Determinants of Disability in Multiple Sclerosis at Various Disease Stages

A Multiparametric Magnetic Resonance Study

Annalisa Pulizzi, MD; Marco Rovaris, MD; Elda Judica, MD; Maria Pia Sormani, PhD; Vittorio Martinelli, MD; Giancarlo Comi, MD; Massimo Filippi, MD

Objective: To investigate whether diffusion-tensor magnetic resonance imaging and whole brain N-acetylaspartate (WBNAA) proton magnetic resonance spectroscopy can provide complementary pieces of information to achieve a better understanding of the factors associated with disability in multiple sclerosis (MS).

Design: Cross-sectional survey.

Setting: Referral hospital-based MS center.

Patients: Ten healthy control subjects, 27 patients with a clinically isolated neurological syndrome, 21 patients with relapsing-remitting MS, and 29 patients with secondary progressive MS.

Main Outcome Measures: Conventional and diffusion-tensor magnetic resonance imaging, as well as WBNAA proton magnetic resonance spectroscopy, of the brain was performed. T2-hyperintense lesion volumes were measured. The mean values of mean diffusivity (MD) and fractional anisotropy of T2-visible lesions were computed. Histograms of MD and fractional anisotropy values were produced for normal-appearing white matter and gray matter (GM).

Results: Patients with a clinically isolated neurological syndrome had a significantly ($P=.002$) lower WBNAA concentration than control subjects. Patients with relapsing-remitting MS had significantly higher T2 lesion volume ($P=.007$), mean lesion MD ($P=.003$), normal-appearing white matter fractional anisotropy peak height ($P=.03$), and a lower WBNAA concentration ($P<.001$) than patients with a clinically isolated neurological syndrome. Patients with secondary progressive MS had significantly higher T2 lesion volume ($P=.01$), lower mean normal-appearing white matter fractional anisotropy ($P=.003$), higher mean GM MD ($P=.004$), and lower GM MD peak height ($P=.01$) than patients with relapsing-remitting MS. Disease duration, GM MD peak height, and WBNAA concentration entered a multivariate model, explaining nearly 70% of the disability variance.

Conclusions: The accumulation of macroscopic lesions and normal-appearing white matter damage seems to occur mainly during the earliest clinical phases of MS, whereas pathological features of GM may be a hallmark of the late progressive stage of the disease. This supports the notion of MS as a “2-stage” disease.

Arch Neurol. 2007;64(8):1163-1168
tivoxel \(^1\)H-MRS techniques, which are limited by partial brain coverage, WBNA1H-MRS can be considered a valid approach to quantify the “global” burden of a diffuse disease like MS, without losing the pathological specificity to its more destructive features. Conversely, diffusion-tensor (DT) MRI is able to quantify the severity of tissue damage within T2-visible lesions in MS, and in the normal-appearing white matter (NAWM) and gray matter (GM) separately,\(^7\) albeit with less pathological specificity than \(^1\)H-MRS.

Against this background, we performed this cross-sectional study to investigate whether DT MRI and WBNA1H-MRS can provide complementary pieces of information to achieve a better understanding of the nature of disability in MS at different disease stages. The aims of the study were to assess how global neuroaxonal damage or loss and tissue-specific injury of the WM and GM are associated with the clinical status of patients with MS and to interrogate the performance of multiparametric MR-based models\(^8,9\) as paraclinical correlates of MS disability.

### METHODS

#### SUBJEC TS

We studied the following groups of individuals: (1) patients with a clinically isolated neurological syndrome (CIS) suggestive of MS, enrolled within 3 months from symptom onset and with MRI evidence for disease dissemination in space\(^10\); (2) patients with clinically definite MS and either a relapsing-remitting (RR) or a secondary progressive (SP) disease phenotype\(^11\); and (3) healthy volunteers with no history of neurological diseases and a normal neurological examination result. Such a selection aimed at studying the different MR features of the main stages of relapsing MS, from the onset of clinical manifestation to the SP phase. In all patients, corticosteroid treatment for clinical relapses, if any, had to be concluded at least 4 weeks before brain MRI acquisition. Within 3 days of the MRI session, patients also underwent a clinical evaluation, including a neurological visit with an Expanded Disability Status Scale\(^12\) rating by a single observer (V.M.) unaware of the MRI results. All patients were consecutively selected from the outpatient MS Clinic population from January 1 to September 30, 2003. Local ethical committee approval and written informed consent from each subject were obtained before study initiation.

#### MR DATA ACQUISITION

Brain MRI and \(^1\)H-MRS were performed using a scanner operating at 1.5 T (Vision; Siemens, Erlangen, Germany). During a single session, the following sequences were collected, without moving the subject from the scanner: (1) dual-echo, turbo spin echo (repetition time, 3300 milliseconds; echo time, 16/98 milliseconds; and echo train length, 5); (2) pulsed-gradient echo-planar (interecho spacing, 0.8 milliseconds; and echo time, 123 milliseconds), with diffusion gradients applied in 8 noncollinear directions (the duration and maximum amplitude of the diffusion gradients were 25 milliseconds and 21 mT/m, respectively, giving a maximum \(b\) factor in each direction of 1044 s/mm\(^2\)); and (3) WBNA1H-MRS, following the scheme originally described by Gonen et al.\(^5\) Twenty-four contiguous, 5-mm-thick, axial sections with a 250 \(\times\) 250 mm matrix size and a 250 \(\times\) 250-mm field of view were obtained for dual-echo images. For pulsed-gradient echo-planar images, 10 contiguous, 5-mm-thick, axial sections with a 128 \(\times\) 128 matrix and a 250 \(\times\) 250-mm field of view were acquired, with the second-last caudal section positioned to match exactly the central sections of the dual-echo image sets.

#### MR DATA POSTPROCESSING

All postprocessing was performed by observers (A.P., M.R., and E.J.) who were blinded to subjects’ identity. The total T2-hyperintense lesion volume (LV) measurement was assessed using a semiautomated segmentation technique.\(^13\) Using SPM2 software (Wellcome Functional Imaging Laboratory, Institute of Neurology, London, England) and maximum image inhomogeneity correction to segment dual-echo images, we obtained 3 maps representing GM, WM, and cerebrospinal fluid, following a procedure that is described elsewhere.\(^14\) After DT calculation from pulsed-gradient echo-planar images, mean diffusivity (MD) and fractional anisotropy (FA) values were derived for every pixel, as previously described.\(^15\) Mean lesion MD and FA were calculated using a method that has been described elsewhere.\(^15\) Eroded NAWM and GM maps were then obtained,\(^14\) and the corresponding normalized histograms of MD and FA values were produced. The FA histograms were derived only for the NAWM, because no preferential direction of water molecular motion is expected to occur in the GM. The WBNA1H-MRS data from each subject were processed offline. To correct for the interindividual brain size variations, the absolute NAA amount from each subject, which was calculated following the original procedure by Gonen et al.,\(^5\) was divided by the subject’s absolute brain volume. This yielded an absolute WBNA1H-MRS concentration, expressed in millimoles.

#### STATISTICAL ANALYSIS

To compare the characteristics of subject subgroups, a 1-way analysis of variance model, corrected for subjects’ age, was used and 3 comparisons were decided a priori (a priori contrasts): CIS vs healthy control subjects, RRMS vs CIS, and SPMS vs RRMS. A univariate logistic regression analysis, in which MR-derived metrics were the independent variables, was used to identify the strongest predictors of the disease phenotype (RRMS vs CIS and SPMS vs RRMS), after correction for patients’ age. The discriminating ability of the final multivariate models was tested with the leave-one-out cross-validation method. With this method, for each observation, the final model is refit after leaving that observation out of the data set and the predicted value for that observation is then computed. This procedure is run \(n\) times (where \(n\) is the sample size), and the discriminating ability of the final model is estimated as the proportion of patients whose clinical phenotype is correctly predicted. Univariate correlations between clinical findings and MR-derived metrics were assessed using the Spearman rank correlation coefficient. A multivariate analysis, corrected for subjects’ age, was used to identify the strongest independent correlates of patients’ clinical characteristics.

#### RESULTS

Seventy-seven patients, of whom 27 had CIS, 21 had RRMS, and 29 had SPMS, and 10 healthy subjects were studied (Table 1). Age heterogeneity among the 4 subgroups of subjects was found (\(P<.001\)); patients with CIS were younger than control subjects (\(P=.02\)), patients with RRMS were older than patients with CIS (\(P=.004\)), and patients with SPMS were older than patients with RRMS (\(P<.001\)).
No T2-visible abnormalities were seen on MRIs from control subjects. Table 2 reports MRI and 1H-MRS findings in the study groups. Heterogeneity among the 4 groups of subjects was found for all MR-derived variables. A priori contrast analysis, adjusted for subject age, showed the following: (1) patients with CIS had a lower WBNAA concentration than control subjects (P < .001); (2) patients with RRMS had a higher T2 LV (P < .001), a higher mean lesion MD (P = .003), a higher NAWM FA peak height (P = .03), and a lower WBNAA concentration (P < .001) than patients with CIS; and (3) patients with SPMS had a higher T2 LV (P = .01), lower mean NAWM FA (P = .003), a higher mean GM MD (P = .004), and a lower GM MD peak height (P = .01) than patients with RRMS. Logistic regression analysis showed that WBNAA (odds ratio [OR], 0.41; 95% confidence interval [CI], 0.23-0.73; P = .002), NAWM MD peak height (OR, 0.95; 95% CI, 0.91-0.99; P = .04), and mean lesion MD (OR, 1.01; 95% CI, 1.00-1.02; P = .048) were independent predictors for the identification of patients with RRMS vs patients with CIS, while mean GM MD (OR, 1.24; 95% CI, 1.06-1.44; P = .006) was the only independent predictor for the identification of patients with SPMS vs patients with RRMS. The discriminating ability of the multivariate model, including WBNAA concentration, NAWM MD peak height, and mean lesion MD, was 79% (22 of 27 patients were correctly classified as having CIS and 16 of 21 patients were correctly classified as having RRMS). The model, including mean GM MD, has a 66% discriminating ability (13 of 21 patients were correctly classified as having RRMS and 20 of 29 patients were correctly classified as having SPMS).

Significant correlations were found between most MR-derived metrics and patient clinical characteristics (Table 3). Disease duration (P < .001), GM MD peak height (P = .02), and WBNAA concentration (P = .02) entered a multivariate model accounting for 67% (R² = .67, P < .001) of the observed Expanded Disability Status Scale variance. Age (P < .001) and T2 LV (P = .001) were independent explanatory variables of patient disease duration (R² = .58, P < .001).

In the whole cohort of 77 patients, T2 LV was correlated with WBNAA concentration (r = -.56, P < .01). T2 LV had stronger correlations with all DT MRI histogram-derived metrics than did WBNAA concentration (Table 4).

Using a cross-sectional design, the MR-based investigation of the correlates of disability in MS may have a subject selection bias, because different patients are chosen to represent the different disease stages. The best design
for such a study would, therefore, be a long-term prospective survey of patients with CIS, which should observe them until the development of established RRMS and the beginning of the SP phase of the disease. In addition to the expected long period needed to perform this type of longitudinal study, the MR surveillance for such a follow-up would also have technical problems related to unavoidable scanner upgrades and measurement drifts. Thus, despite the evident limitations, we believe that our study design is the only feasible way to investigate the MR correlates of the different stages of relapse-onset MS.

In our study, the different subgroups of subjects showed highly heterogeneous features on conventional, DT MRI, and WBNAA 1H-MRS scans. Contrast analyses revealed that the only significant difference between patients with CIS and healthy subjects was a lower mean WBNAA concentration in the former group. This finding confirms the results of previous studies indicating that in patients with CIS and MRI evidence of disease dissemination in space, significant neuroaxonal damage has already occurred. Such damage may, however, go undetected by DT MRI, possibly because the various, and often concomitant, pathological features of the MS-aFFECTED brain may affect the tissue diffusivity properties differently, thereby being associated with a “pseudonormalization” of the resulting DT MRI histogram-derived metrics. In patients with RRMS, the severity of WBNAA 1H-MRS–detectable neuroaxonal damage was significantly greater than in those with CIS, and it was associated with a more pronounced WM disruption within and outside T2-visible lesions, as well as with an increased T2-visible lesion burden. These findings suggest that the evolution from CIS to established MS is accompanied by the accumulation of NAWM damage, which may possibly also lead to the observed WBNAA decrease. Our results are also consistent with previous studies using magnetization transfer MRI and 1H-MRS imaging. In our cohort, patients with SPMS, when compared with those with RRMS, showed a greater T2 LV and more pronounced NAWM and GM damage. These findings indicate that a critical sampling of GM damage using DT MRI might be a rewarding strategy to identify the correlates of irreversible disability in MS, while the measurement of NAA levels may be affected by a “plateauing” effect, limiting its sensitivity. The association of GM damage with irreversible disability accumulation in SPMS has been highlighted by several MR-based studies. Possible explanations for such GM damage include the presence of “discrete” GM lesions, which may go undetected when using cMRI, the apoptotic loss of neurons in the cortex, and the retrograde degeneration of GM neurons secondary to the damage of fibers traversing WM lesions. Admittedly, GM damage may not only cause “intrinsic” tissue diffusivity changes but also cortical atrophy, which

### Table 3. Univariate Correlations Between the MR and Clinical Characteristics of Patients

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>EDSS Score</th>
<th></th>
<th></th>
<th>Disease Duration</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r Value</td>
<td>P Value</td>
<td></td>
<td>r Value</td>
<td>P Value</td>
<td></td>
</tr>
<tr>
<td>T2 LV</td>
<td>0.55</td>
<td>&lt;.001</td>
<td>0.58</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBNAA</td>
<td>-0.49</td>
<td>&lt;.001</td>
<td>-0.74</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean lesion MD</td>
<td>0.50</td>
<td>&lt;.001</td>
<td>0.57</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean lesion FA</td>
<td>-0.36</td>
<td>.005</td>
<td>-0.48</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean NAWM FA</td>
<td>-0.52</td>
<td>.001</td>
<td>-0.51</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAWM FA peak height</td>
<td>0.42</td>
<td>.001</td>
<td>0.47</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean NAWM MD</td>
<td>0.26</td>
<td>.02</td>
<td>0.38</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAWM MD peak height</td>
<td>-0.31</td>
<td>.007</td>
<td>-0.43</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean NAGM MD</td>
<td>0.52</td>
<td>&lt;.001</td>
<td>0.54</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAGM MD peak height</td>
<td>-0.50</td>
<td>&lt;.001</td>
<td>-0.50</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: EDSS, Expanded Disability Status Scale; FA, fractional anisotropy; LV, lesion volume; MD, mean diffusivity; MR, magnetic resonance; NAGM, normal-appearing gray matter; NAWM, normal-appearing white matter; WBNAA, whole brain N-acetylaspartate.

### Table 4. Univariate Correlations Between MR-Derived Metrics in Patients

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>T2 LV</th>
<th></th>
<th>WBNAA</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r Value</td>
<td>P Value</td>
<td></td>
<td>r Value</td>
<td>P Value</td>
</tr>
<tr>
<td>Mean lesion MD</td>
<td>0.67</td>
<td>&lt;.001</td>
<td>-0.35</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Mean lesion FA</td>
<td>-0.61</td>
<td>&lt;.001</td>
<td>0.40</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Mean NAWM FA</td>
<td>-0.61</td>
<td>&lt;.001</td>
<td>0.48</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>NAWM FA peak height</td>
<td>0.54</td>
<td>&lt;.001</td>
<td>-0.44</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Mean NAWM MD</td>
<td>0.66</td>
<td>&lt;.001</td>
<td>-0.44</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>NAWM MD peak height</td>
<td>-0.64</td>
<td>&lt;.01</td>
<td>0.41</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Mean NAGM MD</td>
<td>0.75</td>
<td>&lt;.001</td>
<td>-0.42</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>NAGM MD peak height</td>
<td>-0.65</td>
<td>&lt;.001</td>
<td>0.32</td>
<td>.004</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: See Table 3.
might also result in increased GM MD values through partial volume averaging from the cerebrospinal fluid. The observed changes of MD histogram-derived measures in SPMS may, therefore, reflect tissue loss and abnormal status of the remaining GM. In this study, the MR acquisition protocol did not include 3-dimensional sequences tailored for volumetric measurements of GM atrophy. However, the preliminary erosion of MD maps and the normalization step of the histogram creation process should have minimized the role played by cerebrospinal fluid partial volume effect on our findings, if any at all.

The results of the multivariate analysis highlighted that increased tissue damage within MS lesions and in the NAWM, together with a further decrease of WBNA, are associated with the evolution from CIS to RRMS. The only predictor for the identification of a patient with SPMS vs a patient with RRMS was increased mean GM diffusivity. These findings again support the notion that the evolution of MS after the clinical onset of the disease is initially driven by the worsening of tissue damage within and outside WM lesions, while it is mainly associated with significant GM involvement when the progressive accumulation of irreversible disability begins. The notion of MS as a “2-stage” disease is supported by findings from epidemiological, postmortem, and in vivo cross-sectional studies.

An additional novel finding of the present study, which included many patients with a wide range of disability, is that a multiparametric model (including disease duration, WBNA, concentration, and GM diffusivity peak height) accounted for about 70% of patients’ Expanded Disability Status Scale score variance. Conversely, about 60% of patients’ disease duration was explained by a combination of age and T2 LV. These results suggest that, whereas a longer disease duration is accompanied by a greater T2-visible lesion burden, only more pronounced neuroaxonal dysfunction and GM damage lead to a severe neurological disability in the late stages of MS. Our inability to fully explain the disability variance has 2 possible explanations. First, we did not investigate the patterns of MS-related damage outside the brain. Cord damage seems to play a pivotal role in determining the clinical progression of MS since its early stages, thus, adding measures of cord atrophy or normal-appearing tissue damage to multiparametric MR models might ameliorate their performance as paraclinical predictors of MS disability. Second, our structural MR-based model could not account for the role played by cortical reorganization in limiting the clinical impact of MS-related central nervous system injury.

The analysis of the relationships between different MR-derived metrics in our cohort of patients with MS revealed that T2 LV and, to a lesser extent, WBNA concentration were significantly correlated with all DT MRI histogram-derived quantities. T2 LV explained, on average, about 40% of MD and FA variance, while the weakest correlations were those between WBNA concentration and GM diffusivity features. This indicates that the burden of macroscopic lesions does contribute to “occult” DT MRI-detectable brain damage, which is not fully translated into a decrease of WBNA levels. As highlighted by a recent longitudinal study, the extent of T2-visible lesion load may also significantly influence the long-term accumulation of clinical and MR-measured irreversible MS damage. Thus, therapeutic strategies aimed at reducing the development of macroscopic MS lesions may, at least partially, also exert a beneficial effect on the long-term clinical disease evolution, because 1 factor contributing to the accrual of tissue damage associated with permanent disability would be removed.

Accepted for Publication: February 16, 2007.

Correspondence: Massimo Filippi, MD, Neuroimaging Research Unit, Department of Neurology, San Raffaele Scientific Institute, via Olgettina 60, 20132 Milan, Italy (filippi.massimo@hsr.it).

Author Contributions: Study concept and design: Rovaris and Filippi. Acquisition of data: Pulizzi, Judica, and Martinelli. Analysis and interpretation of data: Pulizzi, Rovaris, Sormani, Comi, and Filippi. Drafting of the manuscript: Pulizzi, Rovaris, and Filippi. Critical revision of the manuscript for important intellectual content: Pulizzi, Rovaris, Judica, Sormani, Martinelli, Comi, and Filippi. Statistical analysis: Sormani. Administrative, technical, and material support: Pulizzi, Rovaris, Judica, and Martinelli. Study supervision: Filippi.

Financial Disclosure: None reported.

REFERENCES


1167


