Chapter 4

Analysis of putative causal relations between white matter and grey matter abnormalities
4.1

Grey and white matter atrophy in a large cohort of multiple sclerosis patients: Relation to MRI parameters and impact on clinical disability

Stefan D Roosendaal
Kerstin Bendfeldt
Hugo Vrenken
Chris H Polman
Stefan Borgwardt
Ernst W Radue

Ludwig Kappos
Daniel Pelletier
Stephen L Hauser
Paul M Matthews
Frederik Barkhof
Jeroen JG Geurts

Submitted for publication
**ABSTRACT**

The abundance of grey matter damage in multiple sclerosis and its clinical relevance is currently fully recognized. Although previous studies have provided information regarding determinants of grey matter atrophy and its relationship with disability in multiple sclerosis, the majority of these studies were limited by their sample size. The objectives of the current study were therefore to: 1) measure grey matter and white matter atrophy across different disease phenotypes in a large cohort (N=927) of multiple sclerosis patients; 2) identify magnetic resonance imaging parameters that determine grey matter atrophy; 3) study grey matter and white matter atrophy as explanatory variables for clinical impairment. Magnetic resonance imaging data of 95 clinically isolated syndrome, 657 relapsing remitting, 125 secondary progressive and 50 primary progressive multiple sclerosis patients from three centres (GeneMSA consortium) were acquired. Grey and white matter volumes, normalized for head size, were determined, together with T2 hyperintense and T1 hypointense lesion volumes. Physical disability was assessed with the Expanded Disability Status Scale (EDSS), cognitive impairment with the Paced Auditory Serial Addition Task (PASAT). Data were analyzed using analysis of covariance and multiple regression. Grey matter atrophy, corrected for age, sex and centre effects, was more prominent in relapsing remitting patients (0.80 ± 0.05L) than in clinically isolated syndrome patients (0.82 ± 0.05L). Greater relative atrophy was found in secondary progressive patients (0.77 ± 0.05L). In contrast, white matter atrophy in secondary progressive patients was comparable to that in relapsing remitting patients. T2 lesion volume was an independent predictor of grey matter atrophy (beta=-0.27; p<0.001). Grey matter atrophy was the strongest independent predictor of physical disability and cognitive impairment. Our findings in a large cohort confirm that grey matter atrophy is greater in the secondary progressive phase than in relapsing remitting disease, whereas white matter atrophy is similar between these groups. Grey matter atrophy explained both physical and cognitive impairment better than white matter atrophy, and is itself partly predicted by T2 lesion volume.
INTRODUCTION

Multiple sclerosis (MS), a disease commonly diagnosed in the prime of life and often leading to chronic disability, has until recently been regarded as a typical white matter (WM) disease. This has changed, however, with the introduction of new immunohistochemical staining procedures, which showed that grey matter (GM) demyelination and axonal loss were frequent and abundant. [1,2]

Clinically, cortical damage is of importance because MRI-visible focal demyelination in the WM cannot explain the entire array of clinical impairment in MS patients. GM atrophy was significantly associated with physical disability [3,4] and with cognitive decline. [5,6] To what extent GM atrophy explains disability better than focal WM (or GM) lesions or than WM atrophy has not yet been clearly investigated in patient samples that are representative of the total MS population.

In vivo imaging of focal GM demyelination has been challenging, despite the use of advanced MRI techniques [7,8] and higher field strengths, [9] and many cortical GM lesions remain undetected on conventional MRI. [10] Besides visualization of focal cortical lesions, quantitative MR techniques like diffusion tensor imaging (DTI), [11] magnetization transfer imaging [12] and MR spectroscopy, [13] have been applied to image the 'normal appearing' GM (as defined with MRI). Atrophy measures are regarded to be the most robust quantitative measures for structural damage in MS. [14]

In a seminal histopathological study of acute RR, SP and PP MS cases, the evolution of grey matter damage over subsequent disease stages was recently illustrated: although GM atrophy exists already early in the disease, it becomes much more prominent in progressive MS. [15] In vivo, the finding of grey matter atrophy in early MS cases, with a significant and disproportionate increase in patients with more advanced disease, has also been shown. [3,16] In contrast, WM atrophy showed a relatively constant accrual over time. [17]

While it has been suggested that accumulation of demyelination and neuro-axonal damage in the GM occurs secondarily to WM tract damage, [18] a number of studies now indicate that an independent and partly overlapping occurrence of GM and WM damage may be more likely. [15,19] This may also be true for PPMS. [3]
Although previous studies have provided clues regarding determinants of GM atrophy and the relationship with clinical disability in MS, the majority of these studies were limited by relatively small and somewhat biased patient samples. A much larger cohort of over 900 MS patients, prospectively assessed through the GeneMSA consortium, [20] provided the opportunity to further investigate the relationship between disease progression and atrophy. More specifically, we test the hypotheses that GM atrophy and not WM atrophy is a characteristic of progressive MS, and that it explains clinical disability better. Additionally, we investigate putative MRI determinants of GM atrophy.

Table 1 | Disease type descriptives.

<table>
<thead>
<tr>
<th></th>
<th>CIS</th>
<th>RR</th>
<th>SP</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (% of total patient group)</td>
<td>95 (10.2)</td>
<td>657 (70.9)</td>
<td>125 (13.5)</td>
<td>50 (5.4)</td>
</tr>
<tr>
<td>mean age, yrs (SD)$^1$</td>
<td>41.8 (9.4)</td>
<td>42.1 (9.8)</td>
<td>50.0 (13.5)</td>
<td>50.4 (8.4)</td>
</tr>
<tr>
<td>male: female ratio</td>
<td>1: 1.8</td>
<td>1: 2.5</td>
<td>1: 1.4</td>
<td>1: 0.9</td>
</tr>
<tr>
<td>proportion originating from UCSF/BAS/AMS</td>
<td>0.84/0.10/0.06</td>
<td>0.53/0.26/0.21</td>
<td>0.33/0.33/0.34</td>
<td>0.28/0.24/0.48</td>
</tr>
<tr>
<td>mean disease duration, years (SD)$^2$</td>
<td>2.2 (3.3)</td>
<td>10.0 (8.0)</td>
<td>18.9 (9.0)</td>
<td>10.3 (8.4)</td>
</tr>
<tr>
<td>mean EDSS (SD)$^3$</td>
<td>1.3 (1.0)</td>
<td>2.2 (1.4)</td>
<td>5.2 (1.4)</td>
<td>4.8 (1.6)</td>
</tr>
<tr>
<td>median PASAT (IQR)$^4$</td>
<td>51 (45-56)</td>
<td>49 (40-56)</td>
<td>43.5 (33.3-51)</td>
<td>50.5 (38.3-55)</td>
</tr>
<tr>
<td>median T2 lesion volume, mL (IQR)$^5$</td>
<td>0.8 (1.4-2.2)</td>
<td>2.3 (0.8-6.1)</td>
<td>4.0 (1.8-12.6)</td>
<td>1.7 (0.5-4.8)</td>
</tr>
<tr>
<td>median T1 black hole volume, mL (IQR)$^6$</td>
<td>0.3 (0-0.9)</td>
<td>0.7 (0.2-2.2)</td>
<td>1.7 (0.5-5.5)</td>
<td>0.4 (0.1-1.5)</td>
</tr>
<tr>
<td>median T1-T2 ratio (IQR)$^7$</td>
<td>0.3 (0.1-0.5)</td>
<td>0.3 (0.2-0.5)</td>
<td>0.4 (0.3-0.5)</td>
<td>0.2 (0.1-0.4)</td>
</tr>
<tr>
<td>no. of patients on DMT (%)</td>
<td>28 (29.5)</td>
<td>373 (56.9)</td>
<td>46 (36.8)</td>
<td>3 (6.0)</td>
</tr>
</tbody>
</table>

CIS: clinically isolated syndrome; RR: relapsing remitting; SP: secondary progressive; PP: primary progressive; EDSS: Expanded Disability Status Scale; PASAT: Paced Auditory Serial Addition Task; UCSF: San Francisco; BAS: Basel; AMS: Amsterdam; SD: standard deviation; IQR: interquartile range. Significant differences between groups (p<0.05): $^1$CIS vs. (SP, PP), RR vs. (SP, PP); $^2$CIS vs. (RR, SP, PP), SP vs. (PP, RR); $^3$CIS vs. (SP, PP), RR vs. (SP, PP); $^4,^5,^6$between all groups except between RR and PP; $^7$SP vs. (CIS, RR, PP).
METHODS

Subjects
Patients were recruited from three clinical centres participating in the GeneMSA consortium: the VU University Medical Centre in Amsterdam; the University Hospital in Basel; and the University of California San Francisco (UCSF). Clinically isolated syndrome (CIS) patients were defined as those with a first neurological event lasting longer than 48 hours and involving optic nerve, spinal cord, brainstem, or cerebellum, with at least two hyperintense lesions present on the T2-weighted MR image. Patients with a diagnosis of clinically definite MS [21] were classified either as relapsing remitting (RR) MS, secondary progressive (SP) MS, defined by at least six months of worsening neurological disability not explained by clinical relapse, or primary progressive (PP) MS, defined by progressive clinical worsening for more than 12 months from disease onset without any relapses and the presence of more than 2 oligoclonal bands or an elevated IgG index in the cerebrospinal fluid (CSF). A group of progressive relapsing patients was excluded from the current analysis because of its small size (n=12). Patients with a clinical relapse or glucocorticosteroids treatment within the month previous to enrolment were excluded. The concomitant use of disease modifying therapies (DMT) for MS was permitted. In all included subjects, disability was assessed with the Expanded Disability Status Scale (EDSS) [22] and the Paced Auditory Serial Addition Task (PASAT), [23] and brain MRI scans were performed. The study protocol was approved by the institutional ethics review boards of the clinical centres and all patients gave written informed consent prior to participation.

Magnetic resonance imaging protocol
MR imaging was performed on two 1.5T MR systems (Amsterdam: Siemens Vision; Basel: Siemens Avanto) and one 3.0T MR system (UCSF: GE Excite). For brain volume measurement, 3D-T1 images were acquired (TR: 7-20.8 ms; TE: 2-4 ms; TI: 300-400 ms), consisting of isotropic 1 x 1 x 1 mm³ voxels. Additionally, dual echo proton density (PD)-T2-weighted images (TR: 2000-4000 ms; TE: 14-20 / 80-108 ms), with interleaved axial 3.0 mm-thick slices and an in-plane resolution of 1.0×1.0 mm² or 0.5×0.5 mm² (UCSF), were acquired. Lastly, post-contrast T1-weighted spin-echo images (TR: 467-650 ms; TE: 8-17
ms; axial 3.0 mm-thick slices with an in-plane resolution of 1.0×1.0 mm² or 0.9×0.9 mm² [UCSF]) were obtained.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>(Multiple) linear regression for NGMV, corrected for centre.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Linear regression</td>
</tr>
<tr>
<td></td>
<td>Beta</td>
</tr>
<tr>
<td>Age</td>
<td>-0.52</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.17</td>
</tr>
<tr>
<td>Disease duration</td>
<td>-0.4</td>
</tr>
<tr>
<td>logT2LV</td>
<td>-0.38</td>
</tr>
<tr>
<td>logT1:T2</td>
<td>-0.27</td>
</tr>
<tr>
<td>NWMV</td>
<td>0.14</td>
</tr>
</tbody>
</table>

NGMV: normalized grey matter volume; NWMV: normalized white matter volume; logT2LV: normal-transformed T2 lesion volume; logT1:T2: normal-transformed T1-T2 ratio. When centre was the only independent variable, R² was 0.20. Adjusted R² of the final model was 0.59.

MRI measurements

Brain volume analyses were performed at the Image Analysis Centre in Amsterdam. SIENAX [24] (version 2.2) from the FMRIB Software Library was used to estimate normalized grey matter volume (NGMV) and normalized white matter volume (NWMV). For this purpose, SIENAX registers each individual scan to MNI-152 standard space, using the skull as a scaling constraint. The volumetric scaling factor is then used to correct GM and WM volumes obtained from the automated tissue segmentation, to volumes normalized for head size. Scans of all subjects and the resulting segmentation maps were visually inspected for scan quality (such as noise, artefacts and tissue contrast) and segmentation quality (tissue misclassification), respectively.

Marking and measurement of focal WM lesions was performed at the University Hospital in Basel, using commercial semi-automatic software (AMIRA 3.1.1; Mercury Computer Systems Inc). T2 hyperintense lesions and T1 hypointense lesions (black holes) were manually outlined on the PD images and on post contrast T1-weighted spin echo images, respectively. Subsequently, volumes were calculated for these lesion categories.
Statistical analysis

Statistical analyses in this study were performed using SPSS version 15.0 for Windows (SPSS, Chicago, USA). Comparisons of the demographical data between the disease types were made using the Mann-Whitney U test or the Student's t-test when appropriate. Values are reported as mean ± SD, unless indicated otherwise. Statistical models in this paper were always corrected for centre.

We examined disease type differences for NGMV and NWMV using an ANCOVA design correcting for age, sex and centre. The estimated marginal means were pair-wise compared between the disease types using post-hoc Bonferroni correction.

Associations between NGMV and the other variables were assessed, correcting for centre. Subsequently, an explanatory multiple linear regression model for NGMV was constructed by entering the other MRI variables together and removing them one by one using manual backward stepwise exclusion until the remaining variables were significant at p<0.1. This multiple linear regression model was corrected for age, sex, disease duration and centre. Before entering the regression models, T2 lesion volume was transformed to improve normality using the logarithm with base 10 (logT2LV). T1 hypointense lesion volume was not used in the models, because a high correlation of rho=0.9 existed between T2 lesion volume and T1 hypointense lesion volume. Instead, to avoid collinearity in the models, the ratio between T1 and T2 lesion volume was used. This ratio was also log-transformed to improve normality (logT1:T2).

The relative predictive value of GM and WM brain atrophy, together with logT2LV and logT1:T2, on EDSS and on PASAT was assessed using ordinal logistic regression and multiple linear regression, respectively. Again, this was done separately for each variable correcting for centre, and subsequently by entering the independent variables together and removing them one by one using manual backward stepwise exclusion until the remaining MRI variables were significant at p<0.1. For this purpose, EDSS was categorized into four classes: <2.0; ≥2.0 and <4.0; ≥4.0 and <6.0; ≥6.0. Odds ratios (OR) resulting from the ordinal logistic regression are reported per standard deviation for each independent variable. PASAT was transformed to normality using a square-root transformation (sqrtPASAT). Similar to the multiple regression
model used for NGMV, age, sex, disease duration and centre were included as covariates for the EDSS and PASAT models.

To rule out that the final models were biased by the phenomenon called ‘pseudo-atrophy’, all model analyses were repeated without the patients who started their DMT within one year before the investigation.

RESULTS

Patient descriptives
A total of 977 MS patients with CIS, RR, PP or SP disease type were included in the study. Slightly more than half of the patients were enrolled at UCSF; Amsterdam and Basel contributed equally. The descriptive data provided in Table 1 account for the 927 patients (95%) for whom reliable brain volume measurements could be obtained. The majority of these patients had an RR disease type (657 patients; 70.9%); there were 125 (13.5%) SP patients, 95 (10.2%) CIS patients and 50 (5.4%) PP patients. In Amsterdam, relatively more progressive patients were included compared to the other sites, whereas 80 of the 95 CIS patients were from UCSF. As expected, mean age and EDSS of SP and PP patients were significantly higher than those of CIS and RR patients. Median PASAT-score was lowest in SP patients, and the mean disease duration of SP patients significantly higher than that of the other patient groups. No significant differences in disease duration existed between sites. Lesion loads differed significantly between all disease types, except between PP and RR patients. DMT (Avonex, Rebif, Betaseron or Copaxone) was received by 450 (48.7%) out of the total of 927 patients; 105 patients had started their DMT within the year before the investigation.

Clinical disease types and brain volumes
Whereas NGMV and NWMV were not significantly different between patients from Amsterdam (NGMV: 0.77 ± 0.06; NWMV: 0.78 ± 0.06) and Basel (NGMV: 0.77 ± 0.07; NWMV: 0.78 ± 0.05), mean NGMV of patients from UCSF was significantly higher (0.83 ± 0.06; p<0.001) and mean NWMV significantly lower (0.69 ± 0.04; p<0.001) compared to the other two sites.
Table 3 | (Multiple) ordinal regression for EDSS, corrected for centre.

<table>
<thead>
<tr>
<th>Ordinal regression</th>
<th>Multiple ordinal regression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PE</strong></td>
<td><strong>OR</strong></td>
</tr>
<tr>
<td>Age</td>
<td>0.72</td>
</tr>
<tr>
<td>Sex (female:0; males:1)</td>
<td>-0.26</td>
</tr>
<tr>
<td>Disease duration</td>
<td>0.8</td>
</tr>
<tr>
<td>NGMV</td>
<td>-0.87</td>
</tr>
<tr>
<td>NWMV</td>
<td>-0.29</td>
</tr>
<tr>
<td>logT2LV</td>
<td>0.44</td>
</tr>
<tr>
<td>logT1:T2</td>
<td>0.38</td>
</tr>
</tbody>
</table>

EDSS: expanded disability status scale; PE: parameter estimate; OR: odds ratio; NGMV: normalized grey matter volume; NWMV: normalized white matter volume; logT2LV: normal-transformed T2 lesion volume; logT1:T2: normal-transformed T1-T2 ratio. PE and OR are reported per standard deviation for each independent variable, except for sex, where the OR of 0.8 indicates a lower EDSS in males compared to females. When centre was the only independent variable, Nagelkerke’s R² was 0.21. Nagelkerke’s R² of the final multiple ordinal regression model was 0.40.

In all centres, uncorrected NGMV values were lower in RRMS patients than in CIS patients, further decreasing in SPMS patients (Figure 1A). While uncorrected NWMV also differed significantly between CIS patients and RRMS patients, it did not differ between RRMS and SPMS patients (Figure 1B). The estimated NGMV (ANCOVA F: 14.5), corrected for age, sex and centre (Figure 1C) was highest in CIS (0.82 ± 0.05 L), decreasing in RR (0.80 ± 0.05 L) and PP (0.80 ± 0.05 L), and further decreasing in SP patients (0.77 ± 0.05 L). Post-hoc pair-wise comparisons showed significance between all disease subtypes, except for PP vs. RR and CIS patients. Estimated NWMV (F: 7.4) was also highest in CIS patients (0.77 ± 0.05 L), in the other disease types the estimated NWMV had the same value and SD (0.75 ± 0.05 L). Differences in NWMV were significant between CIS vs. RR, and CIS vs. SP (Figure 1D).

**Predictors of GM atrophy**

NWMV, logT2LV and logT1:T2 were each significantly associated with NGMV (Table 2). Multiple regression analysis (Table 2) revealed that, after correction for age, gender, disease duration and centre, logT2LV was the only significant explanatory MRI variable of NGMV (beta: -0.27; p<0.001). LogT1:T2 and
NWMV were not included in the model, because their p-values were higher than 0.1. The model accounted for 59% (adjusted $R^2 = 0.59$) of the variance in NGMV.

**Figure 1** | **A**: Uncorrected NGMV is shown for each disease type (CIS: clinically isolated syndrome; RR: relapsing remitting; SP: secondary progressive; PP: primary progressive) and each centre (SF: San Francisco; AMS: Amsterdam; BAS: Basel). In each centre, NGMV can be seen to be lower in RR patients than CIS patients, and to further decrease in SP patients. **B**: Uncorrected NWMV is shown for each disease type and each centre. **C**: Estimated NGMV is shown here, corrected for age, sex and centre. Comparisons that are significant after Bonferroni correction are indicated by ** : $p<0.001$; * $p<0.01$. **D**: The estimated NWMV corrected for age, sex and centre, is shown here. Comparisons that are significant after Bonferroni correction are indicated by ** : $p<0.001$; * $p<0.01$. 
Predictors of disability

Correlations between EDSS and MRI variables (Table 3) were significant for both NGMV and NWMV, as well as for logT2LV and logT1:T2. In the multiple ordinal regression model (Nagelkerke’s R² = 0.40), NGMV was the strongest MRI predictor of EDSS (OR: 0.7; p<0.001), logT2LV was a weaker predictor (Table 3). The OR of NGMV can also be expressed as an odds increase of 1.5 in having greater disability per 1 SD smaller NGMV. NWMV and logT1:T2 were not included in the model.

Table 4 | (Multiple) linear regression for PASAT, corrected for centre.

<table>
<thead>
<tr>
<th></th>
<th>Linear regression</th>
<th></th>
<th>Multiple linear regression</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>P-value</td>
<td>adjusted R²</td>
<td>Beta</td>
</tr>
<tr>
<td>Age</td>
<td>-0.22</td>
<td>&lt;0.001</td>
<td>0.06</td>
<td>-0.05</td>
</tr>
<tr>
<td>Sex</td>
<td>0.06</td>
<td>0.07</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>Disease duration</td>
<td>-0.26</td>
<td>&lt;0.001</td>
<td>0.08</td>
<td>-1.4</td>
</tr>
<tr>
<td>logT2LV</td>
<td>-0.15</td>
<td>&lt;0.001</td>
<td>0.03</td>
<td>-</td>
</tr>
<tr>
<td>logT1:T2</td>
<td>-0.14</td>
<td>&lt;0.001</td>
<td>0.03</td>
<td>-</td>
</tr>
<tr>
<td>NWMV</td>
<td>0.19</td>
<td>&lt;0.001</td>
<td>0.03</td>
<td>0.11</td>
</tr>
<tr>
<td>NGMV</td>
<td>0.28</td>
<td>&lt;0.001</td>
<td>0.08</td>
<td>0.19</td>
</tr>
</tbody>
</table>

PASAT: Paced Auditory Serial Addition Task; NGMV: normalized grey matter volume; NWMV: normalized white matter volume; logT2LV: normal-transformed T2 lesion volume; logT1:T2: normal-transformed T1-T2 ratio. When centre was the only independent variable, R² was 0.01. The adjusted R² of the final model was 0.11.

Also for the PASAT, correlations with all MRI variables (Table 4) were significant. The multiple linear regression model had an adjusted R² of 0.11 and contained NGMV as the strongest predictor (Table 4, beta: 0.19; p<0.001), and contained NWMV as well (beta: 0.11; p=0.02).

Results of all of the above-described model analyses did not change when the 105 patients who started their DMT less than one year before the investigation were excluded.
DISCUSSION

Our study, involving a large cohort of CIS, RR, SP and PP patients, shows that GM atrophy grows significantly in secondary progressive MS, while WM atrophy seems to be relatively constant. Moreover, GM atrophy explains physical disability and cognitive impairment, as measured by the EDSS and the PASAT respectively, better than WM atrophy. T2 lesion volume in the WM was the best MRI predictor of GM atrophy.

Cortical involvement has been demonstrated even in the early phase and was related to conversion from CIS to clinically definite MS. [25,26] Demyelination of axons in the GM was shown in both RRMS and acute MS, [15] although it is more prominent in the later, chronic progressive stages. Recent studies demonstrated that the progression of GM atrophy over time is much larger than the progression of WM atrophy. [16,17] Although cross-sectional, our results support these findings in a large unselected group of patients.

The pathological substrate of GM atrophy as measured with in vivo MRI is largely unknown. Clues that GM atrophy may mostly be based on neuronal and glial damage come from a post mortem study, [27] in which strict regional associations between cortical demyelination and atrophy were not found. Relations between GM atrophy and T2 lesion volume have been reported previously, [3,28] and in our study T2 lesion volume was the strongest independent predictor of NGMV (a beta of -0.27 was found). Although this finding may be an argument for the hypothesis that GM damage results from damage to WM tracts, an equally likely possibility is that GM atrophy and T2 lesion volume occur independently but mirror the general disease process in certain phases of the disease.

GM atrophy explained physical disability as measured by the EDSS and cognitive impairment, measured by the PASAT, better than WM atrophy or T2 lesion volume. This may be somewhat unexpected for the EDSS, given the fact that it is well-known to incompletely cover cortical functions. On the other hand, associations between GM atrophy and EDSS scores have been reported by several previous studies. [17,29,30] The PASAT measures working memory and sustained attention, and is often used as a measure of cognitive impairment in MS. [31,32] Our finding that GM atrophy is the strongest predictor of both
physical disability and cognitive impairment emphasizes the clinical relevance of this measure.

Now that it is becoming increasingly clear that MS not only involves inflammatory demyelination, but also substantial and early neurodegeneration, new neuroprotective and reparative treatments are sought. Clinical trials necessary for these treatments rely on imaging markers as surrogate outcomes. Although whole-brain atrophy has proven to be a reproducible and sensitive marker of disease it lacks pathological specificity. GM atrophy may be a more specific marker of neurodegenerative processes, as it occurs faster as disease progresses than WM atrophy. Furthermore, GM atrophy could be a more specific marker than WM atrophy, since it has been suggested to be less influenced by the so-called pseudo-atrophy phenomenon, which may occur supposedly by lessening of oedema when interferon or corticosteroid therapy is initiated. In our study, this potential nuisance was not likely to have an effect, because our results did not change when patients who started interferon therapy within one year of the investigation were excluded. Furthermore, none of our patients used corticosteroids in the month prior to the investigation.

Measurement of brain volume using SIENAX has been shown to be consistent across centres and is relatively insensitive to different MR systems of the same field strength. In our study, NWMV was lower and NGMV was higher in patients from the UCSF when compared to those of patients from the other two sites. A 3T MR system was used in the UCSF and 1.5T MR systems were used in the other two centres; this field strength difference may be responsible. To limit the influence of this difference, all models in our study were corrected for centre. By visual inspection of SIENAX segmentation results in our study, we avoided influences of artefacts on the volumes. Some studies have reported that frontal and temporal cortical areas may exhibit more severe pathology than others. Such regional atrophy or thickness studies might have the advantage of an increased sensitivity, since the results are not hindered by cortical areas that do not suffer from damage. However, post mortem studies indicate that in advanced disease the cortex is globally affected, up to 68% of the cortical area; it may therefore well be that regional GM measures do not have added value in the progressive phase.

A limitation of our study was that discrimination between cortical and deep GM volumes was not made. In addition to cortical GM atrophy, recent
studies have shown that deep GM structures, [41] such as the thalamus, [42,43] and mixed WM-GM structures such as the hippocampus, [44] are involved in MS. The relation between deep GM atrophy and cortical GM atrophy remains to be elucidated, as well as the accrual of deep GM atrophy over the disease types. Another potential limitation of our study is that we did not correct for lesion misclassification. In the automated segmentation process some of the WM lesions may have been misclassified as cerebrospinal fluid (CSF) or as GM, which can result in too low WM and too high GM volume estimates. However, since our results are opposing the effect one would expect from WM lesion misclassification, i.e., GM volume is further decreased in secondary progressive MS, our results are unlikely to be false-positive due to this phenomenon.

In conclusion, our study confirms in a large group of patients the dominance of GM pathology in (secondary) progressive MS. In addition, we show that GM atrophy is best predicted by T2 lesion volume and that GM atrophy explains clinical disability better than WM atrophy, which is important for future clinical trial design.

Acknowledgements
The authors would like to thank all people involved in the GeneMSa consortium. This work was supported by the Dutch MS Research Foundation [grant numbers 02-358b, 06-592 to S.D.R., 05-358c to H.V. and J.J.G.G.], and partially funded by the Glaxo Smith Kline Clinical Imaging Centre.
REFERENCES


