Regional Gray Matter Atrophy in Early Primary Progressive Multiple Sclerosis

A Voxel-Based Morphometry Study

Jorge Sepulcre, MD; Jaume Sastre-Garriga, MD; Mara Cercignani, PhD; Gordon T. Ingle, MRCP; David H. Miller, FRCP; Alan J. Thompson, FRCP

Background: Gray matter (GM) atrophy has been reported in multiple sclerosis (MS). However, little is known about its regional distribution.

Objective: To investigate the regional distribution of GM atrophy in clinically early primary progressive MS (PPMS).

Design and Patients: Thirty-one patients with PPMS within 5 years of symptom onset (mean age, 43.2 years; median Expanded Disability Status Scale score, 4.5) and 15 healthy control subjects (mean age, 43.7 years) were studied. All subjects underwent a 3-dimensional inversion-recovery fast spoiled gradient-recalled echo sequence that was repeated after 1 year in patients only. Magnetic resonance images underwent an optimized voxel-based morphometric analysis that segments magnetic resonance data volumes in a normalized space and quantifies tissue atrophy on a voxel-by-voxel basis. A lesion mask was created for each patient and used in normalization and segmentation steps to minimize bias from lesions. A multisubject design was used in the cross-sectional study to compare patients with PPMS and controls. A 1-way analysis of variance (within-subjects) design was used in the longitudinal study.

Results: At baseline, patients with PPMS displayed bilateral thalamic atrophy compared with controls. In addition, a significant association between lesion load and decreased GM volume was found for the thalami. Loss of GM in the putamen, caudate, thalami, and cortical and infratentorial areas was observed in patients after 1 year of follow-up.

Conclusions: Atrophy is most obvious in deep GM in clinically early PPMS. This may reflect increased sensitivity of these regions to neurodegeneration. Cortical and infratentorial atrophy developed as the disease evolved.

Arch Neurol. 2006;63:1175-1180
ages an in-house modified version of the protocol, tailored to deal with MR images of patients with MS.

**METHODS**

**PATIENTS AND CONTROL SUBJECTS**

Thirty-one patients with PPMS (definite or probable)\(^2\) in the early clinical stages (within 5 years of symptom onset) and 15 control subjects were recruited at the National Hospital for Neurology and Neurosurgery, London, England (same cohort studied by Sastre-Garriga et al\(^3\)). The control group matched the patient group for age and sex, and no controls had a history of neurological or psychiatric illness. One patient had undergone a course of oral corticosteroid therapy 3 months before the second scan, and no patient was receiving disease-modifying drugs. The ethics committee of the National Hospital for Neurology and Neurosurgery and the Institute of Neurology, London, approved the study, and all subjects gave their written informed consent according to the Helsinki declaration.

**MR IMAGE ACQUISITION**

We acquired MR images using a 1.5-T scanner (GE Signa; General Electric, Milwaukee, Wis). A 3-dimensional (3D) inversion-recovery fast spoiled gradient-recalled echo sequence (FSPGR) (124 contiguous axial slices; echo time, 4.2 milliseconds; repetition time, 13.3 milliseconds; inversion time, 450 milliseconds; number of excitations, 1; field of view, 300 × 225 mm) was acquired on all subjects at baseline and after 1 year of follow-up only in the patients with PPMS. A dual spin-echo sequence (28 contiguous slices; echo time, 30/80 milliseconds; repetition time, 1720 milliseconds; number of excitations, 0.75; slice thickness, 3 mm) was acquired on all patients at baseline.

**MR IMAGING DATA ANALYSIS AND VBM PREPROCESSING**

We analyzed MR images for all subjects on a commercially available workstation (Sun Blade 150; Sun Microsystems Inc, Mountain View, Calif).

**CREATION OF MS LESION MASKS**

First, MS lesions were outlined on the FSPGR scans (to calculate 3D binary masks) and proton density scans (to calculate lesion loads) using a semiautomatic local thresholding contour technique (Dispiimage; DL Plummer, University College London, London). Next, a 3D binary MS lesion mask for every patient was created by nulling the signal within the lesion outlines and setting the remaining to 1. The next stages of image processing were obtained using SPM2 software (Wellcome Department of Cognitive Neurology, University College London [http://www.fil.ion.ucl.ac.uk/spm]) running under MATLAB, version 6.5 (Mathworks Inc, Natick, Mass). The images were processed according to a customized version of the optimized VBM protocol\(^2\) (http://dbm.neuro.uni-jena.de/vbm.html), modified to reduce the influence of MS lesions in the process, which could alter the normalization and segmentation procedures.

**CREATION OF CUSTOMIZED TEMPLATES**

According to the VBM-optimized protocol, we first generated 2 customized whole-brain T1-weighted templates and GM, WM, and CSF prior probability maps. The first set was obtained from the overall population (patients and controls) at baseline, whereas the second was obtained from patients with PPMS only (patients at baseline and after 1 year of follow-up). The creation of customized templates was done to reduce scanner-specific and population-specific biases.

This template creation has been described elsewhere\(^2\); however, because our protocol was slightly modified owing to the presence of WM lesions, we briefly summarize herein the steps involved. All of the FSPGR images were spatially normalized into the same stereotactic space using the Montreal Neurological Institute (MNI) template, and resliced to 1 × 1 × 1 isotropic voxels. The normalization was performed by first estimating the optimal 12-variable affine transformation for matching images and then optimizing the normalization using 16 nonlinear iterations.\(^2\) The patients’ MR images were weighted by the corresponding 3D MS lesion masks. In our case, we assigned a weight of 0 to lesions and 1 to the rest of the brain, and therefore voxels within lesions were not considered during the normalization procedure. The images were then segmented into WM, GM, and CSF. Again, the 3D MS lesion masks, previously normalized using the same transformation variables as the FSPGR images, were later applied to the GM and WM segments to remove any remaining lesion wrongly classified as either tissue. The normalized, whole-brain images were averaged, and the resulting mean was smoothed with an 8-mm full-width at half-maximum isotropic gaussian kernel to form a T1-weighted template.

The GM, WM, and CSF segments were also averaged and smoothed with an 8-mm gaussian kernel to create template images for the 3 classes of tissues. These 3 templates are used as prior images during image segmentation, and the GM template is also used as a target for image normalization.

**NORMALIZATION AND SEGMENTATION FOR CROSS-SECTIONAL COMPARISON**

The FSPGR original images were segmented in native space into GM, WM, and CSF images, using the customized WM, GM, and CSF prior maps. The extracted segmented GM images were normalized (using the same combination of affine and nonlinear functions as described in the “Creation of Customized Templates” subsection of the “Methods” section) to the customized GM template. As we mentioned, when dealing with patients’ images, the 3D MS lesion masks were used to weight images during normalization and to remove any remaining lesion wrongly classified as either tissue. The optimized transformation variables obtained in this step were then applied to the original T1-weighted volumes in native space.

Finally, the optimally normalized T1-weighted images were segmented, again producing GM, WM, and CSF maps in MNI space. Again, normalized 3D MS lesion masks were used to remove MS lesions from these final maps. To preserve the total within-voxel volume, which may have been affected by the nonlinear transformation, every voxel signal intensity in the segmented WM and GM images was multiplied by the Jacobian determinants derived from the spatial normalization.\(^2\) The analysis of these modulated data tests was used for regional differences in absolute tissue volume. In this final step, all images were smoothed using a 12-mm full-width at half-maximum gaussian kernel.

**NORMALIZATION AND SEGMENTATION FOR LONGITUDINAL COMPARISON**

A procedure similar to that described in the preceding section was performed on baseline and follow-up MR images from patients only. In this case, however, we used an affine registra-
tion only in the normalization process, without nonlinear warping, because warping is known to reduce the sensitivity to subtle changes and use of only the affine registration was justified by assuming that the changes in brain shape during the follow-up period are expected to be subtle even in the presence of nonuniform atrophy.

STATISTICAL ANALYSIS

We used $\chi^2$ and unpaired $t$ tests to assess and compare sex and age distribution in the 2 cohorts of subjects (controls and patients with PPMS). For these tests, a significance level of $P < .05$ was considered. Statistical analysis was performed with SPSS software, version 11 (SPSS Inc, Chicago, Ill).

Regional GM atrophy was assessed using SPM2 on a voxel-by-voxel basis. This analysis is based on the general linear model. For the cross-sectional analysis, a multisubject design with age and sex as nuisance variables was used to compare the modulated GM images of the patients with PPMS with those of the controls. For the longitudinal study, a 1-way analysis of variance (within-subjects) design was used to compare GM images of patients with PPMS at baseline and after 1 year of follow-up. We used $P < .05$, after correction for multiple comparisons with the familywise error method (to minimize type I error), for all statistical comparison. We retained only clusters with 100 or more contiguous voxels to obtain consistent results. We tested the association between the lesion load and the regional distribution of GM atrophy at baseline using a simple regression model with no covariates. Finally, the MNI coordinates corresponding to areas of significant differences were converted to Talairach coordinates and localized as a neuroanatomical area.

RESULTS

The demographic and clinical data of the patients with PPMS and the controls are shown in Table 1. There were no differences between the patients and controls in sex and age distribution.

<table>
<thead>
<tr>
<th></th>
<th>Control Subjects (n = 15)</th>
<th>Patients With PPMS (n = 31)</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at scanning, mean ± SD, y</td>
<td>43.2 ± 10.9</td>
<td>43.71 ± 9.87</td>
<td>.88†</td>
</tr>
<tr>
<td>Sex ratio, No. M/F</td>
<td>9/6</td>
<td>18/13</td>
<td>.58‡</td>
</tr>
<tr>
<td>Disease duration, y</td>
<td>3 (2 to 5)</td>
<td>4.5 (3.5 to 7)</td>
<td>. . .</td>
</tr>
<tr>
<td>EDSS score</td>
<td>. . .</td>
<td>−0.26 (−6.16 to 0.79)</td>
<td>. . .</td>
</tr>
<tr>
<td>MSCF score</td>
<td>. . .</td>
<td>−0.79 (−3.32 to 1.43)</td>
<td>. . .</td>
</tr>
<tr>
<td>NHPT z score</td>
<td>. . .</td>
<td>−0.09 (−0.45 to 13.70)</td>
<td>. . .</td>
</tr>
<tr>
<td>TWT z score</td>
<td>. . .</td>
<td>−0.002 (−3.73 to 1.24)</td>
<td>. . .</td>
</tr>
<tr>
<td>PD lesion load, cm$^3$</td>
<td>15.28 (1.4 to 82.2)</td>
<td>15.28 (1.4 to 82.2)</td>
<td>. . .</td>
</tr>
</tbody>
</table>

Abbreviations: EDSS, Expanded Disability Status Scale; MSFC, Multiple Sclerosis Functional Composite; NHPT, Nine-Hole Peg Test; NS, not significant; PASAT, Paced Auditory Serial Addition Test; PD, proton density; PPMS, primary progressive multiple sclerosis; TWT, Timed Walk Test; ellipses, not applicable.

*Unless otherwise indicated, data are expressed as median (range).
†Calculated as unpaired $t$ test.
‡Calculated as $\chi^2$ test.

**CROSS-SECTIONAL ANALYSIS**

In the cross-sectional analysis, compared with controls, patients with PPMS showed reduced GM volume bilaterally in several regions of the thalami (number of voxels, 1617; $P < .001$ [uncorrected]; $z = 3.50$; maximum coordinates [x, y, z], −6, −8, 1); particularly, the anterior, ventral anterior, medial dorsal, and ventral lateral left and right nuclei (Figure 1). However, no cluster of voxels survived after correcting for multiple comparisons. Moreover, a significant association between increased lesion load and reduced GM volume was found for the thalami, specifically in the anterior nuclei, and the association was still significant after multiple comparison correction (number of voxels, 893; $P = .003$ [corrected for multiple comparisons]; $z = 4.95$; maximum coordinates [x, y, z], 4, −9, 12).

**LONGITUDINAL ANALYSIS**

When considering the longitudinal changes (PPMS at baseline vs 1-year follow-up), we observed widespread significant reduction of GM volume at 1 year. This was seen in several deep GM areas (putamen, caudate, and thalami) and in some cortical and infratentorial areas (Figure 2 and Table 2).
The evaluation of the regional distribution of GM atrophy using MR techniques has the potential to improve our understanding of the neurodegenerative process in MS. Few studies have investigated regional GM changes in MS, particularly in early PPMS. We used a VBM protocol specially designed to assess GM atrophy in patients with MS. This approach prevented voxel misclassification of MS lesions as GM, owing to similar signal intensities, in the segmentation process. In addition, given the dearth of information about local GM atrophy in patients with PPMS, the VBM method facilitated the search for GM volume changes without prior knowledge of its location. We report, in cross-sectional and longitudinal studies, a predominant involvement of deep GM in this clinical form of MS.

**CROSS-SECTIONAL COMPARISON**

In an initial approach with an uncorrected P value, we observed a reduction in the volume of the thalami in patients with PPMS in the early stage of the disease compared with controls. The change was most obvious in thalamic regions related with cognitive (anterior nuclei), motor (ventral anterior and ventral lateral nuclei), and sensory (medial dorsal nuclei) functions. Because the statistical significance was lost when correcting for multiple comparisons, the present findings should be considered with caution, and the possibility that they represent a false-positive result cannot be excluded. However, the extent of the area of abnormality and, more importantly, the fact that it is in agreement with previous diffusion, perfusion, and spectroscopy MR studies, which have included patients with PPMS, suggest that our

<table>
<thead>
<tr>
<th>Cluster Location</th>
<th>No. of Voxels</th>
<th>P Value*</th>
<th>z Score</th>
<th>Maximum Coordinates (x, y, z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L sublobar (L caudate head, L medial globus pallidus, L lateral globus pallidus, and L putamen)</td>
<td>5575</td>
<td>&lt;.001</td>
<td>6.81</td>
<td>−11, 10, −2</td>
</tr>
<tr>
<td>L limbic lobe, L and R occipital lobe (L and R Brodmann areas 17, 18, 23, and 30; R Brodmann area 19; and L Brodmann area 31)</td>
<td>3359</td>
<td>&lt;.001</td>
<td>6.60</td>
<td>−1, −84, 1</td>
</tr>
<tr>
<td>R sublobar and frontal lobe (R Brodmann area 13, R caudate head, R lateral globus pallidus, and R putamen)</td>
<td>2094</td>
<td>&lt;.001</td>
<td>6.30</td>
<td>21, 7, −8</td>
</tr>
<tr>
<td>R and L cerebellum (L and R pyramis vermis; L, R, and C tuber vermis; L, R, and C declives vermis; L culmen vermis; L and R uvula; L and R pyramids; L, R, and C declives; and L, R, and C culmen)</td>
<td>2159</td>
<td>&lt;.001</td>
<td>5.93</td>
<td>−6, −74, −31</td>
</tr>
<tr>
<td>R cerebellum (R cerebellar tonsil, R declines, and R culmen)</td>
<td>647</td>
<td>.001</td>
<td>5.82</td>
<td>37, −43, −38</td>
</tr>
<tr>
<td>R parietal lobe (R Brodmann areas 2 and 40)</td>
<td>220</td>
<td>.001</td>
<td>5.69</td>
<td>42, −30, 44</td>
</tr>
<tr>
<td>R sublobar (R ventral posterior medial nucleus, R medial dorsal nucleus, R lateral dorsal nucleus, R pulvinar, and R lateral posterior nucleus)</td>
<td>111</td>
<td>.002</td>
<td>5.54</td>
<td>15, −24, 9</td>
</tr>
<tr>
<td>R parietal and frontal lobe (R Brodmann areas 3, 4, and 6)</td>
<td>276</td>
<td>.003</td>
<td>5.48</td>
<td>49, −14, −41</td>
</tr>
<tr>
<td>L occipital and parietal lobe (L Brodmann areas 7 and 31)</td>
<td>152</td>
<td>.007</td>
<td>5.29</td>
<td>−19, −65, 31</td>
</tr>
</tbody>
</table>

Abbreviations: C, central; L, left; PPMS, primary progressive multiple sclerosis; R, right.

*Corrected for multiple comparisons.
findings represent a real phenomenon. Abnormalities in the thalami of patients with clinically isolated syndromes suggestive of MS have also been detected by a voxel-based comparison of magnetization transfer ratios. \(^{33}\) In patients with relapsing-remitting MS, normalized thalamic volume has also been found to be reduced. \(^{29}\) The thalamus has extensive connections with other subcortical structures and the cerebral cortex, and it may be more sensitive to local inflammatory activity or to changes secondary to WM lesions. \(^{29}\) This would be supported by the significant (albeit moderate) association found in the present study between lesion load and thalamic atrophy. However, other mechanisms, such as secondary effects of damage in normal-appearing WM or a direct mechanism of injury to the thalamus, could also contribute to thalamic atrophy. Thus, we conclude that thalamic atrophy could represent an early stage in the development of GM atrophy in patients with PPMS.

In this study, we have not investigated the association between thalamic changes and specific clinical symptoms. Future studies that focus on patients with defined clinical presentations (pain, fatigue, cognitive impairment, or motor-sensory symptoms) should help to establish associations between GM regional atrophy patterns and clinical findings.

The lack of significant differences for any cortical region might be because atrophy of those areas is patchier and less consistent between patients, thus may not show up in group comparisons.

**LONGITUDINAL COMPARISON**

In 1 year, patients with PPMS showed a decrease in volume in several GM structures, particularly in the basal ganglia. Loss of GM volume was observed not only in thalamic or deep GM areas but also in many other regions such as cortical and infratentorial regions, suggesting ongoing widespread GM damage. The new areas were mainly related to motor control and cognitive function. This temporal evolution might indicate that, after an initial stage of focal deep GM atrophy, extensive GM damage develops as disease evolves. Although it is important to highlight that the effect of normal aging cannot be ruled out, because no longitudinal control data can be used for comparison, a previous study,\(^{22}\) using similar methodology in 465 healthy adults showed that the thalamic areas, amygdala, hippocampus, and entorhinal cortex develop less GM atrophy with aging than other cortical and cerebellar regions. These data would strengthen the value of deep GM findings. However, the lack of longitudinal data in our control group does not allow for a direct comparison between GM loss in the PPMS and control populations. This is especially relevant in the case of cortical GM loss. Furthermore, the lack of control data at follow-up makes it impossible to account for any possible scanner drift over time. However, the scanner used for this study undergoes regular quality assurance tests, ensuring its stability during the follow-up time.

Because the mechanisms of axonal damage in MS are not well established, it is difficult to determine why the deep GM seems to be particularly at risk of developing atrophy. Deep GM enclosed by WM could be a more direct target of local, dying-back, or transsynaptic degeneration evoked by demyelination, at least in this form of the condition. Methodological aspects must also be considered because deep GM changes might be easy to detect and/or GM changes in the cortex might be more difficult to find owing to a patchy and inconsistent distribution between patients but widespread distribution of cortical atrophy.

In conclusion, our findings demonstrate predominant volume loss in the deep GM of patients with PPMS in the early stages of the disease and widespread GM atrophy during follow-up. Controlled longitudinal studies are required to confirm these findings, especially using different methodology and including patients with different disability patterns and other early forms of the condition over longer periods.

**REFERENCES**