Alzheimer's & Dementia 12 (2016) 977-986

Featured Article

Amyloid status imputed from a multimodal classifier including structural MRI distinguishes progressors from nonprogressors in a mild Alzheimer's disease clinical trial cohort

Duygu Tosun[a,*], Yun-Fei Chen[b], Peng Yu[b], Karen L. Sundell[b], Joyce Suhy[c], Eric Siemers[b], Adam J. Schwarz[b], Michael W. Weiner[a], for the Alzheimer's Disease Neuroimaging Initiative

[a]*Department Radiology and Biomedical Imaging, University of California—San Francisco, San Francisco, CA, USA*
[b]*Eli Lilly and Company, Indianapolis, IN, USA*
[c]*BioClinica, CA, USA*

Abstract

Introduction: Mild-Alzheimer's disease (AD) subjects without significant Aβ pathology represent a confounding finding for clinical trials because they may not progress clinically on the expected trajectory, adding variance into analyses where slowing of progression is being measured.
Methods: A prediction model based on structural magnetic resonance imaging (MRI) in combination with baseline demographics and clinical measurements was used to impute Aβ status of a placebo-treated mild-AD sub-cohort (N = 385) of patients participating in global phase 3 trials. The clinical trajectories of this cohort were evaluated over 18 months duration of the trial, stratified by imputed Aβ status within a mixed-model repeated measures statistical framework.
Results: In the imputed Aβ-positive cohort, both cognitive (ADAS-Cog_{14} and MMSE) and functional (ADCS-iADL) measures declined more rapidly than in the undifferentiated population.
Discussion: Our results demonstrate imputing Aβ status from MRI scans in mild-AD subjects may be a useful screening tool in global clinical trials if amyloid measurement is not available.

Keywords: Alzheimer's disease; Aβ pathology; Aβ-positivity; Structural MRI; Florbetapir PET; ADAS-Cog14; MMSE; ADCS-iADL; Clinical decline

1. Introduction

The presence of brain β-amyloid (Aβ) plaques is a defining feature of Alzheimer's disease (AD) and an integral part of recent diagnostic guidelines for clinical research into the disorder [1]. Both positron emission tomography (PET) imaging, using Aβ-specific radiotracers such as [^{11}C]-PIB or [^{18}F]-florbetapir, and measurement of Aβ proteins from cerebral spinal fluid (CSF) samples, are widely used in the research setting to quantify brain Aβ plaque load. Cut points for positivity have been defined for each of these modalities, yielding highly convergent results [2,3]. Analysis of Aβ biomarker subgroups in phase 3 clinical trials from two independent drug development programs revealed that approximately 27% of subjects meeting clinical inclusion criteria for mild-AD were Aβ-negative [4,5]. The presence of Aβ-negative subjects in trial cohorts represents a potential confound for two reasons. First, for putative treatments targeting Aβ pathology, the therapeutic target is not present and hence their clinical trajectories may not be modified in the same way as Aβ-positive subjects. Second, subjects without core AD pathology may not progress clinically on the same trajectory as those that do, even in the absence of treatment. Both these factors may introduce

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.

*Corresponding author. Tel.: +1 415 221 4810x3650.
E-mail address: duygu.tosun@ucsf.edu

http://dx.doi.org/10.1016/j.jalz.2016.03.009

variability and dilute a treatment signal in analyses where a slowing of clinical progression is hypothesized.

Clinical trials of putative therapeutics for AD are increasingly using a baseline measure of brain Aβ as an inclusion criterion. However, PET and CSF methods can be challenging to implement in global clinical trial sites outside the western hemisphere, for reasons including patient acceptance, cost, and availability. In contrast, magnetic resonance imaging (MRI) is more widely available and typically already included in clinical trial protocols for radiologic monitoring [6]. Recent analytical developments have demonstrated that Aβ status can be predicted to a high accuracy in both early and amnestic mild cognitive impairment (MCI) subjects from Alzheimer's Disease Neuroimaging Initiative (ADNI) using a macroscopic pattern of brain structural deformation obtained from structural MRI data [7,8]. Our aim in the present work was to extend this approach along the AD continuum by developing an MRI-based Aβ-positive versus negative classifier applicable to mild-AD subjects and to demonstrate its applicability to structural MRI data obtained from global clinical trials.

We assessed if Aβ status imputed from structural MRIs could distinguish clinically progressing mild-AD subjects from nonprogressors in the placebo-treated sub-cohort of global clinical phase 3 trials. We hypothesized that the clinical trajectories of the subjects imputed Aβ-positive would worsen faster, and those subjects imputed Aβ-negative would worsen substantially slower than the full data set.

2. Methods

2.1. Population and clinical instruments

The primary data set for this study was drawn from the mild-AD subset (Mini-Mental State Examination [MMSE] score between 20 and 26, inclusive, at baseline) from two phase 3 trials EXPEDITION and EXPEDITION 2 (hereafter referred to as the EXPEDITION trials). All subjects in the EXPEDITION trials received a structural MRI scan at baseline.

Classifier development and training were performed using baseline data from the 194 mild-AD subjects from the Aβ PET sub-study of the EXPEDITION trials who had a baseline florbetapir-PET scan (both placebo and treatment arms). The classifier was then applied prospectively to impute Aβ status of N = 385 placebo-treated mild-AD subjects from the EXPEDITION trials who did not have a florbetapir-PET scan at baseline.

An additional 75 cognitively normal (CN) elderly individuals from the ADNI study were included to model normal confounding effects of age, gender, and education. CN subjects had MMSE scores between 24 and 30, a clinical dementia rating (CDR) of 0, no evidence of depression, and no memory complaints. Previous studies reported that the cumulative and regional Aβ burden in CN subjects correlate with regional brain atrophy [9–11]. Furthermore, CN subjects at genetic risk for AD by virtue of the apolipoprotein E (*APOE*) ε4-allele demonstrated regional brain atrophy differences relative to *APOE* ε4-noncarriers [12,13]. Therefore, we included only the *APOE* ε4-noncarrier ADNI CN subjects who were identified as Aβ-negative using PiB-PET and had structural MRI data from 1.5-Tesla scanners. A PiB cutoff of 1.47 from the study of Jagust et al. [14] was applied to identify PiB-negative ADNI CN subjects. Consistency in the positive-negative categorization using PiB and florbetapir radioligands was previously reported [15].

2.2. Instruments for assessing clinical progression

The primary hypothesis of this study was tested on three clinical instruments: MMSE and Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog_{14}) were used to assess global cognitive abilities, and the Alzheimer's Disease Cooperative Study–instrumental Activities of Daily Living Inventory (ADCS-iADL) scale was used to evaluate global function. ADAS-Cog_{14} and ADCS-iADL were assessed at baseline and weeks 12, 28, 40, 52, 64 and 80, and MMSE was assessed at baseline and weeks 28, 52 and 80 from baseline.

2.3. Image data acquisition

Structural MRI scans from the EXPEDITION trials were acquired using a uniform scanning protocol that minimized and accounted for between-site differences in MRI systems. Data from 1.5 Tesla scanners manufactured by Siemens (Symphony, Espree or Avanto), General Electric (Excite), or Philips (Intera or Achieva) were used in the present analysis. A total of 184 1.5 T imaging sites across 16 countries were qualified, of which 44% were in North America, 24% in Europe, 19% in Asia and Russia, and 13% in South America. Data from 3.0 T scanners were excluded as there were insufficient subjects to train and test a classifier (Further details in Supplementary Material).

Structural MRI scans from the ADNI study were downloaded from the official study site: http://adni.loni.usc.edu/. Participants underwent a standardized 1.5 Tesla MRI protocol (http://adni.loni.usc.edu/methods/mri-analysis/mri-acquisition/) (Further details in Supplementary Material).

2.4. Florbetapir-PET processing

N = 194 subjects within the EXPEDITION trials were part of the Aβ PET sub-study and received a florbetapir-PET scan at baseline. The florbetapir-PET scan comprised a 20 minute acquisition period beginning 50–60 minutes following 10 mCi injection of the florbetapir tracer and the images were processed as detailed previously [16,17] and a composite standardized uptake value ratio (SUVR) value calculated using a set of pre-specified cortical regions of interest and the whole cerebellum as the reference region (Further details in Supplementary Material). Aβ-positivity or negativity was then determined using a threshold of 1.1 [16].

2.5. MRI processing

Skull, scalp, and extra-cranial tissue were removed from each structural MRI using the automated Brain Surface Extraction software [18], followed by manual refinement if required. To avoid bias towards a particular subject's geometry in the analysis of anatomical shape variations, we used the data from 75 ADNI CNs to create a study-specific unbiased large deformation brain image template (ULD-template) by applying a framework of large deformation diffeomorphic metric mapping (LDDMM) as described in full elsewhere [19]. ULD-template generation incorporated an unbiased approach where all brain images were first simultaneously affine transformed to adjust for global variations in brain positioning and scale, and then simultaneously deformed. Structural-MRI of each mild-AD subject was first affine aligned and then nonlinearly warped to this ULD-template using the LDDMM framework. The LDDMM was modeled as an evolution in time, with an associated smooth velocity vector field controlling this evolution. A scalar initial momentum map parameterized the entire geodesic with which the optimal trajectory emanated from the ULD-template to reach a subject brain image on a Riemannian manifold of diffeomorphism [19]. These momentum maps uniquely encoded the anatomical shape variations of individual brains relative to the ULD-template. The details of this processing framework is previously described in detail elsewhere [19,20], and are provided in the Supplementary Material.

We used a general linear model–based detrending method to control for any normal confounding effects of age, sex, and education, based on PiB-negative and *APOE* ε3/ε3 CN subjects. Adjusted imaging measures of anatomical shape variation from mild-AD subjects were used for further data analyses.

2.6. Structural MRI-based Aβ score (MRI-BAS)

High-dimensional data, such as voxel-based MR images, based on a relatively small number of participants inherently comes with significant co-dependencies and contain a large number of association patterns, most of which are erroneous or redundant. Our goal was to identify which of these are significant associations, with high predictive power. Partial least squares (PLS) regression [21] has the ability to handle high-dimension, low sample size, multicollinear data, while searching for modes that explain the maximum covariance between the explanatory and response spaces. The PLS regression is a supervised dimensionality reduction technique based on a latent decomposition model. Furthermore, unlike commonly used multivariate latent decomposition approaches such as principal component regression or canonical correlation analysis [22], where the dimensionality reduction of the data is carried out independent of the response variable by maximizing the variance within the regressors alone, PLS models the regression by maximizing the covariance between the regressors and response. The latent components are extracted in the regressor and response data spaces such that the covariance between the two is maximized. We used PLS regression with the anatomical shape variation measures from each and every imaging voxel as regressors and florbetapir-SUVR as the response to assess the patterns of structural MRI-Aβ associations. The statistical significance of the structural MRI-Aβ associations inferred by PLS regression was assessed using the projected data and non-parametric permutation testing.

Furthermore, a MRI-based Aβ score (MRI-BAS) was calculated by projecting each individual's neuroimaging data onto the latent variable (LV) inferred by the corresponding PLS regression. Specifically, the PLS regression is a supervised dimensionality reduction technique based on a latent decomposition model. This is done by extracting a small number of latent components or projection scores that are linear combinations of the original variables to avoid multicollinearity. PLS models the regression by maximizing the covariance between the regressors and response. The latent components are extracted in the regressor and response spaces such that the covariance between the two is maximized. In this study, the response spaces is defined by the florbetapir-SUVR measures and the regressor space is defined by the anatomic shape variation measures from AD subjects with florbetapir-PET scans (i.e., training cohort).

2.7. Derivation of classifier

In the derivation of an Aβ-positivity classifier, the dependent outcome variable was the Aβ-positive versus negative dichotomization, based on florbetapir-PET scans and an established standardized florbetapir-SUVR threshold [16]. The MRI-BAS, age, sex, education, baseline clinical measures including MMSE, ADAS-Cog, Alzheimer's Disease Cooperative Study-Activities of Daily Living Inventory (ADCS-ADL), and CDR-SB, and *APOE* genotype with dose effect were the independent predictor variables. Several of these predictors were highly correlated with each other. In particular, MMSE score was significantly correlated with years of education ($r = 0.22$; $P = .002$), ADCS-ADL ($r = 0.27$; $P = .0001$), and ADAS-Cog ($r = -0.46$; $P < 10^{-10}$); ADCS-ADL score was also significantly correlated with baseline age ($r = -0.16$; $P = .02$) and ADAS-Cog ($r = -0.34$; $P < 10^{-5}$). Traditional linear regression is known to produce unstable parameter estimates in the presence of highly inter-correlated variables, thus leading to regression findings that fail to generalize well to new data sets [23,24]. Regression with least absolute shrinkage and selection operator (LASSO) reduces this instability by pushing regression parameter estimates closer to zero, thus resulting in more conservative parameter estimates. In addition, this conservative parameter biasing obviates the need for multiple comparison correction [24]. To create a mathematical function that best combines the MRI-BAS, demographics (i.e., age, sex, and education), baseline cognitive data, and *APOE* genotype to give a binary prediction of Aβ-

positivity in mild-AD individuals, we based our classification model on a logistic regression with LASSO.

For 10-fold cross-validation of the classifier performance, the data were divided into ten subsets of cases that had similar size and Aβ-positive/negative distributions. Each subset was left out once, whereas the other nine was applied to construct the classifier, which was subsequently validated for the unseen cases in the left-out subset. Classifier performance assessment was based on classification accuracy (CA), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Performance of the best performing classifier with non-MRI factors was referenced to test the added-value of structural MRI in classification of Aβ-positive/negative mild-AD. The significance of the prediction performance of these measures compared with randomly generated variables with N(0,1) distribution was assessed via 50,000 simulations. Once the best non-MRI predictors were identified through these model comparisons, a final model based on elastic net parameter was created with an optimization criterion of maximal PPV.

2.8. Application of classifier to subjects without Aβ scans

The best performing classifier based on training data set was then applied to impute the Aβ status of mild-AD subjects from the placebo arms of the EXPEDITION trials that had an analyzable 1.5-T structural MRI but no FBP-PET scan.

2.9. Statistical analysis for clinical trajectories

Although training of the classifier on baseline (pre-treatment) scans used subjects from both placebo and treatment arms (N = 169), we only assessed cognitive trajectories only from placebo-treated subjects in this study (N = 127 with known florbetapir status and N = 385 with imputed florbetapir status). Declines in cognitive and functional measures over the 18-month study were assessed as least square means (LS means) of change from baseline to 18 months (week 80). Cognitive decline over 18 months was assessed using the ADAS-Cog$_{14}$ and MMSE with between group comparisons made using a mixed-effects repeated-measures (MMRM) model adjusted for study, baseline score, age, Aβ imputation (yes/no), Aβ status (positive/negative), visit, and concurrent use of an acetylcholinesterase inhibitor (AChEI) or memantine. The same model was used to functional decline using ADCS-iADL. All data analyses were carried out using SAS, version 9.2 (SAS Institute Inc, Cary, NC).

3. Results

3.1. Demographic characteristics

The cohort of EXPEDITION mild-AD subjects with baseline 1.5-T structural MRI and florbetapir-PET scans comprised 194 subjects. Of these, 12 failed MRI quality control (unsuitable for whole brain shape analysis due to large cortical stroke, cysts, white matter lesions, or diffused imaging artifacts including ringing, motion, and ghosting) and an additional 13 had no *APOE* status. Consequently, 169 subjects (N = 123 Aβ-positive and N = 46 Aβ-negative) were used as the training cohort to develop the classifier.

Demographic characteristics of these 169 mild-AD subjects along with CN subjects in the training cohort are summarized in Table 1. On average CNs were significantly more highly educated than mild-AD patients (t = 4.9481, df = 178.739, *P* value = 1.722e−06), although still in the same age range with similar gender distribution.

3.2. Structural MRI-Aβ associations

Fig. 1 shows the spatial signature of the LV inferred by PLS regression, that is, anatomic shape variation signature of brain Aβ-burden in mild-AD. Dark red, blue, and white colors indicate greater contribution of the local anatomic shape variations to the latent variable and therefore to the structural MRI-Aβ association. Specifically, increased florbetapir-SUVR was associated with anatomic shape variations largely in the parieto-temporal cortical regions including precuneus, supramarginal, inferior parietal, hippocampus, superior temporal, and to a lesser extend in the frontal lobe regions (r = 0.93; $P < .0001$).

3.3. Selection of predictor variables and classifier training

Estimated performances of the LASSO-penalized logistic regression classifiers with non-MRI variables (i.e., age, sex, years of education, clinical measures, and *APOE*-ε4 genotype) alone and jointly with MRI-BAS are reported in Table 2. All classifiers considered in this study performed significantly better than chance ($P < .01$). Pure MRI-BAS based classifier, with 79% accuracy, 92% sensitivity, 44% specificity, 82% PPV, and 68% NPV, outperformed all the non-MRI classifiers considered in this study. A multidisciplinary classifier combining demographics, baseline clinical measures, *APOE* ε4-genotype,

Table 1
Characteristics of training and testing cohorts

Factors	ADNI CN e3/e3 Aβ negative (training)	EXPEDITION mild-AD with florbetapir-PET scan (training)	EXPEDITION mild-AD (testing)
N	75	169	385
Age (y)	75.18 ± 5.28	73.64 ± 8.77	72.94 ± 7.48
Gender (Female, %)	49.33	57.99	54.80
Education (y)	15.52 ± 2.85	13.38 ± 3.65	12.44 ± 4.06
APOE ε4 status (0/1/2)	75/0/0	78/74/17	138/191/56
Aβ-positive (%)	0 %	72.78%	Unknown
MMSE	29.00 ± 1.12	22.78 ± 1.88	22.48 ± 2.81
ADAS-Cog14	—	29.96 ± 9.74	29.34 ± 8.15
ADAS-Cog13	9.29 ± 4.02	—	29.09 ± 7.96
ADCS-ADL	—	62.83 ± 11.20	64.20 ± 10.74
CDR-SB	0.00 ± 0.00	4.32 ± 1.73	4.36 ± 2.18

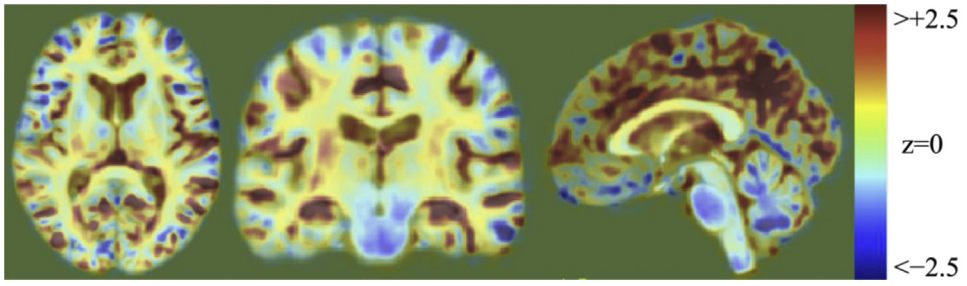

Fig. 1. Pattern of structural MRI–Aβ associations in mild-AD subjects. Dark red, blue, and white colors indicate greater contribution of the local anatomic shape variations to the structural MRI–Aβ association.

and MRI-BAS reached 81% classification accuracy, 95% sensitivity, 45% specificity, 82% PPV, and 80% NPV in identifying Aβ-positive mild-ADs. The biggest improvement compared to MRI-BAS only classifier was in NPV of the classifier. Based on LASSO-penalized regression analysis, among demographic characteristics, age was the best predictor of Aβ-positivity, and among baseline clinical characteristics, MMSE and ADAS-Cog are the best predictors of Aβ-positivity.

With age, MMSE, ADAS-Cog, *APOE* genotype, and MRI-BAS as the predictors, a final classifier model based on elastic net parameter optimization for the best PPV yielded a classifier accuracy of 83%, with 94% sensitivity, 55% specificity, 85% PPV, and 78% NPV. Furthermore, when applied to the test cohort, the classifier imputed 64% of the *APOE* ε4 noncarriers and 97% of the *APOE* ε4 carriers in the mild-AD cohort as florbetapir-positive. We compared the predictive power of the final classifier model to those of hippocampal volume, the most relevant candidate imaging biomarker for the prediction of brain amyloidosis. Intracranial vault volume–adjusted hippocampal volumes, generated by Surgical Navigation Technologies (6) as described in the Supplementary Material, when combined with the non-imaging variables yielded accuracy, sensitivity, specificity, PPV and NPV of 78% (±11%), 88% (±12%), 46% (±27%), 82% (±10%), and 59% (±33%) respectively.

The demographic characteristics of the known and imputed florbetapir-positive cohorts are summarized in Supplementary Material (Supplementary Table 1). On average the mild-AD patients with known florbetapir-positive status were significantly more educated than the mild-AD patients with imputed florbetapir-positive status ($t = 3.86$, $P = .0001526$).

3.4. *Clinical trajectories of subjects imputed Aβ-positive versus Aβ-negative*

With no discrimination based on Aβ status, the test mild-AD cohort exhibited the expected clinical decline on each of the measures, approximately +4.4 points on ADAS-Cog_{14}, −2.0 points on MMSE, and −4.7 points on ADCS-iADL at 80 weeks based on LS means estimation from MMRM (Fig. 2, Table 3). When stratified based on Aβ status as imputed by the classifier, subjects with negative-imputed Aβ status evidenced notably flat trajectories for all clinical instruments evaluated and did not progress significantly over 80 weeks on either the ADAS-Cog_{14}, MMSE, or ADCS-iADL measures. In contrast, subjects with imputed

Table 2
Performance of LASSO-penalized logistic regression classifiers with non-MRI factors jointly with sMRI-BAS

Predictors	ACC	SENS	SPEC	PPV	NPV
MRI-BAS	0.79 ± 0.03	0.92 ± 0.03	0.44 ± 0.10	0.82 ± 0.02	0.68 ± 0.10
Demog	0.71 ± 0.01	0.95 ± 0.01	0.09 ± 0.02	0.73 ± 0.01	0.45 ± 0.03
Demog, *APOE*	0.70 ± 0.01	0.90 ± 0.01	0.17 ± 0.02	0.74 ± 0.01	0.39 ± 0.05
Cogn	0.72 ± 0.01	0.93 ± 0.01	0.20 ± 0.01	0.75 ± 0.01	0.56 ± 0.03
Cogn, *APOE*	0.76 ± 0.01	0.92 ± 0.01	0.33 ± 0.02	0.78 ± 0.01	0.66 ± 0.02
Demog, Cogn	0.76 ± 0.01	0.92 ± 0.01	0.34 ± 0.02	0.78 ± 0.01	0.66 ± 0.02
Demog, Cogn, *APOE*	0.76 ± 0.01	0.91 ± 0.01	0.38 ± 0.02	0.79 ± 0.01	0.65 ± 0.01
MRI-BAS, Demog	0.78 ± 0.02	0.95 ± 0.04	0.33 ± 0.11	0.79 ± 0.03	0.76 ± 0.10
MRI-BAS, Demog, *APOE*	0.80 ± 0.03	0.95 ± 0.03	0.41 ± 0.12	0.81 ± 0.03	0.79 ± 0.09
MRI-BAS, Cogn	0.79 ± 0.04	0.95 ± 0.03	0.36 ± 0.10	0.80 ± 0.03	0.74 ± 0.12
MRI-BAS, Cogn, *APOE*	0.81 ± 0.03	0.94 ± 0.02	0.46 ± 0.09	0.82 ± 0.03	0.76 ± 0.07
MRI-BAS, Demog, Cogn	0.78 ± 0.02	0.95 ± 0.03	0.34 ± 0.08	0.80 ± 0.02	0.77 ± 0.10
MRI-BAS, Demog, Cogn, *APOE*	0.81 ± 0.03	0.95 ± 0.02	0.45 ± 0.10	0.82 ± 0.03	0.80 ± 0.09
Final model: Age, MMSE, ADAS-Cog, *APOE*, and MRI-BAS	0.83 ± 0.03	0.93 ± 0.02	0.61 ± 0.09	0.84 ± 0.02	0.79 ± 0.07

Abbreviations: Demog, Age, gender, education; Cogn, baseline clinical measures including MMSE, ADAS-Cog, ADCS-ADL, and CDR-SB; ACC, Classification accuracy; SENS, classification sensitivity; SPEC, classification specificity; PPV, Positive predictive value; NPV, Negative predictive value.

Aβ-positive status progressed more rapidly than the full, non-stratified cohort, changing by +7.3 points on ADAS-Cog$_{14}$, −3.6 points on MMSE, and −8.1 points on ADCS-iADL based on LS means estimation from MMRM. The differences in the changes in the imputed Aβ-positive and imputed Aβ-negative samples were significant for all three instruments at the 80-week time point based on MMRM ($P < .01$).

The trajectories for the full data set (including both imputed Aβ-positive and Aβ-negative subjects) were intermediate between those different imputed Aβ status subgroups.

3.5. Comparison of trajectories between known and imputed Aβ-positive and negative

The clinical trajectories over 80 weeks observed for subjects imputed Aβ-positive or negative (test set) were very similar to those observed for subjects whose Aβ status was known (training set) as shown in Fig. 3. The least squares mean differences were not statistically significant at any time point for any of the clinical measures.

4. Discussion

The major findings of this study were (1) florbetapir-positive mild-AD subjects could be identified with a classifier accuracy of 83% (cf. 79% with MRI-BAS alone), with 94% sensitivity, 55% specificity, 85% PPV, and 78% NPV by a multidisciplinary classifier based on age, MMSE, ADAS-Cog, *APOE* genotype, and MRI-BAS as the predictors and (2) subjects imputed Aβ-positive declined faster than the full undifferentiated placebo-treated cohort, whereas those imputed Aβ-negative showed stable cognition and function over the 18-month trial duration. Amyloid PET and CSF methods can be challenging to implement outside the western hemisphere, for reasons including patient acceptance, cost, and availability. Taken together, our results demonstrate imputing Aβ status from MRI scans in mild-AD subjects may be a useful screening tool in global clinical trials if Aβ PET or CSF is not available. Furthermore, multimodal approaches including MRI may also offer cost advantages for eventual clinical use, either as a proxy for a direct measure of amyloid status after traditional diagnostic standard of care or as a screening technique before a confirmatory amyloid biomarker test.

The first finding of this study was that using structural MRI, basic demographics, baseline cognitive/clinical measures, and *APOE* genotype data, we could achieve an 83% classification accuracy, with 94% sensitivity, 55% specificity, 85% PPV, and 78% NPV in identifying florbetapir-positive mild-AD patients. This multidisciplinary classification model outperformed other classifiers considered in this study using non-MRI factors.

We previously pursued the use of MRI data in predicting wide-spread brain Aβ deposition in individuals with

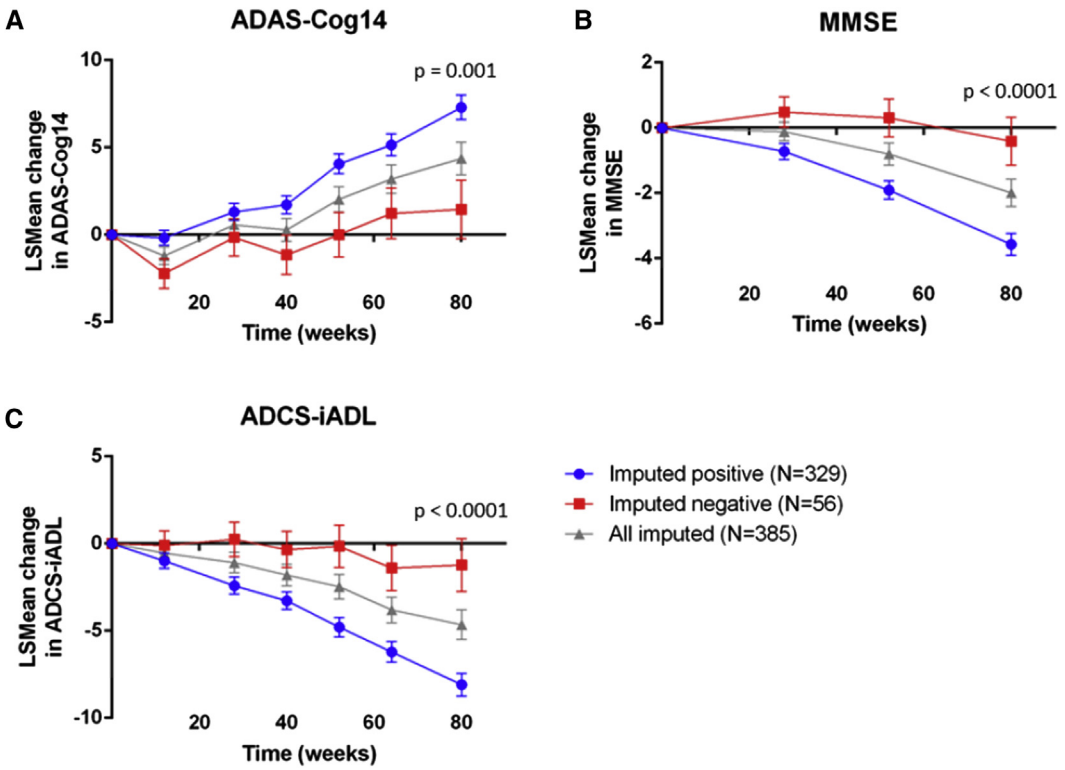

Fig. 2. Change from baseline in clinical variables over 80 weeks for placebo-treated subjects from the EXPEDITION trials imputed as Aβ-negative (red squares, N = 56) or Aβ-positive (blue circles, N = 329), and the full test mild-AD population (gray triangles, N = 385), for (A) ADAS-Cog14, (B) MMSE, and (C) ADCS-iADL. *P* values refer to the comparison between the imputed-positive and imputed-negative subgroups (MMRM model, shown for 80-week time point).

Table 3
Change from baseline in clinical variables >80 weeks based on MMRM for placebo-treated subjects from the EXPEDITION trials with imputed or known Aβ status (Mild-AD)

Instrument	Cohort*	Group	Least square means change at 80 weeks (points)	Amyloid-positive vs. negative at 80 weeks
ADAS-Cog14	Training (N = 127)	Aβ-positive (known, N = 93)	6.78	$P = .0021$
		Aβ-negative (known, N = 34)	−0.63	
	Testing (N = 385)	Aβ-positive (imputed, N = 329)	7.29	$P = .0012$
		Aβ-negative (imputed, N = 56)	1.46	
		All (N = 385)	4.37	—
MMSE	Training (N = 127)	Aβ-positive (known, N = 93)	−3.56	$P < .0001$
		Aβ-negative (known, N = 34)	1.16	
	Testing (N = 385)	Aβ-positive (imputed, N = 329)	−3.57	$P < .0001$
		Aβ-negative (imputed, N = 56)	−0.41	
		All (N = 385)	−1.99	—
ADCS-iADL	Training (N = 127)	Aβ-positive (known, N = 93)	−6.62	$P = .0142$
		Aβ-negative (known, N = 34)	−1.28	
	Testing (N = 385)	Aβ-positive (imputed, N = 329)	−8.09	$P < .0001$
		Aβ-negative (imputed, N = 56)	−1.23	
		All (N = 385)	−4.66	—

Abbreviations: ADAS-Cog14, Alzheimer's Disease Assessment Scale-Cognitive subscale; MMSE, mini-mental state examination; ADCS-iADL, Alzheimer's Disease Cooperative Study–Instrumental Activities of Daily Living Inventory.
*Sample sizes quoted are those at the 80-week time point.

early MCI [7] and amnestic MCI [8]. Although our mild-AD Aβ-positivity prediction model showed similar performance in terms of classification accuracy, PPV, and NPV with slightly better performance in sensitivity compared to previously proposed early-MCI and MCI Aβ-positivity prediction models, its specificity was lower than the specificity of early MCI and MCI Aβ-positivity prediction models (55% versus 87.6% in early MCI and 92.82% in MCI). An explanation for the poor specificity of the proposed mild-AD Aβ-positivity prediction model could be that the class sizes (florbetapir-positive versus florbetapir-negative) in the training data differed considerably (73% Aβ-positive mild-AD versus 50% early MCI and 66% MCI). When the class sizes are very different, most standard classification algorithms may favor the larger (majority) class resulting in poor accuracy in the minority class prediction. One unique aspect of the proposed mild-AD Aβ-positivity prediction model is that unlike the other models, this model was trained and tested on clinical trial data that were generated not just at high-end sites in North America but at trial sites world wide. This makes it feasible for global trial applications.

In practice, given that structural MRI scans are acquired routinely in clinical trials for other purposes, an MRI-based amyloid classifier might be useful as a screening tool that could be applied before confirmation with a CSF or PET biomarker that directly measures amyloid status, to reduce the number of amyloid-negative subjects undergoing those more expensive and/or invasive procedures. In the present study, the inclusion of MRI shape change information yielded increased classification performance when compared to classifiers based solely on non-imaging data (e.g., demographics, cognition, *APOE* genotype). However, the generalizability and reliability of MRI-based classifiers, and their applicability at the individual subject level, warrant further investigation.

The second finding of this study was that the latent variable structure underlying the anatomical shape variations and Aβ burden associations was dominated by the hippocampus, parieto-temporal cortical regions including precuneus, supramarginal, inferior parietal, superior temporal, and to a lesser extent areas in the frontal lobe regions. There is an emerging literature investigating Aβ-related brain changes using structural MRI, describing an association between Aβ burden (e.g., low CSF $A\beta_{1-42}$ or high Aβ-PET binding) and atrophy, especially of the parietal and posterior cingulate regions, extending into the precuneus and medial temporal regions including hippocampus, amygdala, and entorhinal cortex. This parietotemporal dominant pattern of Aβ-atrophy association is even evident at mild stages of cognitive deficits [10,25–33]. Our study does not explain the mechanisms behind the detected Aβ-related structural variations. There are several interpretations to the latent variable structure underlying the anatomic shape variations and Aβ burden associations. One possible interpretation is that the Aβ burden associated anatomic shape variations are due in part to direct Aβ toxicity [34]. The "Aβ hypothesis" [35,36] posits that accumulation of Aβ is primarily responsible for the accumulation of tau tangles, synaptic dysfunction, neurodegeneration, and cognitive decline especially in memory function, which characterizes AD. There is considerable evidence from animal and *in vitro* studies demonstrating that soluble Aβ oligomers, fibril Aβ, and Aβ plaques have adverse effects on neuronal function [34]. The second possible interpretation, not mutually exclusive, comes from pathologic studies, especially those of Braak et al. [37,38], showing that during the development of AD pathology, tau tangles increase, associated with synapse loss

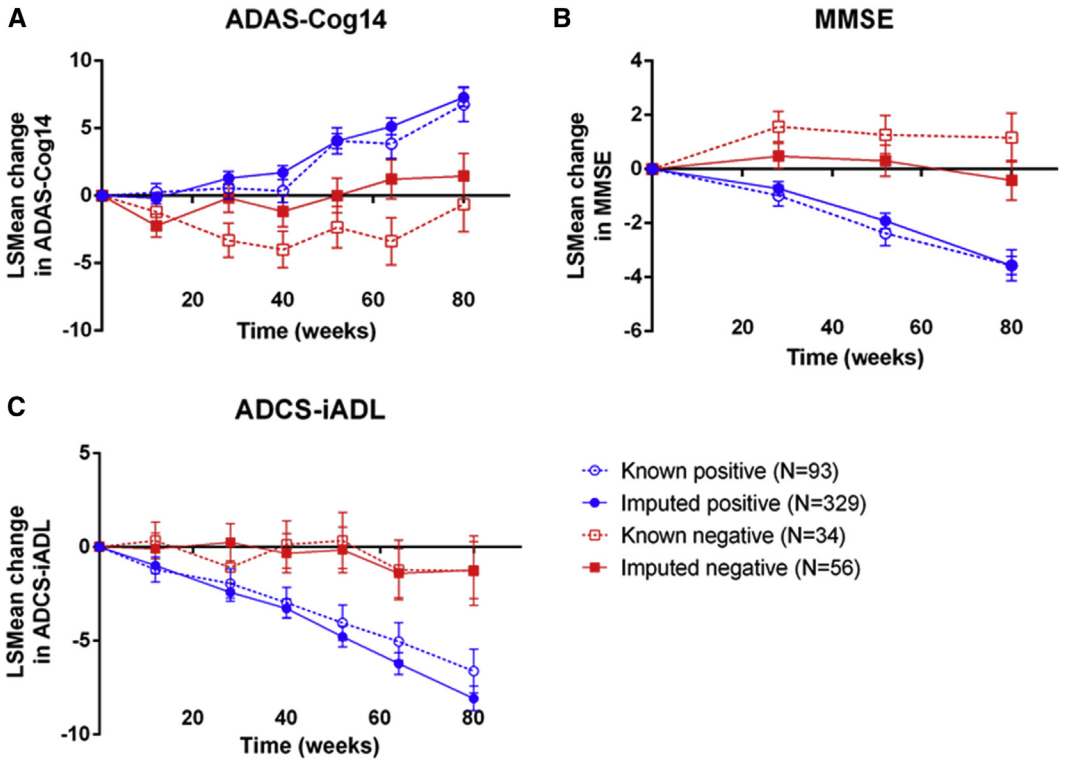

Fig. 3. Comparison of clinical trajectories of (A) ADAS-Cog14, (B) MMSE, and (C) ADCS-iADL for placebo arm subjects from the EXPEDITION trials, stratified by known Aβ status (training set, N = 93 positive, and N = 34 negative) and by imputed Aβ status (test set, N = 329 positive, and N = 56 negative).

and neurodegeneration, while at the same time, widespread neocortical Aβ plaques are developing; therefore, the anatomic shape variation pattern could be an indirect measure of widespread neurodegeneration due to increased Aβ burden.

The third finding of this study was that the presence of Aβ does affect the trajectory of clinical decline, and this occurs independently of the method used to stratify patients based on Aβ status, that is, direct Aβ status assessment via florbetapir-PET imaging versus imputing Aβ status from multidisciplinary data including structural MRI (Figs. 2 and 3 and Table 3). The fact that the brain Aβ accelerated the cognitive and clinical decline regardless of the person's diagnosis has been reported in research cohorts as well [39–41]. Thirty six–month follow-up data from more than 1000 ADNI participants showed that participants whose brains were free of Aβ stayed stable on cognitive tests, including the ADAS-Cog11, MMSE, and CDR-SB [39]. In contrast, people with positive Aβ scans declined, with MMSE scores diverging significantly from Aβ negatives at 6 months, and ADAS-Cog11 scores at 24 months [39], similar to what we observed in these global clinical trial data (Figs. 2 and 3). Moreover, data from Australian Imaging, Biomarker, and Lifestyle Flagship Study of Ageing study also showed that people without brain Aβ remained cognitively stable over 3 years follow-up period, regardless of their diagnosis. On the other hand, people with Aβ all declined at about the same average rate, again regardless of diagnosis [41]. Taken together these results demonstrate that the presence of brain Aβ hastens the trajectory of cognitive or clinical decline. Both research and clinical trial data have shown that patients selected by clinical diagnosis alone may lack underlying AD pathology. This observation leads many current clinical trials and researchers to screen potential participants for Aβ pathology either via PET scanning or CSF sampling to ensure that the patient's clinical presentation is most likely due to AD. It should also be noted that brain Aβ may explain only some of the individual variation in cognitive and/or clinical decline.

There were several limitations of the present study. First, findings from the EXPEDITION trials may not precisely generalize to a general population because the predictive performance of the final classifier model was assessed using cross-validation. This, or related, approaches would need to be prospectively tested and refined for ease of use, especially for applicability in the clinical care setting. Second, the set of candidate independent predictor variables considered in this study for Aβ-positivity may not represent an exhaustive list of factors predictive of Aβ-positivity. Therefore, a certain sub-population might have been overrepresented in the imputed Aβ-positive cohort, biasing the clinical and/or cognitive decline trajectories. Third limitation of this study is that we optimized the classifier based on PPV, not NPV, to maximize the probability that subjects included would be Aβ-positive, thus minimizing the

dilutive contribution of amyloid negative nonprogressors. This strategy does result in some Aβ-positive subjects being excluded but should maximize the fraction of Aβ-negative subjects excluded. Fourth, a methodologic limitation is that the LASSO is only one among several machine-learning algorithms that have been used in medical imaging. It is possible that using other established algorithms, such as support vector machines or random forest, yields a different outcome. We chose the widely used LASSO algorithm because it shrinks some parameters toward zero, stabilizing parameter estimates in the presence of highly intercorrelated variables and allowing an easy interpretation of the models and straightforward identification of those parameters most strongly related with the outcome. Finally, instrument-related factors such as MR scanner manufacturer and model can influence the reliability of MRI-BAS and therefore the performance of the proposed final florbetapir-positivity classifier model. Our preliminary assessment indicated that there might be MR scanner manufacturer-specific differences in predictive performance of the final classifier model. However, the sample sizes available in this study (i.e., N = 78 from GE scanners versus N = 32 from Philips scanners versus N = 59 from Siemens scanners) and the uneven florbetapir-positivity distribution within the manufacturer-specific samples limit our ability to perform a comprehensive assessment of effects of MR scanner manufacturer/model on predictive performance of the final classifier model using data from the EXPEDITION trials alone.

In summary, we defined an Aβ-positivity prediction model for mild-AD patients recruited in a global clinical trial. We detected accelerated clinical decline both in subcohorts with known Aβ status from florbetapir-PET scans and also in patients who were imputed Aβ-positive based on a structural MRI signature of Aβ-positivity and non-MRI factors including demographics, baseline clinical measure, and *APOE* genotype data. The presence of Aβ-negative subjects in clinical trial cohorts represents a potential confound, introducing variability and possibly diluting any treatment signal. For treatments targeting Aβ pathology, Aβ-negative subjects do not have the therapeutic target; hence, their clinical trajectories may not be modified in the same way as Aβ-positive subjects. Furthermore, Aβ-negative subjects may not progress clinically on the same trajectory as those that do even in the absence of treatment.

Acknowledgments

The florbetapir PET scans were processed by Avid.
This work was supported by NIH grant 8P41EB015904-05. 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.; 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 Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jalz.2016.03.009.

RESEARCH IN CONTEXT

1. Systematic review: A low-cost, accessible, and efficient screening tool for Aβ-positivity is needed in global clinical trials. Analysis of Aβ biomarker subgroups in phase 3 clinical trials from two independent drug development programs revealed that approximately 27% of subjects meeting clinical inclusion criteria for mild-AD were Aβ-negative. Aβ-negative mild-AD subjects are not expected to progress clinically on the expected trajectory, adding variance into analyses where a slowing of progression is being measured.
2. Interpretation: This work confirms that in a placebo-treated Aβ-positive mild-AD cohort, either imputed or measured by Aβ-PET, both cognitive and functional measures declined more rapidly than the undifferentiated population, whereas those imputed Aβ-negative showed stable cognition and function.
3. Future directions: Future studies should take into account Aβ-positivity status when assessing the efficacy of treatments targeting Aβ pathology.

References

[1] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:263–9.

[2] Weigand SD, Vemuri P, Wiste HJ, Senjem ML, Pankratz VS, Aisen PS, et al. Transforming cerebrospinal fluid AŒ≤42 measures into calculated Pittsburgh compound B units of brain AŒ≤ amyloid. Alzheimers Dement 2011;7:133–41.

[3] Zwan M, van Harten A, Ossenkoppele R, Bouwman F, Teunissen C, Adriaanse S, et al. Concordance between cerebrospinal fluid biomarkers and [11C]PIB PET in a memory clinic cohort. J Alzheimers Dis 2014;41:801–7.

[4] Salloway S, Sperling R, Fox NC, Blennow K, Klunk W, Raskind M, et al. Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med 2014;370:322–33.

[5] Siemers ER, Sundell KL, Carlson C, Case M, Sethuraman G, Liu-Seifert H, et al. Phase 3 solanezumab trials: Secondary outcomes in mild Alzheimer's disease patients. Alzheimers Dement 2016;12:110–20.

[6] Sperling RA, Jack CR Jr, Black SE, Frosch MP, Greenberg SM, Hyman BT, et al. Amyloid-related imaging abnormalities in amyloid-modifying therapeutic trials: recommendations from the Alzheimer's Association Research Roundtable Workgroup. Alzheimers Dement 2011;7:367–85.

[7] Tosun D, Joshi S, Weiner MW, for the Alzheimer's Disease Neuroimaging Initiative. Multimodal MRI-based imputation of the Aβ+ in early mild cognitive impairment. Ann Clin Transl Neurol 2014;1:160–70.

[8] Tosun D, Joshi S, Weiner MW. Neuroimaging predictors of brain amyloidosis in mild cognitive impairment. Ann Neurol 2013;74:188–98.

[9] Sojkova J, Beason-Held L, Zhou Y, An Y, Kraut MA, Ye W, et al. Longitudinal Cerebral Blood Flow and Amyloid Deposition: An Emerging Pattern? J Nucl Med 2008;49:1465–71.

[10] Becker JA, Hedden T, Carmasin J, Maye J, Rentz DM, Putcha D, et al. Amyloid-β associated cortical thinning in clinically normal elderly. Ann Neurol 2011;69:1032–42.

[11] Bourgeat P, Chetelat G, Villemagne VL, Fripp J, Raniga P, Pike K, et al. β-Amyloid burden in the temporal neocortex is related to hippocampal atrophy in elderly subjects without dementia. Neurology 2010;74:121–7.

[12] Honea RA, Vidoni E, Harsha A, Burns JM. Impact of APOE on the healthy aging brain: a voxel-based MRI and DTI study. J Alzheimers Dis 2009;18:553–64.

[13] Bangen KJ, Restom K, Liu TT, Wierenga CE, Jak AJ, Salmon DP, et al. Assessment of Alzheimer's disease risk with functional magnetic resonance imaging: an arterial spin labeling study. J Alzheimers Dis 2012;31(Suppl 3):S59–74.

[14] Jagust WJ, Landau SM, Shaw LM, Trojanowski JQ, Koeppe RA, Reiman EM, et al. Relationships between biomarkers in aging and dementia. Neurology 2009;73:1193–9.

[15] Landau SM, Breault C, Joshi AD, Pontecorvo M, Mathis CA, Jagust WJ, et al. Amyloid-beta imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. J Nucl Med 2013;54:70–7.

[16] Joshi AD, Pontecorvo MJ, Clark CM, Carpenter AP, Jennings DL, Sadowsky CH, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with alzheimer's disease and cognitively normal subjects. J Nucl Med 2012;53:378–84.

[17] Joshi AD, Pontecorvo MJ, Lu M, Skovronsky DM, Mintun MA, Devous MD Sr. A Semiautomated Method for Quantification of F 18 Florbetapir PET Images. J Nucl Med 2015;56:1736–41.

[18] Shattuck DW, Leahy RM. BrainSuite: an automated cortical surface identification tool. Med Image Anal 2002;6:129–42.

[19] Jiang T, Navab N, Pluim J, Viergever M, Singh N, Fletcher P, et al. Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010. Berlin: Springer; 2010. p. 529–37.

[20] Younes L, Arrate F, Miller MI. Evolutions equations in computational anatomy. Neuroimage 2009;45:S40–50.

[21] Wold S, Geladi P, Esbensen K, Öhman J. Multi-way principal components-and PLS-analysis. J Chemometr 1987;1:41–56.

[22] Jolliffe IT. A note on the use of principal components in regression. J Roy Stat Soc C Appl Stat 1982;31:300–3.

[23] Belsley DA, Kuh E, Welsch RE. Regression Diagnostics: Identifying Influential Data and Sources of Collinearit. New York: John Wiley and Sons; 1980.

[24] van der Kooij AJ. Prediction accuracy and stability of regression with optimal scaling transformations. Leiden: Leiden University; 2007.

[25] Oh H, Mormino EC, Madison C, Hayenga A, Smiljic A, Jagust WJ. β-Amyloid affects frontal and posterior brain networks in normal aging. Neuroimage 2011;54:1887–95.

[26] Whitwell JL, Tosakulwong N, Weigand SD, Senjem ML, Lowe VJ, Gunter JL, et al. Does amyloid deposition produce a specific atrophic signature in cognitively normal subjects? Neuroimage Clin 2013;2:249–57.

[27] Sabuncu MR, Desikan RS, Sepulcre J, Yeo BT, Liu H, Schmansky NJ, et al. The dynamics of cortical and hippocampal atrophy in Alzheimer disease. Arch Neurol 2011;68:1040–8.

[28] Chetelat G, Villemagne VL, Villain N, Jones G, Ellis KA, Ames D, et al. Accelerated cortical atrophy in cognitively normal elderly with high beta-amyloid deposition. Neurology 2012;78:477–84.

[29] Andrews KA, Modat M, Macdonald KE, Yeatman T, Cardoso MJ, Leung KK, et al. Atrophy rates in asymptomatic amyloidosis: implications for Alzheimer prevention trials. PLoS One 2013;8:e58816.

[30] Schott JM, Bartlett JW, Fox NC, Barnes J. Increased brain atrophy rates in cognitively normal older adults with low cerebrospinal fluid Abeta1-42. Ann Neurol 2010;68:825–34.

[31] Chetelat G, Villemagne VL, Bourgeat P, Pike KE, Jones G, Ames D, et al. Relationship between atrophy and beta-amyloid deposition in Alzheimer disease. Ann Neurol 2010;67:317–24.

[32] Tosun D, Schuff N, Mathis CA, Jagust W, Weiner MW, Initiative AsDN. Spatial patterns of brain amyloid-β burden and atrophy rate associations in mild cognitive impairment. Brain 2011;134:1077–88.

[33] Tosun D, Schuff N, Truran-Sacrey D, Shaw LM, Trojanowski JQ, Aisen P, et al. Relations between brain tissue loss, CSF biomarkers, and the ApoE genetic profile: a longitudinal MRI study. Neurobiol Aging 2010;31:1340–54.

[34] Gilbert BJ. The role of amyloid beta in the pathogenesis of Alzheimer's disease. J Clin Pathol 2013;66:362–6.

[35] Karran E, Mercken M, De Strooper B. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. Nat Rev Drug Discov 2011;10:698–712.

[36] Hardy J, Selkoe DJ. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Sci Aging Knowl Environ 2002;2002:or8.

[37] Braak H, Braak E. Neuropathological Assessments Staging of Alzheimer-Related Cortical Destruction. Int Psychogeriatr 1997;9:257–61.

[38] Braak H, Thal DR, Ghebremedhin E, Del Tredici K. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. J Neuropathol Exp Neurol 2011;70:960–9.

[39] Doraiswamy PM, Sperling RA, Johnson K, Reiman EM, Wong TZ, Sabbagh MN, et al. Florbetapir F 18 amyloid PET and 36-month cognitive decline: a prospective multicenter study. Mol Psychiatry 2014;19:1044–51.

[40] Doraiswamy PM, Sperling RA, Coleman RE, Johnson KA, Reiman EM, Davis MD, et al. Amyloid-beta assessed by florbetapir F 18 PET and 18-month cognitive decline: a multicenter study. Neurology 2012;79:1636–44.

[41] Lim YY, Maruff P, Pietrzak RH, Ames D, Ellis KA, Harrington K, et al. Effect of amyloid on memory and non-memory decline from preclinical to clinical Alzheimer's disease. Brain 2014;137:221–31.