Diffusion tensor imaging in the characterization of multiple system atrophy

Aaron Michael Rulseh1–3
Jiri Keller1,4
Jan Ruzs5,6
Michael Syka1
Hana Brozova4
Robert Rusina4,7
Petra Havrankova6
Katerina Zarubova4
Hana Malikova1
Robert Jech6
Josef Vymazal1

1Department of Radiology, Na Homolce Hospital, Prague, Czech Republic; 2Department of Radiology, 1st Faculty of Medicine, General University Hospital, Charles University in Prague, Prague, Czech Republic; 3National Institute of Mental Health, Klicany, Czech Republic; 43rd Faculty of Medicine, Charles University in Prague, Prague, Czech Republic; 5Department of Circuit Theory, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic; 6Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine, Charles University in Prague, Prague, Czech Republic; 7Thomayer Hospital, Prague, Czech Republic; 8Department of Neurology, 2nd Faculty of Medicine, Charles University in Prague, Prague, Czech Republic.

Correspondence: Aaron Michael Rulseh
Department of Radiology, Na Homolce Hospital, Roengenova 2, 15830 Prague, Czech Republic
Tel +420 257 273 200
Fax +420 257 272 604
Email aarulseh@gmail.com

Purpose: Multiple system atrophy (MSA) is a rare neurodegenerative disease that remains poorly understood, and the diagnosis of MSA continues to be challenging. We endeavored to improve the diagnostic process and understanding of in vivo characteristics of MSA by diffusion tensor imaging (DTI).

Materials and methods: Twenty MSA subjects, ten parkinsonian dominant (MSA-P), ten cerebellar dominant (MSA-C), and 20 healthy volunteer subjects were recruited. Fractional anisotropy, mean diffusivity, radial diffusivity, and axial diffusivity maps were processed using tract-based spatial statistics. Diffusion data were additionally evaluated in the basal ganglia. A support vector machine was used to assess diagnostic utility, leave-one-out cross-validation in the evaluation of classification schemes, and receiver operating characteristic analyses to determine cutoff values.

Results: We detected widespread changes in the brain white matter of MSA subjects; however, no group-wise differences were found between MSA-C and MSA-P subgroups. Altered DTI metrics in the putamen and middle cerebellar peduncles were associated with a positive parkinsonian and cerebellar phenotype, respectively. Concerning clinical applicability, we achieved high classification performance on mean diffusivity data in the combined bilateral putamen and middle cerebellar peduncle (accuracy 90.3±9%, sensitivity 86.5±11%, and specificity 99.3±4%).

Conclusion: DTI in the middle cerebellar peduncle and putamen may be used in the diagnosis of MSA with a high degree of accuracy.

Keywords: multiple system atrophy, diffusion tensor imaging, magnetic resonance imaging, neuroimaging, diagnostic imaging

Introduction
Multiple system atrophy (MSA) is a sporadic progressive neurodegenerative disease characterized by autonomic dysfunction with varying degrees of parkinsonian, and/or cerebellar features. The onset of MSA is defined by the manifestation of motor dysfunction occurring on average at 55 years of age and has a mean survival of 9 years. Two forms of MSA are currently recognized: a parkinsonian form (MSA-P, formerly striatonigral degeneration) characterized by parkinsonian features poorly responsive to l-3, 4-dihydroxyphenylalanine and a cerebellar form (MSA-C, formerly olivopontocerebellar atrophy) with predominant cerebellar ataxia. Overlap between these two divisions may be evident, both clinically and histopathologically, with designation to one group determined by the predominant features at evaluation. A purely autonomic form of MSA (MSA-A, or Shy–Drager syndrome) was once considered as a possible separate entity but is not included in the current consensus criteria. MSA-A likely represents MSA-P/MSA-C at an early stage of development as autonomic dysfunction is common to both.
MSA is considered an α-synucleinopathy, and the presence of glial cytoplasmic inclusions containing α-synuclein fibrils and other proteins is currently a primary requirement for definite diagnosis. Early and widespread oligodendrocyte involvement characterized by glial cytoplasmic inclusion accumulation and demyelination has been documented.\(^2\,^3\) As these events correlate topographically and temporally with neuronal degeneration, it has been postulated that MSA is a primary oligodendrogliopathy.\(^4\) Recently, it has been shown that α-synuclein prions may play a role in the development of MSA.\(^5\) However, the pathogenesis remains incompletely understood, and the diagnosis of MSA continues to be challenging.

Qualitative radiological signs may be appreciable on standard magnetic resonance imaging sequences and are routinely used in the workup of MSA patients. Signal changes occurring in a cruciate pattern in the pons (“hot cross bun” sign) and signal changes in the lateral putamen are commonly sought.\(^6\) The hot cross bun sign is associated with the selective depletion of pontine neurons and myelinated transverse pontocerebellar fibers,\(^7\,^8\) while signal changes in the lateral putamen are associated with the accumulation of iron and gliosis.\(^9\,\,\,^10\) However, it has been shown that these changes are not particularly sensitive nor specific in establishing the diagnosis of MSA-C\(^11\,\,\,^13\) or MSA-P.\(^7\,\,\,^14\,\,\,^16\) Indeed, establishing a diagnosis of MSA in vivo is challenging, and currently only “possible” and “probable” levels of certainty are achievable without neuropathological evaluation.

Diffusion tensor imaging (DTI) has the potential to greatly enhance the radiological evaluation of MSA, considering that DTI is sensitive to changes in the white matter that may not be appreciable on standard imaging sequences. Furthermore, as MSA has been postulated to be a primary oligodendrogliopathy, changes in radial diffusivity (RD) and axial diffusivity (AD) may be especially useful in elucidating the origin of white matter changes in the disease process.\(^17\,\,\,^19\) Previous DTI studies in MSA have primarily used region of interest (ROI),\(^20\,\,\,^21\) tractography,\(^22\,\,\,^24\) and voxel-based morphometry\(^25\,\,\,^27\) approaches. Tract-based spatial statistics (TBSS) is a method that is better suited to examine global DTI characteristics in brain white matter. We therefore endeavored to perform an exploratory TBSS analysis in MSA, additionally applying ROIs in the basal ganglia to improve further characterize pathological changes and improve clinical applicability.

**Materials and methods**

Twenty MSA subjects were recruited for the present study (nine male, eleven female, 60.9 years±7 standard deviation [SD]), including ten MSA-P subjects (four male, six female, mean age 62.0 years±SD) and ten MSA-C subjects (five male, five female, mean age 59.9 years±4 SD). The clinical diagnosis of probable MSA was made according to established criteria.\(^1\) Demographic characteristics of the MSA group are presented in Table 1. Twenty healthy, age- and sex-matched subjects (nine male, eleven female, mean age 61.3 years±7 SD) were additionally included as controls. The study was approved by the ethics committee of Na Homolce Hospital, Prague, Czech Republic and all participants provided written informed consent.

Imaging was performed on a Siemens Avanto 1.5T scanner (Siemens Healthcare, Erlangen, Germany). Structural magnetic resonance images acquired using standard proton density-weighted, T1-weighted, T2-weighted, fluid-attenuated inversion recovery, and T2*-weighted sequences were used to exclude pathology in control subjects and to evaluate the presence or absence of radiological signs typically associated with MSA; signal changes in the lateral putamen (Figure 1A), and signal changes occurring in a cruciate pattern in the pons (“hot cross bun” sign; Figure 1B). Diffusion-weighted data were

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Auto*</th>
<th>Park*</th>
<th>Cereb*</th>
<th>Dur (years)</th>
<th>Dx</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>F</td>
<td>59.6</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>4</td>
<td>MSA-C</td>
</tr>
<tr>
<td>02</td>
<td>F</td>
<td>69.4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>2</td>
<td>MSA-P</td>
</tr>
<tr>
<td>03</td>
<td>F</td>
<td>63.3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>MSA-C</td>
</tr>
<tr>
<td>04</td>
<td>M</td>
<td>65.7</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>MSA-P</td>
</tr>
<tr>
<td>05</td>
<td>M</td>
<td>58.7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>5</td>
<td>MSA-C</td>
</tr>
<tr>
<td>06</td>
<td>F</td>
<td>49.7</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
<td>MSA-C</td>
</tr>
<tr>
<td>07</td>
<td>M</td>
<td>68.3</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>2</td>
<td>MSA-P</td>
</tr>
<tr>
<td>08</td>
<td>M</td>
<td>65.3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>10</td>
<td>MSA-C</td>
</tr>
<tr>
<td>09</td>
<td>F</td>
<td>80.3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>MSA-P</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>61.5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>6</td>
<td>MSA-C</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>55.0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>8</td>
<td>MSA-P</td>
</tr>
<tr>
<td>12</td>
<td>M</td>
<td>62.9</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>MSA-C</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>60.2</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>5</td>
<td>MSA-C</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>56.3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>2</td>
<td>MSA-C</td>
</tr>
<tr>
<td>15</td>
<td>M</td>
<td>59.0</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>3</td>
<td>MSA-P</td>
</tr>
<tr>
<td>16</td>
<td>M</td>
<td>55.1</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>5</td>
<td>MSA-P</td>
</tr>
<tr>
<td>17</td>
<td>F</td>
<td>53.4</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>2</td>
<td>MSA-P</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>53.6</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>1</td>
<td>MSA-P</td>
</tr>
<tr>
<td>19</td>
<td>F</td>
<td>60.2</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>6</td>
<td>MSA-P</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>61.1</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>2</td>
<td>MSA-C</td>
</tr>
</tbody>
</table>

Notes: Parkinsonism was considered positive in the presence of bradykinesia and rigidity, while cerebellar signs were documented as positive in the presence of gait ataxia, cerebellar dysartria, limb ataxia, or cerebellar ocular motor dysfunction.

Abbreviations: Auto, autonomic manifestations; Park, parkinsonian manifestations; Cereb, cerebellar manifestations; Dur, duration; Dx, diagnosis; F, female; M, male; MSA-C, multiple system atrophy, cerebellar subtype; MSA-P, multiple system atrophy, parkinsonian subtype; Park, parkinsonian manifestations.
acquired in 12 noncollinear gradient directions (b=1,100; repetition time [TR]=8,839 ms, echo time [TE]=95 ms; number of excitations [NEX]=2; 2.2 mm isotropic voxel). Diffusion data were processed offline in FSL v5 (FMRIB Software Library, Oxford, UK) and subjected to eddy current and motion correction, followed by brain extraction. Diffusion tensor fit was then used to generate fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD) maps (Table 2). Voxel-wise statistical analyses of the diffusion-weighted data were carried out using TBSS. A study-specific mean FA template was created and used in transforming all subjects to standard 1×1×1 mm³ MNI152 space. The white matter skeleton was thresholded at FA >0.4, and statistical tests were performed using 10,000 permutations. The results were thresholded using threshold-free cluster enhancement at the P<0.05 level of significance.

Automated ROI analyses were performed in template space in the basal ganglia (caudate nucleus, putamen, and globus pallidus). The ROIs were created by separating the stated masks from the Harvard-Oxford subcortical structural atlas, transforming them to study space, and finally twice eroding to minimize partial volume effects. The ROIs were then projected against the diffusion tensor data. The ROI data were further processed in R. A one-way analysis of variance test was used to test for group differences, followed by the Tukey–Kramer procedure in those regions and metrics that reached significance. An ROI covering the bilateral middle cerebellar peduncles (MCPs) was additionally extracted from the Johns Hopkins University white matter labels, transformed to study space, and then skeletonized against the mean-FA skeleton. The pROC package was used for receiver operating characteristic analyses of ROI data and leave-one-out cross-validation was performed with the boot package. A linear model was used to assess correlation between DTI metrics and age as well as disease duration in MSA patients. The Bonferroni correction for multiple comparisons was applied to maintain an alpha-level of 0.05 (all significant P-values reported are corrected).

To further evaluate clinical applicability, ROI data were analyzed using a support vector machine in the Matlab v7.13 environment (MathWorks, Natick, MA, USA). To find the best combination of DTI metrics for group classification, we first constructed a feature vector consisting of all measurements for all participants (20 MSA subjects and 20 controls), and subsequently performed an exhaustive search across all possible combinations. A classifier with a Gaussian radial basis kernel was chosen as it enables a smooth, curved decision boundary. The choice of optimal parameters was determined by a grid search over a range of values, and were found to be C=2, p=2. To validate the reproducibility of the method, a cross-validation scheme was used. The original data across 40 subjects were randomly separated into two subsets: a training subset composed of 60% of the data (24 subjects) and a testing subset composed of 40% of the data (16 subjects). This process was repeated 20 times for each combination. The final performance of the model was calculated as the average percentage of subjects correctly classified (MSA vs control) over all 20 cycles.

### Results

Two radiologists experienced in the routine clinical evaluation of MSA and other neurodegenerative diseases (JK, MS)
blindly evaluated the qualitative MSA and control images in the present study. The raters achieved a mean sensitivity of 67.5% and mean specificity of 92.5%. The interrater reliability for the two raters was moderate (Cohen’s kappa =0.68, P<0.001).

In examining differences between all MSA and control subjects reported by TBSS, we observed widespread differences in RD and MD, with similar patterns in FA and AD but in fewer voxels (Figure 2). RD and MD were increased and FA reduced in the superior, middle, and inferior cerebellar peduncles bilaterally (Figure 2A–C), with a smaller volume of increased AD in the bilateral middle and inferior cerebellar peduncles (Figure 2D). No differences in AD between patients and controls were observed in the superior cerebellar peduncles. Changes in RD and MD were also more widespread supratentorially, being increased throughout the central portions of the corona radiata and in the left external capsule (Figure 2D). These voxels generally demonstrated increased RD and MD, but no changes in FA.

We further explored differences between controls and the MSA-P and MSA-C diagnostic subgroups by TBSS. In MSA-P, only supratentorial differences were observed in comparison with controls, which followed a pattern similar to all MSA patients versus controls but in fewer voxels. In MSA-C, differences were also similar to those detected between all MSA subjects and controls, with robust changes detected infratentorially and a similar pattern supratentorially but in fewer voxels. No differences were detected between the MSA-C and MSA-P diagnostic subgroups.

Considering that no differences between the MSA-C and MSA-P diagnostic subgroups were detected by TBSS, we elected to additionally evaluate DTI metrics in the basal ganglia (caudate nucleus, putamen, and globus pallidus) of MSA and control subjects. As no differences in laterality were detected in the basal ganglia, we used the mean of right and left measurements in all further analyses. We then explored group differences between MSA-C, MSA-P, and controls by analysis of variance, followed by the Tukey–Kramer procedure in those regions and metrics that reached significance. We detected that MSA-P significantly differed from other groups (MSA-C, controls) in the putamen only, in AD, MD, and RD. No differences in FA in the putamen were detected (Table 3). No correlation between DTI metrics and age or disease duration was observed in the putamen or the MCP of MSA patients (Table 4).

Support vector machine analysis of MD data from ROIs covering the bilateral putamen and MCP reached a classification performance of 90.3%±9% (sensitivity 86.5%±11%, specificity 99.3%±4%) in discriminating MSA patients from healthy participants (Figure 3). Considering MD in the

Table 3 DTI results in bilateral putamen

<table>
<thead>
<tr>
<th>Metric</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>F(2,37)=16.649</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MD</td>
<td>F(2,37)=15.028</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RD</td>
<td>F(2,37)=13.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FA</td>
<td>F(2,37)=3.5505</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Notes: Results of ANOVA analyses of ROI data in the bilateral putamen in subjects with MSA-C, MSA-P, and controls.

Abbreviations: AD, axial diffusivity; ANOVA, analysis of variance; FA, fractional anisotropy; DTI, diffusion tensor imaging; MD, mean diffusivity; RD, radial diffusivity; MSA-C, multiple system atrophy, cerebellar subtype; MSA-P, multiple system atrophy, parkinsonian subtype; ROI, region of interest.
regions separately, we achieved an accuracy of 79.7%±9% (sensitivity 72.5%±9%, specificity 99.3%±3%) in the MCP and an accuracy of 78.4%±10% (sensitivity 74.6%±11%, specificity 89.6%±12%) in the bilateral putamen.

Leave-one-out cross-validation was additionally performed in the bilateral putamen on MD data to evaluate different classifications: between all MSA subjects and controls, between MSA-P and non-MSA-P, and for a phenotypical classification between parkinsonian-positive subjects and parkinsonian-negative subjects. The data were reclassified in an attempt to improve diagnostic utility, as we expected that subjects with parkinsonian manifestations would likely have detectable changes in MD in the putamen, while those without parkinsonian manifestations would likely not. We observed a prediction error of 15.2% in classifying subjects as patient or control, 13% as MSA-P or non-MSA-P, and 11.6% in classifying subjects as parkinsonian positive or parkinsonian negative. A similar classification experiment was performed in the MCP, resulting in a prediction error of 15.5% in classifying subjects as patient or control, 10.6% as MSA-C or non-MSA-C, and 10.1% in classifying subjects as cerebellar positive or cerebellar negative. Receiver operating characteristic analyses in the bilateral putamen and MCP were performed to identify cutoff values maximizing specificity and were determined as MD >20% of the mean normative value in the MCP and >6% in the putamen.

Discussion

In the present study, we investigated DTI characteristics of brain gray and white matter in MSA subjects in vivo. Our TBSS results showed widespread differences in the white matter of MSA subjects in comparison to controls, reflected primarily in increased RD and MD (Figure 2). While MD is known to be a sensitive but nonspecific marker of morphological changes in the white matter, RD has been shown to correlate with myelin integrity.17,18 Our findings therefore lend support to the hypothesis that MSA is a primary oligodendrogliopathy.

In examining DTI differences between MSA subgroups (MSA-P and MSA-C) and controls, we observed similar differences between MSA-C subjects and controls as in the pooled analysis, while only supratentorial differences were detected between MSA-P subjects and controls. As with all MSA subjects versus controls, the MSA-C and MSA-P subgroups showed the greatest differences in RD and MD. These observations are generally in agreement with those reported previously.20,22,25,36 Approximately half of the patient participants in our study exhibited clinically both cerebellar and parkinsonian manifestations, in agreement with previous studies.37,38 As we did not observe any differences between the MSA-C and MSA-P diagnostic subgroups by TBSS, we hypothesized that this large degree of clinical overlap was likely related to this negative result. To further explore potential differences between the MSA-C and MSA-P subgroups, we additionally evaluated DTI characteristics in the basal ganglia. We detected significant differences between the MSA-P group and other groups (MSA-C, controls) in RD, MD, and AD in the putamen only, with no differences observed in FA.

To evaluate clinical applicability of our findings, support vector machine analysis was performed on MD data extracted from ROIs covering the bilateral putamen and MCP. The selection of MD was made primarily due to the robust results in both regions, as well as the consideration of general accessibility. We reached the highest classification performance
in both regions combined (accuracy 90.3%±9%, Figure 3), indicating that both regions should ideally be evaluated in the workup of MSA subjects. Thus, objective DTI evaluation in these two regions was superior to more traditional subjective evaluation.

We hypothesized that differences in the putamen may be more representative of a parkinsonian phenotype, that is, subjects with clinically evident parkinsonian manifestations are likely to have detectable changes in the putamen. We hypothesized that this more common sense approach would perform better than classification by the traditional subgroups (MSA-C, MSA-P), as half of the MSA subjects had both clinically evident cerebellar and parkinsonian manifestations. To test this possibility, leave-one-out cross-validation was performed on MD data in the putamen. Cross-validation results support this hypothesis, as the prediction error improved from 15.2% in all MSA subjects versus controls, to 13% as MSA-P or non-MSA-P, and 11.6% when classifying subjects by phenotype. We performed a similar analysis in the MCP, and found that a phenotypic classification also performed better than traditional classification, with prediction error improving from 15.5% in classifying subjects as patient or control, 10.6% as MSA-C or non-MSA-C, and 10.1% phenotypically. Thus, the typical designations of MSA-C/MSA-P provide only limited information as, for example, approximately half of MSA-C subjects have clinically evident parkinsonian manifestations and likely have corresponding changes in the putamen detectable by DTI (Figure 1).

Our results suggest that the cerebellar and parkinsonian components of MSA may be separate, but overlapping processes. In a large series of 100 definite MSA cases, Ozawa et al38 reported that pathological changes in striatonigral degeneration likely begin in the substantia nigra, where neuronal cell death appears to occur independent of glial cytoplasmic inclusion accumulation, and concluded that other factors likely contribute to this phenotype. The authors did not observe any cases of pure striatonigral or olivopontocerebellar pathology, which may be related to the postmortem nature of the study; that is, at the terminus of the disease process it is likely that both striatonigral and olivopontocerebellar systems are affected.

In the diagnostic setting, magnetic resonance imaging is performed to assist in the confirmation or exclusion of MSA; therefore, we aimed to maximize specificity by determining diagnostic thresholds using receiver operating characteristic analyses. We found that MD >20% of the mean normative value in the MCP and >6% of the mean normative value in the putamen maximized specificity (100%). However, thresholds and specific cutoff values may be influenced by a number of factors39,40 and should be verified on a per-site basis. As one of the most challenging scenarios in the workup of MSA is differentiating MSA from Parkinson disease, further studies are needed to evaluate the performance of the presented methodology in this setting.

Conclusion
In conclusion, we detected widespread changes in the brain white matter and putamen of MSA subjects. We achieved high classification performance on MD data in both the bilateral putamen and MCP (accuracy 90.3%±9%, sensitivity 86.5%±11%, specificity 99.3%±4%). Finally, we determined that MD >20% of the mean normative value in the MCP and >6% of the mean normative value in the putamen achieved specificity of 100%.

Acknowledgments
This work was supported by the Ministry of Health, Czech Republic – conceptual development of research organization (Nemocnice Na Homolce – NNH, 00023884; IG 154303) and by the grant “Cognitive Predictors of Neurodegeneration from the Czech Science Foundation”, under grant number 16-01781S.

Disclosure
The authors report no conflicts of interest in this work.

References


