NeuroImage: Clinical 14 (2017) 334–343

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Predicting primary progressive aphasias with support vector machine approaches in structural MRI data

Sandrine Bisenius[a,*], Karsten Mueller[a], Janine Diehl-Schmid[b], Klaus Fassbender[c], Timo Grimmer[b], Frank Jessen[d], Jan Kassubek[e], Johannes Kornhuber[f], Bernhard Landwehrmeyer[e], Albert Ludolph[e], Anja Schneider[g], Sarah Anderl-Straub[e], Katharina Stuke[a], Adrian Danek[h], Markus Otto[e], Matthias L. Schroeter[a], &, FTLDc study group:

[a]Max Planck Institute for Human Cognitive and Brain Sciences & Clinic for Cognitive Neurology, University Hospital Leipzig, Germany
[b]Clinic and Polyclinic for Psychiatry & Psychotherapy, Technical University Munich, Germany
[c]Clinic and Polyclinic for Neurology, Saarland University Homburg, Germany
[d]Clinic and Polyclinic for Psychiatry and Psychotherapy, University of Bonn, Germany
[e]Department of Neurology, University of Ulm, Germany
[f]Clinic for Psychiatry and Psychotherapy, Friedrich-Alexander University Erlangen-Nuremberg, Germany
[g]Clinic for Psychiatry and Psychotherapy, University of Goettingen, Germany
[h]Clinic of Neurology, Ludwig Maximilian University of Munich, Germany

ARTICLE INFO

Article history:
Received 14 September 2016
Received in revised form 27 January 2017
Accepted 3 February 2017
Available online 6 February 2017

Keywords:
Grey matter
Multi-center
Primary progressive aphasia
Support vector machine classification
Whole brain approach

ABSTRACT

Primary progressive aphasia (PPA) encompasses the three subtypes nonfluent/agrammatic variant PPA, semantic variant PPA, and the logopenic variant PPA, which are characterized by distinct patterns of language difficulties and regional brain atrophy. To validate the potential of structural magnetic resonance imaging data for early individual diagnosis, we used support vector machine classification on grey matter density maps obtained by voxel-based morphometry analysis to discriminate PPA subtypes (44 patients: 16 nonfluent/agrammatic variant PPA, 17 semantic variant PPA, 11 logopenic variant PPA) from 20 healthy controls (matched for sample size, age, and gender) in the cohort of the multi-center study of the German consortium for frontotemporal lobar degeneration. Here, we compared a whole-brain with a meta-analysis-based disease-specific regions-of-interest approach for support vector machine classification. We also used support vector machine classification to discriminate the three PPA subtypes from each other. Whole brain support vector machine classification enabled a very high accuracy between 91 and 97% for identifying specific PPA subtypes vs. healthy controls, and 78/95% for the discrimination between semantic variant vs. nonfluent/agrammatic or logopenic PPA variants. Only for the discrimination between nonfluent/agrammatic and logopenic PPA variants accuracy was low with 55%. Interestingly, the regions that contributed the most to the support vector machine classification of patients corresponded largely to the regions that were atrophic in these patients as revealed by group comparisons. Although the whole brain approach took also into account regions that were not covered in the regions-of-interest approach, both approaches showed similar accuracies due to the disease-specificity of the selected networks. Conclusion, support vector machine classification of multi-center structural magnetic resonance imaging data enables prediction of PPA subtypes with a very high accuracy paving the road for its application in clinical settings.

1. Introduction

Primary progressive aphasia (PPA) is a neurodegenerative disease with insidious onset mainly characterized by a language dysfunction that remains isolated for at least two years without significant impairment in other cognitive domains (Gorno-Tempini et al., 2011; Mesulam, 1982; Neary et al., 1998). PPA subsumes three gradually progressive language disorders, namely semantic variant PPA (svPPA) or semantic dementia, nonfluent/agrammatic variant PPA (nfvPPA) or

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1A, 04103 Leipzig, Germany.
E-mail address: bisenius@cbs.mpg.de (S. Bisenius).
[1] FTLDc study group: Marie Fischer (Erlangen), Anke Hammer (Erlangen), Manuel Maler (Erlangen), Timo Oberstein (Erlangen), Maryna Polyakova (Leipzig), Tanja Richter-Schmidinger (Erlangen), Carola Roßmeier (Munich), Dorothee Saur (Leipzig), Katharina Schümberg (Leipzig), Elisa Semler (Ulm), Ingo Uttner (Ulm), Christine v. Arnim (Ulm).

http://dx.doi.org/10.1016/j.nicl.2017.02.003

progressive nonfluent aphasia, and logopenic variant PPA (lvPPA) or logopenic progressive aphasia (Gorno-Tempini et al., 2008; Gorno-Tempini et al., 2004; Gorno-Tempini et al., 2011). SvPPA is mainly characterized by impairments in confrontation naming, single-word comprehension, and object-knowledge, as well as surface dyslexia or dysgraphia (Gorno-Tempini et al., 2011). The imaging supported diagnosis of svPPA is given when patients additionally show atrophy and/or hypometabolism in the anterior (ventral and lateral) temporal lobe. Patients suffering from nfvPPA show predominantly agrammatism, effortful halting speech with inconsistent speech sound errors and distortions (apraxia of speech), and impaired comprehension of syntactically complex sentences. These language deficits are often associated with atrophy or hypometabolism in left inferior frontal gyrus, insula, premotor, and supplementary motor areas. LvPPA is characterized by impaired single-word retrieval in spontaneous speech and naming as well as impaired repetition of sentences. Patients suffering from lvPPA furthermore often show phonologic paraphasias in spontaneous speech and naming. The imaging supported diagnosis of lvPPA is given when patients additionally show atrophy and/or hypometabolism in left posterior parietal, supramarginal, and angular gyri (Gorno-Tempini et al., 2011). The suggested imaging criteria have recently been validated by comprehensive meta-analyses (Bisenius et al., 2016).

The prevalence of PPA is roughly estimated to range from 3 to 15/100,000 in the US population (Grossman, 2010; Harvey et al., 2003; Ratnavalli et al., 2002). PPA is thus a rare disease, which makes it very difficult for neurologists outside specialized clinics to correctly recognize and differentiate between the three PPA variants in routine hospital practice (e.g., Wilson et al., 2009). Given that the current imaging criteria are only supportive for the diagnosis of PPA, magnetic resonance imaging (MRI) scans are often not included as standard in the clinical assessment of PPA, but mainly used to exclude differential diagnoses (e.g., Wilson et al., 2009). It has been shown for other types of dementia like for instance, AD that changes in atrophy as visualized by MRI are an especially good biomarker for correct early diagnosis and furthermore even predictive for individuals with mild cognitive impairment to decline into AD (e.g., Frisoni et al., 2010; McEvoy and Brewer, 2010; Schroeter et al., 2009; Weiner et al., 2010). Therefore, it might be highly interesting to investigate whether MRI scans have a similar predictive value for the correct early diagnosis of PPA and to investigate which brain regions contribute the most to the classification of its three variants.

On the one hand, it seems plausible that brain regions proposed in the current diagnostic imaging criteria and based on a large range of imaging studies (Bisenius et al., 2016; Gorno-Tempini et al., 2011) enable the correct diagnosis of the three PPA variants. On the other hand, most of the current imaging studies report comparisons of the three variants of PPA with age-matched healthy controls at a group-level and it has been critically discussed that statistical differences at group level might not necessarily reveal the most important regions to correctly diagnose individual cases (Davatzikos et al., 2008a; Davatzikos et al., 2008b; Wilson et al., 2009). Therefore, it might advance our knowledge in this field crucially to investigate whether brain regions contributing the most to the correct diagnosis of the three PPA variants indeed correspond to regions that are especially atrophied in these three variants. Moreover, it might be highly interesting to explore whether disease-specific regions of interest (ROIs) in comparison with whole brain approaches even enhance the predictive power of MRI scans for the correct PPA classification.

To address these issues, we investigated here atrophy, namely changes in grey matter density, with voxel-based morphometry (VBM) in patients suffering from one of the PPA variants in comparison with healthy controls as well as by comparing patients with different PPA variants at a group level. Subsequently, we used linear support vector machine (SVM) classification of the individual grey matter density maps to investigate their discriminative or predictive power for the correct classification of single subjects as belonging to one of the PPA variants or healthy controls. A number of recent studies have used similar pattern classification methods to classify patients with AD, FTD, and mild cognitive impairment (Davatzikos et al., 2008a; Davatzikos et al., 2008b; Dukart et al., 2013; Dukart et al., 2011; Fan et al., 2008; Klöppel et al., 2008b; Lerch et al., 2008; Misra et al., 2009; Teipel et al., 2007; Vemuri et al., 2008). Wilson et al. (2009) investigated the utility of structural MRI scans for SVM classification in PPA variants in a single center study. Here, we investigated patients included in the multi-center study of the German consortium for frontotemporal lobar degeneration (FTLD; Otto et al., 2011) to replicate and generalize previously reported results, where the multi-center design is a precondition for application in clinical routine in the future. Additionally, we compared a whole-brain approach to a disease-specific ROI approach based on comprehensive anatomical likelihood estimation meta-analyses on the three variants of PPA (Bisenius et al., 2016). These ROIs represent the prototypical networks consistently affected in the three variants of PPA across MRI studies reporting group-level statistics. Note that these ROIs are based on a totally different cohort avoiding circularity. In order to better understand possible differences between the whole brain and the regions-of-interest approach, we furthermore computed and visualized the voxels that contributed the most to the SVM classification in the whole brain approach. To reveal whether the brain regions that contributed the most to the SVM classification in the whole brain approach corresponded to the regions that were especially atrophic in the three PPA variants, we also report pairwise group-level comparisons of grey matter probability maps between patients and healthy controls, respectively between PPA variants. For the pairwise group-level comparisons, we hypothesized that, according to the current imaging criteria and previously published VBM studies, atrophy is focused to left fronto-insular regions in nfvPPA, to the (mainly left) anterior temporal lobe in svPPA, and to the (predominantly left) posterior perisylvian or parietal cortex in lvPPA (e.g., Bisenius et al., 2016; Desgranges et al., 2007; Gorno-Tempini et al., 2011; Grossman et al., 2004; Mummery et al., 2000). Furthermore, we hypothesized that the same brain regions would mainly contribute to the correct SVM classification in PPA variants and healthy controls and that disease-specific ROI approaches would reveal a higher predictive power for the SVM classification than whole-brain approaches.

2. Materials and methods

2.1. Subjects

Patients and healthy controls were recruited within seven centers (located in Ulm, Munich, Leipzig, Homburg, Erlangen, Bonn, and Goettingen) of the German consortium for FTLD (http://www.ftld.de). All subjects gave written consent. The research protocol was in accordance with the latest version of the Declaration of Helsinki and approved by the universities' ethics committees. For each center, the clinical evaluation and the assessment of the MRI scans were done on site according to standard operating procedures. That is, all of these centers used the same study protocol (diagnostic criteria, demographic, neuropsychological and language assessment, and scanning parameters), except for one center, where different scanning parameters were used (see Section 2.2). The diagnosis of PPA required progressive deterioration of speech and that the main deficits were restricted to speech and language for at least two years. Patients were diagnosed more specifically with nfvPPA, svPPA, or lvPPA according to the newest diagnostic criteria (Gorno-Tempini et al., 2011). Note that data from the patient's first visit in the multi-centric FTLD consortium's study was included guaranteeing the relevance of our results for early diagnosis of PPA syndromes. None of the patients included in this study had any comorbid psychiatric or neurodegenerative disease. The degree of clinical impairment of the patients was assessed using the Clinical Dementia Rating scale (CDR) and the FTLD-modified Clinical Dementia Rating scale (FTLD-CDR). We compared 44 right-handed patients suffering

from a variant of PPA (16 nfvPPA, 17 svPPA, and 11 lvPPA) with 20 right-handed healthy controls. We report all possible pairwise comparisons between PPA variants. Subjects from the larger group of a given group comparison were matched as closely as possible to the smaller group for 1) number, 2) scanning parameter, 3) age, and where possible 4) gender.

2.2. Image acquisition

All structural images were acquired on Siemens Magnetom 3 T scanners (2xVerio, 2xSkyra, 2xTrio, 1xAllegra, Erlangen, Germany). 47 T1-weighted images (12 svPPA, 11 nfvPPA, ten lvPPA, 14 healthy controls) were acquired using a magnetization prepared rapid gradient echo sequence with a matrix = 240 × 256 × 176, resolution = 1 × 1 × 1 mm, field of view = 240 mm, repetition time = 2300 ms, echo time = 2.98 ms, inversion time = 900 ms, and flip angle = 9°. For 17 subjects (five nfvPPA, five svPPA, one lvPPA, six healthy controls), T1-weighted images were acquired using a magnetization prepared rapid gradient echo sequence with a matrix = 208 × 256 × 256, resolution = 1 × 1 × 1 mm, field of view = 256 mm, repetition time = 2200 ms, echo time = 4.38 ms, inversion time = 1200 ms, and flip angle = 8°. The distribution of the two sequences (scanning parameters) did not differ significantly, neither between patient groups nor between patient groups and healthy control groups (see Table 1). The very first MRI scans that were assessed as soon as the subjects were enrolled in the study, were used for analyses.

2.3. Data analysis

2.3.1. Clinical characteristics

We used SPSS version 22 (IBM Corporation, Armonk, NY) to compute descriptive group scores (mean and standard deviation) for the overall patient and healthy control groups as well as for the respective subsets after matching for sample size, age, gender, and scanning parameters. Group comparisons for age, disease duration, education, and total grey matter density between all patient and healthy control groups as well as between PPA variants were performed using one-way ANOVAs, Kruskal-Wallis tests, and post-hoc *t*-tests in SPSS. Group comparisons for demographic and clinical characteristics between the matched subsets were performed using independent *t*-tests (normally distributed data) and Mann-Whitney *U* tests (not normally distributed data) in SPSS. Group comparisons for gender and scanning parameter were done using chi-square tests in SPSS.

2.3.2. Voxel-based morphometry

Images were processed with the VBM toolbox (http://dbm.neuro.uni-jena.de/vbm/) in SPM 8 (Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) running in a MATLAB 8.5 environment (Mathworks, Inc., Sherbon, MA, USA) using the default parameters. MRI images were segmented into grey matter, white matter, and cerebrospinal fluid using the unified segmentation module (Ashburner and Friston, 2005) and normalized to the standard Montreal Neurological Institute template including affine and non-linear modulation to account for local compression and expansion during transformation. The normalized segmented grey matter density maps were smoothed with a Gaussian kernel of 8 mm full-width-at-half-maximum. The group comparisons between the three variants of PPA and healthy controls as well as between PPA variants were performed in FSL (FMRIB Analysis Group, Oxford University, UK., http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL) using permutation-based non-parametric testing (5000 permutations) with the Threshold-Free Cluster Enhancement (TFCE) method (Smith and Nichols, 2009; Winkler et al., 2014). Age, gender, and total grey matter were entered as covariates in the general linear model and results are reported at a family-wise error (FWE) corrected $p < 0.05$.

SVM classification (Vapnik, 1995; Vapnik, 1998) was performed using libsvm version 3.18 (Chang and Lin, 2011; https://www.csie.ntu.edu.tw/~cjlin/libsvm) in a MATLAB 8.5 environment (Mathworks, Inc., Sherbon, MA, USA). Analyses were done using a linear kernel and the default solver C-SVC with C = 1. In SVM classification, an optimal separating hyperplane is defined, which maximizes the distance between subjects belonging to different groups. In the training step, SVM assigns

Table 1
Demographic and clinical characteristics of patients and healthy controls.

	nfvPPA	svPPA	lvPPA	HC
Number	16	17	11	20
Gender (m/f)	8/8	11/6	4/7	11/9
Scanning parameter	11/5	12/5	10/1	14/6
Age (years)	67.50 ± 7.42	62.53 ± 7.77	65.36 ± 6.25	67.05 ± 6.61
Education (years)	13.19 ± 4.29	15.35 ± 3.37	13.27 ± 3.35	14.10 ± 3.04
Disease duration (years)	2.19 ± 1.60	3.59 ± 2.45	3.64 ± 2.66	–
Total grey matter density (dm^3)	0.54 ± 0.08	0.52 ± 0.08	0.51 ± 0.09	0.59 ± 0.05
CDR	3.44 ± 3.20	5.32 ± 4.19	4.64 ± 4.43	0.03 ± 0.11
FTLD-CDR	5.94 ± 4.07	7.88 ± 5.44	6.86 ± 5.81	0.05 ± 0.15
CERAD Plus (test battery)				
MMSE	19.94 ± 7.25	19.31 ± 8.35	22.10 ± 6.03	28.70 ± 0.92
Word list memory (trials 1–3)	13.07 ± 6.61	13.92 ± 7.62	11.64 ± 8.93	23.40 ± 3.03
Word list recall	4.33 ± 2.62	3.77 ± 3.30	3.73 ± 3.88	8.20 ± 2.38
Word list recognition (yes)	8.57 ± 2.41	8.85 ± 1.41	9.10 ± 1.20	9.80 ± 0.52
Word list recognition (no)	9.57 ± 0.65	7.69 ± 2.63	8.40 ± 3.34	10.00 ± 0.00
Constructional praxis	9.06 ± 1.95	10.00 ± 2.08	8.18 ± 3.31	11.00 ± 0.00
Constructional praxis recall	6.75 ± 2.86	6.31 ± 4.31	4.55 ± 4.28	9.45 ± 1.91
Trail Making Test A (s)	94.38 ± 46.95	75.69 ± 51.56	75.80 ± 51.33	35.80 ± 9.01
Trail Making Test B (s)	220.18 ± 91.33	123.70 ± 72.24	201.13 ± 84.47	74.50 ± 19.12
Boston Naming Test	9.93 ± 4.76	6.47 ± 4.26	10.18 ± 3.89	14.85 ± 0.49
Verbal Fluency Test	8.06 ± 7.34	8.00 ± 5.01	12.09 ± 8.49	26.75 ± 5.50
Phonemic Fluency Test	3.87 ± 4.09	7.23 ± 5.29	6.80 ± 4.52	18.20 ± 4.63
Repeat and Point Test				
Repeat task	7.93 ± 2.34	8.93 ± 1.98	6.80 ± 3.36	10.00 ± 0.00
Point task	8.53 ± 1.55	6.14 ± 2.85	8.10 ± 1.91	9.88 ± 0.49

CDR clinical dementia rating scale, global score, CERAD Consortium to Establish a Registry for Alzheimer's Disease, FTLD frontotemporal lobar degeneration, HC healthy controls, lvPPA logopenic variant PPA, MMSE Mini-Mental State Examination, nfvPPA nonfluent/agrammatic variant PPA, PPA primary progressive aphasia, svPPA semantic variant PPA. Note age, education, disease duration, CDR, FTLD-CDR, CERAD Plus, and Repeat and Point Test are indicated as mean ± standard deviation. Note that data was missing for a few subjects on some subtests of the CERAD Plus and the Repeat and Point Test.

a weight to the scan of each subject which indicates its importance for the discrimination between groups. This weight is multiplied by a label vector which indicates the group of the scan (e.g., patient or healthy control). The cross-validation of the trained SVM was performed using the leave-one (subject)-out method. This procedure iteratively leaves-out the information of one subject of each group and trains the model on the remaining subjects for subsequent class assignment of the respective subject that was not included in the training procedure. This validation method allows the generalization of the trained SVM to data that have not been presented to the SVM algorithm previously and avoids the danger of inflating accuracies.

In the whole brain approach, we included all voxels that had a probability for grey matter higher than 0.2 (because voxels lying between white matter and ventricular cerebrospinal fluid tend to be misclassified as grey matter (e.g., Ashburner and Friston, 2000; Dukart et al., 2011). In the ROI approach, we used the results from a recently published anatomical likelihood estimation meta-analysis on the three variants of PPA ($p < 0.05$ false discovery rate corrected) across MRI studies as a prototypical disease-specific template (Bisenius et al., 2016). The original meta-analytic clusters were coregistered to the Montreal Neurological Institute template of the VBM results using SPM 8 and dilated by two voxels using the 3D dilation function implemented in the WFU PickAtlas (Maldjian, http://www.nitrc.org/projects/wfu_pickatlas). Non-parametric statistical comparisons were calculated between the performance (as indicated by the area under the receiver operating characteristic curve, AUC) of the ROI approach and the whole brain approach for all pairwise comparisons at $p < 0.05$ in StAR (Vergara et al., 2008; www.melolab.org/star/home.php).

In order to determine and visualize the importance of each voxel for the discrimination between groups in the whole brain approach, we multiplied each grey matter probability map (containing only voxels where $p > 0.2$) by the product of weight and label and summed on a voxel basis (Klöppel et al., 2008b).

3. Results

3.1. Demographic and clinical characteristics

The demographic and clinical characteristics of the overall patient and healthy control groups are shown in Table 1. The patient and healthy control groups did not differ significantly from each other in age, education, or disease duration. The patient and healthy control groups differed however significantly in total grey matter density ($F(3,63) = 4.06$, $p = 0.01$) with svPPA and lvPPA showing lower values than healthy controls. The three PPA variants did not differ significantly from each other in age, education, disease duration, or total grey matter density.

A detailed description of each of the pairwise comparisons between the matched subsets is given in Supplementary Table A.1. As shown in Table A.1, no pair of groups differed significantly in age, gender, and education and none of the patient groups differed significantly from the other patient groups in age, gender, education, duration of disease, CDR, and FTLD-CDR. PPA variants differed significantly from healthy controls in CDR, FTLD-CDR, and most of the subtests of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Plus test battery. NfvPPA and svPPA additionally differed significantly from healthy controls in the Repeat and Point Test. In the pairwise comparisons between PPA variants, svPPA showed a significantly lower test score in the Point task than nfvPPA and a significantly higher test score in the Repeat task than lvPPA (see Supplementary Table A.1).

3.2. Voxel-based morphometry results

Significant results of the statistical comparison between grey matter density maps of healthy controls and patients are shown in red (nfvPPA), light green (svPPA), and blue (lvPPA) color in Figs. 1–3 on the top left (for more details, see Supplementary Table A.2). All results are reported at an FWE corrected significance level of $p < 0.05$. The results of the statistical comparison between grey matter density maps of svPPA and nfvPPA are shown in Fig. 4 on the top left (svPPA < nfvPPA in light green color, nfvPPA < svPPA in red color). The results for the statistical comparison between lvPPA and svPPA are shown in Fig. 5 on the top left (svPPA < lvPPA in light green color, lvPPA < svPPA no significant results at $p < 0.05$). There were no significant results for the comparison between lvPPA and nfvPPA at a FWE corrected significance level of $p < 0.05$ (therefore not shown). More details on the pairwise comparisons between PPA variants are given in Supplementary Table A.2.

3.3. Support vector machine classification results

SVM classification was applied separately to each group comparison: 1) nfvPPA vs. healthy controls, 2) svPPA vs. healthy controls, 3) lvPPA vs. healthy controls, and 4) svPPA vs. nfvPPA. The reported accuracy is the percentage of subjects correctly assigned to the clinical diagnosis (patient/healthy control or svPPA/nfvPPA). Sensitivity refers to the proportion of patients correctly classified as patients and specificity to the proportion of healthy controls correctly classified as healthy controls. Positive predictive value refers to the number of correctly classified patients out of all subjects classified as patients and negative predictive value refers to the number of correctly classified healthy controls out of all subjects classified as healthy controls.

3.4. Group comparisons between patients and healthy controls

The accuracy for the classification between different variants and healthy controls using the leave-one-out approach ranged from 91 to 97% for the whole brain approach and from 82 to 100% for the ROI approach. Details on the respective sensitivity, specificity, and accuracy are given in Figs. 1–3 on the bottom right and details on positive and negative predictive values are shown in Supplementary Table A.3. The results of the SVM classification between patients and healthy controls for the whole brain approach are shown on the top right of Figs. 1–3. Here, values range between −1 and 0 or 0 and 1 and reflect the relative importance of these voxels in the discrimination between both groups. Voxels that contributed the most to the classification of subjects as patients (i.e., had a higher negative value) are depicted in yellow and the voxels that contributed the most to the classification of subjects as healthy controls (i.e., had a higher positive value) are shown in light green. A value near 0 indicates that this voxel was neither indicative for the classification as patient nor as healthy control.

Brain regions that contributed the most to the classification of subjects as nfvPPA vs. control subjects (Fig. 1 on the top right) encompass bilaterally cerebellum, inferior, middle, and superior temporal gyri, middle occipital gyrus, parahippocampal gyrus, crus cerebri, thalamus, precuneus, inferior and superior frontal gyri, as well as in the left hemisphere orbital gyrus, insula, pre- and postcentral gyri, middle frontal gyrus, and angular gyrus. Classification accuracy was 91% for the whole brain approach and 84% for the ROI approach. The statistical comparison between both approaches revealed high AUC values, but without significant differences ($AUC_{ROI} = 0.90$, $AUC_{whole\ brain} = 0.94$, $p = 0.48$).

Regions that contributed the most to the classification of subjects as svPPA vs. control subjects included bilaterally (although predominantly in the left hemisphere) cerebellum, inferior, middle and superior temporal gyri, middle occipital gyrus, parahippocampal gyrus, hippocampus, amygdala, putamen, insula, precentral and postcentral gyri, middle frontal gyrus, inferior parietal gyrus, and cingulate gyrus (see Fig. 2 on the top right). Classification accuracy was very high for both approaches (97% for the whole brain approach and 100% for the ROI approach). The statistical comparison between approaches showed very

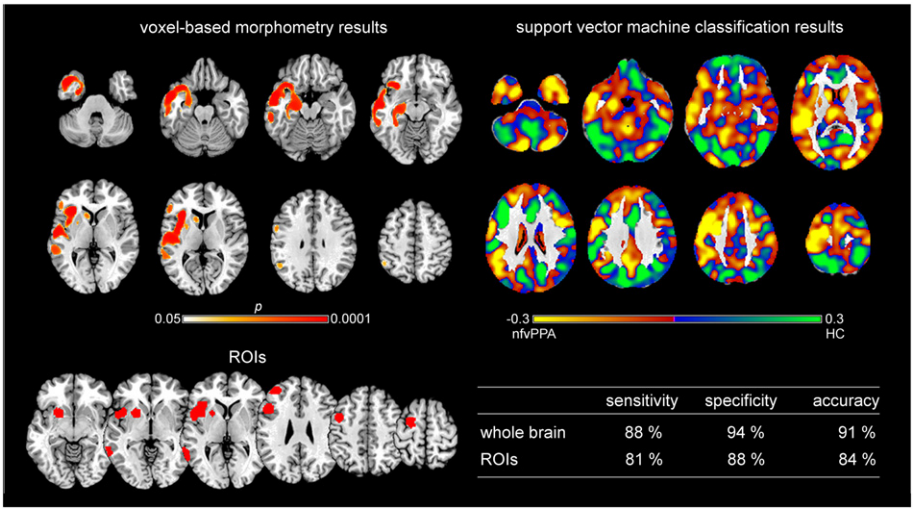

Fig. 1. Voxel-based morphometry and support vector machine classification results for nonfluent/agrammatic variant PPA as compared to healthy controls. Top left: voxel-based morphometry (VBM) results for the comparison between nonfluent/agrammatic variant PPA (nfvPPA) and healthy controls (HC) (family-wise error corrected $p < 0.05$). Bottom left: Regions of interest (ROIs) based on independent meta-analyses. Right: Results of support vector machine classification (SVM) classification. Top: Regions most relevant for classification as patients in yellow, HC in light green. Note that the scale of the distance weights has no applicable units. Bottom: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification.

high AUC values for both, but without significant differences ($AUC_{ROI} = 1.00$, $AUC_{whole\ brain} = 0.97$, $p = 0.32$).

Regions that contributed the most to the classification of subjects as lvPPA patients vs. control subjects are shown in yellow in Fig. 3 on the top right and encompass left inferior temporal gyrus, fusiform gyrus, middle occipital gyrus, parahippocampal gyrus, hippocampus, putamen, insula, thalamus, precentral gyrus, middle and superior frontal gyri, angular gyrus, supramarginal gyrus, and cingulate gyrus as well as bilaterally cerebellum, middle and superior temporal gyri, caudate nucleus, thalamus, middle and superior frontal gyri, precuneus, and superior parietal gyrus. Classification accuracy was high for both, the whole brain approach (95%) and the ROI approach (82%). The statistical comparison between both approaches did show high AUC values without significant differences ($AUC_{ROI} = 0.91$, $AUC_{whole\ brain} = 0.95$, $p = 0.38$).

3.4.1. Group comparisons between PPA variants

Fig. 4 illustrates on top right in yellow the regions that contributed the most to the classification as svPPA and in green the regions that contributed the most to the classification as nfvPPA. Here, sensitivity refers to the ratio of correctly classified svPPA patients and specificity to the ratio of correctly classified nfvPPA patients. Details on positive and negative predictive values are given in Supplementary Table A.3. The regions that contributed the most to the classification of a subject as svPPA included bilaterally cerebellum, inferior, middle and superior temporal gyri, middle occipital gyrus, fusiform gyrus, parahippocampal gyrus, hippocampus, putamen, insula, cuneus, precuneus, inferior frontal gyrus, superior parietal gyrus, cingulate gyrus, and left precentral gyrus. Regions that contributed the most to the classification of nfvPPA included bilateral cerebellum, middle and superior occipital gyrus, superior temporal gyrus, gyrus rectus, posterior orbital gyrus, caudate

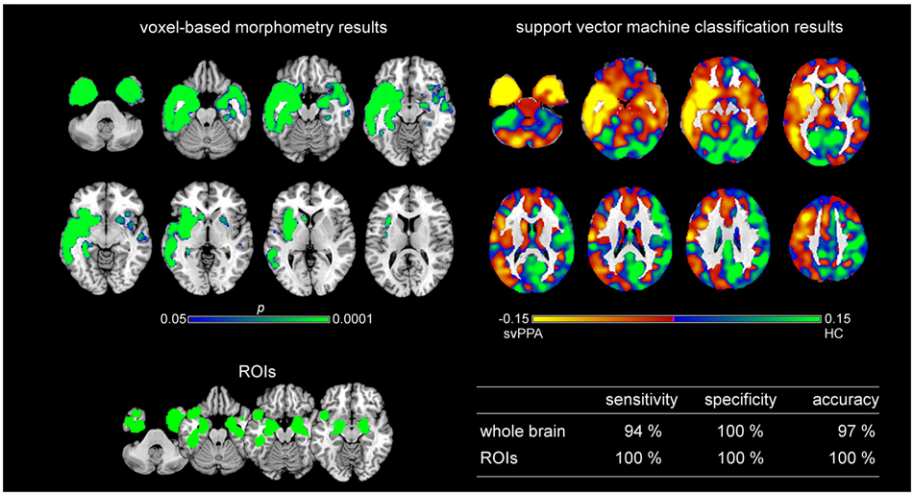

Fig. 2. Voxel-based morphometry and support vector machine classification results for semantic variant PPA as compared to healthy controls. Top left: voxel-based morphometry (VBM) results for the comparison between semantic variant PPA (svPPA) and healthy controls (HC) (family-wise error corrected $p < 0.05$). Bottom left: Regions of interest (ROIs) based on independent meta-analyses. Right: Results of support vector machine (SVM) classification. Top: Regions most relevant for classification as patients in yellow, HC in light green. Note that the scale of the distance weights has no applicable units. Bottom: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification.

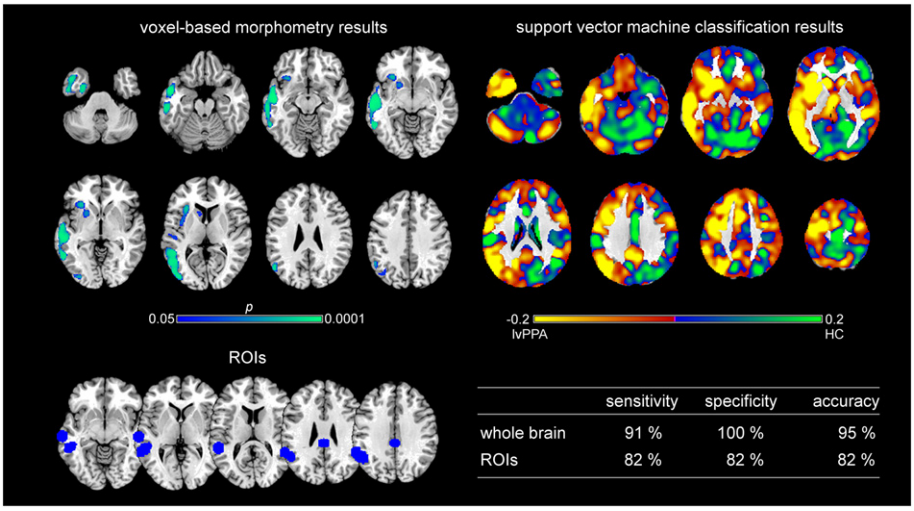

Fig. 3. Voxel-based morphometry and support vector machine classification results for logopenic variant PPA as compared to healthy controls. Top left: voxel-based morphometry (VBM) results for the comparison between logopenic variant PPA (lvPPA) and healthy controls (HC) (family-wise error corrected $p < 0.05$). Bottom left: Regions of interest (ROIs) based on independent meta-analyses. Right: Results of support vector machine (SVM) classification. Top: Regions most relevant for classification as patients in yellow, HC in light green. Note that the scale of the distance weights has no applicable units. Bottom: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification.

nuclei, thalamus, inferior, middle, and superior frontal gyrus, precentral gyrus, postcentral gyrus, inferior parietal gyrus, angular gyrus, supramarginal gyrus, right precuneus, right superior parietal gyrus, and right cingulate gyrus. Classification accuracy was 78% for both, the whole brain and the ROI approach. Both approaches revealed high AUC values without significant differences between them ($AUC_{ROI} = 0.87$, $AUC_{whole\ brain} = 0.88$, $p = 0.72$).

Fig. 5 illustrates on top right in yellow the regions that contributed the most to the classification as lvPPA and in green the regions that contributed the most to the classification as svPPA. Sensitivity refers to the ratio of correctly classified lvPPA patients and specificity to the ratio of correctly classified svPPA patients. Positive and negative predictive values are given in Supplementary Table A.3. The regions that contributed the most to the classification of a subject as lvPPA included bilateral cerebellum, middle occipital gyrus, middle and superior temporal gyri, caudate nuclei, thalamus, superior frontal gyrus, supramarginal gyrus, angular gyrus, precuneus, cingulate gyrus, right lateral orbital gyrus, inferior and middle frontal gyrus, and superior parietal gyrus. Regions that contributed the most to the classification of svPPA included bilateral cerebellum, inferior, middle, and superior temporal gyrus, parahippocampal gyrus, hippocampus, insula, and right putamen. Classification accuracy was 95% for both, the whole brain and the ROI approach. Both approaches reached high AUC values without significant differences between them ($AUC_{ROI} = 0.91$, $AUC_{whole\ brain} = 0.93$, $p = 0.41$).

Fig. 6 illustrates on top in yellow the regions that contributed the most to the classification as lvPPA and in green the regions that contributed the most to the classification as nfvPPA. Sensitivity refers to the ratio of correctly classified lvPPA patients and specificity to the ratio of correctly classified nfvPPA patients. For details on positive and negative predictive values see Supplementary Table A.3. The regions that contributed the most to the classification of a subject as lvPPA included bilateral

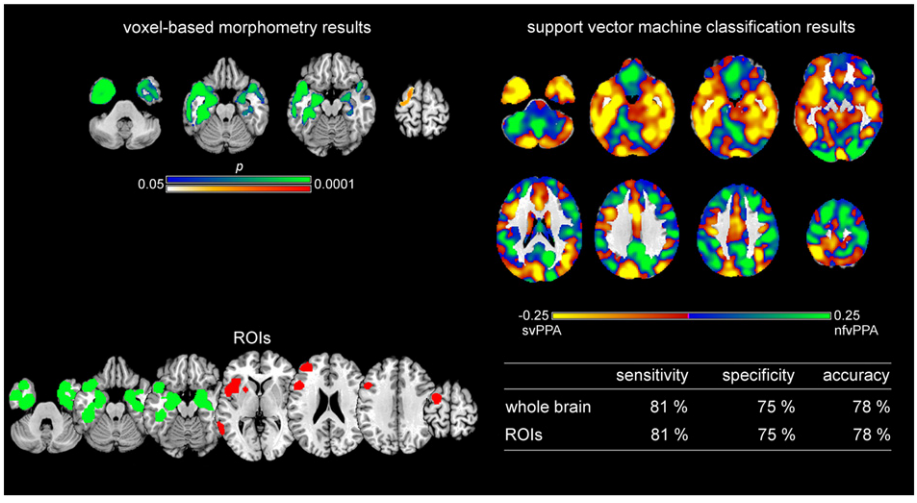

Fig. 4. Support vector machine classification results for the comparison and discrimination between semantic variant PPA and nonfluent/agrammatic variant PPA. Top left: VBM results for the comparison between semantic variant PPA (svPPA) and nonfluent/agrammatic variant PPA (nfvPPA) (svPPA < nfvPPA green, nfvPPA < svPPA red, family-wise error corrected $p < 0.05$). Bottom left: Regions of interest (ROIs) based on independent meta-analyses. Right: Results of support vector machine (SVM) classification. Top: Regions most relevant for classification as svPPA in yellow, nfvPPA in light green. Note that the scale of the distance weights has no applicable units. Bottom: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification.

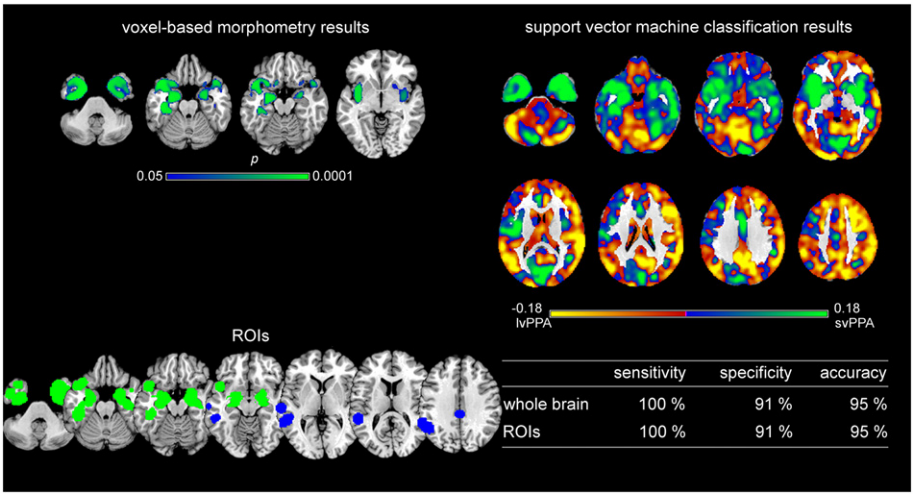

Fig. 5. Support vector machine classification results for the comparison and discrimination between logopenic variant PPA and semantic variant PPA. Top left: VBM results for the comparison between logopenic variant PPA (lvPPA) and semantic variant PPA (svPPA) (svPPA < lvPPA family-wise error corrected $p < 0.05$). Bottom left: Regions of interest (ROIs) based on independent meta-analyses. Right: Results of support vector machine (SVM) classification. Top: Regions most relevant for classification as lvPPA in yellow, svPPA in light green. Note that the scale of the distance weights has no applicable units. Bottom: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification.

cerebellum, inferior, middle occipital gyrus, middle and superior temporal gyri, thalamus, putamen, middle and superior frontal gyrus, supramarginal gyrus, angular gyrus, precentral gyrus, cingulate gyrus, precuneus, superior parietal gyrus. Regions that contributed the most to the classification of nfvPPA included right inferior temporal gyrus, bilateral middle and superior temporal gyri, gyrus rectus, lateral orbital gyrus, insula, caudate nuclei, cuneus, cingulate gyrus, middle and superior frontal gyri, postcentral gyrus, supramarginal gyrus, and superior parietal gyrus. Classification accuracy was low with 55% for the whole brain approach and higher with 64% for the ROI approach. AUC values were comparable, namely higher for the ROI than the whole brain approach, but without significant differences between them ($AUC_{ROI} = 0.64$, $AUC_{whole\ brain} = 0.59$, $p = 0.50$).

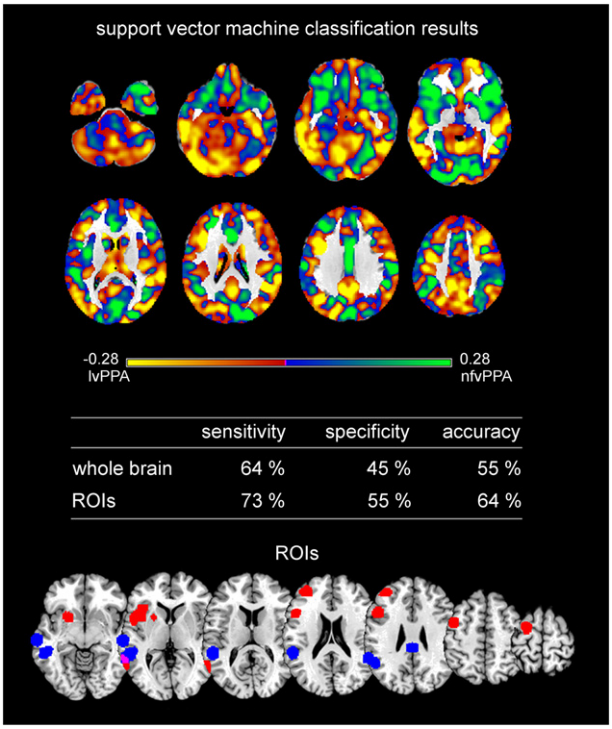

Fig. 6. Support vector machine classification results for the discrimination between logopenic variant PPA and nonfluent/agrammatic variant PPA. Top: Regions most relevant for support vector machine classification as logopenic variant PPA (lvPPA) in yellow, nonfluent/agrammatic variant PPA (nfvPPA) in light green. Note that the scale of the distance weights has no applicable units. VBM results are not shown for the group comparisons, because no significant results were obtained. Middle: Sensitivity, specificity, and accuracy for the ROI approach and the whole brain approach in SVM classification. Bottom Regions of interest (ROIs) based on independent meta-analyses.

4. Discussion

To our knowledge, this is the first study demonstrating that SVM classification in multi-center MRI data can be used to diagnose and dissociate PPA subtypes, where the multi-center design is a precondition for application in clinical routine in the future. Moreover, we compare a whole brain vs. data-driven disease-specific ROI approach for SVM classification. We used ROIs reported in a recent comprehensive meta-analysis on PPA (Bisenius et al., 2016). In order to reveal whether the regions that contributed the most to the whole brain SVM classification of the three variants of PPA corresponded to the regions that were especially atrophic in the respective variant, we additionally conducted statistical group-level comparisons between patient groups and healthy control groups. In the following, we are going to introduce the results of these group-level comparisons, before we discuss in more detail the results of the SVM classification for the whole brain approach and the ROI approach as well as possible further implications.

4.1. Atrophy in the different variants of primary progressive aphasia

The group comparisons in our study revealed regional brain atrophy that included the disease-specific brain areas identified in comprehensive systematic and quantitative meta-analyses across imaging studies from the literature, if one compares this data for each PPA variant (see left top and bottom images in Figs. 1–3). Beyond that our group-level comparisons are in line with studies showing that, with the progression of the disease, the atrophic networks in the three subtypes of PPA partly converge (e.g., Gorno-Tempini et al., 2011; Rogalski et al., 2011). Mild and early svPPA has been shown to involve atrophy in (predominantly left) anterior temporal lobe, with extension to the adjacent temporoparietal junction, hippocampus and amygdala and posterior

orbital cortex as well as in the right anterior temporal lobe (Czarnecki et al., 2008; Grossman, 2010; Krueger et al., 2010; Mesulam et al., 2012; Rohrer et al., 2008) and to progress bilaterally into posterior and superior temporal lobe, left temporoparietal junction, bilateral cingulate cortex and orbitofrontal gyri, left superior orbitofrontal gyrus, left inferior and superior frontal gyri (e.g., Grossman, 2010; Kumfor et al., 2016; Rogalski et al., 2011). Early and mild stages of nfvPPA, on the other hand, have been shown to be characterized by atrophy in left inferior frontal gyrus, temporoparietal junction, anterior superior temporal gyrus, posterior middle frontal gyrus and precentral gyrus (Mesulam et al., 2012) and to progress into left anterior temporal lobe, orbital cortex, dorsolateral prefrontal cortex, anterior cingulate cortex, and along the perisylvian fissure into the parietal lobe (e.g., Grossman, 2010; Rogalski et al., 2011). For lvPPA, atrophy has been shown to progress from (predominantly left) posterior superior temporal cortex, inferior parietal cortex, posterior cingulate cortex and medial temporal cortex into the anterior and lateral temporal cortex, caudate nucleus, insula, inferior frontal gyrus and dorsal frontal cortex as well as into the temporoparietal junction, posterior cingulate and precuneus of the right hemisphere. (e.g., Rogalski et al., 2011; Rohrer et al., 2013).

4.2. Support vector machine classification is a useful tool to differentiate between healthy controls and primary progressive aphasia variants

Accuracies for the whole brain approach in SVM classification between patients and healthy controls ranged from 91% for nfvPPA over 95% for lvPPA to 97% for svPPA. The between-subtype whole brain SVM classification enabled high accuracy of 78 and 95% for the discrimination between svPPA vs. nfvPPA and svPPA vs. lvPPA variant. Only for the discrimination between nfvPPA and lvPPA variants accuracy was low with 55%. These numbers are in line with previously reported accuracies ranging from 58 to 100% in studies on neurodegenerative diseases like AD (e.g., (e.g., Chetelat and Baron, 2003; Davatzikos et al., 2008b; Dukart et al., 2013; Dukart et al., 2011; Klöppel et al., 2008b; Lerch et al., 2008; Teipel et al., 2007; Vemuri et al., 2008), mild cognitive impairment (e.g., Davatzikos et al., 2008a; Teipel et al., 2007), FTD (e.g., Davatzikos et al., 2008b; Dukart et al., 2011; Klöppel et al., 2008a), and PPA (Wilson et al., 2009; Zhang et al., 2013).

Until now, there has only been one study investigating the three variants of PPA with SVM classification (Wilson et al., 2009). These authors performed a principal component analysis on MRI scans from one study center and subsequently used the results for the pairwise SVM classification between patients and healthy controls as well as between the three patient groups. Wilson et al. (2009) found an accuracy of 100% for the discrimination between svPPA patients and healthy controls, 100% for the classification between lvPPA patients and healthy controls, and an accuracy of 89% for the discrimination between nfvPPA and healthy controls. These authors report an accuracy of 89% for the discrimination between svPPA and nfvPPA patients, 93.8% for svPPA vs. lvPPA, and 81.3% for lvPPA vs. nfvPPA. Our SVM results using the whole brain approach on grey matter density maps in the multi-center cohort of the FTLD consortium are thus comparable with the results of Wilson et al. (2009) with regard to the classification between patients and healthy controls showing higher accuracies for svPPA and lvPPA than for nfvPPA and the classification between lvPPA and nfvPPA showing a lower classification accuracy than the other classifications between PPA variants.

Additionally, we performed group-level comparisons on the grey matter density maps between patients and healthy controls as well as between PPA variants in order to investigate whether the regions that contributed the most to the SVM classification of patients also corresponded to the regions mostly atrophied in these patients. Figs. 2 and 3 show that brain regions that were most consistently atrophied in svPPA and lvPPA indeed also contributed the most to the SVM classification of these patients. For nfvPPA, on the other hand, brain regions that contributed the most to the SVM classification as patients were not constrained to the regions that were atrophied in our nfvPPA patients, but also encompassed very similar regions in the contralateral (right) hemisphere (see Fig. 1). A possible explanation for the importance of the additional brain regions in the right hemisphere might be that they were affected to a lesser extent (and thus not significant in the group-level comparison) and that SVM classification as a more sensitive method already took into account early atrophy in these regions. There is a general consensus that the results of group-level statistics might not be applicable to individual scans, because their sensitivity and specificity at early stages of brain pathology is insufficient for the prediction of the status of individual scans (Davatzikos et al., 2008b; Fan et al., 2008; Wilson et al., 2009).

Interestingly, for the discrimination between svPPA and nfvPPA, the regions that contributed to the SVM classification as svPPA patients (see Fig. 4 on the top), corresponded to the regions that were most consistently atrophied in these patients (see Fig. 2 on the top left). The regions that contributed to the SVM classification as nfvPPA, on the other hand, were (except for two characteristic regions in the inferior and middle frontal gyri) rather spread. This might be due to the fact that the group comparisons between patients and healthy controls showed for both, svPPA and nfvPPA, significant atrophy in the superior temporal gyrus, parahippocampal area, hippocampus, insula, and inferior frontal gyrus (see Figs. 1 and 2 on the top left). Although atrophy in the superior temporal gyrus, parahippocampal area, and hippocampus have been discussed to be rather specific for svPPA, while insula, and inferior frontal gyrus have been discussed to be rather characteristic for nfvPPA, it has been shown in longitudinal studies that with the progression of the disease, the atrophic networks in the three variants of PPA partly converge (e.g., Gorno-Tempini et al., 2011; Rogalski et al., 2011). Given that on the one hand several regions might be affected similarly in nfvPPA and svPPA depending upon the current stage of the respective disease and on the other hand structural MRI scans do not provide any information regarding the temporal dynamic pattern of brain atrophy, the SVM classification method, given its high sensitivity, might not always be able to perfectly discriminate between these two variants. For both subtype-specific classifications, svPPA vs. nfvPPA and svPPA vs. lvPPA, we reached a high classification accuracy, although the number of patients was rather low for lvPPA and the respective comparison. The high accuracy might be related to a relatively strong (in the sense of high *t*-values) and regionally focused atrophy in svPPA. This is obvious in the group comparisons revealing much higher atrophy in svPPA than nfvPPA or lvPPA, whereas nfvPPA showed stronger atrophy only in a very small area and lvPPA did not show any atrophy in comparison with svPPA. Note that disease duration and severity did generally not significantly differ between PPA subtypes excluding these factors as explanation for differences in classification accuracy.

As stated before the whole-brain SVM classification between nfvPPA and lvPPA variants reached only a low accuracy. This might be related to relatively small and rather distributed atrophy in these two PPA variants or to conceptual issues. In a prospective data-driven study, Sajjadi et al. (2012) examined to which extent PPA patients would be classifiable according to the revised clinical diagnostic criteria and which linguistic impairments would cluster together (and thus form distinct syndromes) using principal factor analysis. In this cohort, 58.7% of the patients could be assigned to one of the three variants of PPA proposed by Gorno-Tempini et al. (2011), while 41.3% of the patients were classified as mixed PPA because their deficits either extended beyond a single PPA variant or they met the diagnostic criteria for more than one variant. The principal factor analysis identified two clear syndromes corresponding to the proposed syndromes of svPPA and nfvPPA as well as a residual miscellany. Interestingly, impaired sentence repetition, which has been proposed as a cardinal diagnostic feature for lvPPA, aligned with the factor corresponding to nfvPPA. One might therefore speculate that low classification accuracy between nfvPPA and lvPPA in imaging data might not only be related to the rather relatively small and

distributed atrophy, but possibly also to problems in clinically distinguishing both PPA syndromes.

4.3. Regions-of-interest approach or whole brain approach?

We compared the whole brain approach for SVM classification with an ROI approach using ROIs from a recent meta-analysis on the three variants of PPA (Bisenius et al., 2016). A similar approach has already been adopted by Dukart and colleagues who compared the whole brain versus ROI approach for SVM classification between FTD and AD as well as between these patient groups and healthy controls using structural MRI and PET scans (Dukart et al., 2013; Dukart et al., 2011). These authors reported that for MRI scans, the ROI approach was comparable to the whole brain approach for the discrimination between patients and healthy controls, but had a lower accuracy for the discrimination between patient groups (AD vs. FTD) (Dukart et al., 2013; Dukart et al., 2011). In the current study, the ROI approach reached generally a high accuracy in diagnosis and, at least mainly, differential diagnosis/classification of PPA syndromes, comparable to the whole-brain approach. In detail, it showed a higher accuracy as compared to healthy controls for svPPA patients and a slightly lower accuracy for nfvPPA and lvPPA patients, while it showed a similar accuracy for svPPA vs. nfvPPA and svPPA vs. lvPPA patients as compared to the whole brain approach. Remarkably, for the lvPPA vs. nfvPPA comparison the ROI approach showed a higher accuracy than the whole brain approach, may be due to the diffusivity and similar strength (in the sense of *t*-values) of brain atrophy requiring higher regional specificity for the analysis. One might speculate that ROI-based classification might be given preference for special questions in differentiating between syndromes in the future. Given however that none of these trends was statistically significant, we consider both approaches as equally valid.

The visual comparison between the whole brain approach and the ROI approach raises however the question about the optimal method to choose ROIs for SVM classification in PPA. The optimal number of ROIs for SVM classification needs to be such as to accurately capture all subtleties of the structural abnormality in these patients and thus achieve a sufficient predictive accuracy without however reducing predictive accuracy through the increase of noise that possibly accompanies additional ROIs that are less relevant to the classification. Selecting ROIs based on group-level comparison between patients and healthy control groups might for instance provide a higher discriminative power for the SVM classification in the same study sample. These ROIs would however be biased to at least some extent by the specific study sample and might therefore not necessarily lead to similar good results in other study samples. Another possibility to find the optimal ROIs for the SVM classification between nfvPPA (or lvPPA) and healthy controls might consist in rerunning meta-analyses on the three variants of PPA across MRI studies using a less conservative statistical threshold. This methodological approach might no longer exclusively reveal the brain regions that are specific to a given variant (and to some extent possibly even false positive results), but due to the higher sensitivity, also common networks between variants that become usually only visible in longitudinal studies monitoring the progression of the disease (e.g., Rogalski et al., 2011). For the ROI approach of SVM classification between the different variants of PPA, on the other hand, it might be rather promising to only consider ROIs that are either more severely impaired in one variant as in the other variants as for instance the inferior and middle temporal gyri in svPPA, or rather specific to the given variant (e.g., middle frontal gyrus in nfvPPA as compared to svPPA). The considerations regarding the optimal ROI for SVM classification in PPA are however purely hypothetical and need to be investigated in future studies.

Furthermore, it might be interesting to compare the ROI approach to the whole brain approach using combined imaging data as for instance MRI and PET as has already been done for AD and FTD (e.g., Davatzikos et al., 2008b; Dukart et al., 2013; Dukart et al., 2011) or using MRI and diffusion tensor imaging data as has been done by Zhang et al. (2013), who showed, in a small sample, higher accuracies for whole brain SVM classification of diffusion tensor imaging data than of MRI data for nfvPPA and svPPA versus healthy controls. Moreover, the potential of ROI approaches for disease classification has to be validated in longitudinal studies, where one would assume higher accuracy in early stages.

5. Limitations

The relatively small number of subjects might hamper the generalization of the results to the overall population of PPA patients. Given however that our results are very similar to another study including more patients but using a different approach (Wilson et al., 2009) this should not really constitute a major issue. A problem that might occur in pattern classification methods is the risk of overfitting the data due to the high-dimensionality of the data, which can however be reduced by using the leave-one out approach as has been done in the current study. Segmentation and normalization processes are not always perfect which might result in underestimation of atrophy in patients or underestimation of grey matter in healthy controls, which leads to lower accuracies in the SVM classification. This issue should however at least partly be addressed in the current study given that our data have been acquired on different scanners and were well balanced between patient and healthy control groups. Finally, the classification between pairs of groups was a highly idealized situation that does not reflect the problem in the real world of differential diagnosis between several neurological diseases with different prevalence rates – an issue that has to be addressed in future studies validating the application of SVM approaches in every day diagnostic life.

6. Conclusion

Our study aimed at validating the potential of structural multi-center MRI data for disease classification in PPA. We compared the whole brain approach with a disease-specific ROI approach for SVM classification in the three variants of PPA. Generally, both the whole brain and the disease-specific approach reached high classification accuracy in diagnosis and differential diagnosis of PPA syndromes without significant differences. Our results showed that for svPPA, the ROI approach using prototypical disease-related networks as revealed by meta-analyses across MRI studies revealed a higher accuracy (perfect discrimination of 100%) than the whole brain approach. For nfvPPA and lvPPA on the other hand, the SVM classification showed higher accuracies when using the whole brain approach. The regions contributing to the correct SVM classification of patients mostly corresponded to regions that were consistently atrophied in these patients as shown by the VBM results. For the discrimination between svPPA and nfvPPA, and between svPPA and lvPPA the whole brain approach and the ROI approach showed similar results. The ROI approach increased accuracy in classification between lvPPA and nfvPPA in comparison with the whole brain approach, which might be related to the diffusivity and similar strength (in the sense of *t*-values) in these PPA syndromes requiring higher regional specificity for the analysis. Given that the accuracies for SVM classification using the ROI approach were still quite high despite the relatively small size of the chosen ROIs as compared to the regions that were taken into account in the whole brain SVM classification of the respective patients, future studies shall further explore the potential of the ROI approach using different ROIs for SVM classification of PPAs.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nicl.2017.02.003.

Funding

Sandrine Bisenius is supported by the MaxNetAging Research School of the Max Planck Society. The study has been supported by the German

Federal Ministry of Education and Research (BMBF; Grant number FKZ 01GI1007A; German FTLD consortium). Sandrine Bisenius, Karsten Mueller, Katharina Stuke and Matthias L. Schroeter have further been supported by the Parkinson's Disease Foundation (Grant No. PDF-IRG-1307), the Michael J Fox Foundation (Grant No. MJFF-11362), and by LIFE – Leipzig Research Center for Civilization Diseases at the University of Leipzig. LIFE is funded by means of the European Union, by the European Regional Development Fund (ERFD) and by means of the Free State of Saxony within the framework of the excellence initiative.

References

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. NeuroImage 11, 805–821.
Ashburner, J., Friston, K.J., 2005. Unified segmentation. NeuroImage 26 (3), 839–851.
Bisenius, S., Neumann, J., Schroeter, M.L., 2016. Validating new diagnostic imaging criteria for primary progressive aphasia via anatomical likelihood estimation meta-analyses. Eur. J. Neurol. 23 (4), 704–712.
Chang, C.-C., Lin, C.-J., 2011. LIBSVM : a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2 (3), 27:1–27:27.
Chetelat, G., Baron, J.C., 2003. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. NeuroImage 18 (2), 525–541.
Czarnecki, K., Duffy, J., Nehl, C.R., Cross, S.A., Jack Jr., C.R., Shiung, M.M., et al., 2008. Very early semantic dementia with progressive left > right temporal lobe atrophy: an eight-year longitudinal study. Arch. Neurol. 65 (12), 1659–1663.
Davatzikos, C., Fan, Y., Wu, X., Shen, D., Resnick, S.M., 2008a. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol. Aging 29 (4), 514–523.
Davatzikos, C., Resnick, S.M., Wu, X., Parmpi, P., Clark, C.M., 2008b. Individual patient diagnosis of AD and FTD via high-dimensional pattern classification of MRI. NeuroImage 41 (4), 1220–1227.
Desgranges, B., Matuszewski, V., Piolino, P., Chetelat, G., Mezenge, F., Landeau, B., et al., 2007. Anatomical and functional alterations in semantic dementia: a voxel-based MRI and PET study. Neurobiol. Aging 28 (12), 1904–1913.
Dukart, J., Mueller, K., Horstmann, A., Barthel, H., Moller, H.E., Villringer, A., et al., 2011. Combined evaluation of FDG-PET and MRI improves detection and differentiation of dementia. PLoS One 6 (3), e18111.
Dukart, J., Mueller, K., Barthel, H., Villringer, A., Sabri, O., Schroeter, M.L., 2013. Meta-analysis based SVM classification enables accurate detection of Alzheimer's disease across different clinical centers using FDG-PET and MRI. Psychiatry Res. 212 (3), 230–236.
Fan, Y., Batmanghelich, N., Clark, C.M., Davatzikos, C., 2008. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. NeuroImage 39 (4), 1731–1743.
Frisoni, G.B., Fox, N.C., Jack Jr., C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6 (2), 67–77.
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., Brambati, S.M., Ginex, V., Ogar, J., Dronkers, N.F., Marcone, A., et al., 2008. The logopenic/phonological variant of primary progressive aphasia. Neurology 71 (16), 1227–1234.
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.
Grossman, M., 2010. Primary progressive aphasia: clinicopathological correlations. Nat. Rev. Neurol. 6 (2), 88–97.
Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M., et al., 2004. What's in a name: voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. Brain 127, 628–649.
Harvey, R.J., Skelton-Robinson, M., Rossor, M.N., 2003. The prevalence and causes of dementia in people under the age of 65 years. J. Neurol. Neurosurg. Psychiatry 74 (9), 1206–1209.
Klöppel, S., Stonnington, C.M., Barnes, J., Chen, F., Chu, C., Good, C.D., et al., 2008a. Accuracy of dementia diagnosis: a direct comparison between radiologists and a computerized method. Brain 131, 2969–2974.
Klöppel, S., Stonnington, C.M., Chu, C., Draganski, B., Scahill, R.I., Rohrer, J.D., et al., 2008b. Automatic classification of MR scans in Alzheimer's disease. Brain 131, 681–689.
Krueger, C.E., Dean, D.L., Rosen, H.J., Halabi, C., Weiner, M., Miller, B.L., et al., 2010. Longitudinal rates of atrophy in frontotemporal dementia, semantic dementia, and Alzheimer's disease. Alzheimer Dis. Assoc. Disord. 24 (1), 43–48.
Kumfor, F., Landin-Romero, R., Devenney, E., Hutchings, R., Grasso, R., Hodges, J.R., et al., 2016. On the right side? A longitudinal study of left-versus right-lateralize semantic dementia. Brain (epub ahead of print).
Lerch, J.P., Pruessner, J., Zijdenbos, A.P., Collins, D.L., Teipel, S.J., Hampel, H., et al., 2008. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. Neurobiol. Aging 29 (1), 23–30.
McEvoy, L.K., Brewer, J.B., 2010. Quantitative structural MRI for early detection of Alzheimer's disease. Expert. Rev. Neurother. 10 (11), 1675–1688.
Mesulam, M.M., 1982. Slowly progressive aphasia without generalized dementia. Ann. Neurol. 11 (6), 592–598.
Mesulam, M.M., Wieneke, C., Thompson, C., Rogalski, E., Weintraub, S., 2012. Quantitative classification of primary progressive aphasia at early and mild impairment stages. Brain 135, 1537–1553.
Misra, C., Fan, Y., Davatzikos, C., 2009. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: results from ADNI. NeuroImage 44 (4), 1415–1422.
Mummery, C.J., Patterson, K., Price, C.J., Ashburner, J., Frackowiak, R.S., Hodges, J.R., 2000. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. Ann. Neurol. 47 (1), 36–45.
Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., et al., 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology 51 (6), 1546–1554.
Otto, M., Ludolph, A.C., Landwehrmeyer, B., Förstl, H., Diehl-Schmid, J., Neumann, M., et al., 2011. Konsortium zur Erforschung der frontotemporalen Lobardegeneration (German consortium for frontotemporal lobar degeneration). Nervenarzt 82, 1002–1005.
Ratnavalli, E., Brayne, C., Dawson, K., Hodges, J.R., 2002. The prevalence of frontotemporal dementia. Neurology 58 (11), 1615–1621.
Rogalski, E., Cobia, D., Harrison, T.M., Wieneke, C., Weintraub, S., Mesulam, M.M., 2011. Progression of language decline and cortical atrophy in subtypes of primary progressive aphasia. Neurology 76 (21), 1804–1810.
Rohrer, J.D., McNaught, E., Foster, J., Clegg, S.L., Barnes, J., Omar, R., Warrington, E.K., Rossor, M.N., Warren, J.D., Fox, N.C., 2008. Tracking progression in frontotemporal lobar degeneration: serial MRI in semantic dementia. Neurology 71 (18), 1445–1451.
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. 127, 121–126.
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.
Schroeter, M.L., Stein, T., Maslowski, N., Neumann, J., 2009. Neural correlates of Alzheimer's disease and mild cognitive impairment: a systematic and quantitative meta-analysis involving 1351 patients. NeuroImage 47 (4), 1196–1206.
Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44, 83–98.
Teipel, S.J., Born, C., Ewers, M., Bokde, A.L., Reiser, M.F., Moller, H.J., et al., 2007. Multivariate deformation-based analysis of brain atrophy to predict Alzheimer's disease in mild cognitive impairment. NeuroImage 38 (1), 13–24.
Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer, Berlin.
Vapnik, V., 1998. Statistical Learning Theory. Wiley Interscience, New York.
Vemuri, P., Gunter, J.L., Senjem, M.L., Whitwell, J.L., Kantarci, K., Knopman, D.S., et al., 2008. Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. NeuroImage 39 (3), 1186–1197.
Vergara, I.A., Norambuena, T., Ferrada, E., Slater, A.W., Melo, F., 2008. StAR: a simple tool for the statistical comparison of ROC curves. BMC Bioinformatics 9, 265.
Weiner, M.W., Aisen, P.S., Jack Jr., C.R., Jagust, W.J., Trojanowski, J.Q., Shaw, L., et al., 2010. The Alzheimer's disease neuroimaging initiative: progress report and future plans. Alzheimers Dement. 6 (3), 202-11.e7.
Wilson, S.M., Ogar, J.M., Laluz, V., Growdon, M., Jang, J., Glenn, S., et al., 2009. Automated MRI-based classification of primary progressive aphasia variants. NeuroImage 47 (4), 1558–1567.
Winkler, A., Ridgway, G.R., Webster, M., Smith, S., Nichols, T.E., 2014. Permutation inference for the general linear model. NeuroImage 92, 381–397.
Zhang, Y., Tartaglia, M.C., Schuff, N., Chiang, G.C., Ching, C., Rosen, H.J., et al., 2013. MRI signatures of brain macrostructural atrophy and microstructural degradation in frontotemporal lobar degeneration subtypes. J. Alzheimers Dis. 33 (2), 431–444.