



HAL
open science

Contribution of MRI for the Early Diagnosis of Parkinsonism in Patients with Diagnostic Uncertainty

Lydia Chougar, Alice Faucher, Johann Faouzi, François-xavier Lejeune, Gonçalo Gama Lobo, Carna Jovanovic, Florence Cormier, Gwendoline Dupont, Marie Vidailhet, Jean-christophe Corvol, et al.

► **To cite this version:**

Lydia Chougar, Alice Faucher, Johann Faouzi, François-xavier Lejeune, Gonçalo Gama Lobo, et al.. Contribution of MRI for the Early Diagnosis of Parkinsonism in Patients with Diagnostic Uncertainty. *Movement Disorders*, 2024, 39 (5), pp.825-835. 10.1002/mds.29760 . hal-04675671

HAL Id: hal-04675671

<https://hal.science/hal-04675671>

Submitted on 22 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH ARTICLE

Contribution of MRI for the Early Diagnosis of Parkinsonism in Patients with Diagnostic Uncertainty

Lydia Chougar, MD, PhD,^{1,2,3,4*} Alice Faucher, MD,⁵ Johann Faouzi, PhD,^{6,7} François-Xavier Lejeune, PhD,^{8,9} Gonçalo Gama Lobo, MD,¹⁰ Carna Jovanovic, MD,¹¹ Florence Cormier, MD,¹² Gwendoline Dupont, MD,^{13,14} Marie Vidailhet, MD,^{3,8} Jean-Christophe Corvol, MD, PhD,^{8,12} Olivier Colliot, PhD,⁶ Stéphane Lehericy, MD, PhD,^{2,3,4} David Grabli, MD, PhD,^{8,12} and Bertrand Degos, MD, PhD^{5,15}

¹Department of Neuroradiology, Sorbonne Université, Institut du Cerveau–Paris Brain Institute–ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, DMU DIAMENT, Paris, France

²ICM, Centre de Neuroimagerie de Recherche–CENIR, Paris, France

³ICM, Team “Movement Investigations and Therapeutics” (MOVIT), Paris, France

⁴Department of Neuroradiology, Sorbonne Université, Institut du Cerveau–Paris Brain Institute–ICM, CNRS, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, DMU DIAMENT, Paris, France

⁵Assistance Publique Hôpitaux de Paris, Service de Neurologie, Hôpital Avicenne, Hôpitaux Universitaires de Paris Seine-Saint-Denis, Sorbonne Paris Nord, NS-PARK/FCRIN Network, Bobigny, France

⁶Sorbonne Université, Institut du Cerveau–Paris Brain Institute–ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, Paris, France

⁷CREST, ENSAI, Campus de Ker-Lann, Bruz, France

⁸Sorbonne Université, Institut du Cerveau–Paris Brain Institute–ICM, CNRS, Inserm, Paris, France

⁹ICM, Data Analysis Core (DAC), Paris, France

¹⁰Neuroradiology Department, Centro Hospitalar Universitário de Lisboa Central, Lisboa, Portugal

¹¹Neurology Clinic, University Clinical Center of Serbia, Belgrade, Serbia

¹²Département de Neurologie, Hôpital Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris, Clinique des Mouvements Anormaux, Clinical Investigation Center for Neurosciences, Paris, France

¹³Université de Bourgogne, Dijon, France

¹⁴Département de Neurologie, Centre Hospitalier Universitaire François Mitterrand, Dijon, France

¹⁵Dynamics and Pathophysiology of Neuronal Networks Team, Center for Interdisciplinary Research in Biology, Collège de France, CNRS UMR7241/INSERM U1050, Université PSL, Paris, France

ABSTRACT: Background: International clinical criteria are the reference for the diagnosis of degenerative parkinsonism in clinical research, but they may lack sensitivity and specificity in the early stages.

Objectives: To determine whether magnetic resonance imaging (MRI) analysis, through visual reading or machine-

learning approaches, improves diagnostic accuracy compared with clinical diagnosis at an early stage in patients referred for suspected degenerative parkinsonism.

Materials: Patients with initial diagnostic uncertainty between Parkinson’s disease (PD), progressive supranuclear palsy (PSP), and multisystem atrophy (MSA), with

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](#) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

***Correspondence to:** Dr. Lydia Chougar, 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: lydia.chougar@aphp.fr; chougar.lydia@gmail.com

Lydia Chougar and Alice Faucher are co-first authors.

David Grabli and Bertrand Degos are co-last authors.

Relevant conflicts of interest/financial disclosures: No relevant conflict of interest related to the present study.

Funding agencies: L.C. is supported by an Inria/AP-HP “Poste d’accueil”. 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 Neurosciences de Paris, Électricité 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.

Received: 25 October 2023; **Revised:** 16 January 2024; **Accepted:** 16 February 2024

[Correction added on 23 April 2024, after first online publication: The affiliations for the authors Jean-Christophe Corvol and David Grabli were changed from 8,11 to 8,12.]

Published online 14 March 2024 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.29760

brain MRI performed at the initial visit (V1) and available 2-year follow-up (V2), were included. We evaluated the accuracy of the diagnosis established based on: (1) the international clinical diagnostic criteria for PD, PSP, and MSA at V1 (“Clin1”); (2) MRI visual reading blinded to the clinical diagnosis (“MRI”); (3) both MRI visual reading and clinical criteria at V1 (“MRI and Clin1”), and (4) a machine-learning algorithm (“Algorithm”). The gold standard diagnosis was established by expert consensus after a 2-year follow-up.

Results: We recruited 113 patients (53 with PD, 31 with PSP, and 29 with MSA). Considering the whole population, compared with clinical criteria at the initial visit (“Clin1”: balanced accuracy, 66.2%), MRI visual reading

showed a diagnostic gain of 14.3% (“MRI”: 80.5%; $P = 0.01$), increasing to 19.2% when combined with the clinical diagnosis at the initial visit (“MRI and Clin1”: 85.4%; $P < 0.0001$). The algorithm achieved a diagnostic gain of 9.9% (“Algorithm”: 76.1%; $P = 0.08$).

Conclusion: Our study shows the use of MRI analysis, whether by visual reading or machine-learning methods, for early differentiation of parkinsonism. © 2024 The Authors. *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: diagnosis; multisystem atrophy; MRI; Parkinson’s disease; progressive supranuclear palsy

Differential diagnosis of degenerative parkinsonism using international clinical criteria remains challenging at the early stages, with modest diagnostic accuracy at the initial consultation for multisystem atrophy (MSA) (62%–79%)^{1,2} and progressive supranuclear palsy (PSP) (14%–83%).^{3,4} Despite criteria revisions, the time to diagnosis remains long,⁵ resulting in missed opportunities for potential candidates in neuroprotection trials.

Magnetic resonance imaging (MRI) biomarkers have been widely studied for their ability to discriminate parkinsonism. Qualitative signs such as the hummingbird sign and the morning glory sign, indicative of midbrain atrophy, were shown to have high specificity, but low sensitivity in differentiating patients with PSP with Richardson’s syndrome (PSP-RS).⁶ Subsequent development of morphometric measurements such as the mid-brain to pons area ratio and the magnetic resonance parkinsonism index (MRPI) resulted in a significant diagnostic performance improvement.^{7,8} These morphometric measurements were also shown to accurately differentiate patients with the cerebellar subtype of MSA (MSAc) along with cerebellar atrophy and the hot cross bun sign.⁸ Putaminal atrophy, diffusivity, and iron load increase were more accurate for the discrimination of the parkinsonian subtype (MSAp).^{9,10} These results led to the incorporation of imaging into the new diagnostic criteria for MSA.¹¹

However, a major gap in current literature is the limited focus on patients with diagnostic uncertainty or in the early stages of the disease. Prospective studies involving early-stage patients with longitudinal follow-ups are challenging to carry out, but essential for understanding the role of MRI in the initial stages of parkinsonism before meeting diagnostic criteria. Studies suggested the MRPI’s ability to predict PSP progression in clinically unclassifiable parkinsonism¹² and before the onset of vertical gaze palsy in PSP-parkinsonism (PSP-P).⁷ Additionally, early putamen and infratentorial structural atrophy in MSA patients, before meeting

diagnostic criteria, hint at MRI’s potential for early differentiation.¹³ However, the precise contribution of MRI to the early diagnosis of parkinsonism, before diagnostic criteria are met, remains to be quantified.

Promising results were obtained by combining fully automated quantitative MRI biomarkers with machine learning approaches to discriminate parkinsonian syndromes.^{14–21} Although most studies targeted patients with confirmed diagnosis, one study achieved 97% accuracy using a supervised machine learning algorithm trained with volumetry versus 62.9% using clinical criteria at initial visit in patients with initial clinical uncertainty.¹⁶ Another study showed accurate classification of participants with moderate to early-stage parkinsonism, examined on different MRI systems as part of their routine clinical assessment, with accuracies ranging between 0.77 and 0.90, suggesting the potential clinical use of machine learning approaches combined with MRI.²² However, the diagnostic gain of these approaches at the individual level and in uncertain cases, needs further investigation.

In the present work, our objective was to determine whether MRI analyzed by visual reading or using machine-learning approaches improved diagnostic accuracy of parkinsonism compared with clinical diagnosis at the initial clinical consultation as compared to the reference diagnosis established by expert consensus after a 2-year follow-up.

Material and Methods

Population

Data were collected from consecutive patients referred to the movement disorders clinic in the Department of Neurology of the Pitié-Salpêtrière Hospital, Paris, between 2014 and 2019, for diagnostic uncertainty between PD, PSP, and MSA. As a clinical routine procedure, patients underwent a clinical examination by a movement disorders specialist, standardized brain

TABLE 1 Definition of the different types of diagnoses

Type of diagnosis	Definition
Gold standard	Established retrospectively by an expert consensus based on all clinical and paraclinical information collected during the two visits, excluding brain MRI.
Clin1	Clinical diagnosis based on clinical criteria at V1.
Clin2	Clinical diagnosis based on clinical criteria at V2.
MRI	Neuroradiologist's diagnosis blinded to the final diagnosis and to the diagnosis according to clinical criteria at V1.
MRI and Clin1	Neuroradiologist's diagnosis blinded to the final diagnosis and in the light of the diagnosis according to clinical criteria at V1.
Algorithm	Diagnostic prediction given by the algorithm.

Abbreviations: MRI, magnetic resonance imaging; V1, initial visit; V2, 2-year follow-up visit.

MRI, and other investigations depending on the patient clinical profile at the initial visit (V1). Approximately 2 years after the initial visit (V2), patients were re-evaluated as part of their follow-up. The gold standard was the final diagnosis established retrospectively by a consensus of two experts with >15 years of experience in movement disorders (B.D., D.G.) based on all clinical and paraclinical information collected during follow-up, excluding brain MRI (Table 1). Patients were excluded when a final diagnosis could not be provided (Fig. 1).

Patients meeting the following criteria were included: suspicion of parkinsonism; availability of a clinical follow-up; final clinical diagnosis of Parkinson's disease (PD),²³ PSP,⁴ or MSA;¹¹ availability of a brain MRI scan concurrent with the initial visit. Exclusion criteria were a lack of dopaminergic denervation on I-123-FP-CIT single photon emission computed tomography (Datscan) or¹⁸ Fluorodopa positron emission tomography (PET) when available; significant movement artifacts on MRI images; and infarcts in the basal ganglia and/or brainstem or other structural abnormality on MRI images. Patients with a diagnosis of Lewy body dementia or corticobasal syndrome were also excluded because brain MRI does not show any suggestive pattern or lacks sensitivity in these disorders.

The local institutional review board approved the retrospective use of the clinical and imaging data for the present study (Parkatypique: CPP Ile-de-France VI08012015).

Clinical Data

Neurological examinations were performed by expert neurologists and included the Hoehn and Yahr score, the Unified Parkinson Disease Rating Scale part III (UPDRS III) score, the Mini-Mental State Examination (MMSE), the Frontal Assessment Battery (FAB), and blood pressure tests (orthostatic hypotension). Additional tests were performed when considered necessary by the neurologist such as eye movement recordings by direct current electro-oculography, neuropsychological evaluation and/or urodynamic testing, and Datscan. We retrospectively established diagnoses and stratified patients into three levels of diagnostic certainty (probable, possible, and undetermined) at V1 and V2 according to published clinical criteria (Supplementary Table 1, and Supplementary Fig. 1).

MRI Acquisition

MRI scans were acquired at the same time as the initial clinical visit. Patients were scanned under clinical conditions as part of their routine diagnostic workup using four different MRI scanners: 3 T Siemens SKYRA (Siemens Healthineers, Erlangen, DE, 3 T General Electric (GE) SIGNA (GE Healthcare, Chicago, IL), 1.5 T GE OPTIMA, and 3 T GE PET/MRI. The distribution of the different disease groups and certainty categories across the different scanners was not significantly different (Supplementary Table 2). The protocol included the following sequences: three-dimensional (3D) gradient-recalled echo T1-weighted sequence, 3D T2 fluid attenuated inversion recovery (FLAIR), proton density-weighted, susceptibility-weighted or T2* images, diffusion-weighted (b = 1000) or diffusion tensor imaging (b = 1000, 15, or 32 directions) with apparent diffusion coefficient or mean diffusivity maps.

Image Analysis

Radiological Analysis

MRI images were retrospectively analyzed using a standardized reading grid (Supplementary Table 3) by two expert neuroradiologists, blinded to the final diagnosis and to the clinical diagnosis at V1 ("MRI"), then in the light of the clinical diagnosis at V1 ("MRI and Clin1") (Supplementary Table 1). Discrepant cases were reviewed until a consensus was found between the two neuroradiologists.

Automated Classification

We used a supervised machine-learning model (logistic regression) developed in our previous study.²² The model was trained on a pre-existing research cohort consisting of 179 participants with probable diagnosis of PD (n = 63), PSP (n = 21), MSA (n = 23), and healthy controls (n = 72) scanned using a 3 T Siemens Trio system (Siemens Healthineers, Erlangen, Germany) with a 32-channel head coil (Supplementary Table 4).

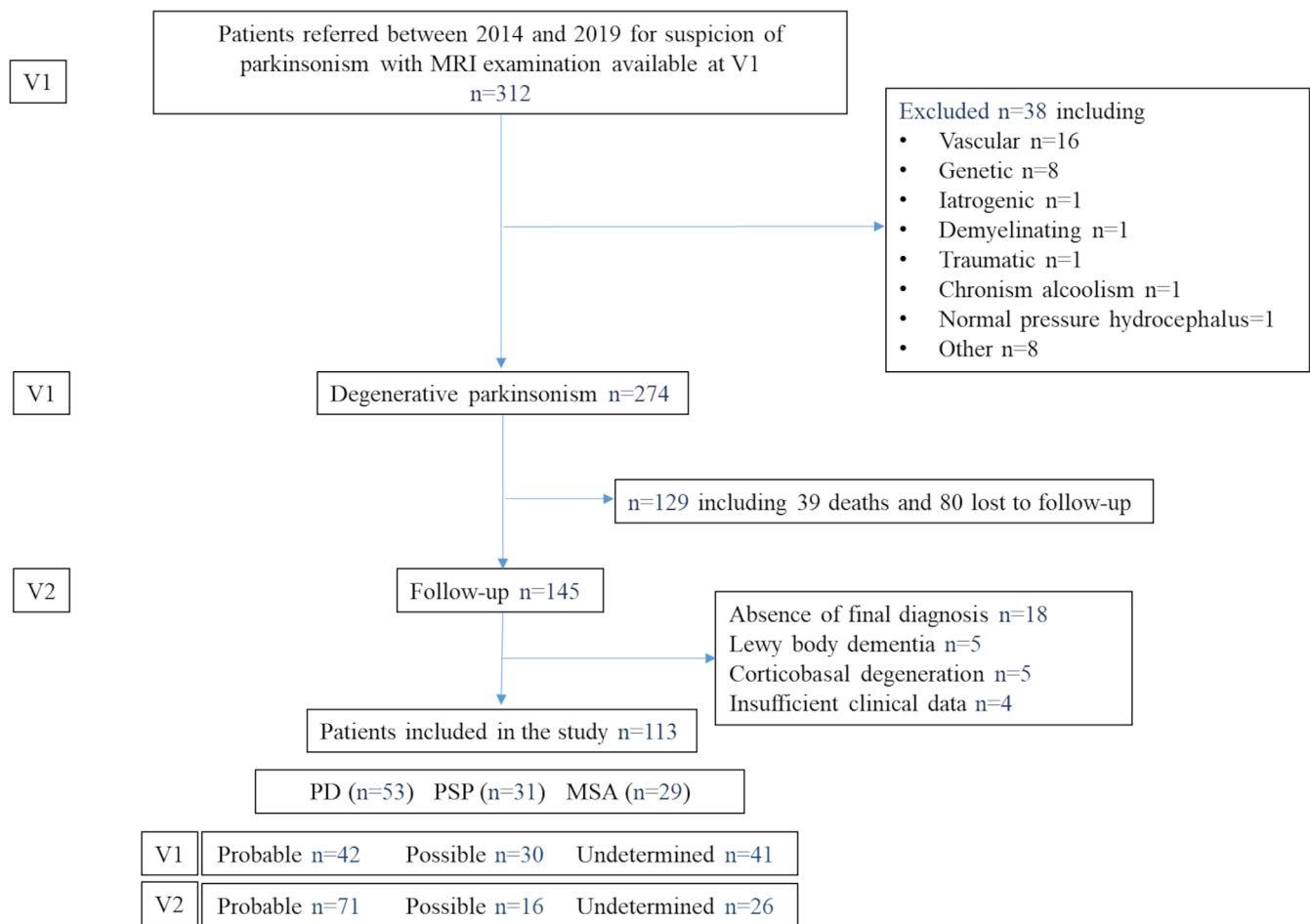


FIG. 1. Distribution of clinical diagnoses. Abbreviations: MSA, multiple system atrophy; PD, Parkinson’s disease; PSP, progressive supranuclear palsy; V1, initial visit; V2, 2-year follow-up visit. [Color figure can be viewed at wileyonlinelibrary.com]

T1-weighted images were automatically segmented using FreeSurfer 6.0 (<https://freesurfer.net/>). Volumes of 12 regions of interest known for being involved in the pathophysiology of parkinsonian syndromes were used as input features for the algorithm. These regions included the brainstem, cerebellar white matter, basal ganglia, thalamus, precentral and insula cortices, and the third and fourth ventricles.²² To remove interindividual variability, all volumes were corrected by the total intracranial volume for each participant. The means of the volumes were used for bilateral regions. The model was then tested on the novel independent population of the current study in each category of diagnostic certainty level (probable, possible, and undetermined). For each subject, a diagnostic prediction defined as the most frequently predicted class using the multiclass classification (PD vs. PSP vs. MSA), was obtained and compared to the gold standard.

Summary of the Different Types of Diagnosis

Overall, we used different types of diagnosis (Table 1):

1. Clinical diagnosis based on international clinical criteria at V1 (“Clin1”).
2. Clinical diagnosis at V2 computed exactly the same way as at V1 (“Clin2”).
3. MRI diagnosis based only on MRI examination at V1 and blinded to the clinical diagnosis (“MRI”).
4. Combined MRI and clinical-based diagnosis at V1 (“MRI and Clin1”).
5. Gold standard (final diagnosis established retrospectively by two experts).

Statistical Analyses

Statistical analyses were conducted using R version 4.2.2 (R Development Core Team, 2022). Quantitative variables were compared using linear regression models with “disease group” and “certainty level category” as factors. The significance of the main or interaction effects was assessed by type II analysis of variance F tests using the function “ANOVA” in the car R package and Tukey post hoc comparisons were performed on statistically significant effects with the emmeans R package. Categorical variables were compared with a

Fisher's exact test for each factor ("disease group" and "certainty category"), and the interaction between the two was assessed with a Cochran–Mantel–Haenszel test. Post hoc analyses using pairwise Fisher's exact tests with false discovery rate correction for multiple comparisons were performed based on a significant effect of a factor (comparisons of levels) or a significant interaction revealed by the Cochran–Mantel–Haenszel test (comparisons of levels on strata of the interacting factor). The level of statistical significance was set at $P < 0.05$.

Interrater reliability between the two neuroradiologists was evaluated using Cohen's κ test.

For each type of diagnosis, accuracy was calculated as the proportion of well-classified subjects out of the total number of subjects with respect to the final reference diagnosis, in each diagnostic certainty category and each disease group. To avoid overestimation of performance because of imbalanced group size, balanced accuracies (BAC) were calculated as the average of the individual-level accuracies across disease groups and certainty categories.

Permutation tests were used to test the magnitude of the differences in BAC achieved by the different diagnoses. Based on 10,000 permutations that randomly exchanged the labels of good and bad classifications between the diagnoses within each disease group, the superiority of one diagnosis over another was, therefore, established with the estimated P -value reporting the proportion of permutations that led to a difference in BAC as large or larger than that observed.

Results

Population

The study population included 113 patients with a final diagnosis of PD ($n = 53$), PSP ($n = 31$ including 21 with Richardson's syndrome [PSP-RS], five with predominant parkinsonism [PSP-P], two with progressive gait freezing [PSP-PGF], two with a mixed form PSP-P/PGF, and one with a frontal presentation [PSP-F]) and MSA ($n = 29$; 17 with a parkinsonian form [MSAp], seven with cerebellar form [MSAc], and five mixed form). At V1, 42 patients met the criteria for probable diagnosis of one of the three types of parkinsonian syndrome (37.2%) and 30 for possible diagnosis (26.5%), whereas the diagnosis remained undetermined for 41 patients (36.3%) (Fig. 1 and Supplementary Fig. 1). Among these 41 patients with an undetermined diagnosis, nine (including five with PD and four with PSP) received a possible or probable diagnosis based on international clinical criteria at V1, but that did not align with the gold standard. Although this category is of interest, we decided to merge these "misdiagnosed" patients with the "undetermined" group to ensure the

statistical analyses' reliability. Indeed, the very low sample size in this "misdiagnosed" category and the absence of representation of this category in the MSA would have biased the computation of balanced accuracies.

Demographic and Clinical Characteristics

There were no significant differences in age, sex, disease duration, or follow-up time between disease groups and between certainty categories. At V1, disease duration was 3.1 ± 2.4 years on average in the whole population. Hoehn and Yahr scores were significantly lower in patients with PD compared with those with PSP and MSA ($P < 0.001$). UPDRS III, MMSE, or FAB scores at V1 did not differ (Table 2). Clinical and imaging characteristics are summarized in Supplementary Tables 5–7.

Diagnostic Accuracies

Clinical Criteria at V1 and V2

Considering all patients meeting the clinical criteria for probable or possible diagnosis, the overall BAC of clinical criteria was 66.2% at V1 ("Clin1"; corresponding to 72 well-diagnosed patients of 113, regardless of disease or certainty group sizes) and significantly increased to 83.5% (87/113) at V2 ("Clin2"; $P = 0.002$). When analyzing by disease group, BAC was equal to 79.3% (23/29) in the MSA group, 64.5% (20/31) in the PSP group, and 54.7% (29/53) in the PD group at V1. At V2, it increased to 94.4% (28/29; $P = 0.002$), 83.1% (25/31; $P = 0.10$), and 73.0% (34/53; $P = 0.02$), respectively (Table 3, Fig. 2).

MRI Analysis

Interrater reliability between the two neuroradiologists was good for both MRI readings, blinded ("MRI"; $\kappa = 0.76$) and in light of the clinical diagnosis ("MRI and Clin1"; $\kappa = 0.78$).

Compared with the clinical criteria at V1, the overall BAC of the MRI analysis alone was 14.3% points higher ("MRI": 80.5% vs. "Clin1": 66.2%; $P = 0.01$) and 19.2% points higher when combined with the clinical diagnosis at V1 ("MRI and Clin1": 85.4%; $P < 0.0001$).

When analyzing by level of diagnostic certainty, the BAC of "MRI" was 89.4% (39/42) in patients with probable form, 82.1% (24/33) in those with possible form, and 69.8% (27/43) in those whose diagnosis was undetermined. Combining MRI analysis with the clinical diagnosis at V1 ("MRI and Clin1") increased the BAC to 93.6% (40/42) in patients with probable form, 86.9% (27/33) in those with possible form, and 75.6% (31/43) in those with undetermined form, although the difference was not statistically significant. When analyzing by disease group, there was a trend toward a better

TABLE 2 Demographic and clinical characteristics at the initial visit (V1)

Disease groups	MSA					Global tests			
	PD	PSP	Total	MSAp	MSAc	Mixed MSA	Status	Certainty by status	Post hoc comparisons
n	53	31	29	17	7	5	—	—	—
Age (y)	64.9 ± 10.6 [40–87]	68.4 ± 7.9 [46–83]	62.5 ± 8.7 [47–81]	63.9 ± 9.5 [47–81]	60.9 ± 7.1 [50–72]	60.0 ± 8.8 [50–71]	0.06	0.8	—
Gender M/F (%M/%F)	40/13 (75.5%/24.5%)	20/11 (64.5%/35.5%)	17/12 (58.6%/41.4%)	10/7 (58.8%/41.2%)	4/3 (57.1%/42.9%)	3/2 (60%/40%)	0.26	0.62	0.19
Disease duration (y)	3.1 ± 2.9 [1–19]	3.2 ± 1.6 [1–7]	3.0 ± 2.1 [1–10]	3.1 ± 2.0 [1–8]	3.1 ± 3.1 [1–10]	2.6 ± 0.6 [2–3]	0.98	0.15	0.4
Follow-up time (months)	27.1 ± 8.1 [15–73]	25.4 ± 8.3 [12–48]	25.4 ± 8.0 [16–59]	24.5 ± 4.0 [16–34]	25.5 ± 6.38 [18–34]	28.6 ± 17.3 [16–59]	0.56	0.15	0.4
LEDD (mg)	322.1 ± 239.5 [0–1064]	285.6 ± 271.6 [0–950]	372.2 ± 371.1 [0–1395]	391.4 ± 308.4 [0–1169]	327.9 ± 523.4 [0–1395]	368.8 ± 409.2 [0–1044]	0.50	0.04	0.54
Hodhn and Yahr score (range, 1–5; 5 = worse)	1.6 ± 0.6 [1–3] (50)	3.1 ± 0.9 [1–5] (28)	2.6 ± 1.1 [1–5] (25)	2.7 ± 1.1 [1–5] (16)	2.4 ± 1.0 [1–3] (4)	2.8 ± 1.1 [1–4]	<0.0001	0.18	0.16
UPDRS III (range, 0–108, higher = worse)	21.5 ± 10.6 [5–49] (26)	28.8 ± 20.4 [6–71] (14)	21.8 ± 14.9 [2–58] (14)	29.1 ± 16.5 [11–58] (7)	14.3 ± 10.6 [2–25] (4)	14.7 ± 9.3 [4–21] (3)	0.32	0.63	0.81
MMSE (range, 0–30; lower = worse cognition)	27.0 ± 3.2 [17–30] (41)	25.2 ± 3.9 [18–30] (24)	26.4 ± 3.7 [16–30] (25)	26.0 ± 4.6 [16–30] (14)	26.7 ± 2.6 [23–30] (6)	27.0 ± 1.9 [25–30]	0.13	0.65	0.3
FAB (range, 0–18; 18 = better outcome)	14.7 ± 2.8 [8–18] (39)	12.1 ± 3.7 [3–17] (22)	14.9 ± 2.5 [7–18] (24)	15.2 ± 3.0 [7–18] (13)	15.0 ± 1.8 [13–18] (6)	14.0 ± 2.2 [11–17]	0.002	0.63	0.12
							0.002	0.63	0.12
									ns

Note: Quantitative variables are summarized as mean ± standard deviation (min–max) (number of available values), and qualitative variables as counts and percentages. Statistically significant effects for global tests are shown in bold. Asterisks indicate the significance level of the post hoc comparisons: adjusted $P > 0.05$ (ns), adjusted $P < 0.05$ (*), adjusted $P < 0.01$ (**), adjusted $P < 0.001$ (***). Significant p values (<0.05) are highlighted in bold. Patients with MSAp, MSAc, and mixed phenotype were analyzed together in a group called “MSA”. Abbreviations: MSA, multiple system atrophy; PD, Parkinson’s disease; PSP, progressive supranuclear palsy; MSAp, parkinsonian form of multiple system atrophy; MSAc, cerebellar form of multiple system atrophy; M, male; F, female; LEDD, levodopa equivalent daily dose; UPDRS, Unified Parkinson Disease Rating Scale; MMSE, Mini-Mental State Examination; FAB, Frontal Assessment Battery; NA, not assigned; ns, not significant;

TABLE 3 Diagnostic accuracies

Groups	All diseases (BAC)				PD		PSP		MSA	
	All	Probable	Possible	Undetermined	Accuracy	BAC	Accuracy	BAC	Accuracy	BAC
Clin1	66.2%	NA	NA	NA	54.7% (29/53)	54.7%	64.5% (20/31)	64.5%	79.3% (23/29)	79.3%
Clin2	83.5%	98.2%	95.2%	57.1%	64.2% (34/53)	73.0%	80.6% (25/31)	83.1%	96.6% (28/29)	94.4%
MRI	80.5%	89.4%	82.1%	69.8%	67.9% (36/53)	69.6%	87.1% (27/31)	87.9%	86.2% (25/29)	83.9%
MRI and Clin1	85.4%	93.6%	86.9%	75.6%	79.2% (42/53)	81.3%	90.3% (28/31)	90.9%	86.2% (25/29)	83.9%
Algorithm	76.1%	81.2%	88.2%	58.8%	89.6% (43/48)	91.9%	62.5% (15/24)	72.3%	65.5% (19/29)	63.9%

Note: Accuracies (number of patients with a correct diagnosis between brackets) and balanced accuracies are reported for the whole population ("All"), by level of diagnostic certainty at V1 and by disease group for each type of diagnosis: clinical criteria at V1 ("Clin1"), clinical criteria at V2 ("Clin2"), MRI alone ("MRI"), MRI combined with clinical criteria at V1 ("MRI and Clin1"), and algorithm ("Algorithm"). Patients with MSaP, MSaC, and mixed phenotype were analyzed together in a group called "MSA". Abbreviations: BAC, balanced accuracy; PD, Parkinson's disease; PSP, progressive supranuclear palsy; MSA, multiple system atrophy; MRI, magnetic resonance imaging; V1, initial visit; V2, follow-up visit; NA, not applicable.

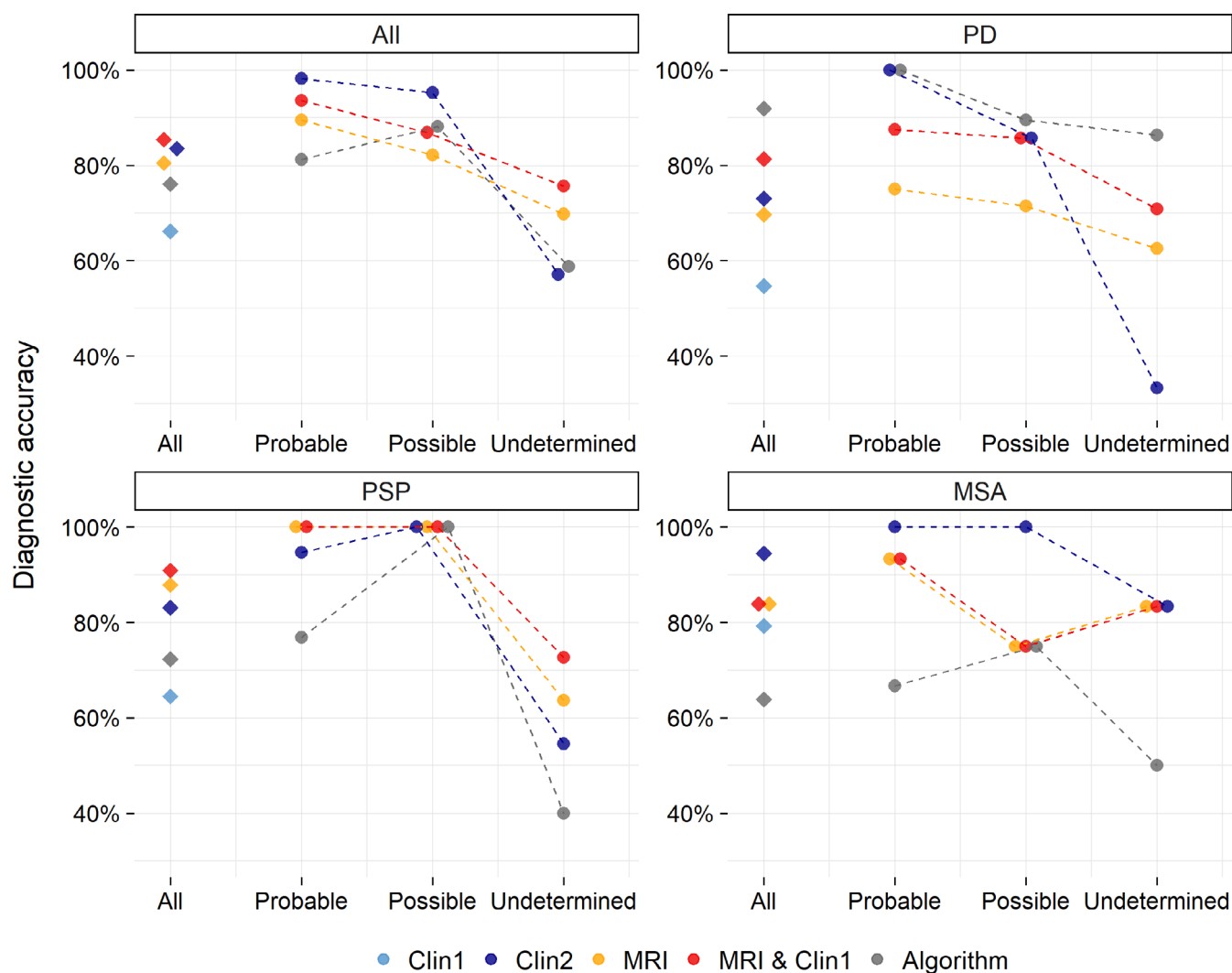


FIG. 2. Diagnostic accuracies. Percentages of the balanced accuracies (y-axis) in the whole population ("All") and in each disease group (PD, PSP, MSA) are given globally (diamonds) and by level of diagnostic certainty at V1 (probable, possible, undetermined; bullet points; x-axis) for each modality (clinical criteria at V1: light blue; clinical criteria at V2: purple; MRI: yellow; MRI combined with clinical criteria at V1 (MRI and Clin1): red; algorithm: gray). Balanced accuracies of clinical criteria at V1 are only reported globally and not by level of diagnostic certainty because clinical criteria were precisely used to define the different levels of diagnostic certainty at V1. Abbreviations: MRI, magnetic resonance imaging; MSA, multiple system atrophy; PD, Parkinson's disease; PSP, progressive supranuclear palsy; V1, initial visit; V2, 2-year follow-up visit. [Color figure can be viewed at wileyonlinelibrary.com]

classification of patients with PSP (BAC = 87.9%) and MSA (83.9%) than those with PD (69.6%) using MRI alone (“MRI”, $P = 0.07$ and 0.08 , respectively). No significant between-group difference was seen using “MRI and Clin1” (BAC: PSP, 90.9%; MSA, 83.9%; PD, 81.3%) (Table 3, Fig. 2).

Algorithm

Automated classification was available for 101 subjects out of 113 including 48 PD, 24 PSP, and 29 MSA (MRI artifacts for 11 subjects—five PD and six PSP; brain segmentation failure for one PSP). The overall BAC using the algorithm increased by 9.9% points compared to clinical criteria at V1 (“Algorithm”: 76.1% vs. “Clin1”: 66.2%), although the difference did not reach significance ($P = 0.08$). It was high in the possible (88.2%, 24/28) and probable (81.2%, 27/35) categories, and low in the undetermined category (58.8%, 26/38). The algorithm classified PD patients (BAC = 91.9%) better than those with MSA (BAC = 63.9%, $P = 0.006$); no difference was seen with the PSP group (BAC = 72.3%) (Table 3, Fig. 2).

Discordances between MRI and Final Diagnosis

Supplementary Table 8 provides the characteristics of the discordant patients.

In the PD group, 11 of 53 patients (20.7%) were discordant. Ten patients (including one with probable, three with possible, and six with undetermined diagnosis at V1) were classified as PSP on MRI because of midbrain atrophy. These patients with a “PSP-like” MRI pattern were significantly older than those with “typical PD” (ie, PD patients in whom MRI was consistent with the final diagnosis) (73.2 vs. 62.7 years, $P = 0.005$). The “typical PD” group had a higher proportion of males (32 males [M]/10 females [F]) compared with the “PSP-like” group (7 M/3F, $P = 0.03$). Disease duration did not differ. There was no difference for patients with “typical PSP” (ie, PSP patients in whom MRI was consistent with the final diagnosis) for age, gender, or disease duration. At V1, the “PSP-like” patients had significantly more postural instability within 3 years (P1) (30%) than those with “typical PD” (4.8%, $P = 0.04$) and less than those with “typical PSP” (75%, $P = 0.03$). They had more rest tremor than patients with “typical PSP” (70 vs. 29%, $P = 0.045$). At V2, they also had progressive freezing of gait within 3 years (A1, 20%) unlike patients with “typical PD” ($P = 0.02$).

The remaining discordant patient with a final clinical diagnosis of PD was classified as MSAP on MRI because of suggestive abnormalities. He had dysautonomia, levodopa (L-dopa)-sensitive parkinsonism, Babinski sign, and no cerebellar syndrome.

In the PSP group, three of 31 patients (9.7%) were discordant. They were all clinically undetermined at V1, and two of them progressed to a probable diagnosis at V2:

- One patient with a clinical phenotype of PSP-P was classified as MSAP on MRI. He had oculomotor disturbances (apraxia of eye-opening or frequent macro square wave jerk-O3), postural instability (repeated unprovoked falls within 3 years-P1; tendency to fall on pull-test within 3 years-P2), and akinesia (parkinsonism with tremor and/or asymmetric and/or L-dopa-responsive-A3), but also dysautonomia, cerebellar syndrome, and L-dopa-sensitive parkinsonism.
- One patient with a clinical phenotype of PSP-PGF was classified as PD on MRI in the absence of mid-brain atrophy. He had O3, P1, A1 (progressive freezing of gait within 3 years), and A2 signs as well as apathy, dysexecutive symptoms, and bradyphrenia.
- One patient with a clinical phenotype of PSP-RS had asymmetric parietal cortical atrophy, but no midbrain atrophy. He had O3, P1, A3 signs, dysautonomia, but no dystonia, apraxia, or sensory cortical deficit.

In the MSA group, four of 26 patients (two MSAP and two MSAC, 15.4%) were discordant. They were all classified as PD on MRI. At V1, three had an undetermined diagnosis and one had a probable diagnosis and all progressed to probable diagnosis at V2. These patients with a “PD-like” MRI pattern had dysautonomia, L-dopa-sensitive parkinsonism (75% vs. 66.7% in the “typical MSA” group and 92.9% in the “typical PD” group) and cerebellar syndrome in the two MSAC cases.

Discussion

This study compared the diagnostic accuracies of clinical criteria at initial and 2-year follow-up visits to MRI analysis by visual reading or by using a machine-learning algorithm. Patients were referred for an uncertain diagnosis of degenerative parkinsonism at a relatively early stage of the disease. Considering retrospective diagnosis by expert consensus as the gold standard, clinical criteria showed limited accuracy at the initial visit (66.2%), which improved after a 2-year follow-up (83.5%). In comparison to clinical criteria at V1, the initial MRI analysis by expert visual reading provided a diagnostic gain of 14.3% points (80.5% vs. 66.2%). This gain increased to 19.2% points when MRI was combined with the clinical diagnosis at V1 (85.4%). This combined approach reached an accuracy close to that achieved with clinical criteria at follow-up. The machine-learning-based classification yielded a diagnostic gain of 9.9% points (76.1%), although the accuracy was lower in patients with an undetermined diagnosis at V1.

Clinical Criteria

The relatively low accuracy of clinical criteria at V1, especially for the diagnosis of PD, could be partly explained by the specific patient population recruited. Our center serves as a tertiary facility, where neurologists refer patients in case of uncertainty regarding the diagnosis of parkinsonism. Patients with PD can present with unusual features (red flags) during the initial stages of the disease. As shown in our study (Supplementary Tables 5 and 6), 39.6% of PD patients exhibited autonomic dysfunction at V1, 19% showed oculomotor signs such as slowed vertical saccades, apraxia of eyelid opening, or frequent macro square wave jerks, and 7.5% had a poor L-dopa response. As patients were referred at an early stage of their disease, the diagnosis was all the more difficult compared with more advanced patients. Our study emphasizes that overlapping clinical features in atypical parkinsonism at early stages can challenge an accurate diagnosis using international clinical criteria. Similarly, using neuropathological examination, a previous study demonstrated that the accuracy of clinical criteria was only 26% for a diagnosis of PD in untreated or not clearly responsive patients, 53% accuracy for early medication-responsive PD (<5 years of duration), and >85% diagnostic accuracy for medication-responsive PD of longer course.²⁴ Several clinico-histological series have also reported discrepancies regarding the diagnosis of MSA. Up to 38% of patients with a postmortem diagnosis of Lewy body dementia, PSP, and PD clinically masqueraded as MSA,² and, conversely, 19% of patients with a postmortem diagnosis of MSA clinically mimicked PD and PSP.²⁵

Further, the accuracy of clinical criteria improved during follow-up, rising from 66.2% to 83.5%. However, it is worth noting that it did not attain 100% accuracy when compared with the gold standard, which was established by considering the opinions of two movement disorders experts, based on all clinical and paraclinical information (except MRI). Indeed, it has been demonstrated that expert opinions yield higher accuracy compared with the application of formal clinical criteria.^{26,27}

MRI Analysis

Our study showed that MRI can enhance the accuracy of parkinsonism diagnosis at the initial visit, upgrading it from “possible” to “probable” or from “undetermined” to “possible”, allowing for diagnosis 2 years earlier than clinical follow-up. MRI holds the potential to significantly shorten the diagnosis process, ultimately aiding in the timely enrollment of eligible patients into neuroprotection trials. It also mitigates the risk of misdiagnosis that could compromise the validity of trial outcomes. However, some discordances were observed between the MRI-based diagnosis and the

gold standard in 16% of patients. Here are the insights from our results (1) when the MRI pattern is consistent with MSA, it is nearly pathognomonic, and we believe that the diagnosis of MSA should be retained, regardless of the clinical diagnosis. (2) If there is a clinical profile indicating MSA, but no suggestive abnormalities on the initial MRI, an MRI should be repeated 2 years later. Results from the literature regarding the sensitivity of MRI for the diagnosis of MSA were discordant. MRI signs associated with MSA were reported to be visible relatively early, especially at 3T, and using the correct sequences along with proper interpretation.¹³ Moreover, in a study including 13 patients with a postmortem diagnosis of MSA, who were at 5 years from disease onset at scan, radiological signs such as the hot cross bun sign or the middle cerebellar peduncles hyperintensity had low sensitivity (33%–67%), but excellent specificity (100%). This suggests that patients with MSA can present with a normal MRI even at a moderate to advanced stage. However, patients in this study were scanned at a low field strength (0.5–1.5 T), whereas it is well-known that MRI abnormalities suggestive of MSA are better detected at high field strength and can be overlooked at lower field strength.¹³ Moreover, of 13 MSA patients, only two had a cerebellar phenotype, which may contribute to the low sensitivity of the cerebellar features observed in this study. Additionally, the study did not report on the accuracy of putaminal signs as seen in MSaP. (3) In our study, the most difficult distinction was between PD and PSP when the MRI criteria were not met or when MRI signs of PSP were present without consistent clinical signs. Previous studies showed that midbrain atrophy is a predictive sign of later conversion to PSP.^{7,12} Indeed, the diagnosis of patients with slowly progressive PSP may be particularly challenging.²⁸ It may be hypothesized that MRI criteria are met before clinical hallmarks of the disease (ie, supranuclear gaze palsy) appear in this subset of patients, preventing reaching a probable or even possible diagnosis of PSP according to clinical criteria.^{7,12} Therefore, these patients may remain with a clinical diagnosis of PD for years. The understanding of the temporal relationship between clinical and MRI progression needs further investigation. Moreover, the specificity of morphometric biomarkers decreases in the elderly population, as midbrain atrophy occurs during normal aging. Therefore, the reliability of literature-based thresholds beyond the age of 80 years remains unclear.²⁹ Conversely, PSP variants generally have little midbrain atrophy compared to PSP-RS, explaining the lack of midbrain atrophy in one patient with PSP-PGF.^{30,31} To our knowledge, such radio-clinical discordances have never been reported in the literature. They may raise questions about the accuracy of clinical diagnoses, even when they are made by experts. In the absence of pathological confirmation, these unsolved challenges

emphasized the urgent need for disease-specific biomarkers such as PET using Tau and alpha-synuclein radiotracers, or fluid biomarkers of alpha-synuclein or Tau pathology.³²⁻³⁴ In the meantime, a longer follow-up could help refine the diagnosis and determine whether patients with discordant diagnoses eventually change their diagnostic trajectory.

Algorithm

Several factors may have contributed to lower the performance of the algorithm, especially in the undetermined category. First, the training and testing cohorts had different characteristics in terms of disease progression duration, diagnostic certainty, and clinical profiles. The training cohort consisted of participants meeting probable disease criteria and scanned in a research setting. In contrast, the test cohort (from the current study) included patients with high diagnostic uncertainty enrolled as part of their routine care. Second, the test cohort underwent MRI scans on different scanners under clinical conditions, potentially introducing inter-scanner variability that can impact the gray/white matter contrast of MRI images and, therefore, the automatic segmentation. These factors could limit the generalizability of the machine-learning model. Additionally, the algorithm demonstrated better prediction for subjects with PD compared to those with PSP and MSA, possibly because of an overrepresentation of PD cases in the training cohort. Nevertheless, our study emphasizes the potential of such algorithms even in real-life conditions, encompassing different MRI scanners and a non-selected patient sample.

To our knowledge, only one previous study has compared the performance of automated classification with clinical consensus criteria.¹⁶ The study targeted a highly selected population of parkinsonian patients at an early to moderate stage of the disease, relatively younger (<74 years) and without significant vascular leukopathy. A decision tree trained with volumetry achieved a 97% accuracy versus 62.9% and 100% using clinical criteria at the initial visit and follow-up, respectively.¹⁶ Our population differed from the previous study as we included patients without strict exclusion criteria on age or vascular leukopathy, which may have introduced confounding factors. However, patients with stroke and/or lacunar infarct in the basal ganglia or brainstem accounting for vascular parkinsonism were systematically excluded. This highlights the challenge of applying research-developed tools from a highly controlled population to a diverse population seen in a clinical setting. There are potential avenues for improving the automated classification performance. Training the algorithm with a combination of MRI biomarkers and clinical variables could enhance its discriminatory ability. Additionally, incorporating other quantitative MRI biomarkers such as

diffusivity, free water, or R2* relaxometry could help improve the performance of the machine-learning classifier. Last, ensuring balanced class representation during training could help reduce misclassification.

Our study has several limitations, including the lack of pathological diagnosis as discussed above. However, cases were systematically excluded when the available clinical data was deemed insufficient. Furthermore, although 2 years of follow-up is not a particularly long time, we believed that it was sufficient in most patients to create a “consensus” diagnosis. Indeed, at the time of inclusion, patients were, on average, ~3 years from the onset of their disease, resulting in a total disease duration of 5 years after 2 years of follow-up. Moreover, atypical parkinsonian syndromes tend to progress rapidly. As a result, we hypothesized that clinical phenotypes could become more distinct within this two-year timeframe (therefore, 5 years after disease onset), allowing for a more reliable diagnosis. Patients for whom the expert consensus could not provide a definitive diagnosis after follow-up were excluded from the study. Further, the follow-up visit was performed by a different clinician than at the baseline consultation. However, a standardized approach listing all motor and non-motor symptoms comprehensively was used, ensuring the reliability of the assessments.

Conclusion

Our study underscores the value of expert visual MRI analysis in the early differentiation of parkinsonism when compared to international clinical criteria at the initial visit. Combining MRI with clinical criteria can help establish an earlier and more accurate diagnosis, as evidenced by the integration of imaging in the new diagnostic criteria for “clinically established” MSA.¹¹ The machine-learning-based classification also enabled the discrimination of parkinsonism at the individual level, which could be beneficial in non-expert centers. Its performance could be further enhanced by incorporating imaging biomarkers sensitive to new contrasts, along with relevant clinical variables of interest. ■

Acknowledgments: L.C. is supported by an Inria/AP-HP “Poste d'accueil.” 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 Neurosciences de Paris, Électricité 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.

Data Availability Statement

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.

References

1. Wenning GK, Stankovic I, Vignatelli L, et al. The Movement Disorder Society criteria for the diagnosis of multiple system atrophy. *Mov Disord* 2022;37(6):1131–1148.
2. Koga S, Aoki N, Uitti RJ, et al. When DLB, PD, and PSP masquerade as MSA: an autopsy study of 134 patients. *Neurology* 2015;85:404–412.
3. Respondek G, Roeber S, Kretzschmar H, et al. 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* 2013;28:504–509.
4. Höglinger GU, Respondek G, Stamelou M, et al. Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria. *Mov Disord* 2017;32:853–864.
5. Mamarabadi M, Razjouyan H, Golbe LI. Is the latency from progressive supranuclear palsy onset to diagnosis improving? *Mov Disord Clin Pract* 2018;5:603–606.
6. Mueller C, Hussl A, Krismer F, et al. The diagnostic accuracy of the hummingbird and morning glory sign in patients with neurodegenerative parkinsonism. *Parkinsonism Relat Disord* 2018;54:90–94.
7. Quattrone A, Morelli M, Williams DR, et al. MR parkinsonism index predicts vertical supranuclear gaze palsy in patients with PSP-parkinsonism. *Neurology* 2016;87:1266–1273.
8. Quattrone A, Nicoletti G, Messina D, et al. MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy. *Radiology* 2008;246:214–221.
9. Chougar L, Pyatigorskaya N, Degos B, Grabli D, Lehericy S. The role of magnetic resonance imaging for the diagnosis of atypical parkinsonism. *Front Neurol* 2020;11:665.
10. Chougar L, Pyatigorskaya N, Lehericy S. Update on neuroimaging for categorization of Parkinson's disease and atypical parkinsonism. *Curr Opin Neurol* 2021;34(4):514–524.
11. Gilman S, Wenning GK, Low PA, et al. Second consensus statement on the diagnosis of multiple system atrophy. *Neurology* 2008;71:670–676.
12. Morelli M, Arabia G, Novellino F, et al. MRI measurements predict PSP in unclassifiable parkinsonisms: a cohort study. *Neurology* 2011;77:1042–1047.
13. Osaki Y, Morita Y, Miyamoto Y, et al. Identification of a prepossible multiple system atrophy phase. *Acta Neurol Scand* 2021;143:313–317.
14. Cherubini A, Morelli M, Nisticó R, et al. Magnetic resonance support vector machine discriminates between Parkinson disease and progressive supranuclear palsy. *Mov Disord* 2014;29:266–269.
15. Huppertz H-J, Möller L, Südmeyer M, et al. Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. *Mov Disord* 2016;31:1506–1517.
16. Scherfler C, Göbel G, Müller C, et al. Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism. *Neurology* 2016;86:1242–1249.
17. Péran P, Barbagallo G, Nemmi F, et al. MRI supervised and unsupervised classification of Parkinson's disease and multiple system atrophy. *Mov Disord* 2018;33:600–608.
18. Salvatore C, Cerasa A, Castiglioni I, et al. Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy. *J Neurosci Methods* 2014;222:230–237.
19. Archer DB, Bricker JT, Chu WT, et al. Development and validation of the automated imaging differentiation in parkinsonism (AID-P): a multi-site machine learning study. *Lancet Digit Health* 2019;1:e222–e231.
20. Morisi R, Manners DN, Gnecco G, et al. Multi-class parkinsonian disorders classification with quantitative MR markers and graph-based features using support vector machines. *Parkinsonism Relat Disord* 2018;47:64–70.
21. Nigro S, Antonini A, Vaillancourt DE, et al. Automated MRI classification in progressive supranuclear palsy: a large international cohort study. *Mov Disord* 2020;35(6):976–983.
22. Chougar L, Faouzi J, Pyatigorskaya N, et al. Automated categorization of parkinsonian syndromes using magnetic resonance imaging in a clinical setting. *Mov Disord* 2020;36(2):460–470.
23. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30:1591–1601.
24. Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease. *Neurology* 2014;83:406–412.
25. Miki Y, Tsushima E, Foti SC, et al. Identification of multiple system atrophy mimicking Parkinson's disease or progressive supranuclear palsy. *Brain* 2021;144:1138–1151.
26. Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain* 2002;125:861–870.
27. Postuma RB, Poewe W, Litvan I, et al. Validation of the MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2018;33:1601–1608.
28. Lukic MJ, Respondek G, Kurz C, et al. Long-duration progressive supranuclear palsy: clinical course and pathological underpinnings. *Ann Neurol* 2022;92:637–649.
29. Ruiz ST, Bakklund RV, Häberg AK, Berntsen EM. Normative data for brainstem structures, the midbrain-to-pons ratio, and the magnetic resonance parkinsonism index. *AJNR Am J Neuroradiol* 2022;43:707–714.
30. Whitwell JL, Jack CR, Parisi JE, et al. Midbrain atrophy is not a biomarker of progressive supranuclear palsy pathology. *Eur J Neurol* 2013;20:1417–1422.
31. Whitwell JL, Höglinger GU, Antonini A, et al. Radiological biomarkers for diagnosis in PSP: where are we and where do we need to be? *Mov Disord* 2017;32:955–971.
32. Zerr I. RT-QuIC for detection of prodromal α -synucleinopathies. *Lancet Neurol* 2021;20:165–166.
33. Iranzo A, Fairfoul G, Ayudhaya ACN, et al. Detection of α -synuclein in CSF by RT-QuIC in patients with isolated rapid-eye-movement sleep behaviour disorder: a longitudinal observational study. *Lancet Neurol* 2021;20:203–212.
34. Nicastro N, Rodriguez PV, Malpetti M, et al. 18F-AV1451 PET imaging and multimodal MRI changes in progressive supranuclear palsy. *J Neurol* 2020;267:341–349.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.