 1871

BRAIN
A JOURNAL OF NEUROLOGY

Brain properties predict proximity to symptom onset in sporadic Alzheimer's disease

Jacob W. Vogel,[1,2] Etienne Vachon-Presseau,[3] Alexa Pichet Binette,[2] Angela Tam,[2,4] Pierre Orban,[2,4,5] Renaud La Joie,[6] Mélissa Savard,[2] Cynthia Picard,[2,7] Judes Poirier,[2,7,8] Pierre Bellec,[4,9] John C. S. Breitner,[2,7] Sylvia Villeneuve[1,2,7] for the Alzheimer's Disease Neuroimaging Initiative* and the PREVENT-AD Research Group

*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

See Tijms and Visser (doi:10.1093/brain/awy113) for a scientific commentary on this article.

Alzheimer's disease is preceded by a lengthy 'preclinical' stage spanning many years, during which subtle brain changes occur in the absence of overt cognitive symptoms. Predicting when the onset of disease symptoms will occur is an unsolved challenge in individuals with sporadic Alzheimer's disease. In individuals with autosomal dominant genetic Alzheimer's disease, the age of symptom onset is similar across generations, allowing the prediction of individual onset times with some accuracy. We extend this concept to persons with a parental history of sporadic Alzheimer's disease to test whether an individual's symptom onset age can be informed by the onset age of their affected parent, and whether this estimated onset age can be predicted using only MRI. Structural and functional MRIs were acquired from 255 ageing cognitively healthy subjects with a parental history of sporadic Alzheimer's disease from the PREVENT-AD cohort. Years to estimated symptom onset was calculated as participant age minus age of parental symptom onset. Grey matter volume was extracted from T_1-weighted images and whole-brain resting state functional connectivity was evaluated using degree count. Both modalities were summarized using a 444-region cortical-subcortical atlas. The entire sample was divided into training ($n = 138$) and testing ($n = 68$) sets. Within the training set, individuals closer to or beyond their parent's symptom onset demonstrated reduced grey matter volume and altered functional connectivity, specifically in regions known to be vulnerable in Alzheimer's disease. Machine learning was used to identify a weighted set of imaging features trained to predict years to estimated symptom onset. This feature set alone significantly predicted years to estimated symptom onset in the unseen testing data. This model, using only neuroimaging features, significantly outperformed a similar model instead trained with cognitive, genetic, imaging and demographic features used in a traditional clinical setting. We next tested if these brain properties could be generalized to predict time to clinical progression in a subgroup of 26 individuals from the Alzheimer's Disease Neuroimaging Initiative, who eventually converted either to mild cognitive impairment or to Alzheimer's dementia. The feature set trained on years to estimated symptom onset in the PREVENT-AD predicted variance in time to clinical conversion in this separate longitudinal dataset. Adjusting for participant age did not impact any of the results. These findings demonstrate that years to estimated symptom onset or similar measures can be predicted from brain features and may help estimate presymptomatic disease progression in at-risk individuals.

1 Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada
2 Centre for the Studies on Prevention of Alzheimer's Disease, Douglas Mental Health University Institute Research Centre, Montreal, Quebec, Canada
3 Department of Physiology, Northwestern University, Chicago, IL, USA

Received September 24, 2017. Revised February 20, 2018. Accepted February 25, 2018. Advance Access publication April 23, 2018

4 Centre de recherche de l'Institut universitaire de gériatrie de Montréal, Montreal, Quebec, Canada
5 Department of Psychiatry, University of Montreal, Montreal, Quebec, Canada
6 Memory and Aging Center, University of California, San Francisco, California, USA
7 Department of Psychiatry, McGill University, Montreal, Quebec, Canada
8 McGill University and Genome Quebec Innovation Centre, Quebec, Canada
9 Department of Computer Science and Operations Research, University of Montreal, Montreal, QC, Canada

Correspondence to: Sylvia Villeneuve
Douglas Mental Health University Institute
Perry Pavilion Room E3417.1
6875 Boulevard LaSalle
Montreal, QC
Canada H4H1R3
E-mail: Sylvia.villeneuve@mcgill.ca

Keywords: functional MRI; structural MRI; biomarkers; MCI; machine learning

Abbreviations: ADNI = Alzheimer's Disease Neuroimaging Initiative; Lasso = Least Absolute Shrinkage and Selection Operator; PREVENT-AD = PResymptomatic EValuation of Experimental of Novel Treatment of Alzheimer's disease

Introduction

Alzheimer's disease is characterized by a lengthy preclinical period involving progressive changes in brain structure and function that precede overt cognitive symptoms (Jack *et al.*, 2013; Sperling *et al.*, 2014; Iturria-Medina *et al.*, 2016). These brain alterations are subtle but measurable, allowing researchers to use neuroimaging to identify individuals at risk for Alzheimer's dementia well before cognitive symptoms manifest (Mathotaarachchi *et al.*, 2017). Few studies have examined how neuroimaging features can predict the timing of symptom onset in preclinical individuals (Oulhaj *et al.*, 2009; Jack *et al.*, 2010; Bateman *et al.*, 2012; Zhang and Shen, 2012; Ten Kate *et al.*, 2017). This is an important topic of research, considering the preclinical period of Alzheimer's disease can last decades (Villemagne *et al.*, 2013), and yet clinical trials typically only last a few years. The ability to predict symptom onset timing could increase the power of such trials by enriching for individuals on the precipice of their symptom onset. As yet, there is no consensus on biomarkers that can accurately reflect the temporality of preclinical disease progression, and no studies exist that have used a prediction/validation study design to establish generalizability of such markers.

A small proportion (<1%) of dementia due to Alzheimer's disease is attributable to autosomal dominant genetic mutations for which a single copy of a pathogenic allele is sufficient to cause symptoms (Mendez, 2013). In such cases, the age at symptom onset is similar within and across generations (Ryman *et al.*, 2014). By comparing a mutation carrier's chronological age with his/her family's typical age of symptom onset (age of the person minus the age of familial onset), one can estimate the number of years to symptom onset (Bateman *et al.*, 2012). The study of disease characteristics related to years to estimated symptom onset has revealed a sequence of Alzheimer's disease biomarkers that become increasingly abnormal as predisposed individuals approach their expected onset timing (Bateman *et al.*, 2012). This pattern has been observed using markers specific to Alzheimer's disease pathology, as well as using MRI data representing structural and functional properties of the brain (Chhatwal *et al.*, 2012; Benzinger *et al.*, 2013).

Most Alzheimer's disease cases are not caused by a single gene mutation and are often referred to as sporadic Alzheimer's disease. Despite the characterization of this illness as 'sporadic', it is known from both parent–offspring and twin studies to be strongly heritable (H estimated at 50–70%) (Gatz *et al.*, 2006). Importantly, age of symptom onset has also been shown to be heritable in sporadic Alzheimer's disease (Li *et al.*, 2002; Dickson *et al.*, 2008; Kamboh *et al.*, 2012; Wingo *et al.*, 2012; Naj *et al.*, 2014). It is unknown whether a measure similar to the estimated onset age calculation might serve in sporadic Alzheimer's disease to indicate advancement of presymptomatic disease. We investigated this topic in asymptomatic subjects with a parental history of sporadic Alzheimer's disease. We took the parent's age of dementia onset as a broad indicator of the age at which such an individual's symptoms might appear, and used this information to calculate each individual's years to estimated symptom onset in the context of sporadic Alzheimer's disease.

Individuals with Alzheimer's disease express a consistent pattern of neurodegeneration that can be detected early on in the disease progression (Buckner *et al.*, 2005; Dickerson *et al.*, 2011). Several lines of evidence also suggest that Alzheimer's disease alters resting state functional MRI connectivity (Sheline *et al.*, 2010; Mormino *et al.*, 2012; Brier *et al.*, 2014). We hypothesized that individuals approaching their parent's onset age would express more severe structural and functional alterations in these Alzheimer's disease-specific regions. To establish the generalizability of these findings, we assessed whether imaging features associated with our construct of years to estimated symptom

onset could predict years to estimated symptom onset in unseen test data from the same dataset. Finally, we generalized the years to estimated symptom onset construct by testing whether the same imaging features could predict the actual timing of conversion events (either to mild cognitive impairment or to dementia) in a well characterized longitudinal dataset.

Materials and methods

Participants

We studied 255 cognitively intact individuals aged 55 or older assembled for PResymptomatic EValuation of Experimental of Novel Treatment of Alzheimer's disease (PREVENT-AD). Enrolment criteria can be found elsewhere (Breitner *et al.*, 2016). All participants had a Clinical Dementia Rating (Perneczky *et al.*, 2006) of 0, and had at least one parent diagnosed clinically with sporadic Alzheimer's disease or a condition suggesting Alzheimer's-like dementia (Tschanz *et al.*, 2013). We estimated parental onset age as the age at which there was unambiguous evidence of cognitive impairment, as reported by the participant and occasionally corroborated by medical records. We then estimated years till parental symptom onset as the age of the participant minus the onset age of the earliest affected parent (Bateman *et al.*, 2012). All participants were fully briefed and gave their explicit consent for participation using procedures and consent forms approved by the Institutional Review Board of the McGill University Faculty of Medicine. Demographic information for this sample can be found in Table 1.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is an open-source public dataset designed to accelerate the discovery of imaging markers for Alzheimer's disease progression and clinical trials. General information about ADNI, as well as specificities of ADNI diagnostic criteria, can be found on the ADNI website (http://adni.loni.usc.edu/methods/documents). One hundred and eighty-eight ADNI subjects with full structural MRI and resting state functional MRI data were identified, 149 of whom passed data quality control procedures (see below). From this group, we identified 26 'converters'. These were individuals who were cognitively normal at baseline but later received a diagnosis of mild cognitive impairment, or those who had mild cognitive impairment at baseline but subsequently developed dementia. Time to conversion was calculated as the number of days between the baseline visit and the date of the session where the subjects were first reported as having converted. This measure was then centred around time of clinical conversion onset to correspond to the scale of estimated symptom onset in the PREVENT-AD (e.g. −3 represents 3 years to clinical conversion). Subjects who had mild cognitive impairment at baseline but converted back to a status of cognitively normal were excluded. Along with the magnetic resonance images, we also downloaded demographic information, Montreal Cognitive Assessment (Nasreddine *et al.*, 2005) scores, *APOE4* carriage information and hippocampal volume. Table 2 contains demographic variables for the ADNI converters.

Table 1 Demographic information for the PREVENT-AD cohort

	Training set	Testing set	Total
n	138	68	206
Age (SD)	62.9 (4.8)	62.2 (5.3)	62.7 (4.9)
% Female	73.5	79.4	74.8
Montreal Cognitive Assessment Score (SD)	28.0 (1.5)	28.2 (1.5)	28.2 (1.5)
Total intracranial volume (SD)	1.40 (0.18)	1.37 (0.11)	1.40 (0.17)
Frame Displacement (SD)	0.22 (0.05)	0.21 (0.05)	0.21 (0.05)
Parental symptom onset (SD)	74.1 (8.1)	73.5 (8.1)	73.9 (8.09)
Years to estimated symptom onset (SD)	−11.1 (8.1)	−11.3 (7.7)	−11.2 (8.0)

SD = standard deviation.
Years to estimated symptom onset was calculated by subtracting the parental symptom onset age from the participant age.

MRI acquisition

MRI acquisition procedures for the PREVENT-AD have been documented elsewhere (Orban *et al.*, 2015). Briefly, MRI data were acquired using a 3 T Magnetom Tim Trio (Siemens) scanner. For structural scans, T_1-weighted images were obtained using a GRE sequence with the following parameters: repetition time = 2300 ms; echo time = 2.98 ms; flip angle = 9°; matrix size = 256 × 256; voxel size = 1 × 1 × 1 mm^3; 176 slices. For resting state functional MRI scans, two consecutive functional T_2*-weighted scans were collected with a blood oxygenation level-dependent (BOLD) sensitive, single-shot echo planar sequence with the following parameters: repetition time = 2000 ms; volumes = 150; echo time = 30 ms; flip angle = 90°; matrix size = 64 × 64; voxel size = 4 × 4 mm^3; 32 slices. ADNI acquisition procedures have previously been described in detail (Jack *et al.*, 2008).

APOE genotyping

Automated DNA extraction was applied for buffy coat samples using the QiaSymphony DNA mini kit (Qiagen). *APOE4* genotype was determined with the help of the PyroMark Q96 pyrosequencer (Qiagen) using the following primers: rs4293 58_amplification_forward 5′-ACGGCTGTCCAAGGAGCT G-3′, rs429358_amplification_reverse_biotinylated 5′-CACCTCGCCGC GGTACTG-3′, rs429358_sequencing 5′-CGGACATGGAGG ACG-3′, rs7412_amplification_forward 5′-CTCCGCGATGCCG ATGAC-3′, rs7412_amplification_reverse_biotinylated 5′-CCCC GGCCTGGTACACTG-3′ and rs7412_sequencing 5′-CGATGA CCTGCAGAAG-3′. Participants were binarized as *APOE4* carriers or non-carriers.

Cognitive scores

As part of the PREVENT-AD battery, all participants were assessed using the Montreal Cognitive Assessment Scale and underwent full cognitive testing using the Repeatable Battery for the Assessment of Neuropsychological Status (Gold *et al.*, 1999). Z-scores of these cognitive tests were aggregated into

Table 2 Demographic characteristics of the ADNI converters

n	26
Age (SD)	72.8 (6.4)
% Female	48.1
% Mild cognitive impairment	77.8
Montreal Cognitive Assessment Score (SD)	23.2 (2.9)
Years to conversion (SD)	1.83 (1.24)

five cognitive composite scores as per recommendations of the battery: Immediate Memory, Attention, Visuospatial Construction, Language and Delayed Memory.

Functional preprocessing

PREVENT-AD and ADNI functional data were preprocessed using the NeuroImaging Analysis Kit version 0.12.17, using GNU Octave version 3.6.1 and the minc-toolkit version 0.3.18. Detailed preprocessing and quality control procedures have been documented elsewhere (Orban *et al.*, 2015). Briefly, for each subject, functional images underwent slice-timing correction, and parameters for rigid realignment were estimated between each run, and between the mean and the T_1 image. T_1 images underwent non-linear structural normalization to the MNI-ICBM152 stereotaxic symmetric template using CIVET (Redolfi *et al.*, 2015). Parameters of the previous transformations were applied to each functional MRI run to bring them into MNI space and resample them to 3 mm isotropic. Slow time drifts, average white matter and CSF signal, and motion artefacts (first six principal components of the six realignment parameters, and their squares) were regressed from the functional MRI time series. Framewise displacement was calculated to assess functional timeframes with excessive motion (Power *et al.*, 2012). All time frames exhibiting frame displacement >0.5 were removed (scrubbed), along with one adjacent frame prior, and two consecutive frames after. Subjects with <70 non-scrubbed frames were considered to have failed quality control and were excluded from analysis. Overall, seven PREVENT-AD subjects had unavailable resting state functional MRI scans, and 42 subjects failed quality control procedures and were removed, leaving 206 for analysis. Similarly, 39 ADNI subjects failed quality control procedures and were excluded, leaving 149 in total (26 converters).

To prepare our data for feature selection, we attempted to strike a balance between biological resolution and dimensionality. We therefore used a freely available medium-high resolution cortical parcellation with 444 regions of interest (Urchs *et al.*, 2017) to summarize the resting state functional MRI data further. One region of interest was removed due to spatial inconsistency. Mean BOLD activity was calculated within each of the 443 regions of interest at each timeframe for each subject. These data were transformed into 443×443 correlation matrix where the timeseries at each brain region was correlated with that of every other brain region, and this matrix underwent Fisher's Z transformation. For PREVENT-AD, the connectivity values were averaged across the two functional MRI runs.

Functional brain connectivity was summarized using degree count, a graph theory metric representing whole-brain high-confidence connection density per region (Bullmore and Sporns, 2009; Power *et al.*, 2011, 2013). This metric was chosen because it effectively reduces the dimensions of the functional connectivity data, but also because this sparsity better represents the properties of real-world graphs (Power *et al.*, 2013). Appropriate calculation of degree count requires a sparse matrix, and so a threshold was applied to each subject's connectivity matrix, and connections were binarized such that those exceeding the threshold were set to 1 (Bullmore and Sporns, 2009). Degree count for each subject was calculated as the sum of binarized connections between each brain region and every other brain region. The end result is one general feature per region, which indicates its level of connectivity with the rest of the brain. Altogether, 443 resting state functional MRI connectivity features were therefore available for each PREVENT-AD and ADNI subject. Degree count for each subject was obtained at several thresholds (90%, 92%, 94%, 95%, 96%, 98%) (Power *et al.*, 2013), and a final threshold was chosen through cross-validation (see below).

Structural image processing

We used identical processing procedures for PREVENT-AD and ADNI subjects. Each participant's T_1-weighted image was segmented into grey matter, white matter and CSF using Statistical Parametric Mapping 12 v.6225, running on MATLAB version 2012a. Total intracranial volume was calculated by adding together the values for all three segmentation volumes. The DARTEL toolbox (Ashburner, 2007) was used to create a subject-specific template (separately for PREVENT-AD and ADNI cohorts), to which all scans were non-linearly normalized. The template then underwent non-linear registration with modulation for linear and non-linear deformations to the MNI-ICBM152 template, and the parameters were applied to each subject's T_1 image to move it into template space. Each image underwent visual quality control after segmentation and after non-linear transformation. Template space grey matter probability (c1) maps were smoothed with an 8 mm^3 isotropic Gaussian kernel, creating grey matter volume images, which were masked with a maximum probability grey matter mask generated from the group-average image. To be consistent with resting state functional MRI analyses, grey matter volume was summarized by averaging volume within each of the 443 regions of interest from the parcellation applied to the resting state functional MRI data, and total intracranial volume was included as an independent feature (Schwarz *et al.*, 2016). Therefore, grey matter volume and connectivity features were available for each subject, for a total of 887 features. Subjects with unavailable resting state functional MRI features (due to exclusion) were removed from analysis leaving a total of 206 subjects. For purposes of comparison, single measures summarizing hippocampal volume were acquired for all subjects in native space using a previously described approach (Coupé *et al.*, 2011).

Statistical analysis

Data were analysed using a training, validation, testing, and generalizing approach with a feature selection routine and nested cross-validation to attain a linear model with improved generalizability (Tavor *et al.*, 2016; Youssofzadeh *et al.*, 2017). We used Least Absolute Shrinkage and Selection Operator (Lasso)

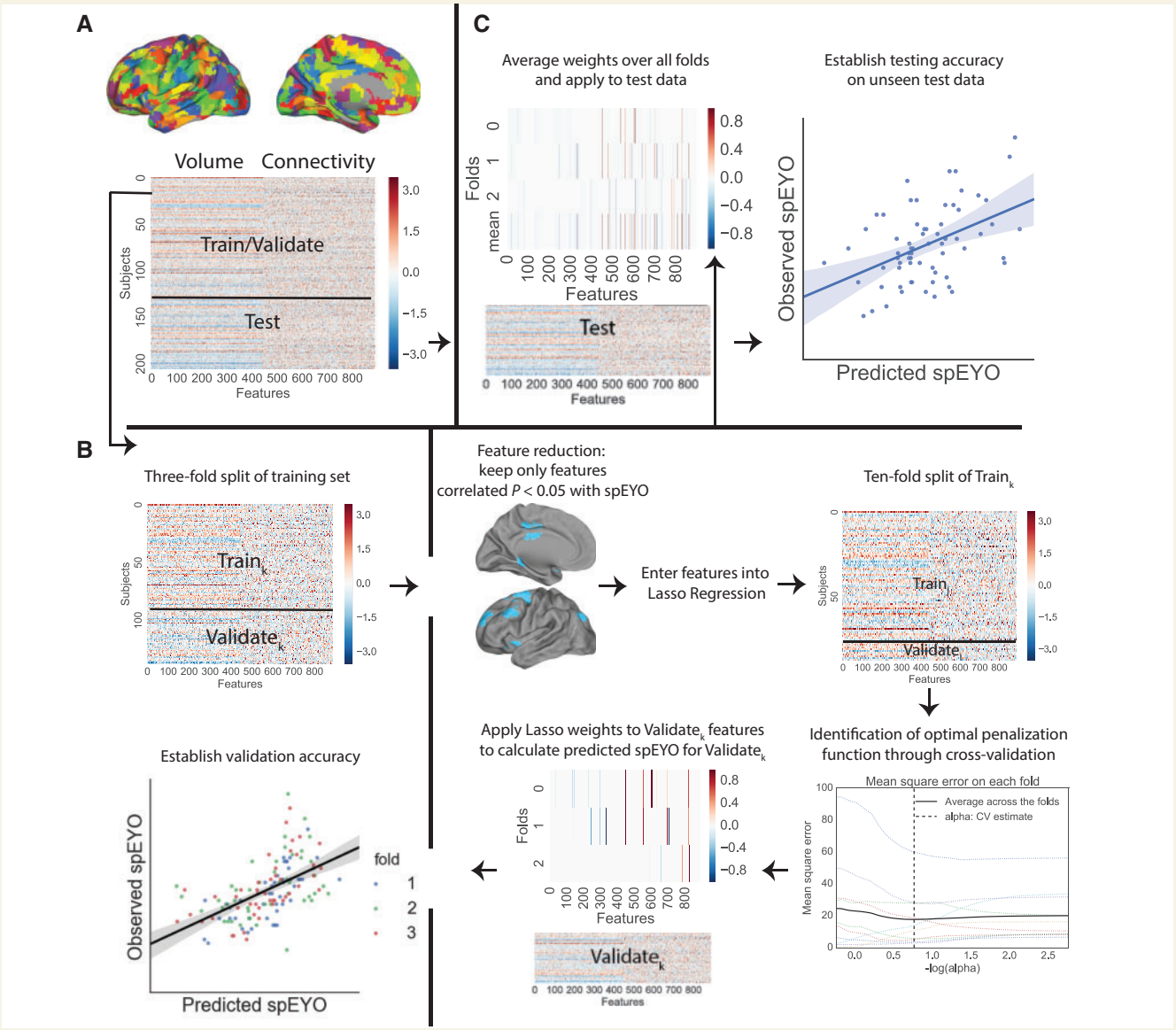

Figure 1 Nested cross-validation pipeline for Lasso Regression-based model optimization. (**A**) Grey matter volume and whole-brain resting state functional connectivity features were summarized using a 443-region cortical atlas (plus total intracranial volume), for a total of 887 brain features. One-third of the sample ($n = 68$) was removed from the pipeline and reserved as unseen testing data (presented in **C**). The other two-thirds ($n = 138$) was entered into the nested cross-validation loop presented in **B**. (**B**) The training set (only) underwent 3-fold cross-validation. For each fold, two-thirds of the sample was extracted and imaging features were filtered such that only those significantly related to years to estimated symptom onset ($P < 0.05$ uncorrected) were retained. These features were entered into a Lasso Regression, which itself used 10-fold cross-validation to define the optimal penalization. The betas (weights) from the Lasso Regression model were applied to the left-out third of the training set to transform them into a single vector of predicted years to estimated symptom onset values. The predicted vectors for each of the $n = 3$ folds were appended together so predicted values were present for all subjects within the training set. Validation accuracy is reported as the r^2 of the relationship between predicted years to estimated symptom onset and observed years to estimated symptom onset. (**C**) Finally, the Lasso Regression weights for all three folds were averaged, and these weights were applied to the unseen test data to generate predicted years to estimated symptom onset in the test dataset. Final accuracy is reported as the r^2 of the relationship between predicted years to estimated symptom onset and observed years to estimated symptom onset in the test dataset. spEYO = years to estimated symptom onset (sporadic Alzheimer's disease).

regression using functional connectivity and grey matter volume to predict years to estimated symptom onset in the PREVENT-AD cohort.

First, PREVENT-AD subjects were split such that two-thirds were used as a training set ($n = 138$) and one-third was kept completely separate as a naïve testing set ($n = 68$). The PREVENT-AD training set was used to select and validate a model through nested cross-validation, which was then applied to the PREVENT-AD testing data. The training and testing sets were balanced so that no significant difference in years to estimated symptom onset or any demographic variable was observable between the two groups (Table 1).

For exploratory purposes, in the training sample only, grey matter volume features significantly related to years to estimated symptom onset are reported, adjusting for age, sex and total intracranial volume, and connectivity features significantly related to years to estimated symptom onset are reported adjusting for age, sex and mean framewise displacement. We examined qualitative relationships between the regions related to years to estimated symptom onset in the PREVENT AD cohort, and regions commonly compromised in Alzheimer's disease. For resting state functional MRI, we compared the results to a meta-analysis map downloaded from Neurosynth using the term 'Alzheimer's disease' (211 studies) with forward inference and a false discovery rate cut-off of Q = 0.01 (Yarkoni *et al.*, 2011; Ashar *et al.*, 2016). For structural features, we compared regions related to years to estimated symptom onset to regions found to be significantly reduced in ADNI normal controls compared to Alzheimer's disease patients, using a *t*-test [$P_{(corrected)} < 0.05$]. Finally, univariate associations between years to estimated symptom onset and several demographic, cognitive and imaging measures are reported, calculated using correlations and *t*-tests.

Our predictive model was defined through a 3-fold nested cross-validation (Fig. 1). Within each fold, models were selected as follows: (i) features were reduced based on having a correlation with years to estimated symptom onset ($P < 0.05$ uncorrected; Emerson *et al.*, 2017); and (ii) these features were normalized to 0 mean and 1 standard deviation (SD) and entered into a Lasso Regression, which used 10-fold cross-validation to tune the penalization parameter. The weights and intercept produced by each model were stored, and were used to predict years to estimated symptom onset for the left-out subjects for the given fold. The predicted values from each fold were appended together, and validation accuracy was calculated as the r^2 between observed years to estimated symptom onset and predicted years to estimated symptom onset from the 3-fold cross-validated Lasso Regression (Fig. 1). Models were trained over each connectivity density threshold (see above), with and without the ($P < 0.05$) feature reduction step. Validation accuracy was used to determine that the features selection step should be included, and that the 98% threshold was the best link density threshold for the creation of functional connectivity features. Confidence intervals for this analysis were determined using 10-fold repeated cross-validation (Supplementary Fig. 1). The nested k-fold predictive model was trained once more with the 98% threshold with features selection for each fold, and a single consensus model was generated averaging the weights across all three folds. This model was applied to the features of the unseen testing sample (normalized to the training sample) by taking the dot product between the individual feature vector and the vector of model weights, plus the model intercept, to predict years to estimated symptom onset in this test group. Testing accuracy was calculated as the r^2 between observed years to estimated symptom onset and predicted years to estimated symptom onset. Confidence intervals (CI, 95%) for the prediction r^2 values were obtained for the test sample using bootstrapping (1000 samples with replacement).

To test the generalizability of this model further, weights determined in the PREVENT-AD cohort were applied to normalized baseline connectivity and grey matter volume features of off-site data from the ADNI cohort. The model trained on time to estimated symptom onset in the PREVENT-AD was used to predict time to conversion in ADNI subjects who eventually converted to mild cognitive impairment or dementia. We standardized the model features within (all 149) ADNI patients, and calculated predicted years to symptom onset values for each ADNI individual as the dot product between the individual ADNI feature vector and the vector of weights from the PREVENT-AD model, plus the model intercept. Note that, because of standardization within ADNI, we are effectively using the pattern of expression from the PREVENT-AD model to predict time to progression in ADNI. We report the r^2 between the predicted and observed values, representing the degree to which pattern expression of the model predicts years to clinical progression. Due to the small sample size, the *P*-value for this analysis was calculated using permutation testing (10 000 samples without replacement). For the purposes of comparison, we also report the effect sizes of other cognitive and biomarker data on years to clinical conversion.

All statistics were performed using the numpy, scipy and scikit-learn (Pedregosa *et al.*, 2012) packages of Python version 3.5.2.

Controlling for age

Given its strong association with both years to estimated symptom onset and imaging features, age represents an important potential confound to this analysis. To ensure the feature weights identified in the previous analysis were not simply an imaging proxy for age, we performed two separate analyses. First, previous analyses predicting years to estimated symptom onset in PREVENT-AD and time to conversion in ADNI were adjusted for age. Second, age, sex and either total intracranial volume (grey matter volume) or mean framewise displacement (connectivity) were regressed out of the imaging features in the training and test set before defining the optimal model weights with Lasso Regression. Predictive analyses were then rerun with the new weights.

Comparing predictive value of imaging features to other clinical markers

Traditionally, many cognitive, demographic and simple imaging metrics have been used for diagnostic and prognostic purposes in the clinic. To put our results into perspective, we trained a new model using only typical demographic and clinical markers (Model 2): age, sex, education, *APOE4* carriage, left and right hippocampal volume, total intracranial volume, Montreal Cognitive Assessment scores, and five cognitive composite scores. Since few of these features were associated with years to estimated symptom onset in the training sample, all features were forced into the model (i.e. the feature preselection step was skipped). Additionally, we trained a third model (Model 3), which combined the imaging features from the original model (Model 1) with all of the features from Model 2, the features from Model 2 being once again forced into the model. The r^2 of these models are reported with confidence intervals established through bootstrapping. The three models were compared directly using bootstrap testing (1000 simulations with replacement).

Table 3 Univariate associations between measures of disease progression and traditional demographic, cognitive and imaging measures

	PREVENT-AD[a]			ADNI[b]		
Variable	**Beta (SE)**	**r**	***P***	**Beta (SE)**	**r**	***P*[c]**
Age	0.51 (0.14)	0.30	<0.001	−0.01 (0.04)	−0.06	0.77
Education	0.15 (0.18)	0.07	0.41	−0.08 (0.11)	−0.15	0.45
Intracranial volume	1.79 (4.99)	0.03	0.72	0.00 (0.00)	0.13	0.53
Montreal Cognitive Assessment	−0.13 (0.05)	−0.02	0.78	0.14 (0.08)	0.33	0.098
Immediate Memory	−0.03 (0.05)	−0.04	0.60	–	–	–
Visuospatial Construction	−0.03 (0.04)	−0.05	0.54	–	–	–
Language	−0.03 (0.07)	−0.04	0.65	–	–	–
Attention	−0.06 (0.05)	−0.11	0.22	–	–	–
Delayed Memory	−0.03 (0.06)	−0.05	0.58	–	–	–
Hippocampal volume	−0.002 (0.002)	−0.08	0.33	0.00 (0.00)	0.17	0.40
Variable		**t**	***P***		**t**	***P*[c]**
Sex	–	0.29	0.77	–	0.29	0.77
APOE4 carrier	–	−2.13	0.03	–	0.15	0.88

[a]Statistics represent associations between each measure and years to estimated symptom onset in the PREVENT-AD cohort.
[b]Statistics represent associations between each measure and time to clinical progression in ADNI.
[c]*P*-values established through permutation testing because of small sample size.

Results

Associations between years to estimated symptom onset and model features

Table 3 lists univariate associations between years to estimated symptom onset and several demographic, cognitive and imaging metrics in the training sample of PREVENT-AD. Older subjects tended to be closer to or beyond their estimated symptom onset, while APOE4 carriers tended to be further from their estimated symptom onset. In ADNI, subjects with worse Montreal Cognitive Assessment scores converted marginally sooner.

After adjustment for age, sex and total intracranial volume, associations between grey matter volume and years to estimated symptom onset emerged in a specific cortical pattern within the training sample of the PREVENT-AD cohort (Fig. 2). Individuals closer to or beyond their parents' age of symptom onset showed reduced grey matter volume in medial parietal and medial frontal cortex, as well as the medial temporal lobe, thalamus, cerebellum and some scattered aspects of lateral temporal, frontal and parietal cortices. Brain regions with significantly reduced grey matter volume associated with years to estimated symptom onset also resembled regions with reduced grey matter volume in ADNI Alzheimer's disease dementia patients compared to controls (Supplementary Fig. 2). Years to estimated symptom onset was not associated with increased grey matter volume in any brain region. In contrast to grey matter volume, both positive and negative relationships emerged between years to estimated symptom and whole-brain resting state functional MRI connectivity, adjusting for age, sex and mean frame displacement (Fig. 2). Negative relationships were observed in certain regions participating in the default mode and frontoparietal networks, and positive relationships emerged mainly in the subcortex, medial temporal lobes and frontal cortex. Brain regions with significantly altered functional connectivity related to years to estimated symptom onset also resembled regions with disrupted functional connectivity in Alzheimer's disease based on a meta-analysis (Supplementary Fig. 2).

Imaging features predict years to estimated symptom onset

Lasso Regression embedded in a nested cross-validation pipeline (Fig. 1) identified a set of weighted imaging features designed to predict years to estimated symptom onset (Fig. 3A). Grey matter volume features included the posterior cingulate most prominently, but also included aspects of the lateral temporoparietal cortex and prefrontal cortex. Functional connectivity features selected included bidirectional effects across default mode, salience and limbic networks. In all, 64 features—17 grey matter volume and 47 functional connectivity features—were selected (beta > 0), and together achieved high predictive power in the PREVENT-AD validation sample [$r^2(138) = 0.256$ (95% CI: 0.18, 0.36), $P < 0.0001$]. The weighted feature set was then applied to unseen PREVENT-AD test data, and significantly predicted a sizable portion of variance in years to estimated symptom onset in this sample [$r^2(68) = 0.221$ (95% CI: 0.077, 0.392), $P < 0.0001$; Fig. 3B].

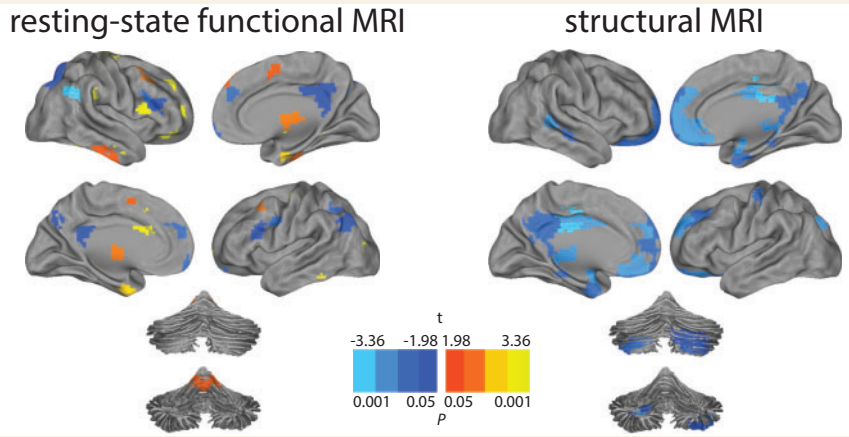

Figure 2 Relationships between brain imaging features and years to estimated symptom onset. Relationships between years to estimated symptom onset and resting state functional connectivity (*left*) and grey matter volume (*right*) features, after adjusting for age, sex and mean frame displacement (connectivity) or total intracranial volume (volume) in the PREVENT-AD training sample. Positive relationships are shown in warm colours, while negative relationships are shown in cool colours. Only significant relationships are visualized ($P < 0.05$ uncorrected).

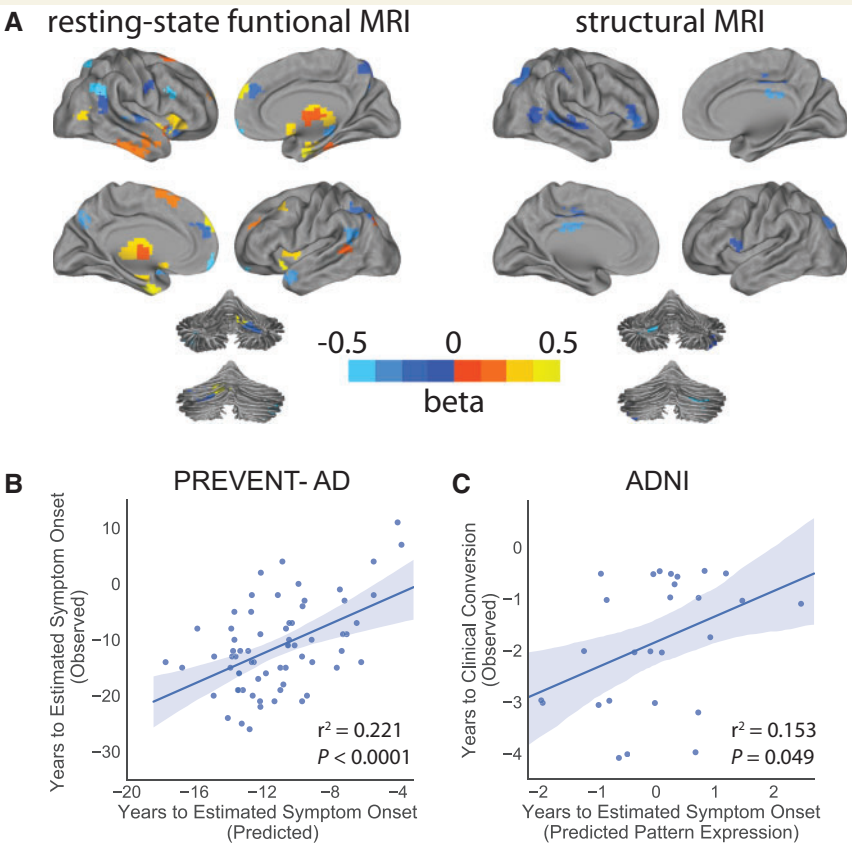

Figure 3 Neuroimaging features predict years to estimated symptom onset and time to clinical conversion. (**A**) Neuroimaging features included in the final predictive model developed in the PREVENT-AD validation group (i.e. brain region with non-zero beta values) are projected onto cortical surfaces. Whole brain resting state functional connectivity features are presented on the *left*, while grey matter volume features are on the *right*. Colours represent the direction and magnitude of beta values. Warmer colours indicate higher positive betas, while cool colours represent lower betas. Higher absolute beta values (brighter colours) indicate increased contribution of the feature in the model. Note that model weights may not accurately represent univariate relationships between predictors and years to estimated symptom onset. (**B**) The predictive model derived from the PREVENT-AD training set significantly predicted years to estimated symptom onset in an independent PREVENT-AD testing set. (**C**) The model derived from the PREVENT-AD training set was generalized to cognitively normal and mild cognitive impairment individuals followed over time from the ADNI cohort. The more the ADNI individuals expressed the pattern predicting years to estimated symptom onset, the closer they were to clinical conversion. The pattern expression was normalized with a mean centre and unit variance of patients from the ADNI cohort. Due to the small sample size, the *P*-value is estimated using permutation testing (1000 samples without replacement).

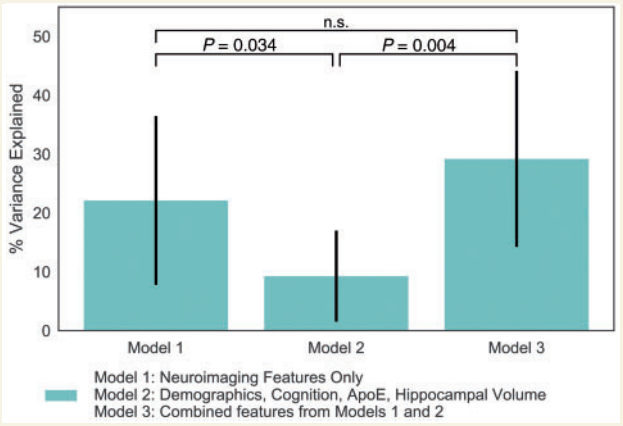

Figure 4 Head-to-head comparison of three different models trained to predict years to estimated symptom onset. Each model was evaluated on unseen PREVENT-AD test data. Bars represent explained variance ($r^2 \times 100$) of each of the three models. 95% confidence intervals were generated through bootstrapping, and bootstrap tests were used to make empirical comparisons between models. Adding neuroimaging features to traditional clinical markers significantly improved the predictive power of the model. Imaging features alone performed significantly better than traditional clinical features alone. n.s. = not significant.

Imaging features predict years to clinical conversion in a separate prospective sample

None of the PREVENT-AD subjects who were used to train the predictive model expressed symptomatic Alzheimer's disease, and the symptom onset was therefore estimated but not known. To validate our estimate, and to establish the generalizability of the model, we applied the model to imaging features of 26 cognitively normal and mild cognitive impairment subjects from an independent dataset, with full resting state functional MRI and structural MRI data, who eventually converted to mild cognitive impairment or dementia, respectively. The model trained on years to estimated symptom onset significantly predicted years to clinical conversion in this group [$r^2(26) = 0.153$ (95% CI: 0.024, 0.360), $P = 0.049$; Fig. 3C]. This association remained significant when only including subjects with mild cognitive impairment [$r^2(21) = 0.201$ (95% CI: 0.028, 0.516), $P = 0.040$]. Specifically, subjects predicted to be closer to or beyond their expected symptom onset tended to convert sooner. This indicates that the model trained to predict years to estimated symptom onset in the PREVENT-AD cohort provides meaningful information to predict actual time to clinical conversion in a subgroup of patients from the ADNI cohort. Therefore, the model may represent a general model of individual progression towards dementia.

Imaging features predict years to estimated symptom onset independently of age

As both imaging features and years to estimated symptom onset are associated with age, a strong relationship between age and the predictive features identified in the previous analysis was expected. This association was indeed present in the PREVENT-AD testing sample [$r^2(68) = 0.140$, $P = 0.0018$] and the group of ADNI converters [$r^2(26) = 0.178$, $P = 0.029$]. Two analyses were undertaken to assess whether the predictive quality of our model was driven by age, summarized in Supplementary Table 1. First, our model significantly predicted years to estimated symptom onset in PREVENT-AD and years to clinical conversion in ADNI, even when adjusting for age. Additionally, new predictive models and feature sets were generated adjusting for age and sex. These models once again significantly predicted years to estimated symptom onset in the PREVENT-AD and years to clinical conversion in ADNI. Effect sizes of these results varied in PREVENT-AD, but not in ADNI.

Multi-modal imaging features improve predictive models using traditional measures

We next compared our model using only multimodal neuroimaging features to a model composed of features used in a traditional clinical setting (Fig. 4). A new model ('Model 2') was trained to predict years to estimated symptom onset in the PREVENT-AD cohort using only traditional demographic, cognitive and imaging measures. This model also predicted years to estimated symptom onset in the unseen test data [$r^2 = 0.093$ (95% CI: 1.87, 17.34), $P = 0.012$]. This model however performed worse ($P = 0.034$) than the model using only neuroimaging features (Model 1), explaining less than half the variance. Finally, a third model (Model 3) was trained combining features from both Models 1 and 2, which again predicted years to estimated symptom onset in unseen data [$r^2 = 0.29$ (95% CI: 0.13, 0.47), $P < 0.0001$]. Model 3 outperformed Model 2 ($P = 0.004$), but not Model 1 ($P = 0.24$). This suggests that adding multimodal high resolution neuroimaging information to models using only traditional markers significantly enhances prediction of preclinical disease progression.

Discussion

In a sample of cognitively unimpaired older individuals with a parental history of Alzheimer's disease-like dementia, we evaluated the construct of estimated years to onset of sporadic Alzheimer's disease. Calculating this metric as the difference between a given participant's current age and the onset age of his/her earliest affected parent (Bateman *et al.*, 2012),

we first identified its relationship to alterations in resting state functional MRI and volumetric structural MRI in a portion of the participant sample. We then used a machine learning-based regression technique (Lasso regression) to construct a model in this training sample that successfully predicted estimated years to onset in a previously unseen portion of the same dataset. Finally, we cross-validated this measure by testing the capacity of the resulting model to predict time to diagnostic progression (conversion to mild cognitive impairment or dementia) in ADNI.

Characterization and evaluation of the presymptomatic stage of Alzheimer's disease is necessary for the development of preventive interventions. The study of presymptomatic sporadic Alzheimer's disease is challenging because, in the absence of symptoms, one cannot be certain whether or when a given individual will develop dementia. Our study identified regionally specific reductions in grey matter volume and altered connectivity as key features that correlated with years to estimated symptom onset. These findings were reminiscent of observations in mutation carriers from families harbouring autosomal dominant Alzheimer's disease (Bateman *et al.*, 2012; Chhatwal *et al.*, 2012; Benzinger *et al.*, 2013). Our study further shows that a single model based on neuroimaging data can predict both an estimate of time to symptom onset in asymptomatic individuals, and actual time to symptom onset in individuals that progressed to mild cognitive impairment and/or Alzheimer's disease dementia, greater than chance. These results suggest that the years to estimated symptom onset measurement contains useful clinical information that may help predict degree and temporal progression of preclinical brain changes in sporadic Alzheimer's disease.

Associations between years to estimated symptom onset and brain function and structure were found in Alzheimer's disease-vulnerable regions. Associations with grey matter volume were found predominantly in the medial temporal lobe and in regions within the default mode network. These regions are characteristically compromised in Alzheimer's disease, even in its preclinical stages (Buckner *et al.*, 2005; Dickerson *et al.*, 2011; Grothe and Teipel, 2016), and resembled grey matter differences between healthy older controls and Alzheimer's disease patients in ADNI. Furthermore, individuals approaching or exceeding their parental symptom onset showed reduced connectivity between certain frontoparietal or default mode network regions and the rest of the brain. By contrast, we observed increased connectivity between medial temporal lobe or salience network regions and the rest of the brain. Abnormal connectivity in these networks has been noted previously by others (Greicius *et al.*, 2004; Sheline *et al.*, 2010; Chhatwal *et al.*, 2012; Badhwar *et al.*, 2017), and the pattern of alterations resembled that of a meta-analysis of resting state functional connectivity disruptions in Alzheimer's disease (Supplementary Fig. 1). The association between years to estimated symptom onset and hippocampal hyperconnectivity is interesting in view of growing evidence that increased hippocampal activity is a prominent feature of preclinical Alzheimer's disease (O'Brien *et al.*, 2010; Bakker *et al.*, 2012; Mormino *et al.*, 2012; Leal *et al.*, 2017), possibly representing a compensatory reaction to deficits elsewhere.

Our work builds on a growing body of literature suggesting that, despite its simplicity, the years to estimated symptom onset measure may contain important information with respect to preclinical trajectories. Accumulation of amyloid-β in the cerebral cortex is an early and specific marker of Alzheimer's disease, and years to estimated symptom onset has been associated with abnormality in amyloid biomarkers across three cohorts of individuals with a family history of Alzheimer's disease, including the present sample (Villeneuve *et al.*, 2018). Similarly, years to estimated symptom onset has been associated with decline in cognition in another, similar cohort (Ritchie *et al.*, 2017). In the current study we show that MRI abnormalities are not only associated with, but can predict years to estimated symptom onset in asymptomatic individuals. We further showed that these same imaging features can predict time to clinical progression in an independent cohort of individuals with Alzheimer's disease. While there may be other brain alterations dynamically associated with clinical progression, at least some of the regions identified in our predictive model appear to be stable across the spectrum of Alzheimer's disease.

The last decade has witnessed substantial progress towards prediction of progression to Alzheimer's disease. Several studies have used neuroimaging data to achieve considerable accuracy toward identifying individuals who will develop dementia within a few years (Mathotaarachchi *et al.*, 2017). Predicting time to symptom onset in sporadic Alzheimer's disease is a much more complex task, and considerably less work has been published on this topic (Oulhaj *et al.*, 2009; Jack *et al.*, 2010; Zhang and Shen, 2012; Ten Kate *et al.*, 2017). At this stage, the performance of our model remains far from the accuracy necessary for clinical utility. Nonetheless, its ability to predict $>20\%$ of the variance in years to estimated symptom onset using only neuroimaging data suggests substantial utility for research of preclinical Alzheimer's disease. While we show that traditional clinical markers can provide reliable information with respect to preclinical disease trajectories, we also provide strong evidence that such models are substantially improved by adding multi-modal neuroimaging measurements, which provide independent relevant information.

In contrast with previous studies (Jack *et al.*, 2010; Ten Kate *et al.*, 2017), we did not find relationships between whole hippocampus volume and years to estimated symptom onset or years to clinical conversion, and hippocampal volume was not selected as a feature in our predictive models. We note that the sample sizes of clinical progressors in these previous studies are four to five times larger than the sample of ADNI progressors in the current study, and were therefore better powered to detect smaller effects. We also note that, while whole hippocampus volume was not associated with years to parental symptom onset in the PREVENT-AD sample, some aspects of the posterior

hippocampus were associated with years to estimated symptom onset in the PREVENT-AD training sample (Fig. 2). It is possible that these features were not selected by our model due to high collinearity with other selected features. Overall, hippocampal volume should not be discounted as an important structural feature of Alzheimer's disease, but greater value should be placed on multimodal, combinatorial biomarkers and models in future studies.

Our study has several important limitations, the most important being its reliance on a crude estimation of years to estimated symptom onset. Many PREVENT-AD participants will never develop Alzheimer's disease dementia. Their presence in the sample is likely to dilute inference available from the work. Prediction accuracy in future studies may be improved by selecting only participants with additional Alzheimer's disease risk factors. In addition, many factors, such as ageing of the population in general (Satizabal *et al.*, 2016), can introduce complexity into the estimation of expected onset. Bias may also be present in the subjective reports of parent onset age, which could be influenced by anxiety over Alzheimer's risk or metacognition. In addition, some important factors that may contribute to both neuroimaging metrics and onset timelines, such as premorbid IQ, are not collected in the PREVENT-AD battery. It is also notable that autopsy confirmation of Alzheimer's disease pathology was not available for the parents of some PREVENT-AD members, or any ADNI individuals. In this sense, we cannot be sure whether we are predicting the onset of Alzheimer's disease symptoms, or the onset of another form of dementia (Ten Kate *et al.*, 2017). Finally, our findings in PREVENT-AD, while mainly cross-sectional at present, could be made substantially more informative with the addition of more extended longitudinal observations. Future studies and the use of additional cohorts would provide an opportunity to corroborate and extend our findings, preferably with additional disease markers.

Acknowledgements

The authors would like to acknowledge Yasser Iturria-Medina, Christian Dansereau, Hazal Özlen, Pierre-Olivier Quiron, Yassine Benhajali, Sebastian Urchs, Cecile Madjar and Jennifer Tremblay-Mercier for technical and administrative support, commentary and advice. We would also like to acknowledge the participants of the PREVENT-AD cohort for dedicating their time and energy to helping us collect these data.

Funding

This work was supported by a Canada Research Chair, a Canadian Institutes of Health Research foundation grant, a Canada Foundation for Innovation, and a joint Alzheimer Society of Canada and a Foundation Brain Canada Research grant to S.V., a Canada Research Chair to J.C.B., and a Vanier Canada Graduate Student Doctoral award to J.W.V. The PREVENT-AD cohort was funded by generous support from McGill University, the government of Canada, an unrestricted gift from Pfizer Canada, the Canada Fund for Innovation, the Douglas Hospital Research Centre, the Levesque Foundation, McGill University and Genome Quebec Innovation Center. 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, Alzheimer's 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; IXICO Ltd.; 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.

Supplementary material

Supplementary material is available at *Brain* online.

References

Ashar YK, Andrews-Hanna JR, Dimidjian S, Wager TD. Empathic care and distress: predictive brain markers and dissociable brain systems. Neuron 2016; 94: 1263–73.e4.

Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage 2007; 38: 95–113.

Badhwar A, Tam A, Dansereau C, Orban P, Hoffstaedter F, Bellec P. Resting-state network dysfunction in Alzheimer's disease: a systematic review and meta-analysis. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 2017; 8: 73–85.

Bakker A, Krauss GL, Albert MS, Speck CL, Jones LR, Stark CE, et al. Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. Neuron 2012; 74: 467–74.

Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012; 367: 795–804.

Benzinger TL, Blazey T, Jack CR, Koeppe RA, Su Y, Xiong C, et al. Regional variability of imaging biomarkers in autosomal dominant Alzheimer's disease. Proc Natl Acad Sci USA 2013; 110: E4502–9.

Breitner JCS, Poirier J, Etienne PE, Leoutsakos JM. Rationale and structure for a new center for studies on prevention of Alzheimer's disease (StoP-AD). J Prev Alzheimers Dis 2016; 3: 236–42.

Brier MR, Thomas JB, Fagan AM, Hassenstab J, Holtzman DM, Benzinger TL, et al. Functional connectivity and graph theory in preclinical Alzheimer's disease. Neurobiol Aging 2014; 35: 757–68.

Buckner RL, Snyder AZ, Shannon BJ, Larossa G, Sachs R, Fotenos AF, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. Neuroscience 2005; 25: 7709–17.

Bullmore ET, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 2009; 10: 186–98.

Chhatwal JP, Schultz AP, Johnson KA, Benzinger TLS, Jack CR, Salloway S, et al. Impaired default network functional connectivity in autosomal dominant Alzheimer's disease: findings from the DIAN study. Ann Neurol 2012; 72: S40.

Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. Neuroimage 2011; 54: 940–54.

Dickerson BC, Stoub TR, Shah RC, Sperling RA, Killiany RJ, Albert MS, et al. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. Neurology 2011; 76: 1395–402.

Dickson MR, Li J, Wiener HW, Perry RT, Blacker D, Bassett SS, et al. A genomic scan for age at onset of Alzheimer's disease in 437 families from the NIMH Genetic Initiative. Am J Med Genet Part B Neuropsychiatr Genet 2008; 147B: 784–92.

Emerson RW, Adams C, Nishino T, Hazlett HC, Wolff JJ, Zwaigenbaum L, et al. Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age. Sci Transl Med 2017; 9: eaag2882. doi: 10.1126/scitranslmed.aag2882.

Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Arch Gen Psychiatry 2006; 63: 168–74.

Gold JM, Queern C, Iannone VN, Buchanan RW. Repeatable battery for the assessment of neuropsychological status as a screening test. Am J Psychiatry 1999; 156: 1944–50.

Greicius MD, Srivastava G, Reiss AL, Menon V. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. Proc Natl Acad Sci USA 2004; 101: 4637–42.

Grothe MJ, Teipel SJ. Spatial patterns of atrophy, hypometabolism, and amyloid deposition in Alzheimer's disease correspond to dissociable functional brain networks. Hum Brain Mapp 2016; 37: 35–53.

Iturria-Medina Y, Sotero RC, Toussaint PJ, Mateos-Perez JM, Evans AC; Alzheimer's Disease Neuroimaging Initiative. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. Nat Commun 2016; 7: 11934.

Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008; 27: 685–91.

Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 2013; 12: 207–16.

Jack CR, Wiste HJ, Vemuri P, Weigand SD, Senjem ML, Zeng G, et al. Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. Brain 2010; 133: 3336–48.

Kamboh MI, Barmada MM, Demirci FY, Minster RL, Carrasquillo MM, Pankratz VS, et al. Genome-wide association analysis of age-at-onset in Alzheimer's disease. Mol Psychiatry 2012; 17: 1340–6.

Leal SL, Landau SM, Bell RK, Jagust WJ. Hippocampal activation is associated with longitudinal amyloid accumulation and cognitive decline. Elife 2017; 6: e22978.

Li YJ, Scott WK, Hedges DJ, Zhang F, Gaskell PC, Nance MA, et al. Age at onset in two common neurodegenerative diseases is genetically controlled. Am J Hum Genet 2002; 70: 985–93.

Mathotaarachchi S, Pascoal TA, Shin M, Benedet AL, Kang MS, Beaudry T, et al. Identifying incipient dementia individuals using machine learning and amyloid imaging. Neurobiol Aging 2017; 59: 80–90.

Mendez MF. Early-onset Alzheimer's disease: nonamnestic subtypes and type 2 AD. Arch Med Res 2013; 43: 677–85.

Mormino EC, Brandel MG, Madison CM, Marks S, Baker SL, Jagust WJ. Aβ Deposition in aging is associated with increases in brain activation during successful memory encoding. Cereb Cortex 2012; 22: 1813–23.

Naj AC, Jun G, Reitz C, Kunkle BW, Perry W, Park YS, et al. Effects of multiple genetic loci on age at onset in late-onset Alzheimer disease: a genome-wide association study. JAMA Neurol 2014; 71: 1394–404.

Nasreddine Z, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Colllin I, et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005; 53: 695–9.

O'Brien JL, O'Keefe KM, Laviolette PS, Deluca AN, Blacker D, Dickerson BC, et al. Longitudinal fMRI in elderly reveals loss of hippocampal activation with clinical decline. Neurology 2010; 74: 1969–76.

Orban P, Madjar C, Savard M, Dansereau C, Tam A, Das S, et al. Test-retest resting-state fMRI in healthy elderly persons with a family history of Alzheimer's disease. Sci Data 2015; 2: 150043.

Oulhaj A, Wilcock GK, Smith AD, De Jager CA. Predicting the time of conversion to MCI in the elderly: role of verbal expression and learning. Neurology 2009; 73: 1436–42.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2012; 12: 2825–30.

Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A. Mapping scores onto stages: mini-mental state examination and clinical dementia rating. Am J Geriatr Psychiatry 2006; 14: 139–44.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 2012; 59: 2142–54.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, et al. Functional network organization of the human brain. Neuron 2011; 72: 665–78.

Power JD, Schlaggar BL, Lessov-Schlaggar CN, Petersen SE. Evidence for hubs in human functional brain networks. Neuron 2013; 79: 798–813.

Redolfi A, Manset D, Barkhof F, Wahlund LO, Glatard T, Mangin JF, et al. Head-to-head comparison of two popular cortical thickness extraction algorithms: a cross-sectional and longitudinal study. PLoS One 2015; 10: e0117692.

Ritchie K, Carriére I, Su L, O'Brien JT, Lovestone S, Wells K, et al. The midlife cognitive profiles of adults at high risk of late-onset Alzheimer's disease: the PREVENT study. Alzheimers Dement 2017; 13: 1089–97.

Ryman D, Morris J, Bateman R. Predicting symptom onset in autosomal dominant Alzheimer's disease: a systematic review and meta-analysis. Alzheimers Dement 2014; 10: P218–19.

Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of dementia over three decades in the Framingham heart study. N Engl J Med 2016; 374: 523–32.

Schwarz CG, Gunter JL, Wiste HJ, Przybelski SA, Weigand SD, Ward CP, et al. Large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. Neuroimage Clin 2016; 11: 802–12.

Sheline YI, Raichle ME, Snyder AZ, Morris JC, Head D, Wang S, et al. Amyloid plaques disrupt resting state default mode network connectivity in cognitively normal elderly. Biol Psychiatry 2010; 67: 584–7.

Sperling RA, Mormino E, Johnson K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. Neuron 2014; 84: 608–22.

Tavor I, Jones OP, Mars RB, Smith SM, Behrens TE, Jbabdi S. Task-free MRI predicts individual differences in brain activity during task performance. Science 2016; 352: 216–20.

Ten Kate M, Barkhof F, Visser PJ, Teunissen CE, Scheltens P, Van Der Flier WM, et al. Amyloid-independent atrophy patterns predict time to progression to dementia in mild cognitive impairment. Alzheimers Res Ther 2017; 9: 73.

Tschanz JT, Norton MC, Zandi PP, Lyketsos CG. The Cache County study on memory in aging: factors affecting risk of Alzheimer's disease and its progression after onset. Int Rev Psychiatry 2013; 25: 673–85.

Urchs S, Armoza J, Benhajali Y, St-Aubin J, Orban P, Bellec P. MIST: A multi-resolution parcellation of functional brain networks. MNI Open Research 2017; 1: 3.

Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol 2013; 12: 357–67.

Villeneuve S, Vogel J, Gonneaud J, Pichette Binette A, Rosa-Neto P, Gauthier S, et al. Proximity to parental symptom onset and amyloid-β Burden in Sporadic Alzheimer's disease. JAMA Neurol 2018, in press. doi: 10.1001/jamaneurol.2017.5135.

Wingo TS, Lah JJ, Levey AI, Cutler DJ. Autosomal recessive causes likely in early-onset Alzheimer disease. Arch Neurol 2012; 69: 59–64.

Yarkoni T, Poldrack R, Nichols T. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 2011; 8: 665–70.

Youssofzadeh V, Mcguinness B, Maguire LP, Wong-lin K. Multi-kernel learning with dartel improves combined MRI-PET classification of Alzheimer's disease in AIBL data: group and individual analyses. Front Hum Neurosci 2017; 11: 380.

Zhang D, Shen D. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS One 2012; 7: e33182.