NeuroImage: Clinical 11 (2016) 90–98

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Clinical and MRI models predicting amyloid deposition in progressive aphasia and apraxia of speech

Jennifer L. Whitwell[a,*], Stephen D. Weigand[b], Joseph R. Duffy[c], Edythe A. Strand[c], Mary M. Machulda[d], Matthew L. Senjem[e], Jeffrey L. Gunter[e], Val J. Lowe[a], Clifford R. Jack Jr. [a], Keith A. Josephs[c]

[a]Department of Radiology, Mayo Clinic, Rochester, MN, USA
[b]Department of Health Sciences Research (Biostatistics), Mayo Clinic, Rochester, MN, USA
[c]Department of Neurology, Mayo Clinic, Rochester, MN, USA
[d]Department of Psychiatry and Psychology (Neuropsychology), Mayo Clinic, Rochester, MN, USA
[e]Department of Information Technology, Mayo Clinic, Rochester, MN, USA

ARTICLE INFO

Article history:
Received 3 September 2015
Received in revised form 11 January 2016
Accepted 18 January 2016
Available online 20 January 2016

Keywords:
Beta-amyloid
Primary progressive aphasia
Apraxia of speech
Volumetric MRI

ABSTRACT

Beta-amyloid (Aβ) deposition can be observed in primary progressive aphasia (PPA) and progressive apraxia of speech (PAOS). While it is typically associated with logopenic PPA, there are exceptions that make predicting Aβ status challenging based on clinical diagnosis alone. We aimed to determine whether MRI regional volumes or clinical data could help predict Aβ deposition. One hundred and thirty-nine PPA (n = 97; 15 agrammatic, 53 logopenic, 13 semantic and 16 unclassified) and PAOS (n = 42) subjects were prospectively recruited into a cross-sectional study and underwent speech/language assessments, 3.0 T MRI and C11-Pittsburgh Compound B PET. The presence of Aβ was determined using a 1.5 SUVR cut-point. Atlas-based parcellation was used to calculate gray matter volumes of 42 regions-of-interest across the brain. Penalized binary logistic regression was utilized to determine what combination of MRI regions, and what combination of speech and language tests, best predicts Aβ (+) status. The optimal MRI model and optimal clinical model both performed comparably in their ability to accurately classify subjects according to Aβ status. MRI accurately classified 81% of subjects using 14 regions. Small left superior temporal and inferior parietal volumes and large left Broca's area volumes were particularly predictive of Aβ (+) status. Clinical scores accurately classified 83% of subjects using 12 tests. Phonological errors and repetition deficits, and absence of agrammatism and motor speech deficits were particularly predictive of Aβ (+) status. In comparison, clinical diagnosis was able to accurately classify 89% of subjects. However, the MRI model performed well in predicting Aβ deposition in unclassified PPA. Clinical diagnosis provides optimum prediction of Aβ status at the group level, although regional MRI measurements and speech and language testing also performed well and could have advantages in predicting Aβ status in unclassified PPA subjects.

1. Introduction

Patients with primary neurodegenerative speech and language disorders can present either with primary progressive aphasia (PPA) (Mesulam, 2001; Gorno-Tempini et al., 2011) or progressive apraxia of speech (PAOS) (Josephs et al., 2012, 2013). The term PPA is reserved for a neurodegenerative disorder in which the most salient feature is language dysfunction (Mesulam, 1982, 2001). Three variants of PPA have been described which are each characterized by different patterns of language impairment (Gorno-Tempini et al., 2011). The agrammatic variant (agPPA) is characterized by written and verbal language that is grammatically flawed and sometimes with apraxia of speech (AOS); the semantic variant (svPPA) is characterized by anomia and loss of single-word knowledge; and the logopenic variant (lvPPA) is characterized by anomia without loss of word knowledge, difficulty with sentence repetition and phonologic errors. In contrast, the term PAOS describes a neurodegenerative disorder in which AOS is the presenting and most dominant clinical feature (Josephs et al., 2012, 2013). These subjects can present with slow speech rate, articulatory distortions, distorted sound substitutions and segmentation of syllables in multisyllabic words or across words. Language impairment can be present, although it must be less severe than the AOS (Josephs et al., 2013).

The pathological underpinnings of the PPA variants and PAOS are variable, typically having either a variant of frontotemporal lobar degeneration (FTLD) or Alzheimer's disease (AD) (Kertesz et al., 2005; Josephs et al., 2006; Harris et al., 2013; Mesulam et al., 2014). Clinical diagnosis is relatively helpful in predicting pathology, with lvPPA subjects usually having AD pathology and PAOS, agPPA and svPPA subjects

* Corresponding author at: Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905, USA.
E-mail address: whitwell.jennifer@mayo.edu (J.L. Whitwell).

http://dx.doi.org/10.1016/j.nicl.2016.01.014

usually having FTLD pathology (Josephs et al., 2006; Mesulam et al., 2014). However, discordant cases are common, with AD pathology, or beta-amyloid (Aβ) deposition on PET, observed in 18% of PAOS(Josephs et al., 2014), 0–30% of agPPA (Rabinovici et al., 2008; Leyton et al., 2011; Chare et al., 2014, Josephs et al., 2014) and 11–21% of svPPA (Rabinovici et al., 2008; Leyton et al., 2011; Chare et al., 2014; Josephs et al., 2014) cases. Conversely, up to 46% of some series of lvPPA subjects do not have AD pathology (Harris et al., 2013). In addition, a large proportion of subjects with language impairment does not fulfill diagnostic criteria for one of the PPA variants, and hence remain unclassified (Sajjadi et al., 2012; Mesulam et al., 2014; Wicklund et al., 2014). Other biomarkers are therefore needed that can help predict the presence of AD in PPA and PAOS and hence help guide potential treatments. The PPA variants and PAOS are each associated with specific patterns of atrophy (Gorno-Tempini et al., 2004; Josephs et al., 2012, 2013), although it is unknown whether these patterns vary according to the presence of AD pathology, particularly within PPA syndromes, and hence whether MRI could provide clues to the underlying pathology.

The aim of this study was to determine whether a model based on regional gray matter volume data measured from MRI could predict the presence of Aβ deposition on PET in PPA variants and PAOS. We also aimed to compare this MRI model to a model based on speech and language test data, and to determine whether either the MRI or clinical model could in fact do better than clinical diagnosis.

2. Methods

2.1. Subject recruitment

All subjects that presented with a predominant speech or language complaint and fulfilled diagnostic criteria for PPA (Mesulam, 2001) or PAOS (Josephs et al., 2012, 2013) were prospectively recruited from the Department of Neurology, Mayo Clinic, between July 1st 2010 and June 1st 2014. All subjects underwent a detailed neurological and speech and language evaluation, as previously described (Josephs et al., 2012), and diagnoses were rendered by consensus between two speech-language pathologists (JRD and EAS) after reviewing video recordings and speech-language test results for each subject. All diagnoses were made blinded to any neurological or neuroimaging findings. All subjects also underwent a volumetric MRI and an Aβ PET scan. The neurological and speech and language evaluations, MRI and PET scans were all performed within 72 h. A total of 143 subjects were recruited into the study. Four subjects were excluded because they either could not perform the MRI or the MRI was of poor quality. Of the remaining 139 subjects, 97 were diagnosed with a PPA variant (Gorno-Tempini et al., 2011) (15 agPPA, 13 svPPA, 53 lvPPA and 16 unclassified PPA, UCPPA) and 42 were diagnosed with PAOS (Josephs et al., 2012, 2013), according to our previously published criteria (Botha et al., 2015). In three of the UCPPA subjects, anomia was the predominant feature, with fluent speech, but these subjects did not meet criteria for lvPPA or svPPA (two of these were Aβ (+)). We have previously classified these subjects as progressive fluent aphasia (Botha et al., 2015). Of the remaining UCPPA subjects, two resembled lvPPA but lacked phonological errors and/or repetition deficits (both Aβ (−)), two resembled lvPPA but had agrammatism (both Aβ (−)), one had impaired comprehension of sentences and loss of word meaning (Aβ (+)), and one subject (Aβ (−)) had prominent anomia with spared word and object meaning, together with AOS and dysarthria. In two UCPPA subjects, impairment was too severe to classify (both Aβ (−)) and, in five, impairment was so mild that discrepancies or patterns of impairment could not be appreciated (all Aβ (−)).

Apolipoprotein genotyping and assessment for the presence of progranulin (Baker et al., 2006) or microtubule associated protein tau mutations (Hutton et al., 1998), and *C9ORF72* repeat expansions (Dejesus-Hernandez et al., 2011) were performed as previously described (Whitwell et al., 2012; Flanagan et al., 2015). The study was approved by the Mayo Clinic IRB. All patients consented for enrolment into the study.

2.2. Speech and language data

Fourteen speech and language tests were entered into the predictive model. These tests were selected to assess the presence or absence of each diagnostic feature of each clinical variant (Wicklund et al., 2014). The fourteen tests include the Token Test Part V (DeRenzi and Vignolo, 1962) to assess comprehension of complex sentences, the auditory word recognition subtest of the Western Aphasia Battery (WAB) (Kertesz, 2007) to assess single word comprehension, the reading and writing irregular and non-word subtests of the WAB to assess surface dyslexia or dysgraphia, the repetition subtest of the WAB to assess repetition, the informational content subtest of the WAB to assess single word retrieval in spontaneous speech, the Pyramids and Palm Trees (PPT) (Howard and Patterson, 1992) test to assess object knowledge, the Boston Naming Test (BNT) (Lansing et al., 1999) to assess confrontational naming, the Apraxia of Speech Rating Scale (ASRS) (Strand et al., 2014) to assess the severity of apraxia of speech, and the Motor Speech Disorders (MDS) (Yorkson et al., 1993) scale to assess motor speech production. The presence of agrammatism in speech and the severity of phonological errors (0 = none, 1 = mild, 2 = moderate, 3 = severe) was determined by consensus between two speech language pathologists (JRD and EAS).

Table 1
Subject demographics.

	Total cohort	Aβ (+)	Aβ (−)	P value Aβ (+) v Aβ (−)
N	139	58	81	NA
PiB SUVR	1.3 (1.2–2.0)	2.1 (2.0–2.3)	1.2 (1.2–1.3)	<0.001
Female gender, no. (%)	70 (50%)	32 (55%)	38 (47%)	0.34
Education, yrs.	16 (13–18)	15 (13–18)	16 (13–18)	0.63
Apolipoprotein e4, no. (%)*	38/120 (32%)	29/53 (55%)	9/67 (13%)	<0.001
Age at exam, yrs.	69 (61–73)	70 (60–74)	68 (62–73)	0.96
Age at onset, yrs.	66 (58–70)	65 (56–70)	66 (59–70)	0.59
Disease duration, yrs.	3.0 (2.0–4.5)	3.5 (3.0–5.0)	3.0 (2.0–4.0)	0.04
Mini-Mental State Examination, (/30)	28 (24–29)	24 (16–28)	29 (27–30)	<0.001
Clinical Dementia Rating Scale sum of boxes, (/18)	1.0 (0–3.0)	3 (1–4.6)	0.5 (0–1.5)	<0.001
Clinical dx., no. (%)				<0.001
agPPA	15 (11%)	1 (2%)	14 (17%)	
svPPA	13 (9%)	2 (3%)	11 (14%)	
lvPPA	53 (38%)	47 (81%)	6 (7%)	
UCPPA	16 (12%)	3 (5%)	13 (16%)	
PAOS	42 (30%)	5 (9%)	37 (46%)	

Data shown as median (inter-quartile range).

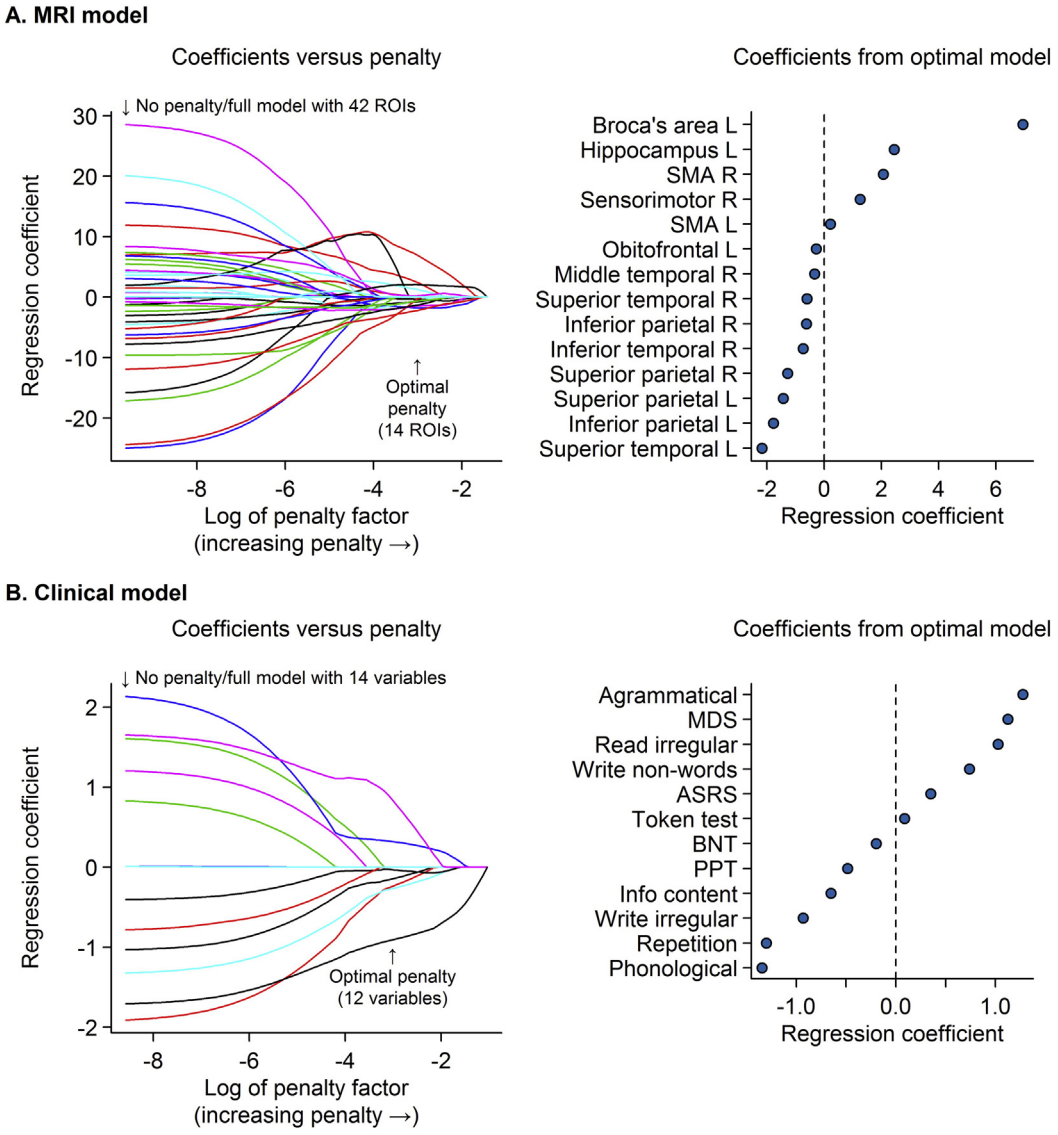

Fig. 1. Results of the penalized logistic regression models. Panel A shows the results of the MRI penalized logistic regression model and panel B shows the results of the clinical penalized logistic regression model. The left panels show the regression coefficients for all variables. The coefficient values on the y-axis are shown as a function of the natural logarithm of the penalty on the x-axis. The coefficients from the optimal model chosen by cross-validation are indicated by an arrow at a penalty of approximately −3.1 on the log scale. The right panels show the coefficient estimates for the selected coefficients from the optimal model. In regions with positive coefficients higher volumes (less atrophy and less clinical impairment) are associated with increased odds of Aβ (+). In regions with negative coefficients lower volumes (more atrophy and more clinical impairment) correspond to increased odds of Aβ (+). SMA = supplementary motor area; MDS = motor speech disorders scale; ASRS = apraxia of speech rating scale; BNT = Boston naming test; PPT = pyramids and palm trees test.

2.3. Beta-amyloid PET analysis

All Aβ PET scans were acquired using a PET/CT scanner (GE Healthcare, Milwaukee, Wisconsin) operating in 3D mode. Subjects were injected with Pittsburgh Compound B (PiB) (average, 614 MBq; range, 414–695 MBq) and after a 40 min uptake period a 20 min PiB scan was obtained consisting of four 5-min dynamic frames following a low dose CT transmission scan. Standard corrections were applied. The automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) was used to calculate median PiB uptake for 6 cortical regions-of-interest (ROIs) which were each divided by median cerebellar uptake to create uptake ratios. A cortical-to-cerebellar (SUVR) ratio was formed by calculating median uptake ratio values across all 6 regions. Patients were classified as Aβ (+) using a SUVR ratio cut-point of 1.5 (Jack et al., 2008). Visual assessment of all Aβ PET scans was also performed (VJL and JLW) to verify Aβ status. The visual assessment and SUVR agreed in all but two cases which were visually read as positive but had low SUVR ratios. These cases were considered Aβ (+) in subsequent analyses and were excluded from analyses that utilized the actual SUVR value.

2.4. MRI analysis

All subjects underwent a standardized MRI imaging protocol at 3.0 Tesla that included a 3D magnetization prepared rapid acquisition gradient echo (MPRAGE) sequence. All images underwent pre-processing that included corrections for gradient non-linearity (Jovicich et al., 2006) and intensity inhomogeneity using both the N3 bias correction (Sled et al., 1998) followed by the SPM5-based bias correction.

An atlas-based parcellation technique was employed using SPM5 and the AAL atlas (Tzourio-Mazoyer et al., 2002), in order to generate regional gray matter volumes. The following 21 mutually exclusive brain regions-of-interest (ROIs) were analyzed: hippocampus, entorhinal cortex, fusiform cortex, inferior, middle and superior temporal gyri,

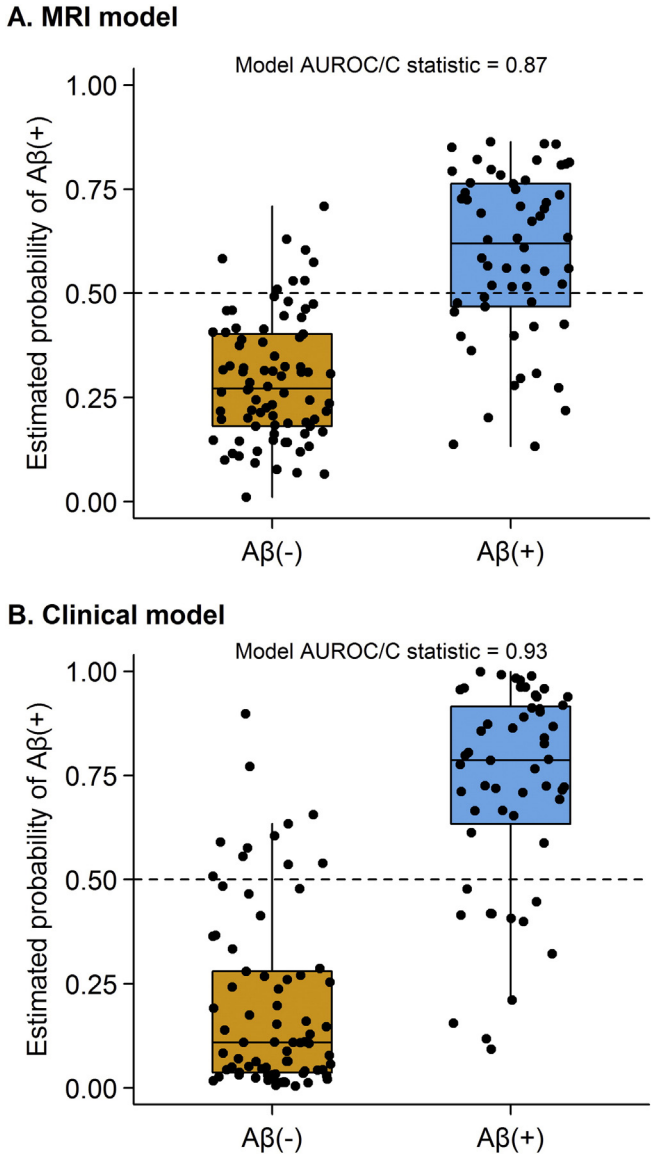

Fig. 2. Estimated probability of Aβ (+) from the optimal MRI and clinical penalized logistic regression models (y-axis) by the observed Aβ status (x-axis). Estimated probabilities >0.5 correspond to a model-based prediction of Aβ (+). The distributions for Aβ (−) and Aβ (+) subjects are summarized with standard box plots indicating the median, quartiles, and having whiskers extending out to 1.5 times the inter-quartile range.

temporal pole, inferior parietal lobe (inferior parietal gyrus + angular gyrus + supramarginal gyrus), superior parietal gyrus, precuneus, posterior cingulate, orbitofrontal cortex, anteromedial frontal lobe (superior medial frontal gyrus + medial orbital gyrus), Broca's area (inferior operculum + inferior triangularis), middle and superior frontal gyri, supplementary motor area (SMA), sensorimotor cortex (precentral gyrus + paracentral lobule + postcentral gyrus + rolandic operculum), occipital lobe (inferior, superior and middle occipital gyri + lingual gyrus + cuneus + calcarine gyrus), insula and striatum (left and right hemispheres were analyzed separately for each region). These regions were selected to cover all major lobes of the brain and, specifically, regions of the brain typically affected by each PPA variant and PAOS.

Each MPRAGE was normalized to custom template space and the inverse transformation was used to warp the custom space AAL atlas to the patient native anatomical space. Each MPRAGE was segmented into gray matter, white matter and cerebrospinal fluid in native space and gray matter volumes were calculated for each region of interest. In addition, total gray matter volume in mm^3 was calculated for each subject by summing the gray matter probabilities from the segmentations and multiplying by the voxel dimensions. All regional volumes were divided by total gray matter volume to correct for differences in global atrophy between subjects. This step was performed because we were interested in assessing the relative involvement of each region rather than just the global severity of atrophy.

2.5. Statistical analysis

All statistical analyses were performed using R version 3.0.2 (http://www.R-project.org) and the glmnet package version 1.9-5. Comparisons of demographic variables of interest among Aβ (+) an Aβ (−) groups were performed using Kruskal–Wallis tests for continuous variables and chi-squared or Fisher's exact tests for categorical variables.

An "elasticnet" penalized logistic regression was used to determine a parsimonious set of ROIs (MRI model) or clinical tests (clinical model) that are predictive of Aβ status (Hastie et al., 2001). The candidate MRI variables were each of 42 ROIs standardized to have mean zero and SD 1. Prior to modeling, we transformed clinical variables as necessary to obtain approximately normal distributions. The following variables were negatively skewed with a maximum of 10 and were transformed via $-\sqrt{(10 - x)}$: informational content, repetition, reading irregular words, reading non-words, writing irregular words, writing non-words, and MDS. ASRS was right-skewed and was square-root transformed. To aid interpretation, we wanted higher values to correspond to "better" performance across all variables and therefore we reversed the sign for agrammatical and phonological errors. All variables were then standardized to have mean zero and SD 1. The clinical model was run using all subjects that had scores for all 14 speech and language variables ($n = 132$). Seven subjects who could complete the MRI were unable to complete the entire clinical battery used in the models due to severity, such as being mute. The alpha tuning parameter was set to 0.8 so that the penalty was primarily constraining the sum of the absolute value of the coefficients which favors a more parsimonious model, but also partly constraining the sum of the squared coefficients which helps stabilize coefficients when there are many are highly correlated variables (Hastie et al., 2001). To obtain a final model, we used ten-fold cross validation and selected the penalty so as to minimize cross-validation error, defined as the change in model deviance. We call this the "optimal" model and use it to calculate predicted probabilities. When the predicted probability was 0.5 or greater, we classify subjects as Aβ (+). The area under the ROC curve (AUROC) for these predictions, also known as the C statistic in logistic regression, was used to evaluate how well the models separated Aβ (+) from Aβ (−) subjects. Multivariable model fitting incorporating penalized coefficients or "shrinkage" is preferable to model selection strategies such as stepwise regression because the latter tends to over-fit and bias coefficients away from the null (Harrell, 2001).

In order to provide an imaging tool that could be useful clinically, we fit a binary classification tree using the rpart package in R to identify a parsimonious model that could serve as a rough surrogate, or first order approximation, to the full penalized logistic regression model described above (Therneau and Atkinson, 1997). In fitting the tree model we considered as predictors only those regions selected by the penalized logistic model. For each of these candidate regions atrophy was classified as mild, moderate, or severe. This was done by first converting the regional volume divided by gray matter volume to a z score and then trichotomizing these z scores based on cut-offs corresponding to 33.3rd and 66.7th percentiles of a standard normal distribution (z scores of −0.43 and +0.43). The classification tree algorithm is recursive such that for any set of subjects it identifies which region can be used to best divide subjects into relatively homogeneous groups in terms of their Aβ status. Cross-validation was used to "prune" the tree model to prevent overfitting.

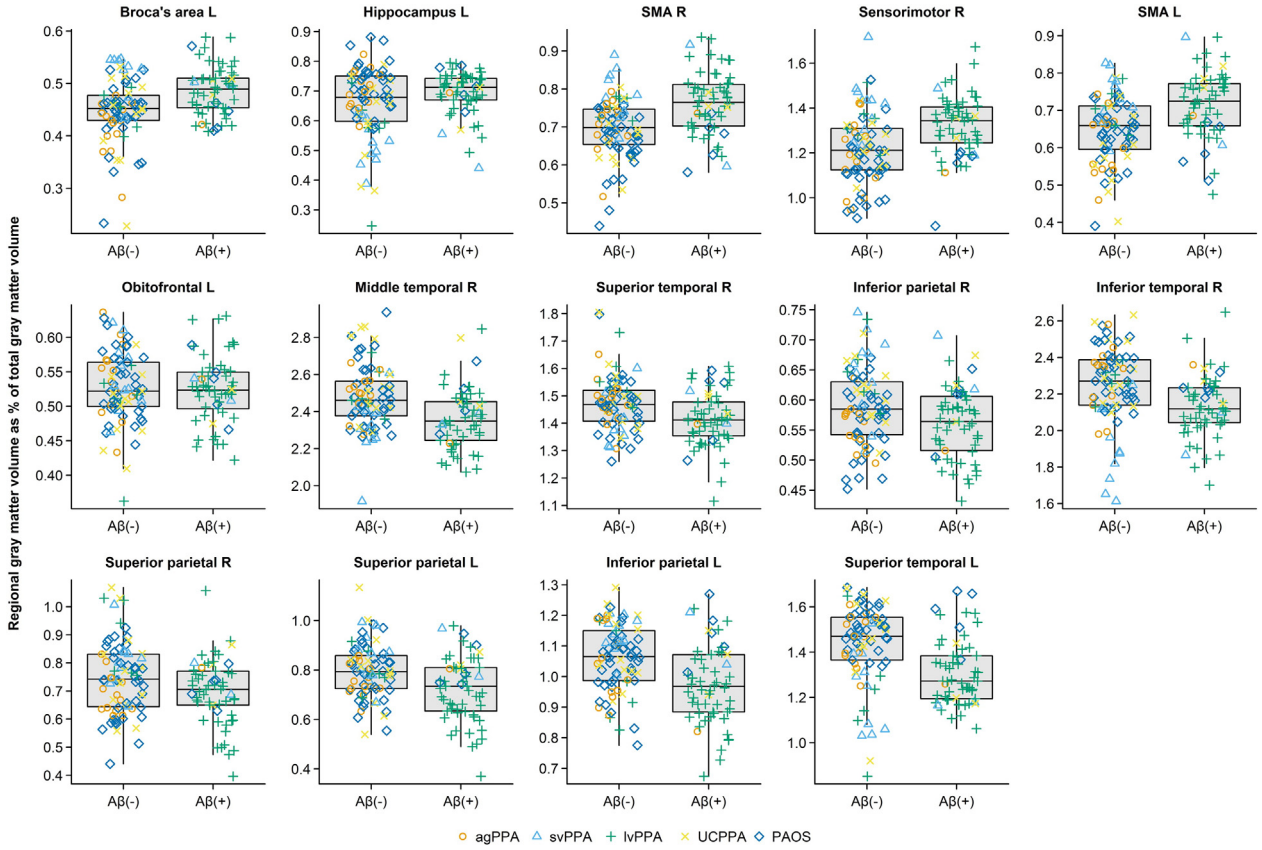

Fig. 3. Dot plots showing regional gray matter volume expressed as a percentage of total gray matter volume stratified by Aβ status for each of 14 regions included in the optimal penalized logistic regression model. The distributions for all Aβ (−) and Aβ (+) subjects are summarized with standard box plots in the middle of each panel indicating the median, quartiles, and having whiskers extending out to 1.5 times the inter-quartile range. Clinical diagnoses are indicated with different symbols.

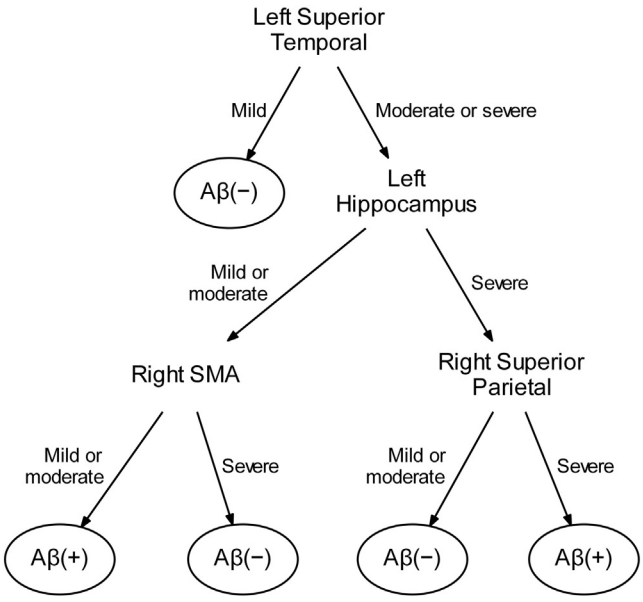

Fig. 4. Binary classification tree to predict Aβ status based on atrophy levels (mild, moderate or severe) in four regions. If the left superior temporal gyrus shows no atrophy or only mild atrophy then the case would be classified as Aβ (−). However, if the left superior temporal gyrus shows moderate or severe atrophy, one should consider the left hippocampus, right SMA and right superior parietal lobe. If the left hippocampus and right superior parietal lobe are both severely atrophic then the case would be classified as Aβ (+). If the left hippocampus is not severely atrophic then the case can only be classified as Aβ (+) if the right SMA is relatively spared. SMA = supplementary motor area.

3. Results

The demographic features of the cohort are shown in Table 1. Of the 139 subjects in the study, 58 (42%) were Aβ (+). There was no difference in age at onset, age at examination, gender or education between the Aβ (+) and Aβ (−) subjects, although the Aβ (+) subjects had slightly longer disease duration. The Aβ (+) subjects performed worse on the Mini-Mental State Examination and clinical dementia rating scale, and had a higher proportion of apolipoprotein e4 carriers than the Aβ (−) subjects. Three Aβ (−) subjects had mutations in progranulin and one Aβ (−) subject had a *C9ORF72* repeat expansion. The majority (81%) of the Aβ (+) subjects were diagnosed with lvPPA, whereas the different clinical diagnoses were more evenly split across the Aβ (−) group. The proportion of Aβ (+) subjects differed across the different clinical syndromes ($p < 0.0001$), with a high proportion of Aβ (+) subjects observed in the lvPPA group (89%), and a lower proportion observed in the UCPPA (19%), svPPA (15%), PAOS (12%) and agPPA (7%) groups.

The optimal penalized binary logistic regression models for MRI and clinical are shown in Fig. 1. Classification accuracy of both models is shown in Fig. 2. The MRI model accurately classified the Aβ status of 113/139 (81%) subjects using 14 variables. Classification was achieved with an AUROC of 0.87, sensitivity of 69% (40/58 Aβ (+) correctly classified) and specificity of 90% (73/81 Aβ (−) correctly classified). The model included a combination of regions in which smaller volumes predicted Aβ (+) status (i.e. negative coefficients) and regions in which larger volumes predicted Aβ (+) status (i.e. positive coefficients). The regions in which small volumes predicted Aβ (+) included left superior temporal gyrus, left inferior and superior parietal lobes, right inferior and superior parietal lobe, right inferior, middle and superior temporal gyri and left orbitofrontal cortex. The regions in which larger volumes

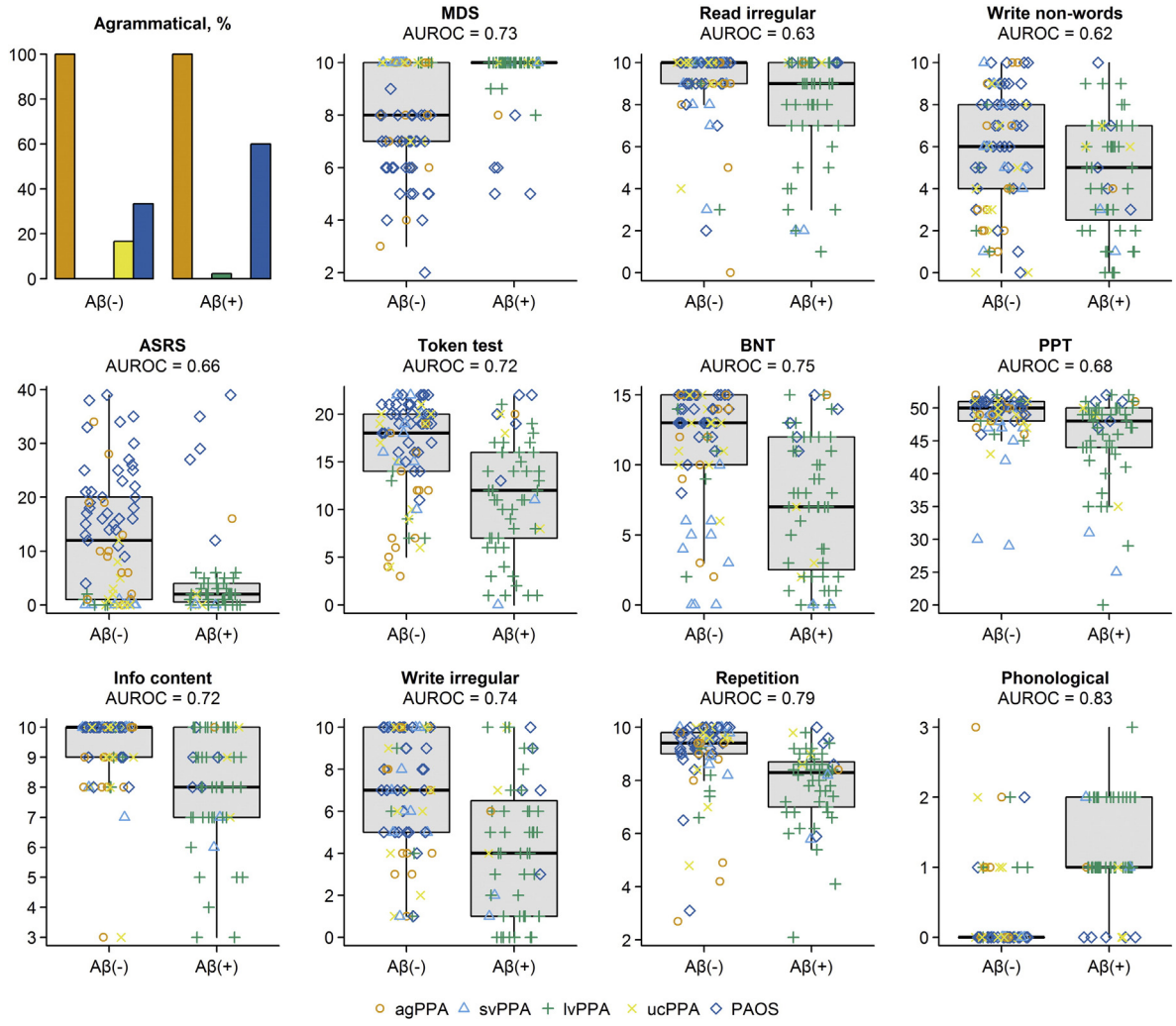

Fig. 5. Dot plots and bar charts showing the 12 speech and language tests included in the optimal penalized logistic regression model stratified by Aβ status. The distributions for all Aβ (−) and Aβ (+) subjects are summarized with standard box plots in the middle of each panel indicating the median, quartiles, and having whiskers extending out to 1.5 times the inter-quartile range. Clinical diagnoses are indicated with different symbols. MDS = motor speech disorders scale; BNT = Boston naming test; PPT = pyramids and palm trees test.

predicted Aβ (+) included left Broca's area, left hippocampus, left and right supplementary motor area and right sensorimotor cortex. Dot-plots showing each of these 14 regions stratified by Aβ status are shown in Fig. 3. In addition, a classification tree based on this model is shown in Fig. 4. As with the full model, this classification tree results in an accuracy of 81% in predicting Aβ status.

The clinical model accurately classified the Aβ status of 109/132 (83%) subjects using 12 clinical tests. Classification was achieved with an AUROC of 0.93, sensitivity of 78% (43/55 Aβ (+) correctly classified) and specificity of 86% (66/77 Aβ (−) correctly classified). More severe phonological errors, repetition deficits, abnormalities in writing irregular words, poor performance on the information content subtest of the WAB, poor performance on the PPT, poor performance on the BNT, preserved motor speech, preserved reading of irregular words and writing non-words, and absence of agrammatism in speech predicted Aβ (+). The Token Test also contributed to the model. Dot-plots showing each of these 12 clinical tests stratified by Aβ status are shown in Fig. 5.

A preliminary analysis assessing performance of the models within each clinical diagnostic group is shown in Fig. 6 and Table 2. Within the lvPPA group, the MRI model correctly classified 5/6 of the Aβ (−) subjects, but falsely classified 10 subjects as Aβ (−). In contrast, the clinical model classified nearly all lvPPA subjects as Aβ (+). Neither model correctly classified the Aβ (+) subjects within the agPPA and PAOS groups. The clinical model performed somewhat better in the svPPA group correctly classifying both Aβ (+) subjects. The MRI model was able to correctly classify 14/16 (87.5%) UCPPA subjects compared to only 10/15 (66.7%) in the clinical model.

Performance of both models can be compared to the performance of clinical diagnosis. If one assumes based on previous studies that lvPPA subjects will be Aβ (+) and the svPPA, agPPA and PAOS subjects will be Aβ (−), then clinical diagnosis accurately classifies 109/123 (89%) subjects. The 16 UCPPA subjects cannot be included in this analysis since we did not know the Aβ status of these subjects a priori and hence could not predict whether they would be Aβ (+) or Aβ (−). In comparison, excluding the UCPPA subjects, the MRI model accurately classified 99/123 (80%) of subjects and the clinical model accurately classified 99/117 (85%) of subjects.

4. Discussion

This study assesses the value of MRI and clinical data in predicting Aβ status in a large cohort of 139 PPA and PAOS subjects. We showed that the optimal MRI and clinical models performed comparably in overall prediction of the presence of Aβ deposition, correctly classifying Aβ status in 81% and 83% of subjects respectively, although the prediction accuracy of clinical diagnosis alone outperformed both models. The models, particularly the MRI model, may, however, still be of value in subjects that cannot be diagnosed clinically.

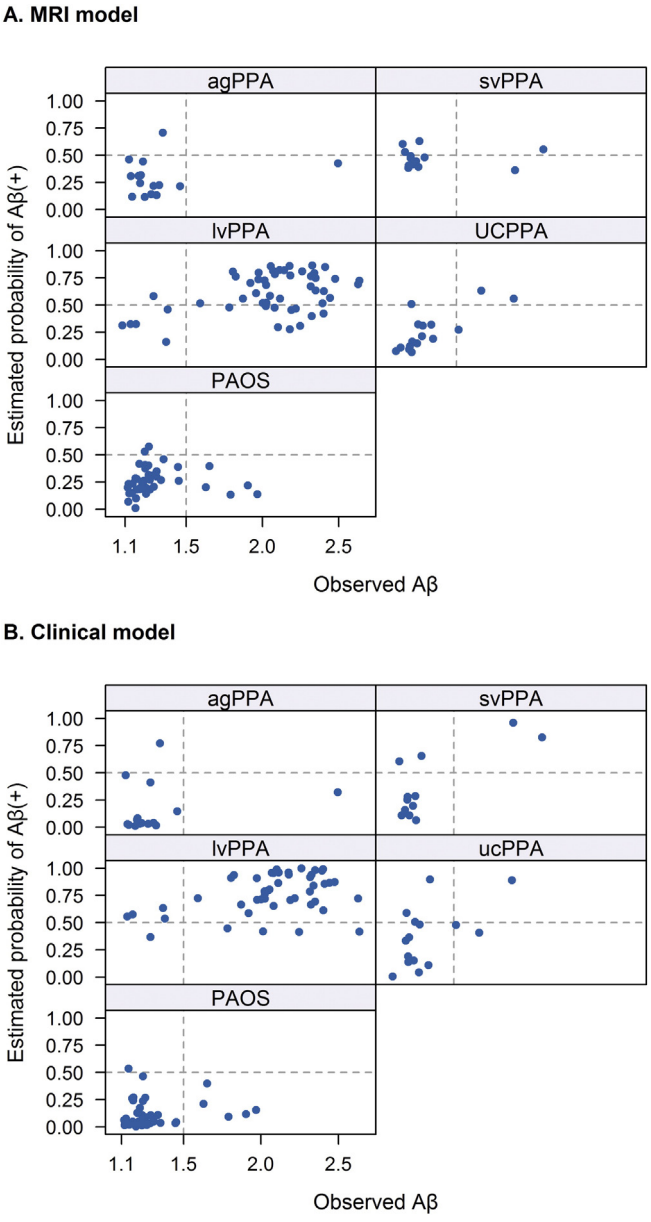

Fig. 6. Scatter plots showing the estimated probability of Aβ (+) from the optimal penalized logistic regression model (y-axis) versus the observed Aβ value (x-axis) by clinical diagnosis. Panel A shows the results for the MRI model and panel B shows the results for the clinical model. A vertical line at 1.5 defines Aβ status while a horizontal line at 0.50 indicates the cut-point for predicting Aβ (+). Subjects in the top right or lower left quadrants are correctly classified by the model.

Table 2
Classification accuracy within each diagnostic group.

Dx	Observed Aβ	MRI model classification		Clinical model classification*	
		Aβ (+)	Aβ (−)	Aβ (+)	Aβ (−)
agPPA	Aβ (+)	0	1	0	1
	Aβ (−)	1	13	1	13
svPPA	Aβ (+)	1	1	2	0
	Aβ (−)	3	8	2	8
lvPPA	Aβ (+)	37	10	40	4
	Aβ (−)	1	5	4	1
UCPPA	Aβ (+)	2	1	1	2
	Aβ (−)	1	12	3	9
PAOS	Aβ (+)	0	5	0	5
	Aβ (−)	2	35	1	35

* The clinical model excluded seven subjects that could not complete all testing.

The MRI model identified 14 regions which provided optimum classification of Aβ status. Involvement of both left and right temporoparietal regions and left orbitofrontal cortex, with a sparing of Broca's area, supplementary motor cortex, left hippocampus and sensorimotor cortex, was predictive of the presence of Aβ deposition. These findings concord with previous smaller pathological studies showing that temporoparietal atrophy is associated with the presence of AD pathology in subjects presenting with atypical clinical syndromes (Lehmann et al., 2010; Lee et al., 2011; Whitwell et al., 2010, 2011). Interestingly, in the current study, both left and right temporoparietal regions were useful for classification, despite the fact that PPA is predominantly associated with involvement of the left hemisphere. Subjects with Aβ deposition are, therefore, more likely to show involvement of both temporoparietal cortices compared to subjects without Aβ deposition. Temporoparietal atrophy has, indeed, been observed in lvPPA subjects with AD pathology (Gorno-Tempini et al., 2004; Madhavan et al., 2013), with atrophy typically starting in the left hemisphere but then spreading to involve the right hemisphere over time (Rohrer et al., 2013). Of note, only lateral temporoparietal regions were predictive of Aβ status, with precuneus and posterior cingulate not adding any additional discriminative power and hence not appearing in the 14 regions that provided optimum classification, despite the fact that they can also be atrophic in lvPPA. It is likely that these regions are less useful because they are typically affected to a lesser degree than lateral temporoparietal regions in lvPPA (Ridgway et al., 2012; Botha et al., 2015). The fact that large volumes of Broca's area, supplementary motor area, sensorimotor cortex and left hippocampus were predictive in the model, may either suggest that when these regions are involved, subjects are less likely to have Aβ deposition, or may suggest that these regions are spared in subjects with Aβ deposition; or perhaps a combination of both. Broca's area is typically atrophic in agPPA (Gorno-Tempini et al., 2004; Whitwell et al., 2013), the supplementary motor area is typically atrophic in PAOS (Josephs et al., 2012, 2013; Whitwell et al., 2013), and the left hippocampus is usually atrophic in svPPA (Chan et al., 2001); all of which are usually associated with non-AD pathology. The sensorimotor cortex can become involved in PAOS (Josephs et al., 2013), but is also a region that is typically spared in subjects with AD pathology (Whitwell et al., 2007). Our model is therefore somewhat dependent on the composition of our Aβ negative cohort. Fig. 4 shows a classification tree that provides an approximation of the larger imaging model to provide a tool that could be clinically useful for predicting Aβ status. Assessing the degree of atrophy of the left superior temporal gyrus, left hippocampus, right supplementary motor area and right superior parietal lobe can help a clinician to predict Aβ status with the same degree of accuracy as the full model.

The clinical model identified 12 speech and language variables which provided optimum classification of Aβ status. The variables which were most strongly associated with Aβ deposition were phonological errors, the presence of deficits in repeating complex sentences and errors in writing irregular words. The absence of agrammatism, the absence of errors in reading irregular words and the absence of motor speech abnormalities were also highly predictive of Aβ deposition. The presence of phonological errors and repetition deficits are features that are typically associated with lvPPA, and absent in agPPA, svPPA and PAOS, as we observed in our cohort (Supplemental Table 1). Deficits in spelling irregular words have been previously observed in PPA (Sepelyak et al., 2011), and the lvPPA subjects in our cohort performed poorly on this task (Supplemental Table 1). Conversely, the presence of agrammatism and motor speech abnormalities are features of agPPA and PAOS that are typically absent in lvPPA, as we observed in our cohort (Supplemental Table 1). Deficits in reading irregular words have been associated with svPPA (Patterson et al., 2001). A couple of other studies have also investigated clinical features that help predict Aβ deposition. Despite the fact that the cohorts in these

studies differed from ours in terms of the clinical diagnoses included, they also found that phonological errors were highly predictive of Aβ deposition (Chare et al., 2014; Leyton et al., 2014).

Performance of the MRI and clinical models differed when each specific clinical diagnosis was examined. In subjects with lvPPA, the clinical model tended to predict that all subjects will be Aβ (+), while the MRI model could identify 5/6 of the Aβ (−) subjects, but also made a number of false Aβ (−) calls. Indeed, we have previously shown that Aβ (−) lvPPA has a different atrophic signature to Aβ (+) lvPPA, with more highly asymmetric atrophy and greater involvement of the anterior temporal lobes (Whitwell et al., 2015). However, the clinical model performed better than the MRI model in svPPA, correctly classifying both Aβ (+) subjects with only two false positives. The Aβ (+) svPPA subjects were likely correctly classified because both showed mild phonological errors and repetition deficits (Supplemental Table 1). The presence of these clinical features could therefore be an important clue to the presence of Aβ deposition in svPPA. The number of Aβ (+) svPPA subjects was, however, small and so these findings will need to be validated in a larger cohort. Neither the MRI nor clinical model was able to predict the few subjects in the agPPA or PAOS groups who were Aβ (+). This could suggest either that the Aβ (+) subjects in these groups do not have a different MRI or clinical signature than the typical Aβ (−) subjects, or that our model was unable to identify subtle differences in these few subjects, possibly due to the small number of subjects. We favor the first explanation given that we could not identify any discernible differences in the patterns of atrophy or clinical performance on visual inspection. It is possible that Aβ deposition in these agPPA and PAOS subjects may represent a secondary pathology, possibly a diffuse Aβ plaque pathology (Kantarci et al., 2012), co-occurring with FTLD pathology (Kertesz et al., 2005; Mesulam et al., 2014), explaining why it does not significantly alter the pattern of neurodegeneration as previously hypothesized (Josephs et al., 2014). Alternatively, the topographic distribution of AD pathology could be atypical in these cases. Autopsy confirmation will be needed to determine the underlying pathology in these subjects.

Interestingly, the MRI model performed well in the UCPPA subjects. Of the 16 UCPPA subjects, three were Aβ (+) and the model correctly predicted two of these three. The third case was, however, only marginally Aβ (+) with a SUVR of 1.51, and hence the Aβ status was borderline. The model also correctly classified 12 (92%) of the 13 Aβ (−) UCPPA subjects, with the final case given a borderline estimated Aβ probability of 0.500 by the model. Therefore, despite the fact that the clinical diagnosis is uncertain, the presence of Aβ is associated with specific MRI imaging signatures in these cases. Patterns of atrophy could therefore be useful to help determine which of these cases will likely have Aβ deposition. This information could then be used to select which individuals should undergo more costly Aβ PET scanning. This was particularly important since in many of the UCPPA cases the clinical dilemma was between lvPPA and the Aβ (−) diagnoses. The clinical model did not perform as well, correctly classifying only one of the Aβ (+) subjects and 75% of the Aβ (−) subjects, suggesting that clinical data will not be as useful in the UCPPA subjects. Since speech and language tests are used to help guide clinical diagnosis this is perhaps unsurprising.

Overall performance of the MRI and clinical models was good with similar predictive accuracy, with the MRI model correctly classifying 81% of subjects and the clinical model correctly classifying 83% of subjects. However, both models were outperformed by clinical diagnosis which could correctly classify 89% of subjects. One thing to consider, however, when using clinical diagnosis alone is that it will be impossible to identify discordant cases, i.e. Aβ (−) lvPPA subjects and Aβ (+) agPPA, svPPA and PAOS subjects. Performance of clinical diagnosis in our cohort was excellent because we had relatively few discordant cases, but other cohorts have reported large proportions of discordant cases. Studies have reported that up to 46% of lvPPA subjects do not have Aβ deposition (Harris et al., 2013) and, for example, up to 30% of agPPA subjects do have Aβ deposition (Chare et al., 2014). In fact, in one recent large clinicopathological study, a clinical diagnosis of lvPPA had sensitivity of only 50% and a specificity of 71% in identifying AD pathology within subjects with PPA (Mesulam et al., 2014). Furthermore, utilizing clinical diagnosis would make it impossible to be able to predict Aβ status in the UCPPA subjects that do not fulfill clinical criteria for one of the PPA variants, which can be a large proportion of subjects (Sajjadi et al., 2012; Wicklund et al., 2014).

A major strength of our study is the large cohort of subjects with Aβ PET imaging and MRI. We included all prospectively recruited subjects that met clinical inclusion criteria in our study and did not exclude left-handed subjects or impose quantitative cut-offs, in order that our results would generalize as much as possible to other cohorts of PPA and PAOS subjects. A number of our subjects showed poor scores on the Mini-Mental State Examination (Supplemental Table 1) suggesting cognitive impairment. However, in many instances performance was affected by naming difficulties or the presence of apraxia of speech. Nevertheless, our cohort is consistent with cohorts from other centers that similarly show low Mini-Mental State Examination scores in PPA, particularly in lvPPA (Sajjadi et al., 2012; Lehmann et al., 2013; Leyton et al., 2013). It will, however, be important to validate our findings in other cohorts to determine whether our findings generalize to cohorts with different diagnostic compositions and different proportions of subjects with Aβ deposition. We also adjusted all regional volumes by total gray matter volume to ensure our results were not driven by differences in severity. A limitation of the study is the small number of discordant cases within each clinical diagnosis. The analysis of how the models perform in these cases should therefore be considered preliminary and will need to be validated in larger cohorts. Another possible limitation was that we did not assess white matter, instead choosing to focus on gray matter structures. Future studies will be needed to determine whether the inclusion of white matter metrics and also, potentially, CSF biomarkers, may help further improve classification.

5. Conclusions

These findings demonstrate that neuroimaging and speech and language data both have the potential to be useful to predict the presence of Aβ deposition, and could prove to be useful if Aβ imaging is unavailable. While clinical diagnosis provided superior prediction at the group level, these models, particularly the neuroimaging model, may be valuable to predict Aβ status in UCPPA.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nicl.2016.01.014.

Acknowledgments

This study was funded by NIH grants R01 DC010367 (PI Josephs) and R01 DC012519 (PI Whitwell). The sponsor played no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

References

Baker, M., Mackenzie, I.R., Pickering-Brown, S.M., Gass, J., Rademakers, R., Lindholm, C., et al., 2006. Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. Nature 442 (7105), 916–919.

Botha, H., Duffy, J.R., Whitwell, J.L., Strand, E.A., Machulda, M.M., Schwarz, C.G., et al., 2015. Classification and clinicoradiologic features of primary progressive aphasia (PPA) and apraxia of speech. Cortex 69, 220–236.

Chan, D., Fox, N.C., Scahill, R.I., Crum, W.R., Whitwell, J.L., Leschziner, G., et al., 2001. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Ann. Neurol. 49 (4), 433–442.

Chare, L., Hodges, J.R., Leyton, C.E., McGinley, C., Tan, R.H., Kril, J.J., et al., 2014. New criteria for frontotemporal dementia syndromes: clinical and pathological diagnostic implications. J. Neurol. Neurosurg. Psychiatry 85 (8), 865–870.

Dejesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., et al., 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron.

DeRenzi, E., Vignolo, L.A., 1962. The token test: a sensitive measure to detect receptive disturbances in aphasics. Brain 85, 665–678.
Flanagan, E.P., Baker, M.C., Perkerson, R.B., Duffy, J.R., Strand, E.A., Whitwell, J.L., et al., 2015. Dominant frontotemporal dementia mutations in 140 cases of primary progressive aphasia and speech apraxia. Dement. Geriatr. Cogn. Disord. 39 (5–6), 281–286.
Gorno-Tempini, M.L., Dronkers, N.F., Rankin, K.P., Ogar, J.M., Phengrasamy, L., Rosen, H.J., et al., 2004. Cognition and anatomy in three variants of primary progressive aphasia. Ann. Neurol. 55 (3), 335–346.
Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., et al., 2011. Classification of primary progressive aphasia and its variants. Neurology 76 (11), 1006–1014.
Harrell, F.E., 2001. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. Springer, New York.
Harris, J.M., Gall, C., Thompson, J.C., Richardson, A.M., Neary, D., du Plessis, D., et al., 2013. Classification and pathology of primary progressive aphasia. Neurology 81 (21), 1832–1839.
Hastie, T., Tibshirani, R., Friedman, J.H., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction: with 200 Full-Color Illustrations. Springer, New York.
Howard, D., Patterson, K., 1992. The Pyramids and Palm Trees Test: A Test of Semantic Access from Words and Picture. Thames Valley Test Company.
Hutton, M., Lendon, C.L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., et al., 1998. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature 393 (6686), 702–705.
Jack Jr., C.R., Lowe, V.J., Senjem, M.L., Weigand, S.D., Kemp, B.J., Shiung, M.M., et al., 2008. 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. Brain 131 (Pt 3), 665–680.
Josephs, K.A., Duffy, J.R., Strand, E.A., Whitwell, J.L., Layton, K.F., Parisi, J.E., et al., 2006. Clinicopathological and imaging correlates of progressive aphasia and apraxia of speech. Brain 129 (Pt 6), 1385–1398.
Josephs, K.A., Duffy, J.R., Strand, E.A., Machulda, M.M., Senjem, M.L., Master, A.V., et al., 2012. Characterizing a neurodegenerative syndrome: primary progressive apraxia of speech. Brain J. Neurol. 135 (Pt 5), 1522–1536.
Josephs, K.A., Duffy, J.R., Strand, E.A., Machulda, M.M., Senjem, M.L., Lowe, V.J., et al., 2013. Syndromes dominated by apraxia of speech show distinct characteristics from agrammatic PPA. Neurology 81 (4), 337–345.
Josephs, K.A., Duffy, J.R., Strand, E.A., Machulda, M.M., Senjem, M.L., Lowe, V.J., et al., 2014. APOE epsilon4 influences beta-amyloid deposition in primary progressive aphasia and speech apraxia. Alzheimers Dement.
Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., et al., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. NeuroImage 30 (2), 436–443.
Kantarci, K., Yang, C., Schneider, J.A., Senjem, M.L., Reyes, D.A., Lowe, V.J., et al., 2012. Antemortem amyloid imaging and beta-amyloid pathology in a case with dementia with Lewy bodies. Neurobiol. Aging 33 (5), 878–885.
Kertesz, A., 2007. Western Aphasia Battery (Revised). San Antonio, Tx, PsychCorp.
Kertesz, A., McMonagle, P., Blair, M., Davidson, W., Munoz, D.G., 2005. The evolution and pathology of frontotemporal dementia. Brain 128 (Pt 9), 1996–2005.
Lansing, A.E., Ivnik, R.J., Cullum, C.M., Randolph, C., 1999. An empirically derived short form of the Boston naming test. Arch. Clin. Neuropsychol. 14 (6), 481–487.
Lee, S.E., Rabinovici, G.D., Mayo, M.C., Wilson, S.M., Seeley, W.W., DeArmond, S.J., et al., 2011. Clinicopathological correlations in corticobasal degeneration. Ann. Neurol. 70 (2), 327–340.
Lehmann, M., Rohrer, J.D., Clarkson, M.J., Ridgway, G.R., Scahill, R.I., Modat, M., et al., 2010. Reduced cortical thickness in the posterior cingulate gyrus is characteristic of both typical and atypical Alzheimer's disease. J. Alzheimers Dis. 20 (2), 587–598.
Lehmann, M., Ghosh, P.M., Madison, C., Laforce Jr., R., Corbetta-Rastelli, C., Weiner, M.W., et al., 2013. Diverging patterns of amyloid deposition and hypometabolism in clinical variants of probable Alzheimer's disease. Brain 136 (Pt 3), 844–858.
Leyton, C.E., Villemagne, V.L., Savage, S., Pike, K.E., Ballard, K.J., Piguet, O., et al., 2011. Subtypes of progressive aphasia: application of the International Consensus Criteria and validation using beta-amyloid imaging. Brain J. Neurol. 134 (Pt 10), 3030–3043.
Leyton, C.E., Hsieh, S., Mioshi, E., Hodges, J.R., 2013. Cognitive decline in logopenic aphasia: more than losing words. Neurology 80 (10), 897–903.
Leyton, C.E., Ballard, K.J., Piguet, O., Hodges, J.R., 2014. Phonologic errors as a clinical marker of the logopenic variant of PPA. Neurology 82 (18), 1620–1627.
Madhavan, A., Whitwell, J.L., Weigand, S.D., Duffy, J.R., Strand, E.A., Machulda, M.M., et al., 2013. FDG PET and MRI in logopenic primary progressive aphasia versus dementia of the Alzheimer's type. PLoS One 8 (4), e62471.
Mesulam, M.M., 1982. Slowly progressive aphasia without generalized dementia. Ann. Neurol. 11 (6), 592–598.
Mesulam, M.M., 2001. Primary progressive aphasia. Ann. Neurol. 49 (4), 425–432.
Mesulam, M.M., Weintraub, S., Rogalski, E.J., Wieneke, C., Geula, C., Bigio, E.H., 2014. Asymmetry and heterogeneity of Alzheimer's and frontotemporal pathology in primary progressive aphasia. Brain 137 (Pt 4), 1176–1192.
Patterson, K., Lambon Ralph, M.A., Hodges, J.R., McClelland, J.L., 2001. Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. Neuropsychologia 39 (7), 709–724.
Rabinovici, G.D., Jagust, W.J., Furst, A.J., Ogar, J.M., Racine, C.A., Mormino, E.C., et al., 2008. Abeta amyloid and glucose metabolism in three variants of primary progressive aphasia. Ann. Neurol. 64 (4), 388–401.
Ridgway, G.R., Lehmann, M., Barnes, J., Rohrer, J.D., Warren, J.D., Crutch, S.J., et al., 2012. Early-onset Alzheimer disease clinical variants: multivariate analyses of cortical thickness. Neurology 79 (1), 80–84.
Rohrer, J.D., Caso, F., Mahoney, C., Henry, M., Rosen, H.J., Rabinovici, G., et al., 2013. Patterns of longitudinal brain atrophy in the logopenic variant of primary progressive aphasia. Brain Lang.
Sajjadi, S.A., Patterson, K., Arnold, R.J., Watson, P.C., Nestor, P.J., 2012. Primary progressive aphasia: a tale of two syndromes and the rest. Neurology 78 (21), 1670–1677.
Sepelyak, K., Crinion, J., Molitoris, J., Epstein-Peterson, Z., Bann, M., Davis, C., et al., 2011. Patterns of breakdown in spelling in primary progressive aphasia. Cortex 47 (3), 342–352.
Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging 17 (1), 87–97.
Strand, E.A., Duffy, J.R., Clark, H.M., Josephs, K., 2014. The Apraxia of Speech Rating Scale: a tool for diagnosis and description of apraxia of speech. J. Commun. Disord. 51, 43–50.
Therneau, T.M., Atkinson, E.J., 1997. An Introduction to Recursive Partitioning Using the Rpart Routines. Department of Health Sciences Research, Mayo Clinic.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15 (1), 273–289.
Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Knopman, D.S., Boeve, B.F., Petersen, R.C., et al., 2007. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain J. Neurol. 130 (Pt 7), 1777–1786.
Whitwell, J.L., Jack Jr., C.R., Boeve, B.F., Parisi, J.E., Ahlskog, J.E., Drubach, D.A., et al., 2010. Imaging correlates of pathology in corticobasal syndrome. Neurology 75 (21), 1879–1887.
Whitwell, J.L., Jack Jr., C.R., Przybelski, S.A., Parisi, J.E., Senjem, M.L., Boeve, B.F., et al., 2011. Temporoparietal atrophy: a marker of AD pathology independent of clinical diagnosis. Neurobiol. Aging 32 (9), 1531–1541.
Whitwell, J.L., Weigand, S.D., Boeve, B.F., Senjem, M.L., Gunter, J.L., DeJesus-Hernandez, M., et al., 2012. Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. Brain 135 (Pt 3), 794–806.
Whitwell, J.L., Duffy, J.R., Strand, E.A., Xia, R., Mandrekar, J., Machulda, M.M., et al., 2013. Distinct regional anatomic and functional correlates of neurodegenerative apraxia of speech and aphasia: an MRI and FDG-PET study. Brain Lang. 125 (3), 245–252.
Whitwell, J.L., Duffy, J.R., Strand, E.A., Machulda, M.M., Senjem, M.L., Schwarz, C.G., et al., 2015. Clinical and neuroimaging biomarkers of amyloid-negative logopenic primary progressive aphasia. Brain Lang. 142, 45–53.
Wicklund, M.R., Duffy, J.R., Strand, E.A., Machulda, M.M., Whitwell, J.L., Josephs, K.A., 2014. Quantitative application of the primary progressive aphasia consensus criteria. Neurology 82 (13), 1119–1126.
Yorkson, K., Strand, E.A., Miller, R., Hillel, A., Smith, K., 1993. Speech deterioration in amyotrophic lateral sclerosis: implications for the timing of intervention. J. Med. Speech-Lang. Pathol. 1 (1), 35–46.