Automated Volumes-of-Interest Identification for Classical and Atypical Parkinsonian Syndrome Differentiation Using T2’ MR Imaging

N. D. Forkert¹; A. Schmidt-Richberg²; A. Treszl³; C. Hilgetag¹; J. Fiehler⁴; A. Münchau³; H. Handels²; K. Boelmans⁶

¹Department of Computational Neuroscience, University Medical Center Hamburg-Eppendorf, Germany; ²Institute of Medical Informatics, University of Lübeck, Germany; ³Department of Medical Biometry and Epidemiology, University Medical Center Hamburg-Eppendorf, Germany; ⁴Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Germany; ⁵Department of Neurology, University Medical Center Hamburg-Eppendorf, Germany; ⁶Department of Psychiatry, University Medical Center Hamburg-Eppendorf, Germany

Keywords
Magnetic resonance imaging, T2’ image sequences, computer-assisted image analysis, Parkinsonian syndrome

Summary
Objectives: In clinical routine, patients with classical Parkinsonian syndromes (CPS) need to be differentiated from those with atypical Parkinsonian syndromes (APS), particularly with respect to prognosis and treatment decision. To date, this diagnosis is mainly based on clinical criteria, leading to failure rates up to 25%, motivating the development of image-based decision support systems. Magnetic resonance imaging (MRI) and in particular T2’ image sequences have been suggested as a potential marker for differential diagnosis. The aim of this study was to investigate whether automatically identified T2’ volumes-of-interest (VOIs) can be used for an automatic differentiation of CPS and APS patients.

Material and Methods: 74 MRI datasets were available for this hypothesis generating trial, including image sequences from 24 healthy subjects, 33 CPS and 17 APS patients. First, a problem-specific reference atlas was generated using the healthy control datasets. Next, patients’ datasets were registered to the atlas. Voxel-wise t-tests, reflecting significance levels of T2’ value differences between CPS and APS patients, were then applied for calculation of a p-map. Finally, the calculated p-map was thresholded and a connected component analysis was performed for final VOI detection. In parallel, manually defined VOIs were determined in grey and white matter for comparison.

Results: Three VOIs in parts of the basal ganglia and the left occipital lobe were automatically identified by the presented method. There was a trend for higher area under the curve on multivariable receiver operating characteristic curves for automatically determined VOIs over manually defined VOIs (0.939 vs. 0.818, p = 0.0572).

Conclusion: The diagnostic role of automatically defined VOIs in differentiation of CPS and APS patients based on T2’ image sequences should be further investigated.

1. Introduction

Parkinson’s disease is one of the most common neurologic diseases at higher ages, with a reported incidence rate of 13.4 per 100,000 [1]. The classical phenotype is characterized by core symptoms such as an asymmetric onset of bradykinesia, rigidity and tremor, and postural instability in later stages of the disease. In the clinical routine, patients with the classical Parkinsonian syndrome (CPS) need to be differentiated from those with atypical Parkinsonian syndromes (APS), such as progressive supranuclear palsy (PSP) or multiple system atrophy (MSA).

To date, the diagnosis of Parkinsonian syndromes is mainly based on clinical criteria, as disease-specific biomarkers are still not used or available in the clinical routine. However, atypical Parkinsonian patients often exhibit similar symptoms compared to CPS, particularly in early disease stages, limiting sensitivity and speci-
ficiency of the diagnosis if solely based on clinical grounds. Clinico-pathological studies have demonstrated failure rates up to 25% for this differentiation in clinical practice [2].

However, making a correct diagnosis, particularly in the early disease stages, is becoming highly relevant due to potential disease-modifying treatment strategies that crucially depend on an accurate diagnosis [3, 4]. Therefore, additional diagnostic tools including standardized neuropsychological assessment, electro-oculography or assessment of postural stability have been employed to improve clinical decision making, with varying results.

To further enhance diagnostic accuracy, automated image-based decision support has been suggested with promising first results [5–9]. In contrast to manually defined brain regions, automated techniques map the morphological and/or metabolic patterns across the entire brain or selected volumes-of-interest (VOIs) and, thus, do not require a subjective rater judgment.

The following section describes selected, previously presented methods for the image-based differentiation of CPS and APS patients.

2. State-of-the-art

The automatic differentiation of patients with Parkinsonian syndromes is an important step towards optimal diagnosis and treatment decisions. Several image-based classification approaches have been presented to improve diagnostic accuracy. Previous approaches can be separated tentatively into morphological and metabolic classification schemes.

Morphological classification schemes are predominantly based on patterns of regional brain atrophy typical for different syndromes, for instance, dorso-rostral midbrain and superior cerebellar peduncle atrophy in PSP patients. Such observations were, for example, used by Duchesne et al. [6], who performed an analysis of T1-weighted MR images from which the deformation information in the midbrain region was extracted using a non-linear registration to a reference atlas. These results were used to train and evaluate a support vector machine classification, which achieved 91% classification accuracy for distinction of Parkinsonian syndromes. A major drawback of this approach was the usage of high-resolution T1-weighted MR image sequences, which are normally associated with long acquisition times.

Metabolic classification approaches make use of the fact that CPS and APS patients exhibit different metabolic patterns within the brain. Tang et al. [5], for example, employed positron emission tomography (PET) imaging for an analysis of metabolic patterns and following automatic differentiation of Parkinsonian syndromes. Using voxel-based spatial covariance mapping, a high specificity (> 90%) for the distinction between Parkinsonian disorders was achieved. However, PET imaging is associated with ionizing radiation and is only available in specialized centers, making its application difficult in routine clinical practice.

Previously, histo-pathological studies have revealed differences in brain iron accumulation and distribution between CPS and APS patients [10]. Therefore, the regional brain iron content and distribution, which is associated with altered cell metabolism to a variable extent, might be suitable to serve as a biomarker. Paramagnetic substances in the brain, such as non-heme iron (e.g. ferritin and hemosiderin), create local magnetic field inhomogeneities producing intra-voxel dephasing and shortened transverse relaxation times. Therefore, an estimation of tissue iron can be obtained from T2-weighted image sequences using magnetic resonance imaging (MRI).

In general, the measured T2 relaxation time is always shorter than what can be theoretically explained by the spin-spin relaxation. This observation can be ascribed to magnetic field inhomogeneities and different tissue properties such as diffusion [7]. The time effectively measured for the transversal relaxation is commonly referred to as the T2* relaxation time. The temporal difference between the T2 and T2* relaxation times, which is referred to as T2’, represents a pure measure of the changes that are mainly caused by iron-induced local field inhomogeneities. Therefore, T2’ changes that are caused by local susceptibility variations are particularly sensitive to tissue iron stores.

T2’ imaging has been employed in a few previous studies of Parkinsonian syndromes. For example, Graham et al. [7] used a T2-weighted PRIME (partially refocused interleaved multiple echo sequence) MR image sequence for an investigation of iron deposits in the basal ganglia in patients with CPS as well as healthy controls. This analysis revealed shortened relaxation rates in the substantia nigra and caudal putamen in patients with CPS compared to healthy subjects. Boelmans et al. [9] employed T2’ imaging to investigate differences in regional brain iron content between CPS and PSP patients. Significant differences of T2’ values between the two groups were found in the caudate nucleus, globus pallidus, and putamen. Apart from these findings, T2’ imaging has also been used for automatic classification of patients with Parkinsonian syndromes [8]. In this approach, three different T2’ atlases were generated: one for healthy controls, one for patients with CPS, and one for patients with PSP. The classification decision was then obtained by determining the T2’ atlas with the minimal mean squared-difference of the T2’ values within the brain between a registered patient dataset and each atlas.

One constraint of the two studies by Graham et al. [7] and Boelmans et al. [9] is that exclusively manually predefined VOIs were used, such that potentially interesting brain structures might have been missed. Furthermore, the results depended to a large extent on the shape and quality of the manually defined VOIs. In contrast to these approaches, only a global measure was used for the automatic classification described in [8]. Thus, focusing on relevant VOIs may aid to further improve the classification accuracy.

Consequently, the aim of the present study was to automatically identify bias-free VOIs that can be used for an automatic differentiation of CPS and APS patients using image sequences.
3. Material and Methods

3.1 Patients and MR Protocol

Overall, 74 MRI datasets were available for automatic identification of VOIs. These 74 datasets included 24 healthy control subjects (62.8 ± 9.9, 41.9 – 77.6; mean age ± SD, range), 33 CPS (63.4 ± 9.6, 41.3 – 79.9) and 17 APS patients (67.7 ± 5.9, 55.4 to 76.8). Three different atypical Parkinsonian syndromes were included in the latter group: MSA (n = 3), PSP (n = 11), and corticobasal syndrome (CBS) (n = 3).

Clinical diagnosis of CPS [11], MSA [12], PSP and CBS [13] was made by movement disorder specialists (K.B., A.M.) according to established consensus criteria. All MR measurements were performed on a 1.5T Avanto scanner (Siemens, Erlangen, Germany). T2´ datasets cannot be measured directly but have to be calculated from quantitative T2 and T2* datasets respectively, which in turn have to be calculated from multi-echo image sequences (▶Figure 1). For T2 determination, a triple-echo sequence using echo times (TE) of 12, 84, and 156 msec was used. Likewise, the T2* weighted images were acquired with an echo-planar imaging sequence with a TE of 20, 52 and 88 msec.

Both the T2 and T2* sequences provided a matrix of 128 × 128 × 24 voxels with a spacing of 1.9 × 1.9 × 5 mm³ (▶Figure 1).

After triple-echo image sequence acquisition, a quantitative T2 map (qT2) was calculated by voxel-wise fitting the exponential function $S(t) = S_0 \exp(-t/T2)$ to the signal intensity decay curve $S(t)$ given by the multiple TE data of the T2-weighted sequence. In analogy to the procedure for qT2 map generation, a quantitative T2* map (qT2*) was calculated using the multiple TE data of the T2* sequence. Finally, the T2´ dataset was calculated from these qT2 and qT2* maps voxel-wisely by the following relationship: $1/T2' = 1/qT2* - 1/qT2$ (▶Figure 1).

3.2 Atlas Generation

The unbiased identification of T2´ VOIs on a voxel-wise level requires a previous registration of all patient datasets to an atlas to achieve a voxel correspondence between the datasets. Various brain atlases have been previously generated for such a purpose in the past. However, these atlases are commonly generated from high-resolution T1- or T2-weighted MR datasets from a young and healthy cohort and may therefore only allow suboptimal registration results in the approach described in this work. Therefore, a case-appropriate brain atlas was generated in this work as a first step, using the 24 datasets from the healthy control subjects matched for age and gender to the group of Parkinsonian patients. Apart from addressing the age issue, this procedure may also lead to improved registration results in the following processing steps, as it assures that the intensities of the atlas and those from a patient dataset are directly comparable.

Due to the fact that the T2´ images are fairly noisy and contain metabolic rather than anatomical information (▶Figure 1), a direct registration would be error-prone. Therefore, the images from the T2 triple echo sequences with the longest TE were used for this purpose, as these images exhibit the best tissue contrast.

The process of atlas generation can be summarized as follows: First, the 24 images $I_j$ of the training set, containing the datasets of the healthy control subjects, ($j = 1, \ldots, 24$) are registered to one arbitrarily chosen reference image, yielding the transformation $\varphi_j$. The images are then transformed to the reference frame by calculating $I_j \circ \varphi_j$ and the mean image $\overline{I_j}$ is com-
However, this procedure introduces a bias due to the choice of reference image. To overcome this drawback, the iterative scheme proposed by Guimond et al. [14] was applied in this work for improved atlas generation results. For this approach, the mean transformation $\varphi^0$ and its inverse are computed. In the next iteration cycle, the registration of all images is repeated with $I^0 \circ (\varphi^0)^{-1}$ as reference image. This process is iterated until the mean transformation converges against the identity and a mean intensity atlas $I^k$ is obtained.

The procedure of Guimond et al. [14] was extended in this work using a diffeomorphic registration approach. Here, transformations are restricted to the group of diffeomorphisms, which means that they are always invertible and differentiable. This allows an intrinsic and very efficient computation of the inverse transformation as required in the atlas generation process. Moreover, statistics such as the computation of the mean transformation are mathematically well-defined on the group of diffeomorphisms [15]. In this work, the diffeomorphic non-linear registration approach as proposed in [16] was applied, which combines an efficient diffusion-based regularization with normalized forces derived from the sum of squared differences distance measure. The final atlas is illustrated in Figure 2.

### 3.3 Volumes-of-Interest Identification

After generation of the problem-specific reference brain atlas, an identification of T2’ volumes-of-interest became possible. Therefore, the image from the T2 triple echo sequence acquired with the longest TE was extracted from each patient dataset and used for calculation of the required transformation field for registration to the generated atlas. The registration of each patient dataset to the atlas was also performed using the aforementioned diffeomorphic diffusion registration approach [16]. Due to the fact that the T2’ dataset is calculated directly from the two triple-echo sequences, the computed transformation field can be used to adapt the corresponding T2’ directly to the atlas without any further conversions.

After registration of all T2’ datasets to the atlas, a voxel-wise statistical analysis was performed to identify T2’ VOIs. Therefore, the T2’ maps of each patient were first preprocessed with a median filter using a kernel of $3 \times 3 \times 1$ voxels to reduce high frequency noise artifacts within the T2’ maps. In a following step, the T2’ values were extracted for each patient and voxel. For each voxel, a two-sided t-test was performed to determine significant differences between the two T2’ value distributions of the 33 CPS and 17 APS patients. In doing so, a p-value map was generated representing the significance level at each voxel of the atlas. The calculated p-value map was then masked using a segmentation of the brain tissue (Figure 3). This brain segmentation was extracted from the calculated reference atlas using a graph-based segmentation approach described in [17], which was modified specifically for this purpose. Global thresholding with an upper threshold of $p < 0.0005$ was then applied to obtain the required voxel candidates for the final VOIs. This threshold was selected to correct for multiplicity and at the same time not to lose much power. Finally, a connected component analysis was performed to identify connected voxels belonging to one VOI. More precisely, this analysis employed a 26-neighbourhood. Connected components with fewer than 25 voxels were excluded from further analyses to ensure a proper average calculation for each VOI for the automatic differentiation of CPS and APS patients in the latter (Figure 4).

In addition to the automatically defined VOIs, manually defined VOIs were determined within the reference atlas in grey and white matter brain tissues as follows:

![Figure 2](image-url) Selected slice from the calculated reference atlas and corresponding slices from selected healthy control subjects used for atlas generation.
Caudate nucleus, thalamus, putamen (anterior and posterior part), pallidum (anterior and posterior part), occipital and frontal white matter and underneath motor cortex white matter (▶ Figure 5). All of these VOIs were separately defined in each hemisphere, resulting in a total of 18 VOIs. The selection of these VOIs was in concordance with manually defined VOIs used in previous studies [9].

3.4 Evaluation and Statistical Analysis

Statistical analysis was performed to quantify a potential benefit of the automatically compared to manually defined VOIs. Therefore, the automatically as well as the manually defined VOIs were used to extract the average T2’ signal for each VOI and patient. Receiver operating characteristic (ROC) analysis was performed for statistical evaluation of the automatically and manually defined VOIs. More precisely, area-under-the-curve (AUC) values with 95% confidence intervals were calculated for the ROC curves. AUCs were compared non-parametrically according to DeLong [18]. Statistical analysis was performed using SAS 9.2 (SAS Institute, Cary, NC, USA).
4. Results

Three different VOIs were automatically identified by the presented method. Two of these VOIs are located nearly symmetrically in the left and right hemisphere in parts of the basal ganglia, in particular in the posterior parts of the pallidum and putamen. The third identified VOI is located in the left lingual gyrus of the occipital lobe (Figure 4). The T2’ values within each of these VOIs are generally lower in APS patients in comparison to CPS patients (Table 1).

The ROC curve analyses revealed AUC values between 0.7763 and 0.8859 for the automatically identified VOIs (Table 1). It can be seen that the ROC AUC values are quite similar for the nearly symmetrically located basal ganglia VOIs (0.8859 and 0.8788), while the third VOI in the left lingual gyrus exhibits the lowest AUC value (0.7763).

For the manually identified VOIs, the results of the ROC analyses are given in Table 2. Here, the ROC AUC values are significantly lower compared to the automatically detected VOIs (p < 0.0001; two-sided t-test). More precisely, the AUC values are between 0.4848 and 0.7219 for the manually defined VOIs. The highest AUC values were achieved for the left (0.6881) and right (0.7219) posterior pallidum, while the lowest AUC values were achieved for the white matter VOIs.

The multiparametric analysis was performed to compare the automatically and manually defined VOIs within one combined analysis also using the ROC AUC metric. Moreover, only the right thalamus, the left anterior pallidum and the left posterior pallidum of the manually defined VOIs were selected using a statistical backward selection and included in this multiparametric ROC analysis. The results of this evaluation are given in Table 3. Overall, the results revealed no significant difference between the automatically identified and manually selected VOIs (p = 0.0572). However, there was a trend for higher area under the curve on multivariable receiver operating characteristic curves for automatically determined VOIs over manually defined VOIs (0.939 vs. 0.818), which corresponds to a 15% improvement (Figure 6).

Due to the fact that this is a hypothesis generating study, the necessary sample size for an adequately powered future study was calculated in a following step. Assuming a

![Figure 5](image-url) Illustration of the manually defined VOIs in each hemisphere. Top row: occipital white matter (left), frontal white matter (centre), and underneath motor cortex white matter (right). Middle row: Pallidum anterior (left), pallidum posterior (centre), thalamus (right). Lower row: Putamen anterior (left), putamen posterior (centre) and caudate nucleus (right)

<table>
<thead>
<tr>
<th>VOI 1</th>
<th>Anatomical structure [hemisphere]</th>
<th>Mean (± STD) T2’ in CPS</th>
<th>Mean (± STD) T2’ in APS</th>
<th>p-value</th>
<th>ROC AUC</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOI 1</td>
<td>putamen and pallidum (posterior parts)(right)</td>
<td>114.36 (± 14.01)</td>
<td>90.88 (± 14.18)</td>
<td>&lt; 0.0001</td>
<td>0.8859</td>
<td>0.7888</td>
</tr>
<tr>
<td>VOI 2</td>
<td>putamen and pallidum (posterior parts)(left)</td>
<td>121.84 (± 14.94)</td>
<td>98.69 (± 12.59)</td>
<td>&lt; 0.0001</td>
<td>0.8788</td>
<td>0.7862</td>
</tr>
<tr>
<td>VOI 3</td>
<td>lingual gyrus of the occipital lobe (left)</td>
<td>180.68 (± 34.53)</td>
<td>145.30 (± 23.86)</td>
<td>&lt; 0.0001</td>
<td>0.7763</td>
<td>0.6327</td>
</tr>
</tbody>
</table>
Table 3 Comparison of the manually and automatically determined VOIs using a multiparametric statistical analysis (p = 0.0572; *right thalamus, left anterior pallidum, left posterior pallidum)

<table>
<thead>
<tr>
<th>VOI</th>
<th>ROC AUC</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left underneath motor cortex</td>
<td>0.4848</td>
<td>0.3077 0.6620</td>
</tr>
<tr>
<td>white matter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right underneath motor cortex</td>
<td>0.6524</td>
<td>0.4992 0.8056</td>
</tr>
<tr>
<td>white matter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left frontal white matter</td>
<td>0.6506</td>
<td>0.4858 0.8155</td>
</tr>
<tr>
<td>Right frontal white matter</td>
<td>0.5561</td>
<td>0.3714 0.7409</td>
</tr>
<tr>
<td>Left occipital white matter</td>
<td>0.5410</td>
<td>0.3444 0.7376</td>
</tr>
<tr>
<td>Right occipital white matter</td>
<td>0.4973</td>
<td>0.3267 0.6680</td>
</tr>
<tr>
<td>Left caudate nucleus</td>
<td>0.5437</td>
<td>0.3658 0.7215</td>
</tr>
<tr>
<td>Right caudate nucleus</td>
<td>0.6078</td>
<td>0.4345 0.7812</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>0.6684</td>
<td>0.5125 0.8244</td>
</tr>
<tr>
<td>Right thalamus*</td>
<td>0.6453</td>
<td>0.4878 0.8027</td>
</tr>
<tr>
<td>Left anterior pallidum*</td>
<td>0.5312</td>
<td>0.3532 0.7092</td>
</tr>
<tr>
<td>Right anterior pallidum</td>
<td>0.6159</td>
<td>0.4449 0.7869</td>
</tr>
<tr>
<td>Left posterior pallidum*</td>
<td>0.6881</td>
<td>0.5333 0.8428</td>
</tr>
<tr>
<td>Right posterior pallidum</td>
<td>0.7219</td>
<td>0.5653 0.8786</td>
</tr>
<tr>
<td>Left anterior putamen</td>
<td>0.5472</td>
<td>0.3613 0.7332</td>
</tr>
<tr>
<td>Right anterior putamen</td>
<td>0.5276</td>
<td>0.3530 0.7023</td>
</tr>
<tr>
<td>Left posterior putamen</td>
<td>0.5588</td>
<td>0.3887 0.7290</td>
</tr>
<tr>
<td>Right posterior putamen</td>
<td>0.6132</td>
<td>0.4341 0.7923</td>
</tr>
</tbody>
</table>

Table 2 Results of the statistical evaluation of the automatically VOIs (* = VOIs identified by the backward selection are highlighted in italic)

<table>
<thead>
<tr>
<th>VOI</th>
<th>ROC AUC</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manually determined VOIs with</td>
<td>0.8182</td>
<td>0.7008 0.9356</td>
</tr>
<tr>
<td>backward selection*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatically determined VOIs</td>
<td>0.9394</td>
<td>0.8754 1.0000</td>
</tr>
</tbody>
</table>

CPS/APC ratio of 10, a sample of 390 CPS and 39 APS patients would achieve 80% power to detect a difference of 0.1212 between two diagnostic tests ROC AUCs of 0.8182 and 0.9394 using a two-sided test at a significance level of 0.05. Both the positive and negative correlations between the two responses of these tests are assumed to be 0.5 [19].

5. Discussion

The main finding of this study is that automatically identified VOIs in parts of the basal ganglia and the left lingual gyrus are promising candidates for an improved differentiation between CPS and APS patients compared to manually defined VOIs based on T2’ image sequences.

One of the main advantages of the proposed method is that T2’ MR imaging is comparably fast and not associated with an exposure to ionizing radiation. The acquisition for the image sequences used in this study requires 107 seconds and is available on nearly all MR scanners, making its use in the clinical routine more realistic.

The results of the present evaluation show that T2’ image sequences may be valuable for an automatic differentiation of patients with Parkinsonian syndromes, which is in concordance with earlier findings [7, 9, 10]. Moreover, the results of this study also suggest that the automatically identified VOIs may improve the differentiation of CPS and APS patients compared to manually defined VOIs. One reason for this result may be that the manual VOIs depend to a large extent on raters’ expertise. However, manual definition may be biased by previous findings, such as histopathological or metabolic studies, for which a direct transfer to these image sequences may not represent an optimal solution.

It needs to be pointed out that the evaluation of automatically determined VOIs in this study has only been performed in a statistical manner, while no classification was performed in the traditional sense. Although the analysis has revealed no statistically significant difference between the two methods, there was a strong trend for superiority of the automatically compared to the manually determined VOIs. An application of high-level classification schemes such as multilayer perceptrons or support vector machines ought to be evaluated in the next steps, to facilitate the usage of the identified VOIs in the clinical practice. In addition to the fast acquisition of T2’ data, the non-linear registration of a patient dataset to the generated reference atlas and its evaluation in terms of average T2’ calculation within the VOIs can be conducted in less than one minute on a standard computer without user interaction, making its application even more appealing.

In addition to the automatic classification of CPS and APS patients using the T2’ image sequences, the presented method may enable an even more sophisticated classification approach, also taking into account morphological information. For this purpose, it needs to be evaluated if the current atlas and image sequences allow analyzing the atrophy within the midbrain similar to previous analyses, for example, based on high resolution T1-weighted MR image sequences [6]. Such a combined approach would enable a joint analysis of morphological as well as metabolic information, which is related to tissue iron content, within one classification approach. This combination may lead to improved classification results compared to previously presented methods focusing only on one parameter. An integration of further MR sequences such as...
diffusion tensor imaging (DTI) (e.g. the apparent diffusion coefficient and fractional anisotropy measurements) may also lead to improved differentiation results.

As a limitation of the present study, it needs to be pointed out that, although all patients were assessed by expert clinicians according to established consensus criteria, none of the patients had a histopathologically proven diagnosis. Thus, there may be still a slight level of uncertainty regarding the clinical diagnosis.

Finally, it should be mentioned that this study represented a hypothesis generating trial rather than a confirmatory trial. Here, the main goal was to investigate the characteristics of automatically identified bias-free VOIs over manually defined VOIs. Therefore, no sample size estimation was performed prior to data analysis and further studies are needed to confirm the diagnostic benefit of the described methods. Within this context it also has to be highlighted that the same datasets have been used for the automatic VOI identification as well as for the evaluation of the associated benefit regarding the differentiation of CPS and APS patients. For this reason, it needs to be evaluated in further studies if the identified VOIs can be used for automatic classification of new datasets without adjustment of the VOIs.

While taking into account these caveats, we conclude that the automatically defined VOIs might improve the differentiation between patients with CPS and those with APS based on image sequences.

**References**