Grey matter damage predicts the evolution of primary progressive multiple sclerosis at 5 years

M. Rovaris,1,2 E. Judica,1,2 A. Gallo,1 B. Benedetti,1,2 M. P. Sormani,1,3 D. Caputo,5 A. Ghezzi,6 E. Montanari,7 A. Bertolotto,8 G. Mancardi,4 R. Bergamaschi,9 V. Martinelli,2 G. Comi2 and M. Filippi1,2

1Neuroimaging Research Unit and 2Department of Neurology, San Raffaele Scientific Institute, Milan, 3Biostatistics Unit, Department of Health Sciences, 4Department of Neurological Sciences, University of Genoa, Genoa, 5Department of Neurology, Don Gnocchi Scientific Institute, Milan, 6Multiple Sclerosis Center, Ospedale di Gallarate, Gallarate, 7Multiple Sclerosis Center, Ospedale di Fidenza, Fidenza, 8Department of Neurology, Ospedale di Orbassano, Orbassano and 9Department of Neurological Sciences, Mondino Scientific Institute, University of Pavia, Pavia, Italy

Correspondence to: Massimo Filippi, Neuroimaging Research Unit, Department of Neurology, San Raffaele Scientific Institute, via Olgettina 60, 20132 Milan, Italy
E-mail: filippi.massimo@hsr.it

Reliable prognostic markers of primary progressive (PP) multiple sclerosis evolution are still needed. Diffusion tensor (DT) MRI can quantify normal-appearing white matter (NAWM) and grey matter (GM) damage in multiple sclerosis patients. We investigated whether conventional and DT-MRI-derived measures can predict the long-term clinical evolution of PP multiple sclerosis. In 54 PP multiple sclerosis patients, conventional and DT-MRI scans of the brain and T1-weighted scans of the cervical cord were acquired at baseline and after a median follow-up of 15 months. Another clinical evaluation was performed, 56 months after baseline, in 52 patients. Measures of lesion load, brain and cord atrophy were obtained. Histograms of the mean diffusivity (MD) and fractional anisotropy (FA) values from the NAWM and GM were analysed. At follow-up, 35 patients (65%) experienced a confirmed disability progression. Baseline expanded disability status scale score and average GM MD were independent predictors of subsequent clinical deterioration in a multivariable model (Nagelkerke $R^2$: 0.44; discriminating ability: 81%). A lower level of disability and a more severe GM damage identify PP multiple sclerosis patients with an increased risk of disease progression over the subsequent 5 years. These data may be relevant to select patients for future exploratory phase II trials.

Keywords: diffusion tensor MRI; disease evolution; grey matter; magnetic resonance imaging; primary progressive multiple sclerosis

Abbreviations: CSA = cross-sectional cord area; DT = diffusion tensor; EDSS = expanded disability status scale; FA = fractional anisotropy; FU1 and FU2 = follow-up visit 1 and 2; GM = grey matter; MP-RAGE = magnetization prepared rapid acquisition gradient echo; MD = mean diffusivity; NAWM = normal-appearing white matter; PP = primary progressive

Received May 25, 2006. Revised July 17, 2006. Accepted July 25, 2006

Introduction

About 15% of patients with multiple sclerosis experience irreversible accumulation of neurological disability since the disease onset: this clinical disease phenotype is named primary progressive (PP) (Thompson et al., 2000). Little is known about the pathobiology of PP multiple sclerosis; however, post-mortem studies suggest that neurodegeneration is predominant over inflammation in these patients (Revesz et al., 1994; Lucchinetti et al., 2000; Kutzelnigg et al., 2005). In addition, despite the advances in multiple sclerosis treatment, the evolution of PP multiple sclerosis does not seem to be favourably influenced by any of the currently available disease-modifying drugs (Kalkers et al., 2002; Leary et al., 2003; Wolinsky et al., 2003).

The unfavourable clinical course of PP multiple sclerosis is in contrast with the well-established observation that the burden and activity of lesions on conventional MRI scans of
the brain are much lower in these patients than in those with other, less disabling forms of the disease (Thompson et al., 1990, 1991; Kidd et al., 1996; Lycklama à Nijeholt et al., 1998; Stevenson et al., 1999, 2000; van Walderveen et al., 2001; Ingle et al., 2002, 2003). The results of quantitative and functional MRI studies underpin that the main factors possibly explaining the clinical/MRI discrepancy observed in PP multiple sclerosis include the presence of a diffuse tissue damage which is beyond the resolution of conventional imaging, the severity of cord damage and the impairment of the adaptive capacity of the cortex to limit the functional consequences of disease pathology (Rovaris et al., 2001, 2002, 2005a, b; Bozzali et al., 2002; Filippi et al., 2002; Rocca et al., 2002; Dehmeshki et al., 2003). Longitudinal studies indicate that newer MR-based techniques, among which is diffusion tensor (DT) MRI, are sensitive to the disease-related accrual of tissue damage in the brain of PP multiple sclerosis patients, which goes undetected when using conventional MRI (Rovaris et al., 2001, 2005a; Schmieri et al., 2004). This is particularly true for grey matter (GM) abnormalities, which seem to be the main factor underlying the presence of ‘fixed’ disability in progressive multiple sclerosis (Miller et al., 2003; Rovaris et al., 2005a). Nevertheless, preliminary data indicate that, over a short-term period, there is no, or just a weak, relationship between the progression of GM damage and that of clinical disability in PP multiple sclerosis (Rovaris et al., 2005a). On the other hand, MRI studies with longer follow-up durations (Losseff et al., 1996a; Ingle et al., 2003; Stevenson et al., 2004; Sastre-Garriga et al., 2005a) suggest that the characteristics of PP multiple sclerosis evolution may be, at least partially, predicted by a composite of clinical and MRI features (Sastre-Garriga et al., 2005a).

Against this background, a long-term clinical assessment of a cohort of PP multiple sclerosis patients (Rovaris et al., 2005a) has been planned. Here we report the results of the first 5-year follow-up, which was performed to investigate whether clinical and MRI parameters at baseline, as well as their 1-year changes, are predictors of the subsequent clinical evolution. This is central to the definition of markers with the potential to enable a reliable selection of PP multiple sclerosis patients with a high risk of unfavourable disease evolution and, therefore, amenable for being enrolled into experimental treatment trials.

Patients and methods

Patients

Fifty-four PP multiple sclerosis patients (women/men: 27/27) were studied (Rovaris et al., 2005a). They were enrolled consecutively from the outpatient populations attending multiple sclerosis clinics, without pre-defined selection criteria for disease duration, progression rate and disability level. The disease phenotype was classified according to the international consensus criteria (Thompson et al., 2000). Other neurological conditions were always carefully excluded by performing in all patients the appropriate investigations, including a CSF examination. A total of 45 patients were affected by definite and 9 by probable PP multiple sclerosis (Thompson et al., 2000). All patients with probable PP multiple sclerosis had negative CSF examination and positive MRI findings. Out of the 54 PP multiple sclerosis patients 43 had had a spinal cord presentation at disease onset, the remaining 11 had had other uni- (10 patients) or multi-focal (1 patient) presentations, with motor (3 patients), visual (1 patient), cerbellar (4 patients), brainstem (3 patients) or sensory (1 patient) disturbances. At study entry, mean age was 51.3 (range = 25–68) years and median disease duration was 10.0 (range = 2–26) years. At this stage, 36 patients were not receiving any disease-modifying treatment, 9 were treated with azathioprine, 4 with pulses of intravenous mitoxantrone and 5 with methotrexate. At final follow-up, two patients had stopped azathioprine, two mitoxantrone and four methotrexate therapy, i.e. 44 patients were not receiving any disease-modifying treatment.

The clinical evaluation consisted of a complete neurological examination, with rating of the Expanded Disability Status Scale (EDSS) scores (Kurtzke, 1983). A first follow-up visit (FU1) was performed in all patients after a median period of 15.0 (range = 12–23) months, within 3 days from the corresponding MRI session (Rovaris et al., 2005a). Fifty-two (96%) patients underwent a second clinical evaluation (FU2) after a median period of 36.0 (range = 35–63) months from study entry. In two patients, FU2 evaluation was not performed because they could not be contacted to schedule the visit. At follow-up evaluations, patients were considered clinically worsened if they had an EDSS score increase ≥1.0, when baseline EDSS was <6.0, or an EDSS score increase ≥0.5, when baseline EDSS was ≥6.0. EDSS changes had always to be confirmed by a second visit after a 3-month interval. In individual patients, all visits were done by the same observer, who was unaware of the MRI results.

All the patients signed a written informed consent prior to study-entry and the study was approved by the local Ethical Committees of all the participating institutions.

Image acquisition

Using a 1.5 T magnet, the following scans of the brain were acquired at baseline and FU1: (i) dual echo turbo spin echo; (ii) T1-weighted conventional spin echo; (iii) echo-planar pulsed-gradient spin-echo (PGSE), with diffusion gradients applied in eight non-collinear directions, chosen in order to cover 3D space uniformly. For dual echo and T1-weighted scans 24 contiguous, axial slices with 5 mm thickness were obtained. For PGSE scans, 10 axial slices with 5 mm thickness were acquired, with the same orientation as the other scans and the second-last caudal slice positioned to match exactly the central slices of these sets. At follow-up, patients were repositioned following published guidelines (Miller et al., 1991). The complete image acquisition scheme is detailed elsewhere (Rovaris et al., 2005a).

In the same scanning sessions, using a tailored cervical cord phased array coil for signal reception, a 3D T1-weighted magnetization prepared rapid acquisition gradient echo (MP-RAGE) sequence was acquired for cervical cord imaging (Rovaris et al., 2001). Due to the length of the acquisition session, three patients at baseline and six patients at FU1 were unable to undergo cord imaging.

All image acquisition was performed at the Neuroimaging Research Unit, San Raffaele University Hospital, using the same magnet, which did not undergo any upgrade during the study period and was under a regular maintenance programme.
GM damage and evolution of PP multiple sclerosis

Image review and analysis

Following the identification of T2-hyperintense and T1-hypointense MS lesions, total lesion volumes (LVs) were measured using a methodology which is described in detail elsewhere (Rovaris et al., 1997). The scans were also reviewed in a chronological order and the numbers of new lesions on FU1 versus baseline scans were counted.

Using T1-weighted images, both longitudinal (two time-points) percentage brain volume change (PBVC) and cross-sectional (baseline) normalized brain volume (NBV) were estimated. PBVC was estimated using structural image evaluation of normalized atrophy (SIENA) (Smith et al., 2001) and NBV was estimated using its cross-sectional version (SIENAx) (Smith et al., 2002).

Mean diffusivity (MD) and fractional anisotropy (FA) maps were produced from the PGSE images, and co-registered with the dual echo images as detailed elsewhere (Filippi et al., 2001). Average lesion MD and FA were computed as previously described (Filippi et al., 2001; Rovaris et al., 2005a). Following the automated segmentation of brain GM, white matter (WM) and CSF from proton density- and T2-weighted images (Rovaris et al., 2005a), the resulting masks for each tissue were superimposed onto the MD and FA maps, where T2-hyperintense lesions were masked out previously, and the corresponding MD histograms of the normal-appearing white matter (NAWM) and GM were produced. FA histograms were derived only for the NAWM, since no preferential direction of water molecular motion is expected to occur in the GM, due to the absence of a micro-structural anisotropic organization of this tissue compartment. For all the histograms, the average MD and FA values were calculated, as well as the heights and locations of their peaks. Only the average MD and FA were a priori chosen to enter the statistical analysis, in order to minimize the number of comparisons and, therefore, reduce the risk of type I errors.

For the cervical cord, reformatting of the original MP-RAGE data was performed using the standard, vendor-supplied multiplanar software available on the operator’s console of the scanner. For each subject, a set of five contiguous, 3 mm thick axial slices (perpendicular to the spinal cord) was reconstructed using the centre of C2–C3 disc as the caudal landmark. Then, the semi-automated technique developed by Losseff et al. (1996b) was used to measure the cross-sectional cord area (CSA) at the level of each slice. Values from the five slices were averaged to obtain a single value for each subject. All the analysis was carried out by a single observer, blinded to the clinical status of the subjects.

Statistical analysis

The values of MRI-derived metrics at baseline and FU1 were compared using a Student t-test for paired data. A univariate logistic regression model adjusted for follow-up duration was used to screen the clinical and MRI variables as independent predictors of the probability to have an EDSS deterioration at follow-up. Those variables with a P-value <0.20 entered a multivariate analysis where the presence or absence of EDSS deterioration was the dependent variable. The multivariate analysis was also adjusted for the actual follow-up duration. The discriminating ability of the final multivariable model was tested with the leave-one-out cross validation method. With this method, for each observation the final model is refitted after leaving that observation out of the dataset and the predicted value for that observation is then computed. This procedure is run n times (n = sample size) and the discriminating ability of the final model is estimated as the proportion of patients whose clinical evolution is correctly predicted. For those MRI-derived variables entering the final model, the predictive value in individual patients was assessed using a receiver operator curve (ROC) analysis. The cut-off value yielding the highest accuracy was computed.

Results

The median value of patients’ EDSS score was 5.5 (range = 2.5–7.5) at study entry, 6.0 (range = 3.0–7.5) at FU1 and 6.5 (range = 3.0–9.0) at FU2. At follow-up evaluations, 35 patients (65%) showed a significant EDSS worsening when compared with study entry. Both patients who did not undergo FU2 visit were clinically stable at FU1. EDSS score showed a significant increase at both FU1 and FU2 in 10 patients and only at FU2 in 25 patients. None of the patients experienced superimposed multiple sclerosis relapses either before or during the study period.

Table 1 reports the conventional and DT-MRI characteristics of patients at study entry and FU1. One or more new T2-hyperintense or T1-hypointense lesions were found on FU1 scans from 27 (50%) and 14 (26%) patients. The mean PBVC was different from zero (P < 0.001). No other MRI-derived metrics showed significant changes at FU1.

Table 2 reports the results of the univariate logistic regression analysis. Baseline EDSS, T1 LV, average lesion and GM MD, average NAWM MD and FA, as well as the number of new T2 and T1 lesions at FU1, entered the multivariate analysis. The final multivariable model included baseline EDSS [odds ratio (OR) = 0.48, 95% confidence interval (CI) = 0.26–0.91; P = 0.03] and average GM MD (OR = 1.21, 95% CI = 1.06–1.38; P = 0.005) as independent predictors of subsequent EDSS deterioration (Nagelkerke $R^2 = 0.44$) (Fig. 1). When only MRI-derived variables were

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Conventional and DT-MRI findings at study entry and after 15 months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study entry</td>
</tr>
<tr>
<td>T2 LV (ml)</td>
<td>17.8 (17.1)</td>
</tr>
<tr>
<td>T1 LV (ml)</td>
<td>6.5 (6.1)</td>
</tr>
<tr>
<td>New T2 lesions (range)</td>
<td>—</td>
</tr>
<tr>
<td>New T1 lesions (range)</td>
<td>—</td>
</tr>
<tr>
<td>NBV (ml)</td>
<td>1374.0 (89.1)</td>
</tr>
<tr>
<td>PBVC (U)</td>
<td>—</td>
</tr>
<tr>
<td>Cervical cord CSA (mm²)</td>
<td>63.9 (10.1)</td>
</tr>
<tr>
<td>Average lesion MD</td>
<td>1.09 (0.11)</td>
</tr>
<tr>
<td>Average lesion FA</td>
<td>0.26 (0.03)</td>
</tr>
<tr>
<td>Average NAWM MD</td>
<td>0.87 (0.05)</td>
</tr>
<tr>
<td>Average NAWM FA</td>
<td>0.26 (0.03)</td>
</tr>
<tr>
<td>Average GM MD</td>
<td>1.08 (0.08)</td>
</tr>
</tbody>
</table>

For abbreviations and statistical analysis, see the text.
Table 2 Univariate logistic regression analysis of the predictive value of clinical and MRI-derived quantities for patients’ EDSS worsening over the study period (dependent variable)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Unit</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1 year</td>
<td>0.96 (0.90–1.03)</td>
<td>0.29</td>
</tr>
<tr>
<td>Gender</td>
<td>Male versus female</td>
<td>1.08 (0.34–3.43)</td>
<td>0.89</td>
</tr>
<tr>
<td>Disease duration</td>
<td>1 year</td>
<td>1.01 (0.89–1.14)</td>
<td>0.44</td>
</tr>
<tr>
<td>EDSS*</td>
<td>1 point</td>
<td>0.52 (0.29–0.94)</td>
<td>0.03</td>
</tr>
<tr>
<td>Baseline T2 LV*</td>
<td>1 ml</td>
<td>1.05 (1.00–1.11)</td>
<td>0.07</td>
</tr>
<tr>
<td>T2 LV percentage change</td>
<td>1%</td>
<td>1.02 (0.98–1.07)</td>
<td>0.29</td>
</tr>
<tr>
<td>Baseline T1 LV</td>
<td>1 ml</td>
<td>1.07 (0.96–1.20)</td>
<td>0.24</td>
</tr>
<tr>
<td>T1 LV percentage change</td>
<td>1%</td>
<td>1.01 (0.98–1.05)</td>
<td>0.48</td>
</tr>
<tr>
<td>Number of new T2 lesions*</td>
<td>1</td>
<td>1.45 (0.95–2.20)</td>
<td>0.08</td>
</tr>
<tr>
<td>Number of new T1 lesions*</td>
<td>1</td>
<td>1.72 (0.79–3.72)</td>
<td>0.17</td>
</tr>
<tr>
<td>NBV</td>
<td>1 ml</td>
<td>0.99 (0.99–1.00)</td>
<td>0.20</td>
</tr>
<tr>
<td>PBVC</td>
<td>1%</td>
<td>1.05 (0.63–1.76)</td>
<td>0.85</td>
</tr>
<tr>
<td>Cervical cord CSA</td>
<td>1 mm²</td>
<td>1.02 (0.96–1.08)</td>
<td>0.58</td>
</tr>
<tr>
<td>Cervical cord CSA percentage change</td>
<td>1%</td>
<td>0.95 (0.85–1.08)</td>
<td>0.45</td>
</tr>
<tr>
<td>Average lesion MD*</td>
<td>0.01 mm²/s × 10⁻³</td>
<td>1.04 (0.98–1.10)</td>
<td>0.19</td>
</tr>
<tr>
<td>Average lesion MD percentage change</td>
<td>1%</td>
<td>0.94 (0.84–1.06)</td>
<td>0.34</td>
</tr>
<tr>
<td>Average lesion FA</td>
<td>0.01</td>
<td>0.92 (0.77–1.08)</td>
<td>0.31</td>
</tr>
<tr>
<td>Average lesion FA percentage change</td>
<td>1%</td>
<td>1.02 (0.93–1.12)</td>
<td>0.68</td>
</tr>
<tr>
<td>Average NAWM MD*</td>
<td>0.01 mm²/s × 10⁻³</td>
<td>1.22 (1.05–1.42)</td>
<td>0.01</td>
</tr>
<tr>
<td>Average NAWM MD percentage change</td>
<td>1%</td>
<td>1.08 (0.87–1.33)</td>
<td>0.49</td>
</tr>
<tr>
<td>Average NAWM FA*</td>
<td>0.01</td>
<td>0.65 (0.47–0.90)</td>
<td>0.01</td>
</tr>
<tr>
<td>Average NAWM FA percentage change</td>
<td>1%</td>
<td>0.99 (0.86–1.14)</td>
<td>0.89</td>
</tr>
<tr>
<td>Average GM MD*</td>
<td>0.01 mm²/s × 10⁻³</td>
<td>1.17 (1.05–1.33)</td>
<td>0.006</td>
</tr>
<tr>
<td>Average GM MD percentage change</td>
<td>1%</td>
<td>1.03 (0.82–1.30)</td>
<td>0.79</td>
</tr>
</tbody>
</table>

*Variables entering the multivariate analysis. Percentage changes were for FU1 versus baseline scans. See the text for abbreviations and further details.

Discussion

During the last decade, quantitative MR-based studies using various techniques (Rovaris et al., 2001, 2002, 2005a, b; Bozzali et al., 2002; Dehmeshki et al., 2003) have highlighted the role of normal-appearing brain tissue (NABT) damage in contributing to the accrual of disability in PP multiple sclerosis. More recently, the severity of GM damage (Bozzali et al., 2002; Rovaris et al., 2002, 2005a) has been identified as one of the main factors associated to a more severe disease evolution. Against this background, we planned a long-term follow-up of a relatively large cohort of PP multiple sclerosis patients (Rovaris et al., 2005a) to ascertain whether a combination of clinical, conventional and DT-MRI features may contribute to the identification of patients with an increased risk of a more severe clinical evolution. A measure of irreversible tissue damage in the cord, which is thought to influence locomotor disability in multiple sclerosis, was also added to the multiparametric MRI evaluation of the brain. The present study reports the results of the first 5 years of follow-up.

As expected from natural history studies of PP multiple sclerosis patients with a relatively long disease duration (Cottrell et al., 1999), >60% of our patients showed a significant and confirmed EDSS worsening when compared with study-entry, independently of any previous or ongoing disease-modifying treatment. Admittedly, the lack of selection criteria for the post-study clinical progression rate may render this cohort not fully representative of PP multiple...
sclerosis patients amenable to be enrolled in experimental treatment trials. However, we believe that the wide range of values for disability present in our sample indicates that this study population is representative of the overall PP multiple sclerosis clinical spectrum. The results of the short-term MRI evaluation indicated that conventional brain MRI activity was detectable in a significant proportion of patients. A progression of brain and cord atrophy, as well as a worsening of DT-MRI derived parameters within T2-visible lesions and in the GM, was also observed in this study cohort, even though it did not reach statistical significance.

The most intriguing and novel finding in the present study is that a lower EDSS and a more severe GM damage at baseline are able to identify PP multiple sclerosis patients with a worse clinical evolution over the subsequent 5 years. The final multivariable model including these two parameters was able to explain about 40% of the observed EDSS variance at follow-up. In addition, ~80% of patients were correctly classified as clinically stable or worsened when the discriminating ability of the model was tested in individual cases.

To the best of our knowledge, only few previous studies (Losseff et al., 1996a; Stevenson et al., 2004; Sastre-Garriga et al., 2005a) investigated whether conventional MRI-derived measures may serve as predictors of long-term PP multiple sclerosis evolution. Clinical disease worsening after ~5 years was modestly associated with lower EDSS and shorter disease duration at baseline (Sastre-Garriga et al., 2005a), previous MRI activity (Stevenson et al., 2004; Sastre-Garriga et al., 2005a), severity (Sastre-Garriga et al., 2005a) and short-term progression (Stevenson et al., 2004; Sastre-Garriga et al., 2005a) of brain or cord atrophy. In the study by Losseff et al. (1996a), there were no correlations between new lesion development on monthly MRI scans obtained for six months and PP multiple sclerosis clinical outcome after 5 years, whereas this was the case for secondary progressive multiple sclerosis patients. Two other studies using magnetization transfer MRI (Santos et al., 2002; Rovaris et al., 2003) have recently shown that the severity of NABT damage may predict the clinical evolution of multiple sclerosis after 5 years, but they did not include patients with PP multiple sclerosis. Our data are consistent with the results of previous studies, since we confirmed that PP multiple sclerosis patients with lower EDSS at baseline have an increased risk of subsequent clinical worsening. This also fits with the notion that EDSS responsiveness to actual clinical changes is higher for the lower scores of the scale.

However, such a different responsiveness may also be the result of a 'ceiling' effect favouring the lower EDSS scores. We also found a significant association between the number of new T2 lesions on FU1 scans and an increased risk of disability worsening at FU2, but the long-term prognostic value of conventional MRI activity after 1 year was lower than that of DT-MRI-derived measures of GM damage at baseline. The latter finding might be explained by the characteristics of our sample of PP multiple sclerosis patients, who, on average, had a relatively long disease duration (10 years). On the contrary, in 'early' PP multiple sclerosis patients (i.e. those within 5 years from the clinical disease onset), MRI-detectable inflammatory activity might play a more important role than later on in the course of the disease (Sastre-Garriga et al., 2004, 2005b, c, Ingle et al., 2005), being associated with more severe clinical disability and brain atrophy.

That DT-MRI is sensitive to GM changes in established multiple sclerosis has been consistently shown by several recent studies (Rovaris et al., 2005c). There is also increasing evidence that DT-MRI-detectable GM damage is more severe in the most disabling multiple sclerosis phenotypes (Bozzali et al., 2002; Rovaris et al., 2002, 2005a; Dehmeshki et al., 2003; Miller et al., 2003). Given the paucity of post-mortem studies correlating DT-MRI findings and multiple sclerosis histopathology (Mottershead et al., 2003), we can just speculate about the substrates underlying diffusion changes in the GM of PP multiple sclerosis
patients. Two possible explanations for GM diffusivity abnormalities might be: (i) the presence of discrete multiple sclerosis lesions, which may go undetected on conventional T_2-weighted imaging (Kidd et al., 1999; Peterson et al., 2001; Geurts et al., 2005), and (ii) the retrograde degeneration of GM neurons secondary to the damage of fibres traversing multiple sclerosis WM lesions (Evangelou et al., 2000). All of these might not only cause 'intrinsic' GM diffusivity changes, but also GM atrophy, which might also result in an increased GM diffusivity through partial volume averaging from the CSF. To take into account both of these factors (tissue loss and status of the remaining tissue), we did not correct GM diffusivity for GM atrophy. That the results of our analyses did not change when MD values were computed from eroded GM maps does, however, support the hypothesis that GM atrophy is likely to account for a minimal part of the observed GM diffusivity changes.

Moreover, a recent post-mortem study (Kutzelnigg et al., 2005), including 14 'late' PP multiple sclerosis cases, showed that diffuse NAWM injury and cortical demyelination are characteristic hallmarks of progressive multiple sclerosis and occur on the background of a global inflammatory response, with a marginal correlation with focal lesion load. DT-MRI findings might, therefore, be the in vivo correlates of such a diffuse GM demyelination, which can then lead to axonal dysfunction and, as a consequence, to irreversible neurological deficits in PP multiple sclerosis.

Our findings indicate that GM damage is the strongest paraclinical predictor of subsequent worsening of disability in PP multiple sclerosis, although T_2 lesion load and NAWM diffusivity characteristics at baseline were also significantly associated with an increased risk of EDSS deterioration after 5 years. An assessment of cognitive functions, which have been recently found to worsen over short periods of time in PP multiple sclerosis without significant correlations with the concomitant accumulation of MRI-visible lesions (Camp et al., 2005), might even disclose a closer relationship than EDSS between GM damage and clinical evolution in these patients. If confirmed by further studies in other cohorts of PP multiple sclerosis and by the prolongation of the present longitudinal study (which is ongoing), GM damage at study entry might be considered among the markers for selecting PP multiple sclerosis patients amenable for experimental, phase II trials of neuroprotective agents. The potential of such an approach is also highlighted by the finding that the vast majority of clinically worsened PP multiple sclerosis patients did have individual GM MD values higher than the cut-off value obtained from the ROC analysis. However, the difficulties for a reliable performance of DT-MRI in multicentre, longitudinal studies, including careful sequence standardization at baseline and adequate measurement stability over time, warrant further consideration and ad hoc methodological studies (Pagani et al., 2006 in press).

The observation that this multiparametric model, when applied to individual cases, leads to a misclassification of 20% of patients (i.e. clinically stable when worsened after

5 years or vice versa) underpins the need of further studies to refine this approach and to investigate whether other prognostic factors might be found by using additional paraclinical tools and by longer follow-up periods. However, a potential explanation for the partially limited prognostic value of any model based upon structural MRI findings when applied to multiple sclerosis is the role which might be played by cortical reorganization in limiting the impact of multiple sclerosis injury on the severity of neurological impairment. Functional MRI studies of PP multiple sclerosis (Filippi et al., 2002; Rocca et al., 2002) have indeed shown that the pattern of cortical activations and the severity of structural changes in the NABT are strongly correlated. The individual ability of PP multiple sclerosis patients in recruiting functionally related cortical areas, in the presence of similar amounts of tissue damage, might, therefore, contribute to a different long-term prognosis.

References


