Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism

Christoph Scherfler, MD*
Georg Göbel, PhD*
Christoph Müller, MD
Michael Nocker, MD
Gregor K. Wenning, MD, PhD
Michael Schocke, MD
Werner Poewe, MD
Klaus Seppi, MD

Correspondence to Dr. Scherfler: christoph.scherfler@i-med.ac.at

ABSTRACT

Objective: To determine whether automated and observer-independent volumetric MRI analysis is able to discriminate among patients with Parkinson disease (PD), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP) in early to moderately advanced stages of disease.

Methods: T1-weighted volumetric MRI from patients with clinically probable PD (n = 40), MSA (n = 40), and PSP (n = 30) and a mean disease duration of 2.8 ± 1.7 y were examined using automated volume measures of 22 subcortical regions. The clinical follow-up period was 2.5 ± 1.2 years. The data were split into a training (n = 72) and a test set (n = 38). The training set was used to build a C4.5 decision tree model in order to classify patients as MSA, PSP, or PD. The classification algorithm was examined by the test set using the final clinical diagnosis at last follow-up as diagnostic gold standard.

Results: The midbrain and putaminal volume as well as the cerebellar gray matter compartment were identified as the most significant brain regions to construct a prediction model. The diagnostic accuracy for PD vs MSA or PSP was 97.4%. In contrast, diagnostic accuracy based on validated clinical consensus criteria at the time of MRI acquisition was 62.9%.

Conclusions: Volume segmentation of subcortical brain areas differentiates PD from MSA and PSP and improves diagnostic accuracy in patients presenting with early to moderately advanced stage parkinsonism.

Classification of evidence: This study provides Class III evidence that automated MRI analysis accurately discriminates among early-stage PD, MSA, and PSP. ***Neurology*® 2016;86:1242–1249**

GLOSSARY

ANOVA = analysis of variance; **AUC** = area under the curve; **DSM-IV** = *Diagnostic and Statistical Manual of Mental Disorders, 4th edition*; **MANCOVA** = multivariate analysis of covariance; **MSA** = multiple system atrophy; **PD** = Parkinson disease; **PSP** = progressive supranuclear palsy; **UPDRS** = Unified Parkinson's Disease Rating Scale.

Signs and symptoms of atypical parkinsonian disorders such as progressive supranuclear palsy (PSP) and multiple system atrophy (MSA) frequently mimic those of Parkinson disease (PD), particularly in early stages of disease.[1] In pathologically proven series of atypical parkinsonian disorders, only 50% of cases were found to be correctly diagnosed by their primary neurologist.[1–3] Early differentiation of PD from atypical parkinsonian disorders is key to adequate patient management and is also crucial for recruitment into academic studies. In recent years, numerous investigators using routine MRI, magnetic resonance spectroscopy, diffusion-weighted MRI, and volumetric MRI have all provided some evidence of diagnostic utility in separating PD from PSP and MSA at the group level; however, attempts to implement and validate such findings in the clinical setting at the single patient level are still limited.[4,5] Recent advances in image analysis algorithms have led to the development of structural MRI-based software tools such as FreeSurfer that enable automatic compartmentalization of the brain into multiple anatomic regions and quantify tissue volume in these regions on an individual case basis.[6]

Supplemental data at Neurology.org

*These authors contributed equally to the manuscript.

From the Departments of Neurology (C.S., C.M., M.N., G.K.W., W.P., K.S.), Medical Statistics, Informatics and Health Economics (G.G.), and Radiology (M.S.), Medical University of Innsbruck, Austria.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Advances in the processing of high dimensional data such as data mining techniques facilitate the development of decision algorithms for the assignment of previously identified signal alterations to disease entities.[7–9] In the present study, such a decision algorithm was built by including 22 segmented subcortical regions and validated for the capability to discriminate patients with early to moderately advanced stage MSA or PSP from PD in the clinical routine workup.

METHODS

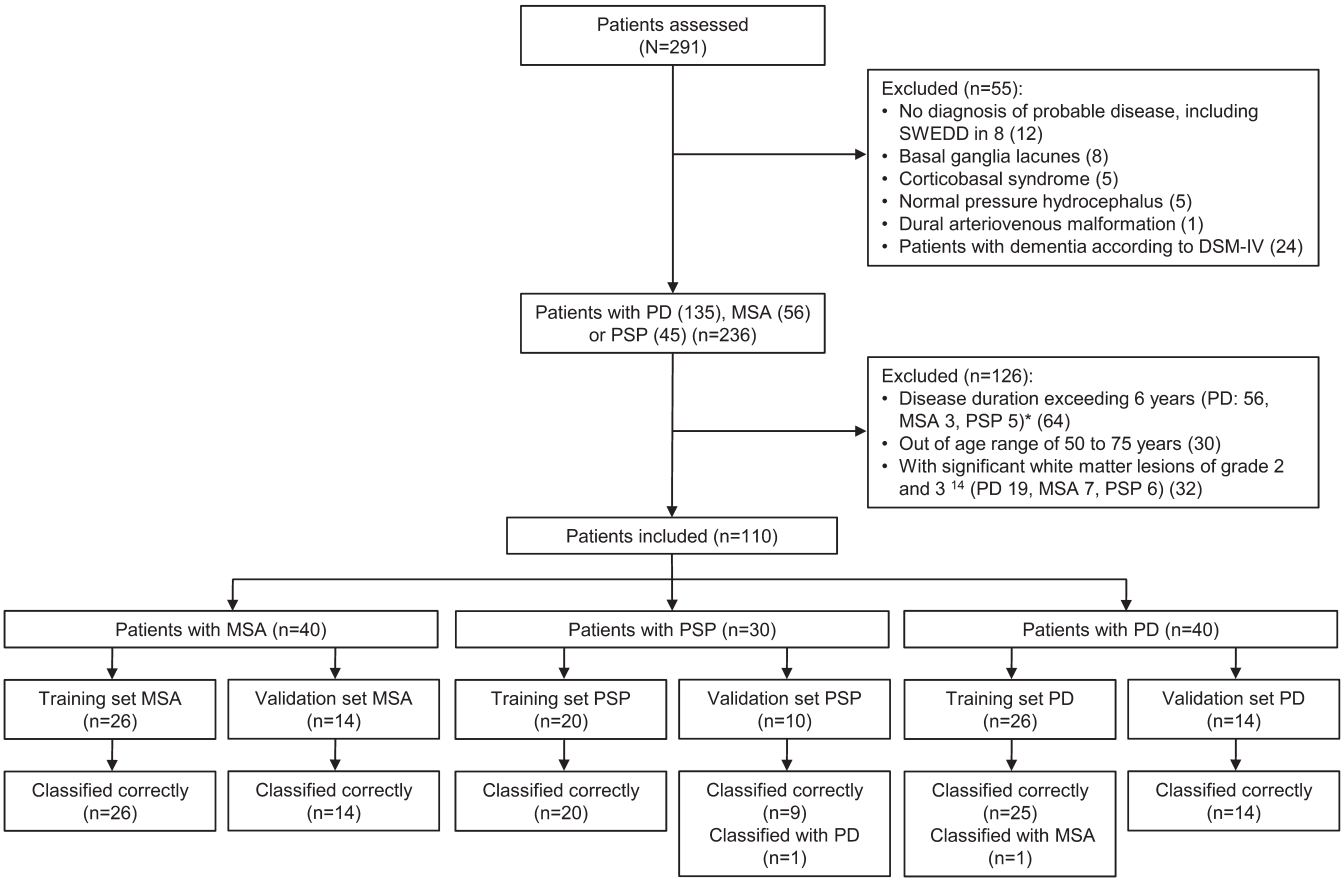
Participants. Participants were identified from the MRI database of the Movement Disorders Clinic of the Department of Neurology at the Medical University of Innsbruck from May 2005 until November 2013. To be eligible, participants had to fulfill several inclusion criteria (figure 1): (1) clinical diagnosis of probable Parkinson variant of MSA, PSP, and PD according to consensus operational criteria made by 2 movement disorders specialists at clinical follow-up of at least 12 months for the Parkinson variant of MSA, PSP, and L-dopa-responsive PD and 42 months for patients with de novo parkinsonism and a final diagnosis of PD[2,10,11]; (2) disease duration of less than 6 years defined by the period of patient-reported experience of symptoms until the time of MRI; (3) age 50–75 years at MRI examination; (4) presynaptic nigrostriatal dopaminergic dysfunction confirmed by either dopamine transporter SPECT or [^{18}F]-dopa PET; and (5) structured clinical assessment of motor disability on regular medication using part III of the Unified Parkinson's Disease Rating Scale (UPDRS) and classification of Hoehn & Yahr stage at the time of MRI (table 1).[12,13] MRI was performed within 1 month of the initial clinical assessment.

Patients with dementia according to the DSM-IV were excluded. Forty-one healthy individuals with no signs of parkinsonism or any other CNS disorders on careful clinical examination and a Mini-Mental State Examination score of >28 served as age-matched control group. Participants with white matter lesions of grade 2 and 3, vascular or space-occupying lesions within the cerebrum, or motion artefacts were excluded.[14]

Study hypotheses/classification of evidence. The hypothesis tested in this proof-of-concept study was that the diagnostic accuracy obtained from the classification algorithm would perform at least equivalent compared to the application of current consensus operational criteria for PD, MSA, and PSP at the initial visit.[2,10,11] This study provides Class III evidence that automated MRI analysis accurately discriminates among early-stage PD, MSA, and PSP.

Standard protocol approvals, registrations, and patient consents. The study was approved by the Ethics Committee of

Figure 1 Study flow chart



*In the excluded group of patients exceeding a disease duration of 6 years, there were 19 patients with significant white matter lesions of grade 2 and 3 (14 Parkinson disease [PD], 2 multiple system atrophy [MSA], 3 progressive supranuclear palsy [PSP]).[14] SWEDDS = scans without evidence of dopaminergic dysfunction.

Table 1 Demographic data of patients with progressive supranuclear palsy, patients with multiple system atrophy, and patients with Parkinson disease

	Training cohort			Validation cohort		
Diagnostic group	**PSP**	**MSA**	**PD**	**PSP**	**MSA**	**PD**
Sample size	20	26	26	10	14	14
Female/male, n	6/14	13/13	15/11	6/4	3/11	8/6
Age at scan, y, mean ± SD	64.6 ± 6.7	60.1 ± 8.2	65.6 ± 5.4	65.2 ± 4.9	62.7 ± 7.8	61.7 ± 7.3
Disease duration to MRI, y, mean ± SD	2.5 ± 1.6	2.6 ± 1.2	3.1 ± 2	2.7 ± 1.9	2.9 ± 1.5	3.3 ± 2.2
No. of patients with disease duration of <2/2-4/<7 y	9/8/3	7/16/3	10/7/9	5/4/1	3/8/3	5/4/5
Follow-up period, y, mean ± SD	1.9 ± 1.7	2 ± 1.9	3.2 ± 2	2.3 ± 2.1	1.9 ± 1.7	3.8 ± 2.1
UPDRS motor score, mean ± SD	29 ± 9.6[a]	31.4 ± 11[a]	20.1 ± 9.1	27.9 ± 8.7[a]	37.5 ± 12[a]	16.2 ± 7.1
Hoehn & Yahr staging, mean ± SD	2.8 ± 0.7[a]	3.1 ± 0.8[a]	2.1 ± 0.5	3 ± 0.5[a]	3.2 ± 0.8[a]	1.8 ± 0.7

Abbreviations: MSA = multiple system atrophy; PD = Parkinson disease; PSP = progressive supranuclear palsy; UPDRS = Unified Parkinson's Disease Rating Scale.
Means are listed with standard deviations (±1 SD).
[a] $p < 0.001$ vs PD.

the Medical University of Innsbruck. Patients' written informed consent was obtained according to the Declaration of Helsinki.

Procedures. ***Structural MRI acquisition and preprocessing.*** Participants were scanned on a 1.5T whole-body MRI scanner (Magnetom Avanto, Siemens Erlangen, Germany) with the help of an 8-channel head coil. High-resolution structural scans were obtained with a T1-weighted magnetization-prepared rapid gradient echo sequence. Images were processed with FreeSurfer (version 5.10) (http://surfer.nmr.mgh.harvard.edu).[6] Acquisition of MRI measurements as well as the segmentation of subcortical regions are described in e-Methods, appendix e-1 on the *Neurology*® Web site at Neurology.org (figure e-1).

Statistical analysis and classification. Demographic data are presented as frequencies and means (±1 SD). The binomial test was applied to test for the distribution of sex. To assess the differences in Hoehn & Yahr stages and UPDRS part III, Kruskal-Wallis test was used, followed by pairwise Mann-Whitney *U* test. One-way analysis of variance (ANOVA) and post hoc least significance difference were applied for the comparison of clinical data, brain region asymmetry indices, ratios of volume measures, and the intracranial volume. Group differences of volumetric measures were analyzed using multivariate analysis of covariance (MANCOVA) adjusted for age, sex, and intracranial volume.

Classification of the dataset into the diagnostic entities MSA, PSP, and PD was performed with the C4.5 decision tree algorithm implemented by Waikato Environment of Knowledge Analysis machine learning software (WEKA 3.6).[15,16] As potential discriminants, the following volumes were entered to the classifier: midbrain, pons, third and fourth ventricle, the putamen/midbrain, and the midbrain/pons ratio, as well as the more affected side of the caudate, putamen, globus pallidus, thalamus, cerebellar gray and white matter, hippocampus, and amygdala. We used a stratified 2:1 holdout procedure to split the entire dataset into a training set for establishing the classification model and a test set to measure the classification performance of the decision tree. The results were compared to the clinical diagnostic accuracy obtained by applying current consensus operational criteria at the initial contact. Both the generation of the decision tree algorithm as well as the classification performance are outlined in the e-Methods, appendix e-1.

RESULTS Patient groups and control subjects were matched for age and disease duration. Disease duration ranged from 0.1 to 6 years. Compared to the PD group, patients in the PSP and MSA groups had more severe motor disability as defined by significantly higher UPDRS motor scores and Hoehn & Yahr stage (UPDRS motor score $p < 0.001$, Hoehn & Yahr stage $p < 0.001$ for both comparisons). No significant difference was evident for sex, duration of clinical follow-up after MRI scanning, and total intracerebral volume (tables 1, e-1, and e-2). The results of MANCOVA of volumetric measures and ANOVA of asymmetry indices and ratios of volumes of brain regions are outlined in table e-1.

Classification by consensus operational criteria. The classification accuracy based on clinical consensus criteria at the initial visit of patients meeting the clinical gold standard of probable MSA and PSP at last follow-up was 62.9%. Out of a total of 40 patients eventually meeting criteria for possible or probable MSA, only 26 were correctly classified at first visit (diagnostic accuracy of 65%). Twenty-three patients with MSA were classified as probable MSA, 3 as possible MSA, 3 as unclassifiable gait disorders, and 11 as clinically uncertain parkinsonian syndromes.[17] Although at follow-up all patients with probable MSA fulfilled criteria for the Parkinson variant of MSA requiring dopaminergic treatment, 5 of them had significant cerebellar dysfunction, also allowing classification to the cerebellar variant of MSA.

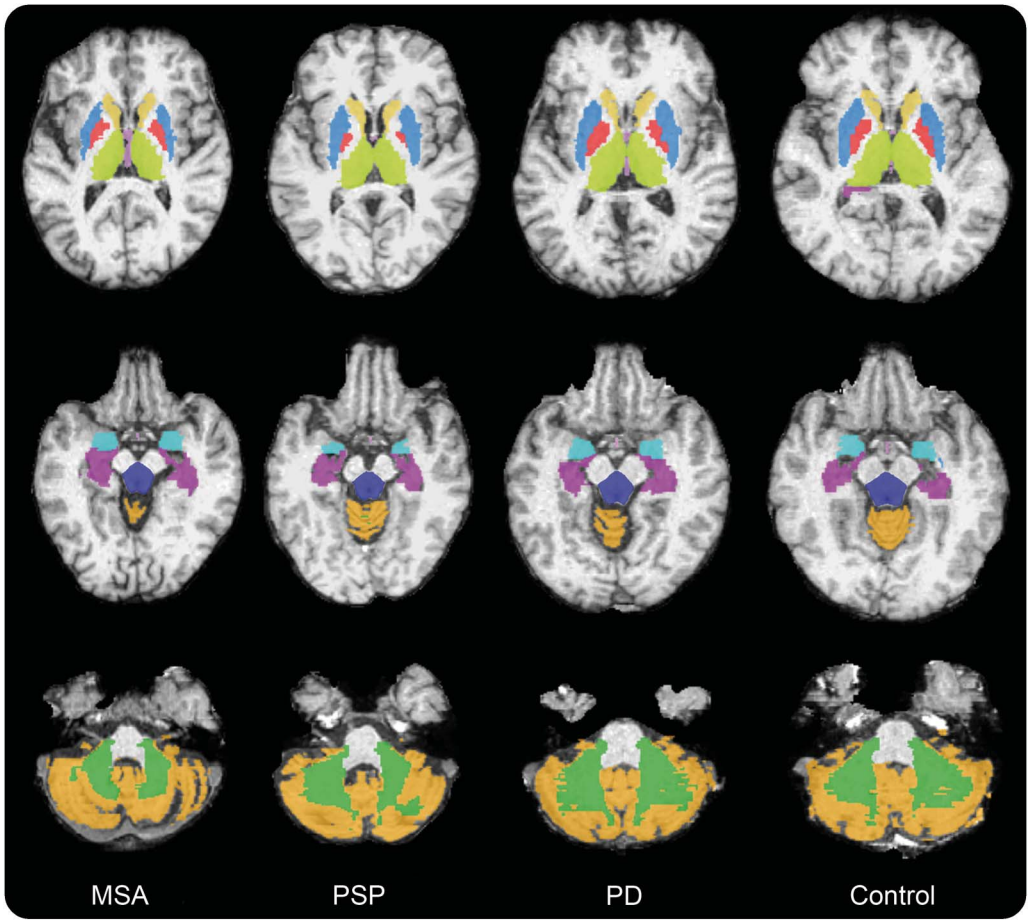
In the PSP group, 18 patients were correctly classified at the first visit. Of those, 15 patients presented

with probable PSP and 3 patients with possible PSP given a diagnostic accuracy of 60%. Out of the remaining 12 participants, 3 presented with PD and 9 patients with a clinically uncertain parkinsonian syndrome. At follow-up investigation, from a total of 30 patients with probable PSP, 20 individuals fulfilled criteria of the Richardson type and 10 of the parkinsonism type.[18] At initial visit, 26 of the 40 patients with PD fulfilled the UK Parkinson's Disease Society Brain Bank Diagnostic Criteria for PD.[10] At last follow-up, all patients with PD were classified according to those criteria. Disease duration at first visit was not significantly different between patients with MSA and patients with PSP with and without reclassification at follow-up.

Classification of volumetric MRI by the C4.5 decision tree. When applying the C4.5 decision tree algorithm to our training set, the most discriminative variables identified were the volume of the midbrain, followed by cerebellar gray matter and putamen (figures 2 and 3). When applying the model to the test set, the total classification accuracy of patients was 97%. Out of 14 patients with clinically probable MSA at follow-up, all patients were correctly classified according to the MRI scan at baseline. In the PSP group, 9 out of 10 patients were correctly classified and 1 patient was incorrectly classified to the PD group. The area under the curve (AUC) as well as sensitivity and specificity for the validation subset are given in table 2.

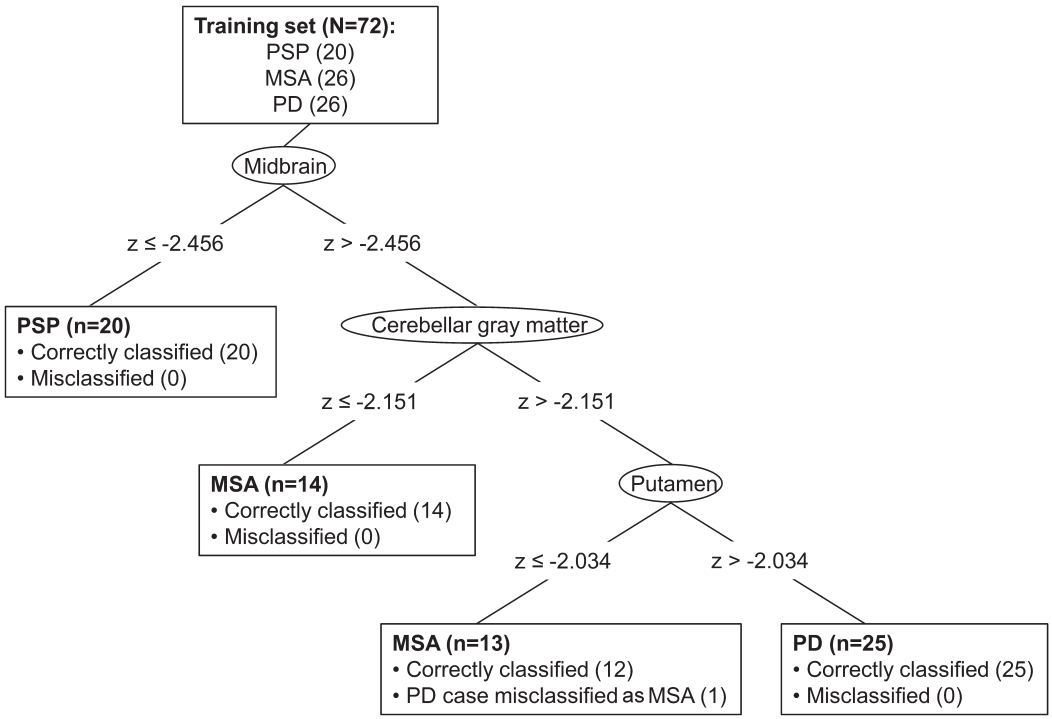
DISCUSSION In the present study, automated segmentation of subcortical brain structures obtained from volumetric T1-weighted MRI and subsequent generation of a decision tree identified the volumes of the midbrain and putamen as well as the

Figure 2 **Subcortical volume reduction in a patient with multiple system atrophy (MSA) and progressive supranuclear palsy (PSP)**



Individual T1-weighted 3D magnetization-prepared rapid gradient echo image of a patient with MSA, a patient with PSP, a patient with Parkinson disease (PD), and a healthy subject (control) and superimposed segmented volumes of the caudate (yellow), putamen (light blue), globus pallidus (red), thalamus (light green), midbrain (dark blue), cerebellar gray matter compartment (orange), cerebellar white matter compartment (green), amygdala (turquois), and hippocampus (purple). Marked volume reduction is visible in the putamen and cerebellar gray and white matter compartment of the patient with MSA and the midbrain and globus pallidus of the patient with PSP.

Figure 3 A 3-node C4.5 decision tree calculated from the training set



The numbers in each box indicate the number of correctly classified samples satisfying the conditions of each tree path followed by the number of misclassified samples. The *z* thresholds represent brain volume normalized volumetric measures processed by *z* transformation using mean centering and unit-variance scaling of the sex-adjusted healthy control cohorts. MSA = multiple system atrophy; PD = Parkinson disease; PSP = progressive supranuclear palsy.

cerebellar gray matter compartment as a highly predictive set of MRI markers yielding an accuracy of 97% for the diagnosis of MSA and PSP at initial clinical visit. By contrast, current diagnostic criteria performed poorly, with a diagnostic sensitivity of only around 40%.[2,11] However, their positive predictive value was reported as greater than 90% in one postmortem controlled series, indicating that although more than half of patients with pathologically confirmed MSA and PSP are misdiagnosed at the first clinical visit, a high proportion of patients meeting clinical diagnostic criteria will be found to harbor the pathology.[19,20] In our study, clinical diagnostic accuracy was 62.9% at the first visit and improved by 34% with the help of the MRI decision algorithm, suggesting that the use of this technique substantially improves early diagnostic accuracy. This is potentially relevant for future interventional therapies in atypical parkinsonian disorders where diagnosis of early disease stages may be critical. Similarly, high AUC values for correct classification of both the training and the validation cohorts indicated good reliability of these measures and the decision algorithm across multiple independent cohorts. This demonstrated that the decision tree is reproducible in cohorts other than that from which the training set was drawn and applicable in the clinical setting.

The segmentation of subcortical structures and subsequent generation of a decision tree identified

Table 2 Imaging classification relative to the final clinical diagnosis: Discriminative measures

	PSP vs MSA/PD	MSA vs PSP/PD	MSA vs PSP
Area under the curve (95% CI)	0.95 (0.85-1.0)	1 (NA)	NA
Sensitivity (95% CI); no. of correctly classified cases	90% (59-98); 9/10	100% (78-100); 14/14	100% (78-100); 14/14
Specificity (95% CI); no. of correctly classified cases	100% (88-100); 28/28	100% (86-100); 23/23	100% (72-100); 10/10

Abbreviations: CI = confidence interval; MSA = multiple system atrophy; PD = Parkinson disease; PSP = progressive supranuclear palsy.
All values were derived from the disease classification algorithm generated by the C4.5 decision tree model.

the midbrain volume as the brain area that classified patients with PSP from MSA and PD with the highest accuracy. The fact that other parameters like the volume of the thalamus, the volume of the third ventricle, the putamen/midbrain ratio, or the midbrain/pons ratio also showing highly significant differences in the group analysis of PSP vs MSA and PD were not included in the decision algorithm indicates the superior specificity of the volume loss in the midbrain as a surrogate marker for the diagnosis of early-stage PSP. Midbrain atrophy as identified by 3D T1-weighted MRI and FreeSurfer is consistent with neuropathologically proven atrophy and neuronal loss.[21,22] In the past 15 years, several different imaging measurements like the anterior posterior midbrain diameter, the midbrain area measured at the midsagittal view, and the ratios between midbrain and pons areas were identified as diagnostic markers with a sensitivity ranging from 63% to 100% and a specificity ranging from 86% to 100% vs PD and MSA.[23] Due to the lack of sophisticated volume measurement tools to quantify the majority of subcortical brain regions, few studies have assessed the spatial component of areas such as the midbrain, pons, cerebellum, thalamus, globus pallidus, and putamen. Applying solely the midbrain volume for the diagnosis of PSP in a group of 18 patients yielded a sensitivity of 72.2% and a specificity of 91.9%.[24] Diagnostic accuracy increased when adding the volume of the cerebellum, pons, and superior cerebellar peduncle.

In our study, one patient with the initial diagnosis of PD and the final clinical diagnosis of PSP parkinsonism at follow-up was falsely classified to the PD group. Interestingly, a second MRI scan performed 1 year later allowed the correct classification. This instance suggests that although 9 out of 10 patients with PSP parkinsonism would have been correctly classified, there is still a remaining overlap with volumetric midbrain measures of patients with PD, which is in line with reports of midbrain distance and area measures observed in PSP parkinsonism and PD.[25] Due to the small number of patients with Parkinson-type PSP, we could not enter these patients as an independent class to the generation of the decision tree. In further studies, the criteria of the decision algorithm to differentiate patients with Parkinson-type PSP and Richardson-type PSP or other PSP subtypes like primary progressive freezing of gait or the cortical predominant type will have to be extended in larger cohorts.[26]

Based on the criteria of volume loss of the more affected cerebellar gray matter compartment and putamen as well as the absence of severe midbrain atrophy, all patients with MSA of the test set were correctly classified by the decision algorithm. In addition, group analysis revealed significant volume reduction of the cerebellar white matter compartment and the pons compared to PSP and PD. As these parameters show some overlap among patient groups, they were not considered as discriminates for the classification algorithm. Cerebellar and nigrostriatal atrophy are the main sites of neuronal fallout in MSA and significant volume loss of those brain areas was reported by several investigators.[6,27–30] Most studies used categorical study designs and did not test for diagnostic accuracy as sample sizes were too small to split for training and test sets. By using an unsupervised approach, a diagnostic accuracy of 93% of semiautomated MRI volumetry was reported for the midbrain, pons, cerebellum, and superior cerebellar peduncles of patients with MSA and PSP and a disease duration of 4.5 years.[24] Interestingly, our decision algorithm stratified patients with MSA into those with predominant cerebellar atrophy and those with predominant putaminal atrophy. In this study, only patients with a parkinsonian syndrome requiring dopaminergic treatment were included. Nevertheless, 5 out of 40 patients with the Parkinson variant of MSA had significant cerebellar ataxia allowing also classification to the cerebellar variant of MSA at the second clinical consultation. Although those 5 patients with MSA were correctly classified to the subgroup of cerebellar gray matter volume loss, a substantial number of patients with the Parkinson variant of MSA were also classified to that subgroup. This finding suggests that in a proportion of patients with the Parkinson variant of MSA, cerebellar atrophy can be measured in early disease stages even prior to the onset of clinical signs related to that brain area.

Limitations. The development and validation of the classification algorithm focused on participants of an age range of 50–75 years and a maximum disease duration of 6 years. The upper age limit for this proof-of-concept study was mainly chosen to minimize error introduced by the covariance of age-related brain volume loss, but these ranges of age and disease duration nevertheless capture the majority of patients with early to moderately advanced parkinsonism presenting at movement disorder outpatient clinics. Including patients with longer disease duration would be expected to strengthen the performance of our classifier, but this has to be tested in future studies. In order to minimize associated subcortical atrophy, we excluded patients with severe white matter lesions in the basal ganglia and surrounding white matter areas with Fazekas scores of 2 or 3. In clinical practice, cerebrovascular comorbidity is not uncommon in PD or other forms of parkinsonism, but it has not been studied how commonly this induces degrees of white matter lesions with Fazekas

scores of 2 or above. Therefore, we cannot be certain whether cerebrovascular comorbidity would have changed the performance of the MRI classifier algorithm.[14,31] In subsequent studies, this classification algorithm needs to be evaluated for patients with atypical parkinsonian presentations, with young disease onset, and in their very early disease stages. In this respect, further independent validation in larger cohorts is needed to overcome potential overfitting originating from the small sample size.

So far, a single study reported on imaging features of postmortem confirmed PSP and found a MRI midbrain to pons diameter ratio smaller than 9.35 mm.[32] The lack of definite diagnostic confirmation by neuropathologic assessment is a potential limitation of the current study. Nevertheless, we stringently applied validated clinical criteria with consensus from 2 movement disorder specialists coupled with evidence for nigrostriatal dopaminergic dysfunction using tracer imaging in all patients and all patients with MSA, PSP, or L-dopa-responsive PD were followed up clinically for at least 12 months, and patients with de novo parkinsonism for 42 months. Further, secondary causes were eliminated through structural MRI and the final clinical classification was anchored on the last visit after extended follow-up. Still, we cannot entirely exclude that patients with mild disease courses were misdiagnosed based on the criteria mentioned above.

This study provides a validated, objective, and easily applicable MRI algorithm to accurately discriminate early to moderately advanced stage MSA, PSP, and PD from each other at the individual patient level using a conventional MRI sequence. Using the software tools described previously, individual volumetric T1-weighted MRI scans can be processed, with little investigator-dependent intervention, in a fairly short time. While maintaining high specificity, this image-based classification approach substantially improves the diagnostic sensitivity at first attendance of patients with idiopathic or atypical parkinsonian syndromes and short disease duration. This will be of particular advantage for adequate patient counseling, predicting disease progression and follow-up consultations, tailoring the appropriate treatment, and providing correct diagnosis for patients to be enrolled in clinical studies. Finally, by identifying patients earlier compared to clinical diagnostic criteria, our MRI tool widens the therapeutic window for future candidate neuroprotective interventions in atypical parkinsonian disorders.

AUTHOR CONTRIBUTIONS

C.S., W.P., and K.S. were responsible for the study concept and design and together with G.W. for data interpretation. C.M., N.M., G.W., and K.S. acquired the clinical data. M.S. was responsible for acquisition and processing of MRI data. C.S. did the analysis of MRI data and produced a draft of the report, which was reviewed and revised by all other authors. G.G. performed the statistical analysis of clinical and MRI data as well as the computation of the decision algorithm. C.S., W.P., and K.S. obtained funding. W.P. and K.S. supervised the study.

STUDY FUNDING

Funded by Medizinische Universität Innsbruck grant (GZ 200626 and 2007152), Jubiläumsfonds der Österreichischen Nationalbank grant (14174), and Fonds zur Förderung der wissenschaftlichen Forschung grant (KLI82-B00).

DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

Received May 15, 2015. Accepted in final form December 14, 2015.

REFERENCES

1. 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.
2. Litvan I, Agid Y, Calne D, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. Neurology 1996;47:1–9.
3. Litvan I, Goetz CG, Jankovic J, et al. What is the accuracy of the clinical diagnosis of multiple system atrophy? A clinicopathologic study. Arch Neurol 1997;54:937–944.
4. Mahlknecht P, Hotter A, Hussl A, Esterhammer R, Schocke M, Seppi K. Significance of MRI in diagnosis and differential diagnosis of Parkinson's disease. Neurodegener Dis 2010;7:300–318.
5. Messina D, Cerasa A, Condino F, et al. Patterns of brain atrophy in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. Parkinsonism Relat Disord 2011;17:172–176.
6. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci USA 2000;97:11050–11055.
7. Verma R, Zacharaki EI, Ou Y, et al. Multi-parametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images. Acad Radiol 2008;15:966–977.
8. Zacharaki EI, Wang S, Chawla S, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med 2009;62:1609–1618.
9. Kuhn M, Johnson K. Applied Predictive Modelling. New York: Springer; 2013.
10. Gibb WR, Lees AJ. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. J Neurol Neurosurg Psychiatry 1988;51:745–752.
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. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology 1967;17:427–442.
13. Fahn S, Elton RL. Members of the Unified Parkinson's Disease Rating Scale Development Committee: Unified Parkinson's Disease Rating Scale. In: Fahn S, Marsden CD, Calne DB, Goldstein M, eds. Recent Developments in Parkinson's Disease, vol. 2. Florham Park, NJ: Macmillan Healthcare Information; 1987:153–164.

14. Schmidt R, Enzinger C, Ropele S, Schmidt H, Fazekas F; Austrian Stroke Prevention study. Progression of cerebral white matter lesions: 6-year results of the Austrian Stroke Prevention study. Lancet 2003;14:2046–2048.
15. Quinlan JR. C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann, Publishers; 1993.
16. Hall M, Frank E, Holmes G, Pfahringer B, Reutermann P, Witten ICH. The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 2009; 11:10–18.
17. Catafau AM, Tolosa E; DaTSCAN Clinically Uncertain Parkinsonian Syndromes Study Group. Impact of dopamine transporter SPECT using 123I-Ioflupane on diagnosis and management of patients with clinically uncertain Parkinsonian syndromes. Mov Disord 2004;19:1175–1182.
18. Williams DR, de Silva R, Paviour DC, et al. Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. Brain 2005;128: 1247–1258.
19. Osaki Y, Ben-Shlomo Y, Lees AJ, et al. Accuracy of clinical diagnosis of progressive supranuclear palsy. Mov Disord 2004;19:181–189.
20. Osaki Y, Ben-Shlomo Y, Lees AJ, Wenning GK, Quinn NP. A validation exercise on the new consensus criteria for multiple system atrophy. Mov Disord 2009;24: 2272–2276.
21. Steele JC, Richardson JC, Olszewski J. Progressive supranuclear palsy: a heterogeneous degeneration involving the brain stem, basal ganglia and cerebellum with vertical gaze and pseudobulbar palsy, nuchal dystonia and dementia. Arch Neurol 1964;10:333–359.
22. Morris HR, Gibb G, Katzenschlager R, Wood NW, Hanger DP, Strand C. Pathological, clinical and genetic heterogeneity in progressive supranuclear palsy. Brain 2002;125:969–975.
23. Liscic RM, Srulijes K, Gröger A, Maetzler W, Berg D. Differentiation of progressive supranuclear palsy: clinical, imaging and laboratory tools. Acta Neurol Scand 2013; 127:362–370.
24. Paviour DC, Price SL, Jahanshahi M, Lees AJ, Fox NC. Regional brain volumes distinguish PSP, MSA-P, and PD: MRI-based clinico-radiological correlations. Mov Disord 2006;21:989–996.
25. Longoni G, Agosta F, Kostić VS, et al. MRI measurements of brainstem structures in patients with Richardson's syndrome, progressive supranuclear palsy-parkinsonism, and Parkinson's disease. Mov Disord 2011;26:247–255.
26. Dickson DW, Ahmed Z, Algom AA, Tsuboi Y, Josephs KA. Neuropathology of variants of progressive supranuclear palsy. Curr Opin Neurol 2010;23:394–400.
27. Fanciulli A, Wenning GW. Multiple-system atrophy. N Engl J Med 2015;372:249–263.
28. Schulz JB, Skalej M, Wedekind D, et al. Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. Ann Neurol 1999;45:65–74.
29. Ghaemi M, Hilker R, Rudolf J, Sobesky J, Heiss WD. Differentiating multiple system atrophy from Parkinson's disease: contribution of striatal and midbrain MRI volumetry and multi-tracer PET imaging. J Neurol Neurosurg Psychiatry 2002;73:517–523.
30. Shigemoto Y, Matsuda H, Kamiya K, et al. In vivo evaluation of gray and white matter volume loss in the parkinsonian variant of multiple system atrophy using SPM8 plus DARTEL for VBM. Neuroimage Clin 2013;2:491–496.
31. Appelman AP, Vincken KL, van der Graaf Y, et al; SMART Study Group. White matter lesions and lacunar infarcts are independently and differently associated with brain atrophy: the SMART-MR study. Cerebrovasc Dis 2010;29:28–35.
32. Massey LA, Jäger HR, Paviour DC, et al. The midbrain to pons ratio: a simple and specific MRI sign of progressive supranuclear palsy. Neurology 2013;80:1856–1861.