Individual Voxel-Based Subtype Prediction can Differentiate Progressive Supranuclear Palsy from Idiopathic Parkinson Syndrome and Healthy Controls

Niels K. Focke,^{1*} Gunther Helms,² Sebastian Scheewe,¹ Pia M. Pantel,¹
Cornelius G. Bachmann,¹ Peter Dechent,² Jens Ebentheuer,³
Alexander Mohr,⁴ Walter Paulus,¹ and Claudia Trenkwalder³

¹Department of Clinical Neurophysiology, Georg-August University, Göttingen, Germany

²MR-Research in Neurology and Psychiatry, Georg-August University, Göttingen, Germany

³Paracelsus Elena Klinik, Center of Parkinsonism and Movement Disorders, Kassel, Germany

⁴Department of Neuroradiology, Georg-August University, Göttingen, Germany

Abstract: Voxel-based morphometry (VBM) shows a differentiated pattern in patients with atypical Parkinson syndrome but so far has had little impact in individual cases. It is desirable to translate VBM findings into clinical practice and individual classification. To this end, we examined whether a support vector machine (SVM) can provide useful accuracies for the differential diagnosis. We acquired a volumetric 3D T1-weighted MRI of 21 patients with idiopathic Parkinson syndrome (IPS), 11 multiple systems atrophy (MSA-P) and 10 progressive supranuclear palsy (PSP), and 22 healthy controls. Images were segmented, normalized, and compared at group level with SPM8 in a classical VBM design. Next, a SVM analysis was performed on an individual basis with leave-one-out cross-validation. VBM showed a strong white matter loss in the mesencephalon of patients with PSP, a putaminal grey matter loss in MSA, and a cerebellar grey matter loss in patients with PSP compared with IPS. The SVM allowed for an individual classification in PSP versus IPS with up to 96.8% accuracy with 90% sensitivity and 100% specificity. In MSA versus IPS, an accuracy of 71.9% was achieved; sensitivity, however, was low with 36.4%. Patients with IPS could not be differentiated from controls. In summary, a voxel-based SVM analysis allows for a reliable classification of individual cases in PSP that can be directly clinically useful. For patients with MSA and IPS, further developments like quantitative MRI are needed. *Hum Brain Mapp* 32:1905–1915, 2011. © 2011 Wiley Periodicals, Inc.

Key words: voxel-based morphometry; support vector machine; individual classification; Parkinson syndrome; progressive supranuclear palsy; multiple systems atrophy

INTRODUCTION

Voxel-based morphometry (VBM) is an automated method to statistically analyze changes in local brain morphology as measured by whole-brain structural MRI data. Various studies have applied VBM in idiopathic Parkinson syndrome (IPS), yet changes were mostly subtle, did not survive correction for multiple comparisons, or were only present in later stages of the disease particularly with additional dementia (Burton et al., 2004; Feldmann et al., 2008). In contrast to IPS, which usually responds well to

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Niels K. Focke, Department of Clinical Neurophysiology, Georg-August University, Göttingen, Germany. E-mail: nfocke@uni-goettingen.de

Received for publication 23 April 2010; Revised 6 August 2010; Accepted 9 August 2010

DOI: 10.1002/hbm.21161

Published online 18 January 2011 in Wiley Online Library (wileyonlinelibrary.com).

TABLE I. Demographic details of controls and patient groups

	<i>N</i>	Age (mean ± SD)	Sex (female/male)
Controls	22	69.3 ± 9.1	9/13
IPS	21	65.2 ± 8.0	6/15
MSA-P	11	62.5 ± 8.0	5/6
PSP	10	67.0 ± 4.0	3/7
Total	64	66.4 ± 8.2	23/41

medical treatment, atypical Parkinson syndromes (APS), e.g., multiple system atrophy (MSA) or progressive supranuclear palsy (PSP) are clinically similar yet different entities and are characterized by a poor efficacy of dopaminergic treatment and more rapid disease progression. In MSA, neuronal loss has been described histopathologically in the nigrostriatal and olivopontocerebellar systems (Wenning et al., 1997). In PSP, typical neurofibrillary tangles and neurophil threads are often located in the brainstem and basal ganglia (Hauw et al., 1994). VBM studies in MSA and PSP have identified corresponding volumetric changes in patients compared to controls (Josephs et al., 2008; Minnerop et al., 2007). So far, all these studies have been conducted on a group level. This approach is useful to understand more about the neurobiology of the disease, but less informative in a clinical context focusing on the differential diagnosis of single patients where imaging techniques are of utmost importance. On an individual basis, (semi-) automated measurements have been proposed as surrogate parameters, such as midbrain diameter, manual volumetry, and area measurements (Quattrone et al., 2008; Schulz et al., 1999; Warmuth-Metz et al., 2001). These methods have limitations including the sole use of a preselected region and relying on manual input that is inevitably subject to operator bias and/or are time consuming. To overcome these shortcomings, a fully automated approach for individual classification based on whole-brain structural data would be valuable. Recently, support vector machines (SVM) have been applied to VBM data for this purpose and high predictions rates have been reported for dementia, Huntington disease, or autism spectrum disorder (Ecker et al., 2010; Klöppel et al., 2008, 2009). In the current study, we aimed to evaluate the potential of SVM for the differential diagnosis of patients with IPS, MSA, and PSP using whole-brain voxel-based data. To this end, we focused on diagnostically meaningful differentiations, comparing patients with IPS against controls and patients with APS. The latter, in particular, can be very challenging even for movement disorder specialists.

METHODS

Subjects

We enrolled 42 patients (21 IPS, 11 MSA-P [parkinsonian subtype of MSA], and 10 PSP) and 22 healthy elderly sub-

jects into the study between 2006 and 2009 after written informed consent. Demographic details are shown in Table I; a flow chart of the enrolled subjects is displayed in Figure 1. There was no statistically significant difference (two-tailed $P < 0.05$) between the groups concerning age, total intracranial volume (TIV) (independent sample t test), or sex (Mann-Whitney U test). Disease duration was significantly shorter ($P = 0.015$) for the patients with PSP (mean 2.4 ± 2.3 years) compared with patients with IPS (mean 5.5 ± 3.4 years); compared with patients with MSA-P (mean 4.4 ± 2.6 years) a trend ($P = 0.083$) was noted. Patients with MSA-P and IPS showed no significantly different disease duration. All patients with atypical Parkinson syndrome (MSA-P and PSP) were recruited from the Paracelsus Elena Klinik, Kassel (Germany), a specialized movement disorders hospital. Patients with IPS were recruited from the University Hospital Göttingen and the Paracelsus Elena Klinik Kassel (both Germany). Patients with IPS were classified according to UK Brain Bank Criteria (Hughes et al., 1992), MSA-P according to the second consensus statement (Gilman et al., 2008), and PSP according to the clinical research criteria (Litvan et al., 1996). Clinical details are shown in the supplementary material. All cases were rereviewed before the final analysis

Figure 1.

Flow chart of enrolled subjects. IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy.

for proof of consistency. As control group, neurologically normal, elderly subjects were recruited from the general population in the Göttingen area by advertisements, from congregations, and leisure groups. The study was approved by the University Hospital of Göttingen ethics committee.

Imaging Protocol

MR examinations were carried out on a 3T whole-body MR system (Magnetom Trio, Siemens Healthcare, Erlangen, Germany) using the body coil for transmission and an eight-channel head coil for signal reception (Invivo, Gainesville, FL). A T1-weighted 3D dataset with 1 mm isotropic resolution (MPRAGE, TI = 900 ms, $\alpha = 9^\circ$, TE = 3.26, TR = 2,250 ms) was acquired as recommended in the Alzheimer's disease neuroimaging initiative (Jack et al., 2008).

Image Processing

The original DICOM images were transferred to a Linux-based image server and converted to 3D NIFTI format using MRIConvert (<http://lcnj.uoregon.edu/~jolinda/MRIConvert/>). The further image processing was done with Statistical Parametric Mapping version 2008 (SPM8, <http://www.fil.ion.ucl.ac.uk/spm/>) running on a Matlab 7.7 platform (The MathWorks, Natick, MA). The 3D-T1 weighted images were segmented into tissue classes (grey matter, white matter, and CSF/non-brain) using the SPM8 unified segmentation with default settings. TIV was estimated by summing up all probability maps for CSF, grey, and white matter in native space. The grey and white matter maps were then spatially normalized with the DARTEL toolbox and resampled to MNI space in 1.5 mm cubic resolution with modulation (including the affine part of the normalization) to preserve the local tissue volumes (Ashburner, 2007). The resulting maps were smoothed with an 8 mm FWHM Gaussian kernel. Group comparisons were set up as two-sample t tests including the TIV as regressor of no interest, with the analysis region defined by an absolute probability threshold of 0.05. The contrast was generated as one-tailed t contrast with correction for multiple comparisons (peak-level family-wise error rate [FWE] $P < 0.05$) and, as an exploratory test, with a threshold of $P < 0.001$ (uncorrected) and a cluster extent threshold of 50 voxels. Resulting t maps were superimposed on the averaged, normalized, and bias-corrected T1-weighted images of all subjects for visualization with MRIcron (<http://www.cabiatl.com/mricro/mricron/>).

Support Vector Machine

The SVM analysis was performed with libsvm v. 2.9 (www.csie.ntu.edu.tw/~cjlin/libsvm/) (Chang and Lin, 2001). In brief, a SVM is an algorithm that is trained with a sample of known target values, i.e., classes. Then, the SVM will estimate a hyperplane that separates these

classes; this hyperplane is defined by the so-called support vectors which relate to the observations closest to the hyperplane and thus most difficult to classify. New datasets can subsequently be presented to the trained model and are classified depending on which side of the hyperplane they fall. An SVM allows for several different ways to generate the hyperplane defined by the support vectors. To prevent overfitting of the data and improve generalizability, we used a linear kernel in a C-SVM (with libsvm default $C = 1$). The kernel matrix was generated from all normalized and smoothed grey or white matter maps by calculating the dot product for all possible pairings resulting in a matrix with $N \times N$ elements (N = number of subjects). The dot product was generated by multiplying each voxel of each image pair and calculating the sum over all voxels as one element of the matrix. Although all voxels of the image were used, only those with a value of more than 0 in both images in question will contribute to the total sum thereby eliminating the effect of remote/non-brain voxels. The final $N \times N$ matrix was then used to train the SVM. Although the SVM is capable of working with a large number of features, in our case one feature per voxel, not all of these will be equally relevant for the classification. Therefore, in addition, a local weighting was imposed on the modulated maps to incorporate a-priori anatomical information about the groups. This was achieved by generating an unthresholded SPM F-contrast map (= weighting map) between the groups in comparison. The advantage of an SPM F-map is that it has no negative values and increases and decreases are treated equally. This weighting map was then multiplied with each individual tissue map before the dot product was calculated. Thus, each voxel was individually weighted proportionally to the difference on the group level. Prediction accuracy (overall rate of correct classification), sensitivity (percentage of true positives/all positives), and specificity (percentage of true negatives/all negatives) was assessed by a leave-one-out strategy: each subject was in turn removed from the training set as well as the generation of the weighting map. The left out subject was then classified with the trained model to evaluate the accuracy of the algorithm. Each of the resulting models (training set and weighting map) was consequently naive to the subject in question; thus, prediction accuracy can be seen as a measure for generalizability. The 95% confidence intervals were calculated using the efficient-score method (corrected for continuity) (Newcombe, 1998).

As additional visualization, we generated an image reflecting the importance of each voxel for the SVM classification. This was performed by calculating the primal variable w which is derived from the dual weight vector α assigned to each scan during the SVM training step (compare libsvm documentation). This (signed) parameter w was then multiplied with each normalized and smoothed tissue map and summed on a voxel basis (Klöppel et al., 2008). The resulting image was color coded and overlaid on the averaged T1-weighted image of all subjects (see Fig. 5).

TABLE II. SPM results of grey matter analysis

<i>P</i> (FWE)	<i>P</i> (FWE/ cluster)	<i>k</i>	<i>T</i>	MNI	Atlas
IPS < controls					
0.949	0.715	163	3.8 3.4	2 -48 72 -9 -60 72	Right precuneus
IPS > controls					
No clusters					
MSA-P < IPS					
0.921	0.911	89	4	21 3 6	Right putamen
MSA-P > IPS					
0.577	0.731	161	4.5	24 -42 69	Right superior parietal lobule
0.737	0.96	61	4.3	2 -73 49	Right precuneus
0.839	0.481	258	3.5	0 -69 37	Left precuneus
			4.2	-6 -49 67	
			3.9 3.7	-16 -46 64 -9 -58 69	
PSP < IPS					
0.047	<0.001	5101	5.9	16 -39 -30	(Right cerebellum)
			5.3	15 -61 -20	
			5.3	22 -57 -17	
0.158	<0.001	4216	5.4	-28 -76 -36	(Left cerebellum)
			5.3	-20 -70 -23	
			4.3	-42 -78 -33	
0.826	0.84	118	4.3	14 -91 37	Right occipital pole
0.895	0.948	71	4.2	54 5 21	right precentral gyrus
0.984	0.953	68	3.9	16 -46 -45	(Right cerebellum)
PSP > IPS					
0.498	0.923	84	4.7	-14 -57 72	Left lateral occipital cortex, superior division
0.971	0.516	229	3.9	22 -45 69	Right superior parietal lobule
MSA-P > PSP					
No clusters					
MSA-P < PSP					
No clusters					

Clusters and local maxima (more than 8 mm apart) of the grey matter SPM analysis thresholded at $P < 0.001$ (uncorrected) are shown. Peak and cluster level corrected P values (FWE) are also displayed. Clusters surviving peak level FWE correction are printed in **bold**. Anatomical position was determined using FSL view atlas tools and the Harvard-Oxford Cortical/Subcortical Structural Atlas; only the structure with the highest probability is reported. When no label was found in the atlas, the region was estimated visually and printed in brackets.

IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy. k = size of the cluster in voxels.

RESULTS

VBM revealed grey matter volume loss in patients with PSP which was strongest in the cerebellum, and reduced white matter in the midbrain. Changes in patients with IPS and MSA-P were more subtle and did not survive full-brain error correction. Detailed results of the SPM8 group

comparison are shown in Tables II and III and Figures 2, 3, and 4.

IPS Versus Controls

In the comparison of patients with IPS against controls, a grey matter reduction in the right precuneus was found.

TABLE III. SPM results of white matter analysis

<i>P</i> (FEW)	<i>P</i> (FWE/ cluster)	<i>k</i>	<i>T</i>	MNI	Atlas
IPS < controls					
No clusters					
IPS > controls					
0.465	0.798	90	4.1	52 -46 25	Right superior longitudinal fasciculus/right angular gyrus
MSA-P < IPS					
0.522	0.78	103	4.3	14 -15 75	Right corticospinal tract/right precentral gyrus
0.73	0.468	246	4.0	32 -7 -0	Right external capsule/right putamen
0.868	0.843	76	3.8	-20 -27 73	Left corticospinal tract/Left precentral gyrus
MSA-P > IPS					
No clusters					
PSP < IPS					
<0.001	<0.001	19,990	8.6	-2 -21 -2	Left anterior thalamic radiation/left thalamus
			8.4	-2 -19 -11	
			6.7	2 -43 -30	
PSP > IPS					
0.237	0.803	95	4.8	10 -45 64	No label found/right postcentral gyrus
MSA-P > PSP					
0.087	<0.001	5,032	6.3	0 -21 0	No label found/midline thalamus
			5.8	0 -16 -11	
			5.2	-12 -10 12	
MSA-P < PSP					
No clusters					

Clusters and local maxima (more than 8 mm apart) of the white matter SPM analysis thresholded at $p < 0.001$ (uncorrected) are shown. Peak and cluster level corrected *P* values (FWE) are also displayed. Clusters surviving peak level FWE correction are printed in **bold**. Anatomical position was determined using FSL view atlas tools and the JHU White-Matter Tractography (first label) and Harvard-Oxford Cortical/Subcortical Structural Atlas (second label); only the structure with the highest probability is reported. IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy. *k* = size of the cluster in voxels.

In the white matter analysis, a small cluster of increased volume was detected in the right superior longitudinal fasciculus/angular gyrus. None of these were significant after full-brain error correction. There were no significant grey matter increases or white matter decreases.

MSA-P Versus IPS

SPM found significantly reduced grey matter in the right putamen of patients with MSA-P; a contralateral cluster did not surpass the chosen cluster threshold ($k = 14$ voxels, $T = 3.55$, MNI -20 3 4) and is therefore not shown. Patients with MSA-P compared with IPS had an increased grey matter volume in the right superior parietal lobule and the pre-

cuneus bilaterally. In the white matter analysis, patients with MSA-P had a reduced volume in the right external capsule/right putamen and bilateral corticospinal tract/precentral gyrus; the cluster in the left external capsule/putamen, again, did not reach the set cluster threshold ($k = 11$ voxels, $T = 3.15$, MNI -30 -7 0). No significant white matter increases were found. Again none of these clusters were significant after full-brain error correction.

PSP Versus IPS

Patients with PSP showed a significantly reduced grey matter volume in the cerebellar hemispheres, right occipital pole, and right precentral gyrus. The cluster in the left cerebellum survived full-brain error correction with FWE

Figure 2.

Results of grey matter SPM8 analysis. Significant grey matter reductions at $P < 0.001$ (uncorrected) with a cluster threshold of $k > 50$ voxels are overlaid on the brain extracted normalized, averaged T1-weighted image of all subjects. The color scale represents SPM t scores. **A**, patients with IPS versus controls (red-yellow = IPS < controls, blue-green = IPS > controls), axial cut plane. **B**, patients with MSA-P versus IPS (red-yellow = MSA < IPS, blue-green = MSA-P > IPS), axial cut plane. **C**, patients with PSP versus IPS (red-yellow = PSP < IPS, blue-green = PSP > IPS), axial cut plane. **D**, patients with IPS versus controls (red-

yellow = IPS < controls, blue-green = IPS > controls), sagittal cut plane. **E**, patients with MSA-P versus IPS (red-yellow = MSA < IPS, blue-green = MSA-P > IPS), sagittal cut plane. **F**, patients with PSP versus IPS (red-yellow = PSP < IPS, blue-green = PSP > IPS), sagittal cut plane. No suprathreshold clusters were detected in the comparison of patients with MSA-P and PSP; this analysis is therefore not shown. IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy.

$P < 0.05$. In the white matter analysis, a highly significant volume reduction was detected encompassing the dorsal aspect of the pons, the mesencephalon, dorsal basal ganglia, and cerebellar peduncles. This cluster was still highly significant after full-brain error correction.

MSA-P Versus PSP

In the final comparison of patients with MSA-P versus PSP, no significant grey matter differences but a reduced white matter volume in the PSP group was detected in the mesencephalon. This cluster was similar to the decrease found in the PSP versus IPS comparison but was smaller and less significant. In addition, it narrowly failed to survive full-brain error correction (FWE corrected $P = 0.087$).

SVM Analysis

A synopsis of accuracy, sensitivity, and specificity of the SVM analysis both with and without local weighting is

given in Table IV. Overall, structural MRI did not permit a separation of patients with IPS from controls by any of the applied models. Accuracy rates were significantly better than chance for the classification of patients with MSA-P against IPS with grey matter maps but the sensitivity was still poor. Patients with PSP, however, were reliably separated from patients with IPS with an accuracy of up to 96.8%, 90% sensitivity, and 100% specificity (using white matter analysis and local weights). The 95% confidence intervals were 57.1%–99.5% (sensitivity), 79.9%–100% (specificity), and 82.0%–99.8% (accuracy) in this analysis. Prediction was somewhat worse for the grey matter based classification but did reach 87.1% accuracy. Results of the comparison against controls were very similar in the white matter stream with an accuracy of 93.8%, 90% sensitivity, and 95.5% specificity with local weights. For the grey matter-based analysis, accuracy and especially sensitivity were lower than in the PSP/IPS comparison. A whole-brain image displaying the influence of each voxel for the classification in the unweighted PSP versus IPS comparison is shown in Figure 5. The separation of patients with MSA-P

Figure 3.

Results of white matter SPM8 analysis. Significant white matter reductions at $P < 0.001$ (uncorrected) with a cluster threshold of $k > 50$ voxels are overlaid on the brain extracted normalized, averaged T1-weighted image of all subjects. The color scale represents SPM t scores. **A**, patients with IPS versus controls (red-yellow = IPS < controls, blue-green = IPS > controls), axial cut plane. **B**, patients with MSA-P versus IPS (red-yellow = MSA < IPS, blue-green = MSA-P > IPS), axial cut plane. **C**, patients with PSP versus IPS (red-yellow = PSP < IPS, blue-green = PSP > IPS), axial cut plane. **D**, patients with MSA-P versus PSP (red-yellow = MSA-P > PSP, blue-green = MSA-P > PSP), axial cut

plane. **E**, patients with IPS versus controls (red-yellow = IPS < controls, blue-green = IPS > controls), sagittal cut plane. **F**, patients with MSA-P versus IPS (red-yellow = MSA < IPS, blue-green = MSA-P > IPS), sagittal cut plane. **G**, patients with PSP versus IPS (red-yellow = PSP < IPS, blue-green = PSP > IPS), sagittal cut plane. **H**, patients with MSA-P versus PSP (red-yellow = MSA-P > PSP, blue-green = MSA-P > PSP), sagittal cut plane. IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy.

from PSP was not possible with the grey matter maps; the white matter-based classification with local weights, however, still gave 76.2% accuracy. In the comparison of patients with MSA-P against controls classification results were different. In this case, the best accuracy (78.8%) could be achieved with white matter images without local weights. Overall, local weights improved the classification for all comparisons of patients with PSP against IPS; and, also, in the patients with PSP against controls and PSP against MSA-P white matter-based analysis. Weighting had little to no effect in the other MSA-P against IPS and the IPS against controls comparisons. In the comparison of

MSA-P against controls, weighting had a negative impact on the classification.

DISCUSSION

We found highly significant white matter reductions in PSP that allowed for a remarkably accurate individual separation from IPS and healthy controls. For patients with MSA-P, separation from the IPS and the control cohort was only marginally better than chance; for patients with IPS, no separation from controls was possible. Overall, VBM findings are in keeping with the previous literature.

Figure 4.

PSP versus IPS white matter analysis after error correction. Significant white matter reductions of patients with PSP versus IPS (red-yellow = PSP < IPS, blue-green = PSP > IPS) at $P < 0.05$ (corrected for multiple comparison with family-wise error correction) are overlaid on the brain extracted normalized, aver-

aged T1-weighted image of all subjects. The color scale represents SPM t scores. **A**, axial cut plane; **B**, sagittal cut plane. IPS, idiopathic Parkinson syndrome; PSP, progressive supranuclear palsy.

For PSP, we found a significant white matter loss in the brainstem reaching up to the level of the basal ganglia and cerebellar grey matter loss. Although the grey matter segmentation of the cerebellum can be technically difficult due to the highly folded microstructure of the cerebellar folia, this is in keeping with previous VBM studies (Boxer et al., 2006; Brenneis et al., 2004; Price et al., 2004) and the neuropathological description of the disease (Hauw et al., 1994; Steele et al., 1964). The unsupervised SVM algorithm allowed for an excellent, individual prediction with best results in the white matter stream with local weights both in comparison with patients with IPS as well as to healthy controls. In the grey matter-based analysis, the classification results were better in the comparison against IPS than against controls. This could be an indication that grey matter changes in patients with IPS are a relevant factor. In the clinical setting, the need to separate patients with PSP from IPS will be the much more common scenario. Compared with SVM studies in other neurological or psychiatric entities, accuracy was comparable with what could be achieved in Alzheimer disease and higher than in Huntington disease or autism spectrum disorder (Ecker et al., 2010; Klöppel et al., 2008, 2009).

Grey matter loss of patients with MSA-P has been shown by previous VBM (Minnerop et al., 2007), manual

volumetry (Schulz et al., 1999), and neuropathological studies (Wenning et al., 1997). We could confirm this finding for the grey and white matter analysis, although this did not survive full-brain error correction in our relatively small cohorts. Additionally, we detected a white matter loss in the precentral gyrus bilaterally that has also been previously described (Minnerop et al., 2007). The SVM analysis yielded a statistically significant accuracy in the grey matter stream with local weights in comparison with patients with IPS, however, with low sensitivity this is unlikely to be of direct clinical value. In the comparison against controls, a slightly higher accuracy could be found in the white matter-based analysis without local weights. Again, the specificity was high (90.1%) but the sensitivity was unsatisfactory (54.6%). Given the consistently reported volumetric differences in MSA, e.g., the putaminal volume loss this lack of sensitivity is surprising. One explanation could be that segmentation of T1-weighted images can be inaccurate due to the iron content of the basal ganglia (Helms et al. 2009) that is even more pronounced in MSA. This also increases susceptibility effects, especially at higher field strength. Other imaging modalities like T2* mapping or diffusion (tensor) imaging may be better suited to classify patients with MSA (Schocke et al., 2004; von Lewinski et al., 2007).

Figure 5.

Relevance of each voxel for the SVM classification. The color scale symbolizes the importance of each voxel in the PSP versus IPS SVM classification (unweighted, white matter based). This SVM training was done with the whole sample (without the leave-one out design) for visualization only. The blue color indi-

cates voxels where a volume decrease is associated with a classification toward PSP, red symbolizes the opposite. Details about the generation of this “w-map” are given in the method section of the article. **A**, axial cut plane; **B**, sagittal cut plane. IPS, idiopathic Parkinson syndrome; PSP, progressive supranuclear palsy.

TABLE IV. SVM classification results

Local weights	Grey matter		White matter	
	No	Yes	No	Yes
IPS vs. controls				
Accuracy (<i>P</i> value)	39.53% (NS)	39.53% (NS)	41.86% (NS)	37.31% (NS)
Sensitivity	28.57%	33.33%	38.10%	28.57%
Specificity	50.00%	45.45%	45.45%	45.45%
MSA-P vs. IPS				
Accuracy (<i>P</i> value)	62.50% (NS)	71.87% (0.010)	65.63% (NS)	65.63% (NS)
Sensitivity	27.27%	36.36%	27.27%	27.27%
Specificity	80.95%	90.48%	85.71%	85.71%
PSP vs. IPS				
Accuracy (<i>P</i> value)	87.10% (<0.001)	87.10% (<0.001)	90.32% (<0.001)	96.77% (<0.001)
Sensitivity	60.00%	70.00%	80.00%	90.00%
Specificity	100.00%	95.24%	95.24%	100.00%
MSA-P vs. PSP				
Accuracy (<i>P</i> value)	38.10% (NS)	33.33% (NS)	61.9% (NS)	76.19% (0.013)
MSA-P vs. controls				
Accuracy (<i>P</i> value)	69.70% (0.017)	63.63% (NS)	78.78% (<0.001)	51.52% (NS)
Sensitivity	27.27%	27.27%	54.55%	18.18%
Specificity	90.91%	81.82%	90.91%	68.18%
PSP vs. controls				
Accuracy (<i>P</i> value)	71.88% (0.010)	65.63% (NS)	84.37% (<0.001)	93.75% (<0.001)
Sensitivity	40.00%	20.00%	70.00%	90.00%
Specificity	86.36%	86.36%	90.91%	95.45%

Overall accuracy, sensitivity, and specificity of the SVM classification based on grey and white matter maps and with/without local weights are shown (leave-one-out design). Accuracy *P* value is given when the likelihood of achieving this binominal distribution by chance is <0.05.

NS, not significant; IPS, idiopathic Parkinson syndrome; MSA-P, multiple systems atrophy (parkinsonian variant); PSP, progressive supranuclear palsy.

VBM studies of patients with IPS compared with healthy controls have shown a heterogeneous pattern, with significant, widespread atrophy only late in the disease especially with concomitant dementia (Beyer et al., 2007; Burton et al., 2004; Feldmann et al., 2008; Martin et al., 2009). In our study, with chronic (mean disease duration 5.5 years) but nondemented and mild to moderately affected patients with IPS (Hoehn&Yahr stages mostly < = 3), VBM showed a grey matter atrophy in the precuneus/superior parietal lobule. This finding was evident in comparison with controls and was reproduced indirectly (as relative increase) in the analyses against both patients with MSA-P and PSP. This area has also been reported by other studies (Burton et al., 2004; Pereira et al., 2009), however, in these, it was not the most significantly affected region. Moreover, it did not survive full-brain error correction and was located very close to the midline reaching into non-brain areas which could be explained by the lower residual variance in these regions. Following the

concept of Braak et al. (2003), widespread cortical atrophy is expected in the late stages of the disease, when a cognitive decline occurs. In a mixed cohort (disease duration 5.5 ± 3.4 years), however, the common denominator should be expected in the brainstem and mesocortex. Why this is not detected consistently by VBM remains unclear but could have technical reasons due to segmentation difficulties. With inconsistent findings in VBM, it is unsurprising that the SVM was unable to separate patients with IPS from controls. Interestingly, the accuracy was yet lower than the 50% expected by chance. This is an indication for a strong heterogeneity between patients and may explain the lack of differentiation in this stage of the disease; the grey and white matter pattern of each individual patient with IPS seems to be only poorly predicted by the remaining group.

Our data and the previous literature indicate that volumetric tools based on standard T1-weighted images may not be the ideal choice to separate patients with IPS from

controls. Other (quantitative) MRI methods, e.g., diffusion tensor imaging may be better suited in this scenario (Menke et al., 2009; Vaillancourt et al., 2009). Also ultrasound of the substantia nigra is an established method for this classification (Berg et al., 2008). Our overall findings are in keeping with another recent study using an SVM for the classification of Parkinson syndromes (Duchesne et al., 2009). The authors of this study have pooled the patients with MSA and PSP, restricted the analysis to the hindbrain and reported an overall classification accuracy of 91%; our data indicate that this could be an average between the good separation of PSP and the mediocre classification of patients with MSA compared with IPS.

The diagnosis in our study was based on clinical information only; so far, no neuropathological information was available. Nevertheless, the assessment was based on clinical criteria, all available technical methods such as ultrasound of the substantia nigra and follow-up observations and was performed by experienced specialists in movement disorders; also, no ambiguous cases were included. A recent study comparing the clinical MSA classification to neuropathological validation showed a high positive predictive value of 95% for “possible” and 100% for “probable” cases (Osaki et al., 2009). The patients enrolled in our study were in the chronic phase of the disease with mean disease duration of 4.5 years over all groups. Therefore, our findings may not be fully representative for the very early stage, where clinical classification is most difficult necessitating auxiliary tools like an SVM prediction. Also, the SVM algorithm currently does not permit to include confounding regressor like TIV; this, however, was not significantly different between the groups and therefore should not have had a strong impact on the results.

The effect of local weights was not consistent; it improved the classification for patients with PSP against IPS and partially for patients with PSP against MSA-P. It, however, had little to no effect in the other comparisons and even deteriorated the MSA-P against controls comparison. The exact reason for this remains unclear, but it is possible that patients with PSP showed a more localized and homogeneous atrophy pattern. In theory, local weighting would be most beneficial in this scenario, whereas in cases with widespread/distributed changes with higher within-group variability, the effect would be smaller or even detrimental (Davatzikos, 2004).

Overall, we could show that for the differential diagnosis of PSP from IPS, a SVM is a useful tool. It can offer a fully automated, whole-brain classification with high accuracy and, in our sample, up to 100% specificity. We envisage a training dataset of pathologically confirmed PSP and, with further technical improvements, MSA cases to improve the reliability of this prediction as early in the clinical course as possible. To this end, a longitudinal study of de novo cases with retrospective analysis is needed. Our findings will, however, have to be reproduced by other groups and in larger samples before these methods can be applied routinely.

ACKNOWLEDGMENTS

We thank Stefan Klöppel for his advice about the support vector machines.

REFERENCES

- Ashburner J (2007): A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113.
- Berg D, Godau J, Walter U (2008): Transcranial sonography in movement disorders. *Lancet Neurol* 7:1044–1055.
- Beyer MK, Larsen JP, Aarsland D (2007): Gray matter atrophy in Parkinson disease with dementia and dementia with Lewy bodies. *Neurology* 69:747–754.
- Boxer AL, Geschwind MD, Belfor N, Gorno-Tempini ML, Schauer GF, Miller BL, Weiner MW, Rosen HJ (2006): Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Arch Neurol* 63:81–86.
- Braak H, Tredici KD, Rüb U, de Vos RAI, Jansen Steur ENH, Braak E (2003): Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol Aging* 24:197–211.
- Brenneis C, Seppi K, Schocke M, Benke T, Wenning GK, Poewe W (2004): Voxel based morphometry reveals a distinct pattern of frontal atrophy in progressive supranuclear palsy. *J Neurol Neurosurg Psychiatry* 75:246–249.
- Burton EJ, McKeith IG, Burn DJ, Williams ED, O'Brien JT (2004): Cerebral atrophy in Parkinson's disease with and without dementia: A comparison with Alzheimer's disease, dementia with Lewy bodies and controls. *Brain* 127:791–800.
- Chang C-C, Lin C-J (2001): LIBSVM: A library for support vector machines. Available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Davatzikos C (2004): Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. *Neuroimage* 23:17–20.
- Duchesne S, Rolland Y, Verin M (2009): Automated computer differential classification in Parkinsonian Syndromes via pattern analysis on MRI. *Acad Radiol* 16:61–70.
- Ecker C, Rocha-Rego V, Johnston P, Mourao-Miranda J, Marquand A, Daly EM, Brammer MJ, Murphy C, Murphy DG (2010): Investigating the predictive value of whole-brain structural MR scans in autism: A pattern classification approach. *Neuroimage* 49:44–56.
- Feldmann A, Illes Z, Kosztolanyi P, Illes E, Mike A, Kover F, Balas I, Kovacs N, Nagy F (2008): Morphometric changes of gray matter in Parkinson's disease with depression: A voxel-based morphometry study. *Mov Disord* 23:42–46.
- Gilman S, Wenning GK, Low PA, Brooks DJ, Mathias CJ, Trojanowski JQ, Wood NW, Colosimo C, Durr A, Fowler CJ, Kaufmann H, Klockgether T, Lees A, Poewe W, Quinn N, Revesz T, Robertson D, Sandroni P, Seppi K, Vidailhet M (2008): Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 71:670–676.
- Hauw J-J, Daniel SE, Dickson D, Horoupian DS, Jellinger K, Lantos PL, McKee A, Tabaton M, Litvan I (1994): Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Neurology* 44:2015–2019.
- Helms G, Draganski B, Frackowiak R, Ashburner J, Weiskopf N (2009): Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps. *Neuroimage* 47:194–198.
- Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992): Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 55:181–184.

- Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Whitwell JL, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DLG, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Deb-bins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW, Study A (2008): The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 27:685–691.
- Josephs KA, Whitwell JL, Dickson DW, Boeve BF, Knopman DS, Petersen RC, Parisi JE, Jack CR Jr (2008): Voxel-based morphometry in autopsy proven PSP and CBD. *Neurobiol Aging* 29:280–289.
- Klöppel S, Chu C, Tan GC, Draganski B, Johnson H, Paulsen JS, Kienzle W, Tabrizi SJ, Ashburner J, Frackowiak RS (2009): Automatic detection of preclinical neurodegeneration: Pre-symptomatic Huntington disease. *Neurology* 72:426–431.
- Klöppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, Fox NC, Jack CR Jr, Ashburner J, Frackowiak RSJ (2008): Automatic classification of MR scans in Alzheimer's disease. *Brain* 131:681–689.
- Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, Goetz CG, Golbe LI, Grafman J, Growdon JH, Hallett M, Jan-kovic J, Quinn NP, Tolosa E, Zee DS (1996): Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): Report of the NINDS-SPSP international workshop. *Neurology* 47:1–9.
- Martin WRW, Wieler M, Gee M, Camicioli R (2009): Temporal lobe changes in early, untreated Parkinson's disease. *Mov Disord* 24:1949–1954.
- Menke RA, Scholz J, Miller KL, Deoni S, Jbabdi S, Matthews PM, Zarei M (2009): MRI characteristics of the substantia nigra in Parkinson's disease: A combined quantitative T1 and DTI study. *Neuroimage* 47:435–441.
- Minnerop M, Specht K, Ruhlmann J, Schimke N, Abele M, Weyer A, Wullner U, Klockgether T (2007): Voxel-based morphometry and voxel-based relaxometry in multiple system atrophy—A comparison between clinical subtypes and correlations with clinical parameters. *Neuroimage* 36:1086–1095.
- Newcombe RG (1998): Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat Med* 17:857–872.
- Osaki Y, Ben-Shlomo Y, Lees AJ, Wenning GK, Quinn NP (2009): A validation exercise on the new consensus criteria for multiple system atrophy. *Mov Disord* 24:2272–2276.
- Pereira JB, Junqué C, María-Jo, sacute, Martíe, Ramirez-Ruiz B, Bargalló N, Tolosa E (2009): Neuroanatomical substrate of visuospatial and visuoperceptual impairment in Parkinson's disease. *Mov Disord* 24:1193–1199.
- Price S, Paviour D, Scahill R, Stevens J, Rossor M, Lees A, Fox N (2004): Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and Parkinson's disease. *Neuroimage* 23:663–669.
- Quattrone A, Nicoletti G, Messina D, Fera F, Condino F, Pugliese P, Lanza P, Barone P, Morgante L, Zappia M, Aguglia U, Gallo O (2008): MR Imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy1. *Radiology* 246:214–221.
- Schocke MFH, Seppi K, Esterhammer R, Kremser C, Mair KJ, Czermak BV, Jaschke W, Poewe W, Wenning GK (2004): Trace of diffusion tensor differentiates the Parkinson variant of multiple system atrophy and Parkinson's disease. *Neuroimage* 21:1443–1451.
- Schulz JB, Skalej M, Wedekind D, Luft AR, Abele M, Voigt K, Dichgans J, Klockgether T (1999): Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. *Ann Neurol* 45:65–74.
- Steele JC, Richardson JC, Olszewski J (1964): Progressive supranuclear palsy: a heterogeneous degeneration involving the brain stem, basal ganglia and cerebellum with vertical gaze and pseudobulbar palsy, nuchal dystonia and dementia. *Arch Neurol* 10:333–359.
- Vaillancourt DE, Spraker MB, Prodoehl J, Abraham I, Corcos DM, Zhou XJ, Comella CL, Little DM (2009): High-resolution diffusion tensor imaging in the substantia nigra of de novo Parkinson disease. *Neurology* 72:1378–1384.
- von Lewinski F, Werner C, Jörn T, Mohr A, Sixel-Doring F, Trenk-walder C (2007): T2*-weighted MRI in diagnosis of multiple system atrophy. A practical approach for clinicians. *J Neurol* 254:1184–1188.
- Warmuth-Metz M, Naumann M, Csoti I, Solymosi L (2001): Measurement of the midbrain diameter on routine magnetic resonance imaging: A simple and accurate method of differentiating between Parkinson disease and progressive supranuclear palsy. *Arch Neurol* 58:1076–1079.
- Wenning GK, Tison F, Shlomo Yb, Daniel SE, Quinn NP (1997): Multiple system atrophy: A review of 203 pathologically proven cases. *Mov Disord* 12:133–147.