Association of Neocortical Volume Changes With Cognitive Deterioration in Relapsing-Remitting Multiple Sclerosis

Maria Pia Amato, MD; Emilio Portaccio, MD; Benedetta Goretti, PhD; Valentina Zipoli, MD; Marco Battaglini, MSc; Maria Letizia Bartolozzi, MD; Maria Laura Stromillo, MD; Leonello Guidi, MD; Gianfranco Siracusa, MD; Sandro Sorbi, MD; Antonio Federico, MD; Nicola De Stefano, MD

Background: We previously reported selective decreases of neocortical volumes in patients with early relapsing-remitting (RR) multiple sclerosis (MS) with mild cognitive impairment, with a good correlation between cortical volumes and cognitive measures.

Objective: To assess the relevance of gray matter changes over time to changes in cognition in RRMS.

Design: A longitudinal survey after 2.5 years. Each patient underwent a magnetic resonance imaging (MRI) protocol identical to that performed at baseline; cognitive performance was reassessed with the Rao Brief Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis.

Setting: Two university MS clinics.

Patients: Of 41 patients with RRMS who participated in the original cross-sectional study, 28 were available for the follow-up evaluation (18 women; mean±SD age, 37.1±8.9 years; mean±SD MS duration, 7.3±2.9 years; mean±SD Expanded Disability Status Scale score, 1.8±1.5).

Main Outcome Measures: We measured the percentage of brain volume changes, normalized cortical volume (NCV) changes, and normalized deep gray matter volume changes on conventional T1-weighted MRIs and changes in lesion load on T2-weighted MRIs. The number of tests failed on the Rao Brief Repeatable Battery were used to classify the patients as cognitively deteriorating or stable or improving.

Results: We identified 12 of 28 cognitively deteriorating and 16 of 28 stable or improving patients. These subgroups did not differ in the mean±SD percentage of brain volume changes (–2.1%±1.2% vs –1.3%±1.3%; P = .11), normalized deep gray matter volume changes (–2.1±2.8 mL vs –0.6±3.1 mL; P = .60), and changes in lesion load on T2-weighted MRIs (1.9±2.6 mL vs 1.6±2.3 mL; P = .73). However, NCV changes were significantly higher in deteriorating than in stable or improving patients (–43.0±18.9 mL vs –17.8±26.6 mL; P = .007). In deteriorating patients, NCV changes were correlated with performance in a verbal fluency test (r = 0.73; P < .001). In a regression model, only NCV changes were significantly associated with deteriorating cognitive performance (odds ratio, 0.8; 95% confidence interval, 0.7-0.9).

Conclusion: Progressive neocortical gray matter loss is relevant to MS-associated cognitive impairment and may represent a sensitive marker of deteriorating cognitive performance in RRMS.

Arch Neurol. 2007;64(8):1157-1161

Background: Cognitive impairment is a core feature of multiple sclerosis (MS), with prevalence estimates ranging from 40% to 65%.1 Deficits typically involve memory, attention, information processing speed, and executive functions. Although there is great intersubject variability, cognitive deficits may be evident from the early stages of the disease and tend to progress over time.2-4 Several studies have associated cognitive dysfunction in MS with white matter (WM) disease. However, correlation with tissue damage as measured by WM lesion load on T2-weighted magnetic resonance imaging (MRI) and magnetization transfer imaging or with metabolic changes as assessed by magnetic resonance spectroscopy has been modest.1 In contrast, the cognitive deterioration seems to increase in parallel with the decrease of brain parenchymal volume rather than with the increase of brain lesion load.5

The recent development of computerized techniques that allow accurate measurements of regional brain volumes has revealed a specific role of gray matter (GM) disease in the development of cognitive impairment in MS.6-8 In particular, in a previous cross-sectional study,6 we reported selective decreases of neocortical volumes in patients with early relapsing-remitting (RR) MS with mild cognitive im-
impaired at time 1 (13 (54%) and 16 of 28 (57%) patients classified as cognitively

cognitive performance was comparable in the 2 groups, with 7 of

follow-up (5.1±5.8 years vs 7.3±2.9 years;

which was significantly shorter in the group of patients lost at

characteristics, with the exception of the disease duration,

mean±SD treatment duration, 1.8±0.6 years).

patients were treated with interferon beta-1a (Rebif, 22 or 44 µg;

No patient was taking psychoactive drugs or substances that might

28 patients were relapse free and were not taking steroids for at

tests at both evaluations. Finally, depression was reassessed with

the Montgomery and Asberg Rating Scale for Depression.11

MRI EXAMINATION

All patients were examined on a Philips Gyroscan operating at 1.5 T (Philips Medical Systems, Best, the Netherlands) using

the same MRI protocol. This protocol was identical to that per-

formed at baseline14 and included (1) a transverse dual-echo,

turbo spin-echo sequence (repetition time, 2075 milliseconds;

echo time 1, 30 milliseconds; echo time 2, 90 milliseconds;

256× 256 matrix; 1 signal average; 250 × 250-mm field of view) that yielded proton density (PD) and T2-weighted im-

ages with 50 contiguous 3-mm-thick slices, acquired parallel to

the line connecting the anterior and posterior commis-

sures; and (2) transverse T1-weighted images (repetition time,

35 milliseconds; echo time, 10 milliseconds; 256×256 ma-

trix; 1 signal average; 250×250-mm field of view) that yielded images of 50 contiguous 3-mm-thick slices, oriented to match

exactly the PD/T2-weighted images acquisition. No major hard-

ware upgrades were performed on the scanner during the study,

and monthly quality assurance sessions confirmed the stabili-

ty of measurements throughout the study.

MRI DATA ANALYSIS

Lesion Volumes

Classification of T2-weighted lesion volume (LV) was performed on each patient by a single observer (M.L.S.), who was

unaware of the patients’ identity, using a segmentation tech-

nique based on user-supervised local thresholding. Lesion bor-

ders were determined primarily on PD-weighted images, but in-

formation from T2- and T1-weighted images was also considered.

Such information was considered because the software used (Jim

3.0; Xinapse System, Leicester, England) offered the ability to

switch among the PD, T2-weighted, and T1-weighted images,

providing the operator with convenient access to the informa-

tion in both data sets, while defining lesions and facilitating the

discrimination of cerebrospinal fluid from the periventricular

plaques. The value of total brain LV was calculated by multiply-

ing lesion area by slice thickness. The coefficient of variation mean

was 5% or less in serial measurements.

Total Brain Volumes

In all patients, longitudinal (2 time points) normalized per-

centage of brain volume changes (PBVs) were estimated on

T1-weighted images with SIENA (structural image evaluation

using normalization of atrophy).14 To estimate changes be-

tween the images, SIENA finds all brain surface edge points using

tissue-type segmentation and then correlates differentiated 1-di-

mensional perpendicular profiles taken around the position of

these points in both images. This technique gives edge motion
to subvoxel accuracy. Thus, the method is relatively insensi-
tive to changes in intensity of tissues from one scan to the next.

Brain atrophy is quantified by taking the mean perpendicular
dge motion over all edge points and converting this to the PBVC.

A variety of validation tests showed the accuracy in measuring

the PBVC to be approximately 0.2%.14

GM Volumes

Because the cross-sectional version of the SIENA method

(SIENAX) allows for regional measures of tissue volume,15 nor-

malized values of cortical GM were separately assessed on both

time points as previously described.16 The difference between

PATIENT POPULATION

Twenty-eight of the original sample of 41 patients with RRMS9

were available for the follow-up evaluation, which took place af-

after a mean±SD period of 2.5±1.1 years. The study sample con-

sisted of 18 women and 10 men whose mean±SD age was

37.1±8.9 years and mean±SD educational level was 11.2±3.3

years. At time 2, all patients still had an RR disease form, the

mean±SD disease duration was 7.3±2.9 years, and the mean±SD

score on the Expanded Disability Status Scale10 was 1.8±1.5. All

patients were treated with interferon beta-1a (Rebif, 22 or 44 µg;

mean±SD treatment duration, 1.8±0.6 years).

Comparing the 28 patients available for the follow-up study

with the 13 patients lost at follow-up, we did not find any sig-

nificant difference in terms of the main demographic and clini-
cal characteristics, with the exception of the disease duration,

which was significantly shorter in the group of patients lost at

follow-up (3.1±5.8 years vs 7.3±2.9 years; P=.03). Baseline cog-
nitive performance was comparable in the 2 groups, with 7 of

13 (54%) and 16 of 28 (57%) patients classified as cognitively

impaired at time 1 (P=.80).

The study was approved by the Ethics Committee of the Fac-

ulty of Medicine of the University of Siena, where patients with

MS underwent the MRI protocol. An informed consent was ob-

tained from all participating patients at baseline and fol-

low-up evaluations.

NEUROPSYCHOLOGICAL ASSESSMENT

For each patient, neuropsychological testing at time 2 was per-

formed within 1 week of MRI examination by the same neu-

ropsychologist who performed the test at time 1 and who was

blinded to the MRI results. The neuropsychological perfor-

mance was reassessed by using the alternate version B of the

Rao Brief Repeatable Battery of Neuropsychological Tests in Mul-
	iple Sclerosis.11 This battery incorporates tests of verbal memory

acquisition and delayed recall (Selective Reminding Test and

Selective Reminding Test–Delayed Recall), spatial memory ac-

quisition and delayed recall (10/36 Spatial Recall Test and 10/36

Spatial Recall Test–Delayed), sustained attention, concentra-

tion and speed of information processing (Paced Auditory Se-

rial Addition Test at 3 and 2 seconds; Symbol Digit Modalities

Test), and verbal fluency on semantic stimulus (Word List Gen-

eration). Consistent with baseline evaluation, a test was con-

sidered failed when the score was at least 2 SDs below the mean

normative values.13 The cognitive performance of the patient

was considered deteriorating when the number of tests failed

at time 2 was greater than at time 1, improving when it was

lower, and stable when the patient failed the same number of

tests at both evaluations. Finally, depression was reassessed with

the Montgomery and Asberg Rating Scale for Depression.13

Downloaded From:  on 10/16/2018

(Reprinted) Arch Neurol/Vol 64 (No. 8), Aug 2007   www.archneurol.com

©2007 American Medical Association. All rights reserved.

©2007 American Medical Association. All rights reserved.
measurements of 2 time points of the same patients provided changes of normalized cortical volume (NCV) and percentage of changes in NCV. Also, selective measures of deep GM were obtained by using the standard space-based masks that selectively included the GM present in the basal ganglia. In each patient, the difference between these measures at the 2 time points provided the changes in normalized deep GM.

STATISTICAL ANALYSIS

Group analyses were performed using the t test for unpaired samples, the nonparametric Mann-Whitney test, and a χ² test when appropriate. Relationships between MRI and cognitive variables were assessed using the nonparametric Spearman rank order correlation. Data below the .05 level were considered statistically significant. To assess possible predictors of the patient’s cognitive outcome at time 2 (deteriorating or stable or improved), we used a stepwise logistic regression model that included age, educational level, number of tests failed at time 1, PBVC, T2-weighted LV changes, NCV changes, and normalized deep GM changes as possible predictors. The SPSS software version 12.0 running on Windows (SPSS Inc, Chicago, Illinois) was used for the analysis.

RESULTS

On the basis of their neuropsychological performance on the Rao Brief Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis, 12 patients (43%) were classified as deteriorating, 11 as stable (39%), and 5 as improving (18%). The number and type of tests failed at time 2 and comparisons with time 1 are reported in Table 1 and Table 2. Comparing patients with a deteriorating cognitive performance with patients with a stable or improving performance, we did not find any significant difference in terms of the main clinical and demographic characteristics (Table 3). Moreover, 4 patients in the cognitively stable or improving subgroup and 3 in the cognitively deteriorating subgroup (25%) were classified as depressed on the Montgomery and Asberg Rating Scale for Depression.

When comparing patients with MS and a deteriorating cognitive performance and those with a stable or improving performance, quantitative MRI analysis showed no significant differences in mean ± SD T2-weighted LV changes (1.9 ± 2.6 mL vs 1.6 ± 2.3 mL; P = .73), PBVC (−2.1% ± 1.2% vs −1.3% ± 1.3%; P = .11), and mean ± SD normalized deep GM changes (−2.1 ± 2.8 mL vs −0.6 ± 3.1 mL; P = .60). However, both values of NCV changes (Figure) and percent-

![Table 1. Number of Tests Failed at Time 1 and Time 2](image1)

<table>
<thead>
<tr>
<th>No. of Tests Failed</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time 1</td>
</tr>
<tr>
<td>0</td>
<td>12/28</td>
</tr>
<tr>
<td>1</td>
<td>10/28</td>
</tr>
<tr>
<td>2</td>
<td>2/28</td>
</tr>
<tr>
<td>3</td>
<td>2/28</td>
</tr>
<tr>
<td>4</td>
<td>2/28</td>
</tr>
<tr>
<td>5</td>
<td>0/28</td>
</tr>
<tr>
<td>6</td>
<td>0/28</td>
</tr>
<tr>
<td>7</td>
<td>0/28</td>
</tr>
</tbody>
</table>

![Table 2. Type of Tests Failed at Time 1 and Time 2](image2)

<table>
<thead>
<tr>
<th>Test</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRT-LTS</td>
<td>3</td>
</tr>
<tr>
<td>SRT-CLTR</td>
<td>0</td>
</tr>
<tr>
<td>SPART</td>
<td>3</td>
</tr>
<tr>
<td>SDMT</td>
<td>2</td>
</tr>
<tr>
<td>PASAT 3</td>
<td>3</td>
</tr>
<tr>
<td>PASAT 2</td>
<td>6</td>
</tr>
<tr>
<td>SRT-D</td>
<td>4</td>
</tr>
<tr>
<td>SPART-D</td>
<td>6</td>
</tr>
<tr>
<td>WLG</td>
<td>1</td>
</tr>
</tbody>
</table>

Abbreviations: PASAT, Paced Auditory Serial Addition Test; SDMT, Symbol Digit Modalities Test; SPART, 10/36 Spatial Recall Test; SPART-D, 10/36 Spatial Recall Test-Delayed; SRT-CLTR, Selective Reminding Test–Consistent Long-Term Retrieval; SRT-D, Selective Reminding Test–Delayed Recall; SRT-LTS, Selective Reminding Test–Long-Term Storage; WLG, Word List Generation.

![Table 3. Clinical and Demographic Characteristics of Deteriorating and Stable or Improving Patients](image3)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cognitively Deteriorating</th>
<th>Cognitively Stable or Improving</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients (M/F)</td>
<td>12 (6/10)</td>
<td>16 (4/8)</td>
<td>.82</td>
</tr>
<tr>
<td>Age, mean ± SD, y</td>
<td>38.8 ± 10.8</td>
<td>35.8 ± 7.4</td>
<td>.42</td>
</tr>
<tr>
<td>Education, mean ± SD, y</td>
<td>12.3 ± 2.1</td>
<td>10.3 ± 3.9</td>
<td>.11</td>
</tr>
<tr>
<td>Disease duration, mean ± SD, y</td>
<td>7.5 ± 2.9</td>
<td>7.1 ± 2.9</td>
<td>.68</td>
</tr>
<tr>
<td>EDSS score, mean ± SD</td>
<td>1.7 ± 0.8</td>
<td>1.9 ± 0.9</td>
<td>.51</td>
</tr>
</tbody>
</table>

Abbreviation: EDSS, Expanded Disability Status Scale.

![Figure. Mean normalized cortical volume (NCV) changes in cognitively stable or improving and deteriorating patients. Error bars indicate SD.](image4)

Figure. Mean normalized cortical volume (NCV) changes in cognitively stable or improving and deteriorating patients. Error bars indicate SD.
were not significantly different between cognitively deteriorating or stable or improving patients, and these MRI parameters did not seem to correlate with cognitive scores. Moreover, in a stepwise logistic regression model, changes in neocortical volumes were the only significant correlate of deteriorating cognitive performance. Similar results were also obtained by using a more conservative approach in the definition of cognitive change (ie, failure of 2 or more tests compared with baseline performance). This finding further supports our conclusions on the relevance of GM volume changes to cognitive impairment in patients with early RRMS. Therefore, in our sample, cognitive deterioration seemed to proceed in parallel with the decrease of neocortical volumes.

Several recent pathological and imaging studies have recognized that neocortical abnormalities are relevant in MS, can be detected even in the earliest stages of the disease, and may be, at least in part, unrelated to WM lesion accumulation. It is not clear, however, whether the cortical volume loss is associated with a diffuse neurodegeneration of the neocortex or the presence of neocortical inflammatory lesions that are rarely detected by conventional MRI. Independently of the underlying pathological mechanisms, the present data add to prior evidence in supporting the relevance of GM disease in MS. They also extend previous information by reporting evidence that MS-related cognitive dysfunction is closely associated with progressive neocortical changes in early RRMS, even after a relatively short follow-up period.

In previous cross-sectional studies, smaller GM volumes were associated with poorer performances on specific neuropsychological tests. In particular, in one of these studies, performance on the Paced Auditory Serial Addition Test was correlated with GM volume in brain regions associated with working memory and executive functions. In the present longitudinal study, we found only a significant correlation between NCV and changes in scores on a verbal fluency test. This finding may be due to the small sample size and limitations of serial administration of neuropsychological tests during a short follow-up period. Moreover, our GM volume estimates did not allow assessment of specific brain regions that may also explain the paucity of correlations found in our study. In this context, studies of larger cohorts of patients during a longer period and the assessment of specific GM regions may improve correlations between neuropsychological and MRI findings.

Interestingly, the results of our study did not show a clear involvement of the deep GM in the pathological process, since we observed comparable deep GM volume decreases in patients with stable or deteriorating MS. However, structural and metabolic abnormalities of the deep GM have been documented in a number of previous MS studies and its pathological features have been indicated as a possible factor contributing to MS-related cognitive impairment. It is possible that the small sample size used in the present study and the technical difficulties in performing an accurate segmentation in this brain region with the present postprocessing imaging procedures may have contributed to our negative results. Probably, studies on larger patient populations using model-based parcellation or diffusion tractography for a specific evaluation of subcortical structures could provide a more appropriate answer to this interesting issue.
In this study, we used quantitative MRI to investigate the underlying pathological basis of cognitive dysfunction in MS. From our results, neocortical atrophy emerges as a significant determinant of progressive cognitive dysfunction to a greater extent than can be explained by conventional WM lesion assessment. Therefore, MS-related cognitive impairment may be, at least in part, due to mechanisms that are not related to focal WM lesion genesis. However, the relationship between brain disease and cognition remains complex, and it likely involves multiple factors, brain structures, and reciprocal connections. Further research is needed to clarify these complexities.

Accepted for Publication: January 29, 2007.
Correspondence: Maria Pia Amato, MD, Department of Neurology, University of Florence, Viale Morgagni, 85-50134 Florence, Italy (maria.pia.amato@unifi.it).

Author Contributions: Study concept and design: Amato and De Stefano. Acquisition of data: Amato, Portaccio, Goretti, Zipoli, Bartolozzi, Stromillo, Guidi, Siracusa, and De Stefano. Analysis and interpretation of data: Amato, Portaccio, Battaglini, Stromillo, Sorbi, Federico, and De Stefano. Drafting of the manuscript: Amato, Portaccio, Goretti, Zipoli, Battaglini, Bartolozzi, Stromillo, Guidi, Siracusa, and De Stefano. Critical revision of the manuscript for important intellectual content: Amato, Sorbi, and Federico. Statistical analysis: Portaccio and De Stefano. Administrative, technical, and material support: Guidi. Study supervision: Amato, Goretti, Zipoli, Stromillo, Sorbi, and Federico.

Financial Disclosure: None reported.

REFERENCES