Comparison of mean diffusivity, R2* relaxation rate and morphometric biomarkers for the clinical differentiation of parkinsonism

Lydia Chougar, François-Xavier Lejeune, Johann Faouzi, Benjamin Morino, Alice Faucher, Nadine Hoyek, David Grabli, Florence Cormier, Marie Vidailhet, Jean-Christophe Corvol, et al.

Parkinsonism and Related Disorders 108 (2023) 105287

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

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

Comparison of mean diffusivity, R2* relaxation rate and morphometric biomarkers for the clinical differentiation of parkinsonism

Lydia Chougar [a,b,c,d,*], François-Xavier Lejeune [e,f], Johann Faouzi [e,g], Benjamin Morino [d], Alice Faucher [h,i], Nadine Hoyek [j], David Grabli [k,l], Florence Cormier [k,l], Marie Vidailhet [c,e,k,l], Jean-Christophe Corvol [e,k,l], Olivier Colliot [e,g], Bertrand Degos [h,i], Stéphane Lehéricy [b,c,d]

[a] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, DMU DIAMENT, Department of Neuroradiology, F-75013, Paris, France
[b] ICM, Centre de NeuroImagerie de Recherche–CENIR, Paris, France
[c] ICM, Team "Movement Investigations and Therapeutics" (MOV'IT), Paris, France
[d] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, DMU DIAMENT, Department of Neuroradiology, F-75013, Paris, France
[e] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inserm, F-75013, Paris, France
[f] ICM, Data and Analysis Core, Paris, France
[g] Sorbonne Université, Institut du Cerveau - Paris Brain Institute - ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France
[h] Dynamics and Pathophysiology of Neuronal Networks Team, Center for Interdisciplinary Research in Biology, Collège de France, CNRS UMR7241/INSERM U1050, Université PSL, Paris, France
[i] Service de Neurologie, Hôpital Avicenne, Hôpitaux Universitaires de Paris Seine-Saint-Denis, APHP, Bobigny, France
[j] Department of Radiology, Hotel Dieu de France University Hospital, Faculty of Medicine, Saint Joseph University, Beirut, Lebanon
[k] Clinique des mouvements anormaux, Département de Neurologie, Assistance Publique Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, Paris, France
[l] ICM, Centre d'Investigation Clinique Neurosciences, Paris, France

ARTICLE INFO

Keywords:
Atypical parkinsonism
Parkinson's disease
Progressive supranuclear palsy
Multisystem atrophy
Diagnosis
MRI

ABSTRACT

Introduction: Quantitative biomarkers for clinical differentiation of parkinsonian syndromes are still lacking. Our aim was to evaluate the value of combining clinically feasible manual measurements of R2* relaxation rates and mean diffusivity (MD) in subcortical regions and brainstem morphometric measurements to improve the discrimination of parkinsonian syndromes.

Methods: Twenty-two healthy controls (HC), 25 patients with Parkinson's disease (PD), 19 with progressive supranuclear palsy (PSP) and 27 with multiple system atrophy (MSA, 21 with the parkinsonian variant -MSAp, 6 with the cerebellar variant -MSAc) were recruited. R2*, MD measurements and morphometric biomarkers including the midbrain to pons area ratio and the Magnetic Resonance Parkinsonism Index (MRPI) were compared between groups and their diagnostic performances were assessed.

Results: Morphometric biomarkers discriminated better patients with PSP (ratio: AUC 0.89, MRPI: AUC 0.89) and MSAc (ratio: AUC 0.82, MRPI: AUC 0.75) from other groups. R2* and MD measurements in the posterior putamen performed better in separating patients with MSAp from PD (R2*: AUC 0.89; MD: AUC 0.89). For the three-class classification "MSA vs PD vs PSP", the combination of MD and R2* measurements in the posterior putamen with morphometric biomarkers (AUC: 0.841) outperformed each marker separately. At the individual-level, there were seven discordances between imaging-based prediction and clinical diagnosis involving MSA. Using the new Movement Disorder Society criteria for the diagnosis of MSA, three of these seven patients were clinically reclassified as predicted by quantitative imaging.

Conclusion: Combining R2* and MD measurements in the posterior putamen with morphometric biomarkers improves the discrimination of parkinsonism.

* Corresponding author. Centre de NeuroImagerie de Recherche–CENIR, Institut du Cerveau–ICM, Hôpital Pitié-Salpêtrière, 47 Boulevard de l'Hôpital, 75651 Paris Cedex 13, France.
E-mail addresses: lydia.chougar@aphp.fr, chougar.lydia@gmail.com (L. Chougar).

https://doi.org/10.1016/j.parkreldis.2023.105287
Received 9 August 2022; Received in revised form 15 December 2022; Accepted 14 January 2023
Available online 18 January 2023

1. Introduction

Early and accurate diagnosis of parkinsonism including Parkinson's disease (PD), progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) remains challenging despite its importance for patient care and for recruitment into disease-modifying clinical trials. Clinical criteria were reported to have suboptimal overall accuracy for the diagnosis of MSA (62–79%) [1,2] and PSP (14–83%) [3,4] at the first visit. The diagnostic criteria for MSA were recently revised by the Movement Disorder Society and introduced a new category of clinically established MSA in which at least one brain MRI marker is required [1]. In the cerebellar variant of MSA (MSAc), MRI qualitative biomarkers include atrophy in the pons, cerebellar peduncles and cerebellum as well as increased T2/proton density signal in the middle cerebellar peduncles and the hot-cross-bun sign in the pons [5]. In the parkinsonian variant of MSA (MSAp), qualitative markers combine atrophy, signal decrease on iron-sensitive images and signal increase in diffusivity maps predominantly in the posterior putamen [5]. The cerebellar and parkinsonian patterns are often associated in patients. A number of quantitative MRI markers derived from the analysis of the same regions were also proposed in a research setting for the differentiation of MSA from PSP and PD. For MSAp, these markers include increased apparent diffusion coefficient or mean diffusivity (MD) using diffusion imaging, reflecting microstructural alterations due to cell death and myelin changes, and increased R2* relaxation rate, which is a proxy for iron deposition within tissues [5]. Changes were reported in MSAp versus PD [6–8], but results were less consistent between MSAp and PSP patients in the entire putamen with overlapping diffusivity [9,10] and iron deposition [11,12]. Measurements targeting the posterior putamen may be more discriminative [8]. Morphometric measurements including the midbrain to pons area ratio [13] and the Magnetic Resonance Parkinsonism Index (MRPI) [13–15] were shown to be efficient in distinguishing MSAc from healthy subjects, PSP and PD with accuracy values usually greater than 0.90 [13]. These measurements were less efficient to categorize MSAp from PD with accuracy values around 0.74 [13,14]. Previous studies have reported that combining markers in a multiparametric approach could improve the classification accuracy [16,17]. However, these methods are not yet usable in clinical practice.

In an effort to provide a transferable approach to clinical practice, we evaluated the value of combining clinically feasible standard brainstem morphometric measurements with manual measurements of R2* relaxation rates and mean diffusivity (MD) in subcortical regions to improve the differentiation of patients with MSA (and particularly MSAp), PSP and PD.

2. Material & methods

2.1. Population

Participants were prospectively and consecutively enrolled between 2017 and 2020 in the movement disorders clinic of the Pitié-Salpêtrière Hospital, Paris. Inclusion criteria were probable or possible diagnosis of PD [18], PSP [4] and MSA according to the previous consensus criteria [19] established by movement disorders specialists. Patients with MSA were subsequently evaluated according to the new Movement Disorder Society diagnostic criteria [1]. They were separated into parkinsonian (MSAp) and cerebellar (MSAc) subtypes depending on the clinically predominant subtype at diagnosis. Patients with a mixed pattern were assigned to the MSAp group. The clinical examination included the Unified Parkinson's Disease Rating Scale Part III (UPDRS III) scores. Healthy controls (HC) with no history of neurological or psychiatric disease were included. Subjects were excluded if they had any additional neurological disorder. Local institutional review boards approved the study (CPP Ile-de-France VI, 08012015).

2.2. MRI acquisition

Participants were scanned under clinical conditions for diagnostic purposes in the Neuroradiology Department of the hospital using a 3T Siemens Skyra system with a 64-channel head coil. The MRI protocol included three-dimensional (3D) high-resolution T1-weighted gradient-recalled echo sequence (magnetization prepared rapid acquisition with gradient-recalled echo, MPRAGE, 0.9-mm isovoxel size), diffusion tensor imaging (DTI) with 32 gradient-encoding directions (voxel size: 2 $\times$ 2 $\times$ 2.6) and gradient echo T2 acquisition with nine echo times (from 4.0 to 50.0 ms, voxel size: 0.8 $\times$ 0.8 $\times$ 2.4) for R2* relaxometry. Acquisition parameters are provided in Supplementary Table S1. Quality control was performed by visual inspection and images with significant motion artifacts or image distortions were discarded.

2.3. Data analysis

MD and R2* maps were calculated by the Siemens scanner as part of the routine clinical examination with no further image post-processing. Data were subsequently analyzed in the Picture Archiving and Communication System (PACS) system of the hospital. MD and R2* maps were reformatted in the intercommissural plane. Measurements were performed in standardized rounded shape regions of interest (ROIs) of 8-mm^2 surface area. ROIs were manually drawn on MD and R2* maps by an experienced radiologist blinded to the clinical status on six bilateral subcortical grey matter structures including the posterior putamen (PuP), the anterior putamen (PuA), the head of the caudate nucleus (CN), the globus pallidus (GP), the medial thalamus (Th) (row A) and the red nucleus (RN) as this region was shown to be particularly affected in PSP [12]. The regions were placed on the slice that best showed the structure of interest (i.e. with the largest area), avoiding Virchow-Robin spaces (Supplementary Fig. S1). The ratios of midbrain to pons sagittal areas and MRPI were measured on 3D T1-weighted images as previously explained [14] (Supplementary Fig. S2). Measurements (R2*, MD, midbrain to pons area ratio, MRPI) were performed twice by the same observer on two different sessions to assess the intra-rater reliability and by a second observer to assess the inter-rater reliability.

2.4. Statistical analysis

All statistical analyses were performed using R version 3.6.1 (R Core Team 2019).

2.4.1. Clinical data

Clinical and demographic data were compared between groups using the Kruskal-Wallis test, followed by pairwise Dunn's tests with Bonferroni correction, or the Fisher's exact test.

2.4.2. Reproducibility analysis

For each imaging modality (MD, R2*, midbrain to pons ratio, MRPI), intraclass correlation coefficients (ICCs) were calculated to assess intra-(ICC [1,3]: two-way mixed effects, absolute agreement, single rater) and inter-rater reliability (ICC [2,3]: two-way mixed effects, absolute agreement, multiple raters) using the psych package (v2.0.9). ICCs values of reliability were reported with 95% confidence interval (95% CI), and were interpreted as follows: $<$0.50, poor; 0.50 to 0.75, moderate; 0.75 to 0.90, good; $>$0.90, excellent.

2.4.3. Between-group comparisons

R2* and MD measurements were compared between groups using linear mixed-effect models (LMMs, one model for each biomarker) with covariate adjustment for age and sex. In these models, the predictors of interest included Group (disease groups), Region (ROIs), Side (left or right hemisphere), and their interaction terms as fixed effects, while the subject identifier was assigned as a random (intercept) effect to account

for the repeated measurements acquired in both sides of the ROIs for the same subject.

We also considered the relatives values of MD and R2* measurements in the posterior putamen to each of the other regions (PuP/PuA, PuP/GP, PuP/CN, PuP/Th, PuP/RN) to test whether they might be more sensitive in detecting between-group differences compared with the absolute values in the posterior putamen. These ratios were compared between groups using a different LMM.

All LMMs were fitted using restricted maximum-likelihood estimation (REML) from the function lmer in the lme4 package (v1.1-21). Significance for the main effects and the interactions was assessed based on Type II Wald chi-square tests using the function Anova in the car package (v3.0-7). Post hoc pairwise comparisons were performed on a significant interaction or main factor effect with the emmeans package (v1.4.5) to further determine where the differences occurred across the study groups and the ROIs. All p-values from the post hoc tests were obtained using Kenward-Roger's approximation for degrees of freedom (df), and after adjustment for multiple testing by Tukey's method. For each fitted model, the assumptions of normality and constant variance of residuals were checked afterwards. The level of statistical significance was defined as a two-sided p-value or adjusted p-value <0.05 for all tests.

2.5. Classification performances

Using either binary or multinomial logistic regression, we studied the classification performance of each individual biomarker and combinations of selected biomarkers for separating two (MSAp vs PD, MSAp vs PSP, PD vs PSP, MSAc vs PD, MSAc vs PSP) or three classes (MSAp vs PD vs PSP and MSA vs PD vs PSP). Receiving operating curves (ROC) were generated and area under the curve (AUC), sensitivity, specificity and balanced accuracy (BA) were used as performance metrics. For the three-class classification, we used a multiclass definition of AUC called "AU1U″ as defined by Ferri et al. [20]. Differences between AUCs (two classes) and BAs (three classes) were tested following the method described in Robin et al. [21] with 100 times repeated 5-fold cross validation to compute the mean differences between all pairs of AUCs or BAs, and bootstrap standard-error estimates of these differences calculated from 200 bootstrap samples. For each imaging marker, an optimal cutoff value that best discriminated the different groups was determined from the point on the ROC curve that is closest to the top-left corner. Performance evaluation was carried out with the pROC (v1.16.2) and caret (v6.0-86) packages in binary classification and with the nnet (v7.3-13) and mlr (v2.19.0) packages in three-class classification.

A further analysis was performed at the subject level. Diagnostic predictions defined as the most frequent predicted group over the 100 cross-validation iterations using the three-class classification were compared to the clinical diagnosis, and, for patients with MSA, to the new diagnosis according to the recently published criteria [1].

2.5.1. Multivariate analysis

Visualization of subject group separation based on the most discriminative imaging markers was performed using a sparse partial least squares discriminant analysis (sPLS-DA) with the mixOmics package (v6.10.9) [22]. sPLS-DA is a supervised machine learning approach allowing for dimension reduction, feature selection and multiclass classification. The model based on sPLS-DA then consists of a small number of orthogonal components, where each component is calculated as weighted sums of the manifest variables under a covariance maximization criterion with the group labels. The "sparse" approach used by the method allows setting the least significant weights of each component to zero, thus selecting only the most relevant biomarkers. The weight values (or loadings) on the resulting components are also indicative of the importance of each variable for group discrimination on the different dimensions. Optimal number of components and imaging variables to keep per component in the final model was determined using the "tune.sPLS-DA" function to minimize the balanced error rate for group classification with 20 × 5-fold cross-validation.

3. Results

3.1. Participants demographic and clinical characteristics

In total, 93 participants were analyzed, including 22 HC, 25 patients with PD, 19 with PSP (16 probable PSP with Richardson syndrome (PSP-RS), 3 possible with pure gait freezing (PSP-PGF), 21 with MSAp (16 probable, 5 possible), and 6 with MSAc (3 probable, 3 possible). There was a difference in age (Kruskal-Wallis test, $p < 0.001$), PSP patients being older than other groups (all $p < 0.024$). There was no significant difference in terms of gender, UPDRS III scores or disease duration (Table 1).

3.2. Measurements reproducibility

Intra-rater agreement was moderate for MD (ICC 0.77; 95% CI: 0.74–0.79), good for R2* measurements (0.87; 0.86–0.89), excellent for midbrain-to-pons ratio (0.97; 0.91–0.99) and excellent for MRPI (0.99; 0.98–1). Inter-rater agreement was moderate for MD (ICC 0.77; 0.74–0.81), good for R2* measurements (0.82; 0.78–0.85), excellent for midbrain-to-pons ratio (0.94; 0.80–0.98) and excellent for MRPI (0.96; 0.86–0.99).

3.3. Between-group comparison of biomarkers

For both MD and R2* biomarkers, there was a significant effect of the Group ($p < 0.001$) and Region ($p < 0.0001$) factors, with a Group by Region interaction ($p < 0.0001$), without effect of the Side factor. Thus, means of the right and left values were used in the following analyses.

MD values were significantly higher in the **posterior putamen** in MSAp patients versus all groups ($p < 0.0001$) and in PSP vs HC ($p < 0.05$), in the **globus pallidus** in PSP patients versus all other groups ($p < 0.0001$) and in PD vs HC ($p < 0.05$), in the **red nucleus** in PSP patients versus HC ($p < 0.0001$) and PD ($p < 0.01$) (Fig. 1). All **ratios** were higher in MSAp patients vs all other groups ($p < 0.001$).

R2* values were significantly higher in the **posterior putamen** in MSAp patients versus all groups ($p < 0.001$), in the **anterior putamen** in MSAp versus PD ($p < 0.05$) and in the **red nucleus** in PSP versus HC ($p < 0.001$) and PD ($p < 0.05$) (Fig. 1). All **ratios** were higher in MSAp patients vs all other groups ($p < 0.001$).

PSP patients had significantly higher **MRPI** ($p < 0.0001$) and lower **midbrain to pons area ratio** values ($p < 0.01$ vs PD, $p < 0.0001$ vs other groups).

MSAc patients had higher **midbrain-to-pons ratio** values than MSAp ($p < 0.05$) and all other groups ($p < 0.0001$). MSAp patients also had higher **midbrain to pons ratio** than PD ($p < 0.05$) (Fig. 1, Table S2).

For all models, a visual inspection of the residual distributions did not show any important deviation from the normality and variance-variance assumptions.

3.4. Classification performances

Results are provided in Supplementary Table S3. For all classifications, there was no difference in AUC values for both MD and R2* measurements between values measured in the posterior putamen and the different ratios (PuP/PuA, PuP/GP, PuP/CN, PuP/Th, PuP/RN). Similarly, there was no difference between the MRPI and the midbrain to pons area ratio.

For "HC vs PD", diagnostic performances were low (AUC: 0.514–0.658).

For "HC vs PSP", MRPI (AUC: 0.933) and midbrain to pons ratio (AUC: 0.912) had similar performances than MD (AUC: 0.804) and R2*

Table 1
Clinical characteristics.

	HC	PD	PSP	MSAp	MSAc	P
Participants, n	22	25	19	21	6	–
Age at MRI, years (mean±SD, range)	64.7 ± 7.3 (47.3–76.1)	66.6 ± 10.1 (57–87)	73.6 ± 6.1 (63–86.7)	65.3 ± 8.5 (47.4–82.1)	60.2 ± 7.3 (51–71.3)	0.0009* PSP > HC, PD, MSAp, MSAc
Gender (F/M, %F)	10/12 (45.5%)	8/17 (32%)	7/12 (36.8%)	11/10 (52.4%)	1/5 (16.7%)	0.44
UPDRS III (mean ± SD, range) (number of participants with available data)	0.1 ± 0.3 (0–1) (22)	18.9 ± 9.3 (1–34) (15)	33.6 ± 17.4 (12–71) (8)	26.0 ± 13.4 (11–38) (5)	17.8 ± 6.0 (9–22) (4)	0.07
Disease duration, years (mean±SD, range)	–	4.84 ± 3.0 (1–11)	4.7 ± 3.2 (1–10)	4.1 ± 2.1 (1–10)	2.5 ± 1.4 (1–5)	0.28

There was a difference in age (Kruskal-Wallis test, $p = 0.001$), PSP patients being older than other groups ($p < 0.024$). There was no significant difference in gender, UPDRS III scores or disease duration. Disease duration was calculated using the date of first symptoms as the starting point. Of note, UPDRS III scores and disease duration were compared between the patient groups only (HC excluded). p values corresponding to the Kruskal-Wallis tests are reported.
*$p \leq 0.05$.
Abbreviations: F, female; M, male; HC, healthy control; MSAc, cerebellar form of multiple system atrophy; MSAp, parkinsonian form of multiple system atrophy; PD, Parkinson's disease; PSP, progressive supranuclear palsy; SD, standard deviation; UPDRS III, Unified Parkinson's Disease Rating Scale Part III.

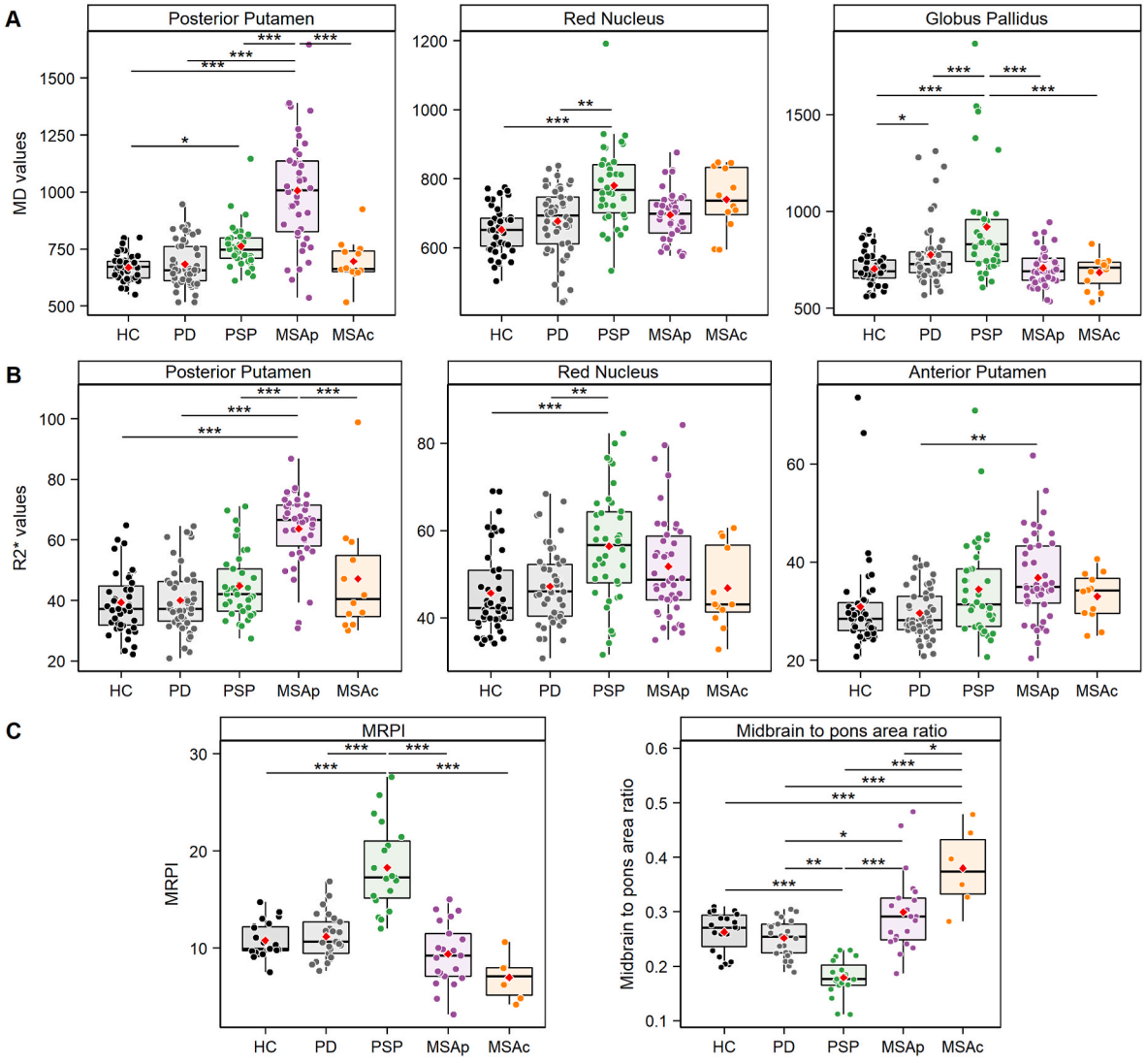

Fig. 1. Between-group comparison of MD (A), R2* relaxation rates (B) and morphometric biomarkers (C).
Only regions with significant differences were represented. The red diamond marker represents the mean value for each group.
***$p < 0.001$; **$0.001 < p \leq 0.01$; *$0.01 < p \leq 0.05$.
Abbreviations: MD, mean diffusivity; HC, healthy control; MSAc, cerebellar form of multiple system atrophy; MSAp, parkinsonian form of multiple system atrophy; PD, Parkinson's disease; PSP, progressive supranuclear palsy.

(AUC: 0.911) in the posterior putamen, MD (AUC: 0.905) in the red nucleus and MD (AUC: 0.871) in the pallidum ($p < 0.05$).

For "MSAp vs PD", both MD (AUC: 0.886) and R2* (AUC: 0.894) in the posterior putamen performed better than morphometry (MRPI, AUC: 0.540; midbrain to pons ratio, AUC: 0.623) (all $p < 0.01$). There was no difference between MD and R2* measurements. Similarly, for "MSAp vs HC", MD (AUC: 0.902) and R2* (AUC: 0.887) in the posterior putamen performed better.

For all other binary classifications, morphometric markers showed the best performance: "MSAp vs PSP", MRPI, AUC: 0.951, ratio, AUC: 0.940; "MSAc vs PD", MRPI, AUC: 0.823, ratio, AUC: 0.750; "PSP vs PD", MRPI, AUC: 0.889, ratio, AUC: 0.893; "MSAc vs PSP", MRPI, AUC: 0.988, ratio, AUC: 1; "MSAc vs HC", MRPI, AUC: 0.819 ratio, AUC: 0.813.

For all binary classifications, combining two or more markers did not significantly improve diagnostic performance.

For the multi-class classification "MSAp vs PD vs PSP", the combination of MD and R2* in the posterior putamen with morphometric biomarkers (AUC: 0.853, BA: 0.806) improved the overall accuracy compared to each marker separately (midbrain to pons ratio, AUC: 0.714, BA: 0.631, $p < 0.01$; MRPI, AUC: 0.693, BA: 0.580, $p < 0.01$; R2* in the posterior putamen, AUC: 0.778, BA: 0.576, $p < 0.001$; MD in the posterior putamen, AUC: 0.804, BA: 0.683, $p = 0.10$). Combining three or more biomarkers gave higher accuracy values than the two-biomarker classifiers (MRPI and MD in the posterior putamen, AUC: 0.752, BA: 0.750; midbrain to pons ratio and R2* in the posterior putamen, AUC: 0.759, BA: 0.742), although the difference did not reach significance. Results were similar for the "MSA vs PD vs PSP", although slightly lower.

Diagnostic cutoff values for the binary classifications are provided in Supplementary Table S4.

3.5. Multivariate analysis

The optimization step of the sPLS-DA parameters led to a model with two components. Component 1 separated patients with PSP from all other groups and those with MSAc from HC, PD and PSP, whereas Component 2 separated patients with MSAp from those with HC, PD and MSAc. PD subjects were not distinguishable from HCs (Fig. 2A, Supplementary Table S5).

Component 1 was correlated with four variables (in descending order: MRPI, midbrain to pons area ratio, MD in the globus pallidus, age) and Component 2 was correlated seven variables (in descending order: R2* and MD in the posterior putamen, R2* and MD in the anterior putamen, R2* and MD in the red nucleus and R2* in the thalamus) (Fig. 2B), in line with the above results using the logistic regression model. Loading values for each variable are given in Supplementary Table S6 and Supplementary Fig. S3.

3.5.1. Discordance between clinical status and diagnostic predictions at the individual-level

Concordance between imaging-based diagnostic prediction and clinical diagnosis was improved when the new international criteria for MSA [1] were used as compared with the previous criteria [19]. Using the previous criteria [19], clinical diagnoses and imaging-based predictions were discordant in 11 out of 71 patients (15.5%), including five patients with PD (5/25, 20%), three with PSP (3/19, 15.8%) and three with MSA (3/27, 11.1%). Compared with the new criteria for MSA [1], clinical diagnoses and imaging-based predictions were discordant in eight patients only (8/71, 11.3%), including the same five patients with PD (5/25, 20%), two with PSP (2/19, 10.5%) and one MSA (1/27, 3.7%) (Fig. 2A, Table 2).

The most frequent discordance was observed between PD and PSP with two PD patients according to the clinical criteria classified as PSP using imaging (subjects 1 and 2 in Table 2 and Fig. 2A) and two PSP patients classified as PD (subjects 6 and 8). In all four patients, the midbrain to pons ratio and the MRPI were close to the pathological threshold. There were seven discrepancies involving MSA with three patients with clinical PD (subjects 3, 4 and 5) and one with PSP-PGF (subject 7), classified as MSA using imaging and three patients with clinical MSA classified as PSP (subject 9) or PD (subjects 10 and 11) using imaging. Using the new diagnostic criteria [1], three of these seven patients were clinically reclassified as predicted using imaging. One patient diagnosed MSA according to both previous and new criteria was classified as PD using imaging and had cerebellar atrophy, but no cross-bun sign, which might be compatible with early MSAc at visual reading of the images (subject 11).

4. Discussion

Here, we compared the diagnostic performances of R2*, MD and morphometry and their combination using an approach that could be

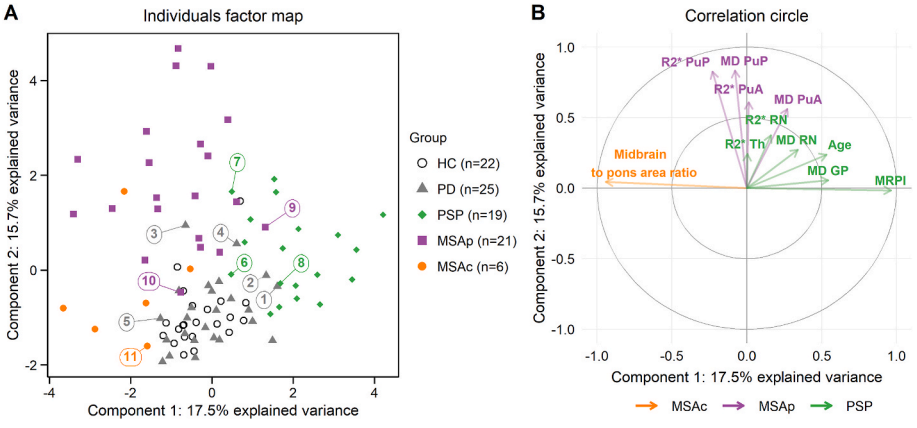

Fig. 2. Multivariate analysis using sparse partial least square discriminant analysis.
The individuals factor map (A) shows the separation between the different subjects. The two components explained 17.5% and 15.7% of the variance. Component 1 separated patients with PSP from all other groups and those with MSAc from HC, PD and PSP. Component 2 separated patients with MSAp from those with HC, PD and MSAc. PD subjects were not distinguishable from HCs. Subjects with discordance between imaging-based prediction and clinical diagnosis were numbered 1 to 11. For more details, see Table 2.
The correlation circle (B) shows that Component 1 was correlated with four variables (in descending order: MRPI, midbrain to pons area ratio, MD in the globus pallidus, age) while Component 2 was correlated with seven variables (in descending order: R2* and MD in the posterior putamen, R2* and MD in the anterior putamen, R2* and MD in the red nucleus and R2* in the thalamus). Colors indicated which disease group tends to have a higher probability of prediction with increasing values of the selected biomarker. Loading values for each variable are given in Supplementary Table S6 and Supplementary Fig. S3.
Abbreviations: GP, globus pallidus; MD, mean diffusivity; PuA, anterior putamen; PuP, posterior putamen; RN, red nucleus; HC, healthy control; MSAc, cerebellar form of multiple system atrophy; MSAp, parkinsonian form of multiple system atrophy; PD, Parkinson's disease; PSP, progressive supranuclear palsy; sPLS-DA, sparse partial least square discriminant analysis. . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Discordance between imaging-based prediction and clinical diagnosis at the individual-level. Subjects with discordance between clinical diagnosis and imaging-based prediction were numbered from 1 to 11 (see also Fig. 2A).

Subject number	Clinical diagnosis[a]	Clinical diagnosis for MSA according to the new criteria[b]	Predicted class using imaging	MRI pattern by visual reading	Age (years)	Sex	Disease duration (years)	Midbrain to pons ratio	MRPI	MD [f] (mm2/s)	R2* [g](s-1)
1	PD CE	NA	PSP	PSP	87	Male	8	0.20	14.5	826	35
2	PD CE	NA	PSP	PSP	71	Male	1	0.22	15.4	722	41
3	PD CE	no MSA	MSA	PD	75	Male	5	0.28	8.4	893	46
4	PD CE	no MSA	MSA	PD	67	Male	7	0.22	13.1	839	42
5	PD CE	no MSA	MSA	PD	50	Male	3	0.30	8.3	594	48
6	PSP-RS probable	NA	PD	PSP	72	Male	1	0.23	12.0	761	41
7	PSP-PGF possible	MSAc	MSA	PD	67	Male	4	0.23	13.0	924	64
8	PSP-RS probable	NA	PD	PSP	83	Male	10	0.22	14.9	652	48
9	MSAp probable	no MSA[c]	PSP	PSP	70	Male	1	0.19	15.0	714	56
10	MSAp probable	no MSA[d]	PD	PD	63	Male	4	0.29	9.0	796	35
11	MSAc probable	MSA	PD	PD[e]	51	Male	2	0.33	8.0	658	31

Abbreviations: MD, mean diffusivity; CE, clinically established; HC, healthy control; MSA, multiple system atrophy; MSAp, parkinsonian form of multiple system atrophy; MSAc, cerebellar form of multiple system atrophy; PD, Parkinson's disease; PSP, progressive supranuclear palsy; PSP-PGF, PSP with pure gait freezing; PSP-RS, PSP with Richardson syndrome.

[a] Diagnosis of MSA according to the previous consensus criteria [19].
[b] Diagnosis of MSA according to the new consensus criteria [1].
[c] MRI sign of PSP (exclusion criterion for MSA according to the new diagnostic criteria [1]).
[d] Presence of anosmia (exclusion criterion for MSA according to the new diagnostic criteria [1]).
[e] Presence of cerebellar atrophy without cross-bun sign.
[f] MD values in the posterior putamen.
[g] R2* values in the posterior putamen.

part of a routine clinical assessment. Such manual measurements have the advantage of being reproducible and easy to perform, which facilitates their implementation in clinical routine unlike techniques used in research that are time-consuming and require high expertise [9–17]. The main results of this study may be summarized as follows. Firstly, manual measurements of MD and R2* in the posterior putamen performed better than brainstem morphometric measurements to distinguish MSAp from PD, but not from PSP. Secondly, brainstem morphometric biomarkers were the most efficient to distinguish PSP and MSAc from all other groups. Thirdly, combining the three biomarkers improved the classification performances for the three-group comparison. Finally, we showed an improvement in diagnostic categorization when the new diagnostic criteria for MSA [1] were used.

Our results confirmed that morphometric markers reflecting midbrain atrophy were the best markers for the differentiation of PSP from the other groups [13,14]. In our work, the MRPI did not perform better than the midbrain to pons area ratio in line with a previous study [13].This is likely due to the small size of the superior and middle cerebellar peduncles, limiting inter- and intra-rater reliability [13]. Another study showed that the MRPI was more accurate than the midbrain to pons ratio in differentiating patients with possible PSP from those with PD, but not with probable PSP [23]. Morphometric measurements, which also reflect atrophy in the pons and middle cerebellar peduncles, discriminated well MSAc from PD and PSP as shown previously [13]. As expected in MSAc, MD and R2* measurements in the putamen did not improve the classification accuracy which is consistent with the mostly normal visual aspect of the putamen in patients with a pure cerebellar phenotype.

Conversely, morphometric biomarkers performed poorly in differentiating MSAp and PD. Unlike MSAc, most patients with MSAp have little pontocerebellar atrophy whereas they show marked putaminal alterations that predominate in the posterior part [5–7]. In our study, MD and R2* measurements in the posterior putamen yielded the best discrimination between MSAp and PD. These measurements also showed good performance for the separation of MSAp from PSP subjects although slightly lower than the MRPI. Several studies have shown that MSAp can be distinguished from PD based on putaminal diffusivity [24–27], some studies targeting the posterior putamen [25], with a meta-analysis showing an overall sensitivity of 90% and an overall specificity of 93% [28]. For MSAp versus PSP, results were less clear, a few studies showing a good separation [10], especially when focusing on the posterior putamen [8], while others did not [9,11]. Using relaxometry, MSAp could be distinguished from PD using R2* [11,26] and quantitative susceptibility mapping (QSM) [12]. Again, results were discordant between MSAp and PSP, some studies showing between-group differences [11] while others did not [12].

Our study showed that the combination of MD, R2* and morphometric markers improved the performances of the three-class classification. Automated extraction of MD, fractional anisotropy (FA) and R2* measured in striatal, midbrain, limbic and cerebellar regions allowed good differentiation of parkinsonian syndromes in a previous study [17]. Their combination improved the classification of PD and MSAp, but not PD and PSP or MSAp and PSP [17]. The combination of R2* and MD allowed 95% discrimination between MSAp and PSP [26]. In another study, 95% accuracy was reached for MSAp vs PD using a combination of grey matter, MD and FA changes in several brain regions [16]. Patients with MSAc were distinguished from those with PD using grey matter, MD and R2* in the cerebellum, and R2* in the left cerebral peduncle [16]. Diagnostic accuracies obtained in a previous study using diffusion measurements in the entire putamen were low for the categorization of MSAp versus PD and PSP patients and did not significantly improve the performances when combined with volumetry [29].

Relatives values of MD and R2* measurements in the posterior putamen to each of the other regions did not improve the categorization performance. These ratios were tested to determine if they could be more sensitive in detecting between-group differences compared to the absolute values in the posterior putamen. Indeed, values being higher in the posterior putamen in subjects in MSAp and higher in the red nucleus in those with PSP, we hypothesized that the posterior putamen to red nucleus ratio could perform better that the posterior putamen alone,

which was not the case in our study. Nevertheless, these ratios could be interesting to normalize the data to limit the scanner-effect in multicenter studies.

As expected, the new international diagnostic criteria for MSA improved the agreement between the clinical diagnosis and the imaging-based prediction by reclassifying three out of seven subjects with disagreement. Two participants with a diagnosis of MSA according to the previous criteria were excluded using the new criteria, in favor of an alternative diagnosis, and one participant with a clinical diagnosis of PSP-PGF was classified as MSA by the new criteria, which was in agreement with the imaging-based prediction for all three subjects. On the other hand, our model misclassified one participant with clinical MSAc as PD (despite the presence of pons atrophy) and three with PD as MSA (because of pons atrophy in two cases). Therefore, although morphometry was effective in differentiating MSAc from PD, these results suggest that other biomarkers such as diffusivity in the cerebellum, middle cerebellar peduncles and pons may help improve this differentiation. Regarding the classification PD vs PSP, subjects misclassified had borderline morphometric values with midbrain to pons area ratios and MRPI values close to the pathological thresholds. For these discordant subjects, only a longitudinal follow-up would confirm the diagnosis in the absence of brain pathological examination. Lastly, although age was included as a covariate of no interest in our logistic regression model, it could still be a confounding factor explaining, for instance, disagreement between the clinical diagnosis of PD and the imaging prediction of PSP in an elderly patient. Indeed, midbrain atrophy occurs during normal aging and threshold values may not apply at this age [30].

Our study had some limitations. Analyses were restricted to six subcortical regions that are small and where measurements are easy to perform in clinical practice. Other regions could have been interesting such as the dentate nucleus and the superior cerebellar peduncles for PSP and the white matter of the cerebellum and middle cerebellar peduncles for MSAc [5]. Quantitative biomarkers such as fractional anisotropy, free water, R2* relaxation rates and magnetic susceptibility measured in the substantia nigra (SN) have been shown to differentiate parkinsonian patients from healthy subjects. Studies reported differences between PD and atypical parkinsonism and between MSA and PSP based on free-water measurements [31,32] and QSM [33], between PSP and PD using fractional anisotropy [34], free-water measurements [32] and QSM [33]. However, we decided not to include SN measurements for the following reasons. Results for the differentiation between parkinsonian syndromes were not always consistent across studies. Some studies showed no differences using R2* relaxation rates [16] and QSM [12]. In the study by Sjöström et al., 2019 [33], magnetic susceptibility in the SN had lower performance than the red nucleus for the separation of PSP vs PD and of PSP vs MSA and than the putamen for the separation of MSA vs PD. For diffusion measurements, free-water and FA values measured in the SN had lower accuracies for the separation of parkinsonian patients compared to other regions [32]. Moreover, FA showed lower performances than free water in the SN [35]. Unlike FA, MD and R2* biomarkers, free-water and QSM are not yet useable in a clinical setting. On the other hand, we decided to use a standardized round shape ROI with an area of 8 mm^2 suitable for small and round nuclei to make measurements more reproducible and easy to perform in a clinical setting. Such a region is more difficult to place reproducibly on the SN, given its ovoid shape, than on the red nucleus. In the future, methods using fully automated extraction of biomarkers might become accessible for a clinical use. Cutoffs values provided are usually scanner-specific. Multicenter studies using a standardized MRI protocol are needed to derive normalized cutoff values that would be applicable on a large scale. Finally, there was no neuropathological confirmation of the diagnosis.

In conclusion, this study provides accurate quantitative biomarkers usable in clinical routine for the differentiation of parkinsonian syndromes. R2* and MD measurements in the posterior putamen are robust biomarkers for the discrimination of MSAp, and their combination with brainstem morphometry increases the performance of multiclass classification. Our study also confirms that the joint use of clinical criteria and imaging data improves diagnostic accuracy.

Funding

This work was supported by grants from Agence Nationale de la Recherche (grant numbers ANR-11-INBS-0006 [France Life Imaging], ANRMNP 2009 [Nucleipark], ANR-11-INBS-0011 [NeurATRIS, Investissements d'Avenir], ANR-19-P3IA-0001 [PRAIRIE 3IA Institute] and ANR-10-IAIHU-06 [IHU–Paris Institute of Neurosciences]), Association France Parkinson, Ecole Neuroscience de Paris, Electricité de France (Fondation d'Entreprise EDF), Institut National de la Santé et de la Recherche Médicale, DHOS-Inserm (2010, Nucleipark), PSP France, and the Fondation Thérèse and René Planiol pour l'étude du Cerveau. Lydia Chougar is supported by a poste d'accueil Inria/AP-HP.

Data availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Contributors

Lydia Chougar designed and conceptualized the study, analyzed the data and drafted the manuscript.

François-Xavier Lejeune designed and conceptualized the study, analyzed the data, and drafted the manuscript.

Johann Faouzi analyzed the data and revised the manuscript for intellectual content.

Benjamin Morino collected and analyzed the data.

Alice Faucher collected the data.

Nadine Hoyek analyzed the data.

David Grabli collected the data.

Florence Cormier collected the data.

Marie Vidailhet collected the data.

Jean-Christophe Corvol collected the data and revised the manuscript.

Olivier Colliot revised the manuscript.

Bertrand Degos collected the data and revised the manuscript.

Stéphane Lehéricy designed and conceptualized the study, analyzed the data, and revised the manuscript.

Ethics approval.

The study was approved by La Pitié Salpêtrière University Hospital ethical standards committee.

Declaration of competing interest

No conflicts of interest related to the present study.

Competing financial interests unrelated to the present work:

Jean-Christophe Corvol has served in scientific advisory boards for Biogen, Denali, Ever Pharma, Isdorsia, Prevail Therapeutics, and UCB; and has received grants from Sanofi, the Michael J Fox Foundation, ANR, France Parkinson, the French Ministry of Health.

Olivier Colliot received consulting fees from AskBio (2020), received fees for writing a lay audience short paper from Expression Santé (2019), received speaker fees for a lay audience presentation from Palais de la découverte (2017). His laboratory received grants (paid to the institution) from Air Liquide Medical Systems (2011–2016) and Qynapse (2017–present). Members from his laboratory have cosupervised a PhD thesis with myBrainTechnologies (2016–present). Colliot's spouse is an employee of myBrainTechnologies (2015–present). Colliot has submitted a patent to the International Bureau of the World Intellectual Property Organization (PCT/IB2016/052699 3, J.-B.

Schiratti, S. Allassonniere, O. Colliot, S. Durrleman, "A method for determining the temporal progression of a biological phenomenon and associated methods and devices," 2016).

Bertrand Degos received research support grants from Fondation de France, Inserm, and ANR; speech honoraria from Ipsen, Merz Pharma, and Orkyn; and received travel funding from Merz Pharma, Elivie, and Orkyn.

Stéphane Lehéricy received grants from 'Investissements d'avenir' (gran t numbers ANR-10-IAIHU -06 and AN R-11-INBS-0006) and Biogen Inc.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.parkreldis.2023.105287.

References

[1] G.K. Wenning, I. Stankovic, L. Vignatelli, A. Fanciulli, G. Calandra-Buonaura, K. Seppi, J. Palma, W.G. Meissner, F. Krismer, D. Berg, P. Cortelli, R. Freeman, G. Halliday, G. Höglinger, A. Lang, H. Ling, I. Litvan, P. Low, Y. Miki, J. Panicker, M.T. Pellecchia, N. Quinn, R. Sakakibara, M. Stamelou, E. Tolosa, S. Tsuji, T. Warner, W. Poewe, H. Kaufmann, The movement disorder society criteria for the diagnosis of multiple system atrophy, Mov. Disord. (2022), 29005, https://doi.org/10.1002/mds.29005 mds.

[2] S. Koga, N. Aoki, R.J. Uitti, J.A. van Gerpen, W.P. Cheshire, K.A. Josephs, Z. K. Wszolek, J.W. Langston, D.W. Dickson, When DLB, PD, and PSP masquerade as MSA: an autopsy study of 134 patients, Neurology 85 (2015) 404–412, https://doi.org/10.1212/WNL.0000000000001807.

[3] G. Respondek, S. Roeber, H. Kretzschmar, C. Troakes, S. Al-Sarraj, E. Gelpi, C. Gaig, W.Z. Chiu, J.C. van Swieten, W.H. Oertel, G.U. Höglinger, Accuracy of the national institute for neurological disorders and stroke/society for progressive supranuclear palsy and neuroprotection and natural history in Parkinson plus syndromes criteria for the diagnosis of progressive supranuclear palsy: PSP Diagnostic Criteria, Mov, Disord 28 (2013) 504–509, https://doi.org/10.1002/mds.25327.

[4] G.U. Höglinger, G. Respondek, M. Stamelou, C. Kurz, K.A. Josephs, A.E. Lang, B. Mollenhauer, U. Müller, C. Nilsson, J.L. Whitwell, T. Arzberger, E. Englund, E. Gelpi, A. Giese, D.J. Irwin, W.G. Meissner, A. Pantelyat, A. Rajput, J.C. van Swieten, C. Troakes, A. Antonini, K.P. Bhatia, Y. Bordelon, Y. Compta, J.-C. Corvol, C. Colosimo, D.W. Dickson, R. Dodel, L. Ferguson, M. Grossman, J. Kassubek, F. Krismer, J. Levin, S. Lorenzl, H.R. Morris, P. Nestor, W.H. Oertel, W. Poewe, G. Rabinovici, J.B. Rowe, G.D. Schellenberg, K. Seppi, T. van Eimeren, G. K. Wenning, A.L. Boxer, L.I. Golbe, I. Litvan, Movement Disorder Society-endorsed PSP Study Group, Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria, Mov. Disord. Off. J. Mov. Disord. Soc. 32 (2017) 853–864, https://doi.org/10.1002/mds.26987.

[5] L. Chougar, N. Pyatigorskaya, S. Lehéricy, Update on neuroimaging for categorization of Parkinson's disease and atypical parkinsonism, Curr. Opin. Neurol. (2021), https://doi.org/10.1097/WCO.0000000000000957. Publish Ahead of Print.

[6] K. Seppi, M.F.H. Schocke, K. Prennschuetz-Schuetzenau, K.J. Mair, R. Esterhammer, C. Kremser, A. Muigg, C. Scherfler, W. Jaschke, G.K. Wenning, W. Poewe, Topography of putaminal degeneration in multiple system atrophy: a diffusion magnetic resonance study, Mov. Disord. Off. J. Mov. Disord. Soc. 21 (2006) 847–852, https://doi.org/10.1002/mds.20843.

[7] M.T. Pellecchia, P. Barone, C. Mollica, E. Salvatore, M. Ianniciello, K. Longo, A. Varrone, C. Vicidomini, M. Picillo, G. De Michele, A. Filla, M. Salvatore, S. Pappatà, Diffusion-weighted imaging in multiple system atrophy: a comparison between clinical subtypes, Mov. Disord. Off. J. Mov. Disord. Soc. 24 (2009) 689–696, https://doi.org/10.1002/mds.22440.

[8] K. Tsukamoto, E. Matsusue, Y. Kanasaki, S. Kakite, S. Fujii, T. Kaminou, T. Ogawa, Significance of apparent diffusion coefficient measurement for the differential diagnosis of multiple system atrophy, progressive supranuclear palsy, and Parkinson's disease: evaluation by 3.0-T MR imaging, Neuroradiology 54 (2012) 947–955, https://doi.org/10.1007/s00234-012-1009-9.

[9] K. Seppi, M.F.H. Schocke, R. Esterhammer, C. Kremser, C. Brenneis, J. Mueller, S. Boesch, W. Jaschke, W. Poewe, G.K. Wenning, Diffusion-weighted imaging discriminates progressive supranuclear palsy from PD, but not from the Parkinson variant of multiple system atrophy, Neurology 60 (2003) 922–927, https://doi.org/10.1212/01.wnl.0000049911.91657.9d.

[10] G. Nicoletti, R. Lodi, F. Condino, C. Tonon, F. Fera, E. Malucelli, D. Manners, M. Zappia, L. Morgante, P. Barone, B. Barbiroli, A. Quattrone, Apparent diffusion coefficient measurements of the middle cerebellar peduncle differentiate the Parkinson variant of MSA from Parkinson's disease and progressive supranuclear palsy, Brain J. Neurol. 129 (2006) 2679–2687, https://doi.org/10.1093/brain/awl166.

[11] N.K. Focke, G. Helms, P.M. Pantel, S. Scheewe, M. Knauth, C.G. Bachmann, J. Ebentheuer, P. Dechent, W. Paulus, C. Trenkwalder, Differentiation of typical and atypical Parkinson syndromes by quantitative MR imaging, Am. J. Neuroradiol. 32 (2011) 2087–2092.

[12] H. Sjöström, T. Granberg, E. Westman, P. Svenningsson, Quantitative susceptibility mapping differentiates between parkinsonian disorders, Park. Relat. Disord. 44 (2017) 51–57, https://doi.org/10.1016/j.parkreldis.2017.08.029.

[13] L. Möller, J. Kassubek, M. Südmeyer, R. Hilker, E. Hattingen, K. Egger, F. Amtage, E.H. Pinkhardt, G. Respondek, M. Stamelou, F. Möller, A. Schnitzler, W.H. Oertel, S. Knake, H.-J. Huppertz, G.U. Höglinger, Manual MRI morphometry in Parkinsonian syndromes, Mov. Disord. Off. J. Mov. Disord. Soc. 32 (2017) 778–782, https://doi.org/10.1002/mds.26921.

[14] A. Quattrone, G. Nicoletti, D. Messina, F. Fera, F. Condino, P. Pugliese, P. Lanza, P. Barone, L. Morgante, M. Zappia, U. Aguglia, O. Gallo, MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy, Radiology 246 (2008) 214–221, https://doi.org/10.1148/radiol.2453061703.

[15] A. Quattrone, M. Morelli, S. Nigro, A. Quattrone, B. Vescio, G. Arabia, G. Nicoletti, R. Nisticò, M. Salsone, F. Novellino, G. Barbagallo, E. Le Piane, P. Pugliese, D. Bosco, M.G. Vaccaro, C. Chiriaco, U. Sabatini, V. Vescio, C. Stanà, F. Rocca, D. Gullà, M. Caracciolo, A new MR imaging index for differentiation of progressive supranuclear palsy-parkinsonism from Parkinson's disease, Park. Relat. Disord. 54 (2018) 3–8, https://doi.org/10.1016/j.parkreldis.2018.07.016.

[16] P. Péran, G. Barbagallo, F. Nemmi, M. Sierra, M. Galitzky, A.P.-L. Traon, P. Payoux, W.G. Meissner, O. Rascol, MRI supervised and unsupervised classification of Parkinson's disease and multiple system atrophy, Mov. Disord. Off. J. Mov. Disord. Soc. 33 (2018) 600–608, https://doi.org/10.1002/mds.27307.

[17] G. Du, M.M. Lewis, S. Kanekar, N.W. Sterling, L. He, L. Kong, R. Li, X. Huang, Combined diffusion tensor imaging and apparent transverse relaxation rate differentiate Parkinson disease and atypical parkinsonism, AJNR Am. J. Neuroradiol. 38 (2017) 966–972, https://doi.org/10.3174/ajnr.A5136.

[18] R.B. Postuma, D. Berg, M. Stern, W. Poewe, C.W. Olanow, W. Oertel, J. Obeso, K. Marek, I. Litvan, A.E. Lang, G. Halliday, C.G. Goetz, T. Gasser, B. Dubois, P. Chan, B.R. Bloem, C.H. Adler, G. Deuschl, MDS clinical diagnostic criteria for Parkinson's disease, Mov. Disord. Off. J. Mov. Disord. Soc. 30 (2015) 1591–1601, https://doi.org/10.1002/mds.26424.

[19] S. Gilman, G.K. Wenning, P.A. Low, D.J. Brooks, C.J. Mathias, J.Q. Trojanowski, N. W. Wood, C. Colosimo, A. Dürr, C.J. Fowler, H. Kaufmann, T. Klockgether, A. Lees, W. Poewe, N. Quinn, T. Revesz, D. Robertson, P. Sandroni, K. Seppi, M. Vidailhet, Second consensus statement on the diagnosis of multiple system atrophy, Neurology 71 (2008) 670–676, https://doi.org/10.1212/01.wnl.0000324625.00404.15.

[20] C. Ferri, J. Hernández-Orallo, R. Modroiu, An experimental comparison of performance measures for classification, Pattern Recogn. Lett. 30 (2009) 27–38, https://doi.org/10.1016/j.patrec.2008.08.010.

[21] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinf. 12 (2011) 77, https://doi.org/10.1186/1471-2105-12-77.

[22] K.-A. Lê Cao, S. Boitard, P. Besse, Sparse PLS discriminant analysis: biologically relevant feature selection and graphical displays for multiclass problems, BMC Bioinf. 12 (2011) 253, https://doi.org/10.1186/1471-2105-12-253.

[23] S. Nigro, M. Morelli, G. Arabia, R. Nisticò, F. Novellino, M. Salsone, F. Rocca, A. Quattrone, Magnetic Resonance Parkinsonism Index and midbrain to pons ratio: which index better distinguishes Progressive Supranuclear Palsy patients with a low degree of diagnostic certainty from patients with Parkinson Disease? Park. Relat. Disord. 41 (2017) 31–36, https://doi.org/10.1016/j.parkreldis.2017.05.002.

[24] M.F.H. Schocke, K. Seppi, R. Esterhammer, C. Kremser, K.J. Mair, B.V. Czermak, W. Jaschke, W. Poewe, G.K. Wenning, Trace of diffusion tensor differentiates the Parkinson variant of multiple system atrophy and Parkinson's disease, Neuroimage 21 (2004) 1443–1451, https://doi.org/10.1016/j.neuroimage.2003.12.005.

[25] K. Seppi, M.F.H. Schocke, K.J. Mair, R. Esterhammer, C. Scherfler, F. Geser, C. Kremser, S. Boesch, W. Jaschke, W. Poewe, G.K. Wenning, Progression of putaminal degeneration in multiple system atrophy: a serial diffusion MR study, Neuroimage 31 (2006) 240–245, https://doi.org/10.1016/j.neuroimage.2005.12.006.

[26] G. Barbagallo, M. Sierra-Peña, F. Nemmi, A.P.-L. Traon, W.G. Meissner, O. Rascol, P. Péran, Multimodal MRI assessment of nigro-striatal pathway in multiple system atrophy and Parkinson disease, Mov. Disord. Off. J. Mov. Disord. Soc. 31 (2016) 325–334, https://doi.org/10.1002/mds.26471.

[27] P. Péran, A. Cherubini, F. Assogna, F. Piras, C. Quattrocchi, A. Peppe, P. Celsis, O. Rascol, J.-F. Démonet, A. Stefani, M. Pierantozzi, F.E. Pontieri, C. Caltagirone, G. Spalletta, U. Sabatini, Magnetic resonance imaging markers of Parkinson's disease nigrostriatal signature, Brain 133 (2010) 3423–3433, https://doi.org/10.1093/brain/awq212.

[28] S. Bajaj, F. Krismer, J.-A. Palma, G.K. Wenning, H. Kaufmann, W. Poewe, K. Seppi, Diffusion-weighted MRI distinguishes Parkinson disease from the parkinsonian variant of multiple system atrophy: a systematic review and meta-analysis, PLoS One 12 (2017), e0189897, https://doi.org/10.1371/journal.pone.0189897.

[29] L. Chougar, J. Faouzi, N. Pyatigorskaya, L. Yahia-Cherif, R. Gaurav, E. Biondetti, M. Villotte, R. Valabrègue, J.-C. Corvol, A. Brice, L.-L. Mariani, F. Cormier, M. Vidailhet, G. Dupont, I. Piot, D. Grabli, C. Payan, O. Colliot, B. Degos, S. Lehéricy, Automated categorization of parkinsonian syndromes using magnetic resonance imaging in a clinical setting, Mov. Disord. (2020) 12.

[30] S.T. Ruiz, R.V. Bakklund, A.K. Håberg, E.M. Berntsen, Normative data for brainstem structures, the midbrain-to-pons ratio, and the magnetic resonance parkinsonism index, Am. J. Neuroradiol. 43 (2022) 707–714, https://doi.org/10.3174/ajnr.A7485.

[31] D.B. Archer, J.T. Bricker, W.T. Chu, R.G. Burciu, J.L. Mccracken, S. Lai, S. A. Coombes, R. Fang, A. Barmpoutis, D.M. Corcos, A.S. Kurani, T. Mitchell, M.

L. Black, E. Herschel, T. Simuni, T.B. Parrish, C. Comella, T. Xie, K. Seppi, N. I. Bohnen, M.L.T.M. Müller, R.L. Albin, F. Krismer, G. Du, M.M. Lewis, X. Huang, H. Li, O. Pasternak, N.R. McFarland, M.S. Okun, D.E. Vaillancourt, Development and validation of the automated imaging differentiation in parkinsonism (AID-P): a multi-site machine learning study, Lancet Digit. Health. 1 (2019), https://doi.org/10.1016/s2589-7500(19)30105-0 e222–e231.

[32] P.J. Planetta, E. Ofori, O. Pasternak, R.G. Burciu, P. Shukla, J.C. DeSimone, M. S. Okun, N.R. McFarland, D.E. Vaillancourt, Free-water imaging in Parkinson's disease and atypical parkinsonism, Brain J. Neurol. 139 (2016) 495–508, https://doi.org/10.1093/brain/awv361.

[33] H. Sjöström, Y. Surova, M. Nilsson, T. Granberg, E. Westman, D. van Westen, P. Svenningsson, O. Hansson, Mapping of apparent susceptibility yields promising diagnostic separation of progressive supranuclear palsy from other causes of parkinsonism, Sci. Rep. 9 (2019) 6079, https://doi.org/10.1038/s41598-019-42565-4.

[34] N. Pyatigorskaya, L. Yahia-Cherif, R. Gaurav, C. Ewenczyk, C. Gallea, R. Valabregue, F. Gargouri, B. Magnin, B. Degos, E. Roze, E. Bardinet, C. Poupon, I. Arnulf, M. Vidailhet, S. Lehericy, Multimodal magnetic resonance imaging quantification of brain changes in progressive supranuclear palsy, Mov. Disord. Off. J. Mov. Disord. Soc. 35 (2020) 161–170, https://doi.org/10.1002/mds.27877.

[35] E. Ofori, F. Krismer, R.G. Burciu, O. Pasternak, J.L. McCracken, M.M. Lewis, G. Du, N.R. McFarland, M.S. Okun, W. Poewe, C. Mueller, E.R. Gizewski, M. Schocke, C. Kremser, H. Li, X. Huang, K. Seppi, D.E. Vaillancourt, Free water improves detection of changes in the substantia nigra in parkinsonism: a multisite study, Mov. Disord. Off. J. Mov. Disord. Soc. 32 (2017) 1457–1464, https://doi.org/10.1002/mds.27100.