Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability

Stefan D Roosendaal¹, Kerstin Bendfeldt², Hugo Vrenken¹, Chris H Polman¹, Stefan Borgwardt², Ernst W Radue³, Ludwig Kappos², Daniel Pelletier³, Stephen L Hauser³, Paul M Matthews⁴,⁵, Frederik Barkhof¹ and Jeroen JG Geurts¹

Abstract

Background: Although grey matter damage in multiple sclerosis is currently recognized, determinants of grey matter volume and its relationship with disability are not yet clear.

Objectives: The objectives of the study were to measure grey and white matter volumes across different disease phenotypes; identify MRI parameters associated with grey matter volume; and study grey and white matter volume as explanatory variables for clinical impairment.

Methods: This is a cross-sectional study in which MRI data of 95 clinically isolated syndrome, 657 relapsing–remitting, 125 secondary-progressive and 50 primary-progressive multiple sclerosis patients from three centres were acquired. Grey and white matter volumes were determined, together with T2 and T1 lesion volumes. Physical disability was assessed with the Expanded Disability Status Scale, cognitive impairment with the Paced Auditory Serial Addition Task. Data were analysed using multiple regression.

Results: Grey matter volume was lower in relapsing–remitting patients (mean [SD]: 0.80 [0.05] L) than in clinically isolated syndrome patients (0.82 [0.05] L), and even greater relative atrophy was found in secondary-progressive patients (0.77 [0.05] L). In contrast, white matter volume in secondary-progressive patients was comparable to that in relapsing–remitting patients. Grey matter volume was the strongest independent predictor of physical disability and cognitive impairment, and was associated with both T2 and T1 lesion volume.

Conclusions: Our findings show that grey matter volume is lower in secondary-progressive than in relapsing–remitting disease. Grey matter volume explained physical and cognitive impairment better than white matter volume, and is itself associated with T2 and T1 lesion volume.

Keywords

Cognitive decline, grey matter, multiple sclerosis, volumetric MRI

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\(^1\textsuperscript{3,4}\) and with cognitive decline.\(^1\textsuperscript{5,6}\) 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\(^1\textsuperscript{7,8}\) and higher field strengths,\(^1\textsuperscript{9}\) and many cortical GM lesions remain undetected on conventional MRI.\(^1\textsuperscript{10}\) Besides visualization of focal cortical lesions, quantitative MR techniques like diffusion tensor imaging (DTI),\(^1\textsuperscript{11}\) magnetization transfer imaging\(^1\textsuperscript{12}\) and MR spectroscopy\(^1\textsuperscript{13}\) 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.\(^1\textsuperscript{14}\)

While it has been suggested that accumulation of demyelination and neuro-axonal damage in the GM occurs secondarily to WM tract damage,\(^1\textsuperscript{15}\) a number of studies now indicate that an independent and partly overlapping occurrence of GM and WM damage may be more likely.\(^1\textsuperscript{16,17}\) This may also be true for primary-progressive MS (PPMS).\(^1\textsuperscript{3}\) In a seminal histopathological study of acute relapsing–remitting (RR), secondary-progressive (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.\(^1\textsuperscript{16}\) 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.\(^1\textsuperscript{3,18}\) In contrast, WM atrophy showed a relatively constant accrual over time.\(^1\textsuperscript{19}\)

These interesting in vivo findings require further investigation. This in particular accounts for clinically isolated syndrome (CIS) patients in comparison to the other MS patient groups, because previous studies included either CIS patients with a non-representative disease duration of over 20 years\(^1\textsuperscript{18}\) or a small sample of seven CIS patients.\(^1\textsuperscript{19}\) Furthermore, a comparison with a PPMS patient group has not been made. A large cohort of over 900 MS patients, including representative CIS, RR, SP and PP patients, prospectively assessed through the GeneMSA consortium,\(^1\textsuperscript{20}\) provided the opportunity to further investigate the relationship between disease progression and brain volume. More specifically, we test the hypotheses that decreased GM volumes and not WM volumes are a characteristic of progressive MS, and that it explains clinical disability better. Additionally, we investigate putative MRI determinants of GM volume.

**Materials and 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). 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\(^1\textsuperscript{21}\) were classified either as RRMS, SPMS, defined by at least six months of worsening neurological disability not explained by clinical relapse, or PPMS, defined by progressive clinical worsening for more than 12 months from disease onset without any relapses and the presence of more than two 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 (DMTs) for MS was permitted. In all included subjects, disability was assessed with the Expanded Disability Status Scale (EDSS)\(^1\textsuperscript{22}\) and the Paced Auditory Serial Addition Task (PASAT),\(^1\textsuperscript{23}\) 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 1.5 T MR systems (Amsterdam: Siemens Vision; Basel: Siemens Avanto) and one 3.0 T 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 × 1 × 1 mm\(^3\) 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 mm\(^2\) or 0.5 × 0.5 mm\(^2\) (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 mm\(^2\) or 0.9 × 0.9 mm\(^2\) [UCSF]) were obtained.
MRI measurements

Brain volume analyses were performed at the Image Analysis Centre in Amsterdam. SIENAX24 (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. Statistically adjusted brain volumes are specifically referred to as ‘adjusted volumes’. Statistical models in this paper were always adjusted for centre.

We examined disease type differences for NGMV and NWMV using two multiple regressions adjusting for age, sex and centre. The adjusted means were pairwise compared between the disease types; both post-hoc Bonferroni corrected and the p-values without Bonferroni correction are provided. Possible effects of centre were further explored by performing additional multiple regression analyses separately for each centre.

Associations between NGMV and the other variables were assessed, statistically adjusting 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 adjusted for age, sex, disease duration and centre, and it was repeated for the disease types separately. A high correlation of $\rho = 0.9$ existed between T2 lesion volume and T1 hypointense lesion volume. To avoid collinearity in the models T2 lesion volume and T1 hypointense lesion volume were separated in the multiple regression models, whereby it is not possible to establish their relative strength. Before entering the regression models, T2 and T1 hypointense lesion volumes were transformed to improve normality using the logarithm with base 10 (logT2LV and logT1LV).

The relative predictive value of GM and WM brain volumes, together with logT2LV, on EDSS and on PASAT was assessed using ordinal logistic regression and multiple linear regression, respectively. Again, this was done separately for each variable adjusting 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: $<2.0; \geq 2.0$ and $<4.0; \geq 4.0$ and $<6.0; \geq 6.0$. Odds ratios (ORs) 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. The regression models were repeated for each disease type separately.

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 with the other sites, whereas 80 of the 95 CIS patients were from UCSF. 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] L; NWMV: 0.78 [0.06] L) and Basel (NGMV: 0.77 [0.07] L; NWMV: 0.78 [0.05] L), mean NGMV of patients from UCSF was significantly higher (0.83 [0.06] L; \( p < 0.001 \)) and mean NWMV significantly lower (0.69 [0.04] L; \( p < 0.001 \)) compared with the other two sites.

In all centres, NGMV values were lower in RRMS patients than in CIS patients, further decreasing in SPMS patients (Figure 1A). While NWMV also differed significantly between CIS patients and RRMS patients, it did not differ between RRMS and SPMS patients (Figure 1B). The NGMV adjusted 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 surviving Bonferroni’s correction between all disease subtypes, except for PP vs. RR and CIS patients. Bonferroni corrected \( p \)-values are as follows (without Bonferroni correction): CIS vs. RR: 0.05 (0.008); CIS vs. PP: 0.5 (0.08); CIS vs. SP: <0.001 (<0.001); PP vs. SP: 0.006 (0.001); RR vs. SP: <0.001 (<0.001); RR vs. PP: 1.0 (0.7). Adjusted NWMV was also highest in CIS patients (0.77 [0.05] L), in the other disease types the adjusted NWMV had the same value and SD (0.75 [0.05] L). Differences in NWMV were significant surviving Bonferroni’s correction between CIS vs. RR, and CIS vs. SP (Figure 1D). Bonferroni corrected \( p \)-values are as follows (without Bonferroni correction): CIS vs. RR: 0.005 (0.001); CIS vs. PP: 0.3 (0.05); CIS vs. SP: 0.03 (0.005); PP vs. SP: 1.0 (0.5); RR vs. SP: 1.0 (0.7); RR vs. PP: 1.0 (0.3). The analyses repeated for the centres separately showed similar results; importantly, in each centre adjusted NGMV of SP patients is significantly lower than that of RR patients, whereas in none of the centres does adjusted NWMV differ significantly between these disease types.

Predictors of GM volume

NWMV, logT2LV and logT1LV were each significantly associated with NGMV (Table 2). Multiple regression analysis (Table 2) revealed that, after adjusting for age, gender, disease duration and centre, logT2LV was a significant explanatory MRI variable of NGMV (beta: \(-0.27\); \( p < 0.001 \)). The multiple regression model with logT1LV instead of logT2LV did not differ substantially. NWMV was not included in the model, because its \( p \)-value was higher than 0.1. The model accounted for 59% (adjusted \( R^2 = 0.59 \)) of the variance in NGMV. Subsequent analyses for the disease types separately showed no substantial differences.

Predictors of disability

Associations between EDSS and MRI variables (Table 3) were significant for both NGMV and NWMV, as well as for logT2LV and logT1LV. In the

<table>
<thead>
<tr>
<th>Table 1. Disease type descriptives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>N (% of total patient group)</td>
</tr>
<tr>
<td>Mean age, years (SD)</td>
</tr>
<tr>
<td>Male:female ratio</td>
</tr>
<tr>
<td>Proportion originating from UCSF/BAS/AMS</td>
</tr>
<tr>
<td>Mean disease duration, years (SD)</td>
</tr>
<tr>
<td>Mean EDSS (SD)</td>
</tr>
<tr>
<td>Median PASAT (IQR)</td>
</tr>
<tr>
<td>Median T2 lesion volume, ml (IQR)</td>
</tr>
<tr>
<td>Median T1 black hole volume, ml (IQR)</td>
</tr>
<tr>
<td>No. of patients on DMT (%)</td>
</tr>
</tbody>
</table>

Significant differences between groups (\( p < 0.05 \)):

- \( ^a \)CIS vs. (SP, PP), RR vs. (SP, PP)
- \( ^b \)CIS vs. (RR, SP, PP), SP vs. (PP, RR)
- \( ^c \)CIS vs. (SP, PP), RR vs. (SP, PP)

\( ^d, ^e, ^f \)Between all groups except between RR and PP

multiple ordinal regression model (Nagelkerke’s $R^2=0.40$), NGMV was the strongest MRI predictor of EDSS (OR: 0.67; $p<0.001$), logT2LV was a weaker predictor (Table 3). A similar model was obtained with logT1LV instead of logT2LV. The OR of NGMV can also be expressed as an odds increase of 1.49 in having greater disability per 1 SD smaller NGMV. NWMV was excluded from the model, because its $p$-value was higher than 0.1. Subsequent analyses for each disease type showed that in the RR patient group all MRI variables were included in the final model with NGMV being the strongest predictor, whereas in the other disease types none of the MRI variables were retained.

Also for the PASAT, correlations with all MRI variables (Table 4) were significant. The multiple linear regression model had an adjusted $R^2$ 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; $p=0.02$). LogT2LV was (just) not significant in
the multiple regression model, whereas logT1LV was a significant predictor (beta: \(0.07; p = 0.05\)). Subsequent analyses showed that in both the RR and the SP cohorts NWMV and NGMV are retained in the final multiple regression model, with NGMV as the strongest predictor. In contrast, in the PPMS group the multiple regression model contained only logT2V, and in the CIS group none of the MRI variables were retained. Results of all of the above-described model analyses did not change materially when the 105 patients who started their DMT less than one year before the investigation were excluded.

### Discussion

Our cross-sectional study, involving a large cohort of CIS, RR, SP and PP patients, shows that in contrast to WM volume, GM volume is significantly lower in secondary-progressive MS than in relapsing–remitting MS. Moreover, GM volume explains physical disability and cognitive impairment, as measured by the EDSS and the PASAT respectively, better than WM volume. Both T2 and T1 hypointense lesion volumes in the WM were better MRI predictors of GM volume than WM volume.

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

<p>| Table 2. (Multiple) linear regression for NGMV, statistically adjusted for centre |</p>
<table>
<thead>
<tr>
<th>Linear regression</th>
<th>Multiple linear regression&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Multiple linear regression&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>p-value</td>
<td>Adjusted (R^2)</td>
</tr>
<tr>
<td>Age</td>
<td>(-0.52)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>Sex (female: 0; male: 1)</td>
<td>(-0.17)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>Disease duration</td>
<td>(-0.4)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>LogT2LV</td>
<td>(-0.38)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>LogT1LV</td>
<td>(-0.39)</td>
<td>(&lt;0.001)</td>
</tr>
<tr>
<td>NWMV</td>
<td>(0.14)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>LogT2LV is used as the single lesion variable in the model because of the collinearity problem.

<sup>b</sup>LogT1LV is used as the single lesion variable. When centre was the only independent variable, Nagelkerke’s \(R^2\) was 0.20. Adjusted \(R^2\)-values of both final multiple ordinal regression models was 0.40.


<p>| Table 3. (Multiple) ordinal regression for EDSS, adjusted for centre |</p>
<table>
<thead>
<tr>
<th>Ordinal regression</th>
<th>Multiple ordinal regression&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Multiple ordinal regression&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>OR</td>
<td>p-value</td>
</tr>
<tr>
<td>Age</td>
<td>0.72</td>
<td>2.05</td>
</tr>
<tr>
<td>Sex (female: 0; male: 1)</td>
<td>0.26</td>
<td>1.30</td>
</tr>
<tr>
<td>Disease duration</td>
<td>0.80</td>
<td>2.23</td>
</tr>
<tr>
<td>NWMV</td>
<td>0.14</td>
<td>1.70</td>
</tr>
</tbody>
</table>

<sup>a</sup>LogT2LV is used as the single lesion variable in the model because of the collinearity problem.

<sup>b</sup>LogT1LV is used as the single lesion variable. When centre was the only independent variable, Nagelkerke’s \(R^2\) was 0.21. Nagelkerke’s \(R^2\) of both final multiple ordinal regression models was 0.40.


PE and OR are reported per standard deviation for each independent variable, except for sex, where the OR of 1.30 indicates a lower EDSS in females compared with males.
these findings in a large group of patients. An interesting and more novel aspect of our study is the inclusion of patients with PPMS. They appear to have higher GM volume, but a similar low WM volume, compared with SPMS patients. Interestingly, in contrast to RRMS and SPMS patients, NGMV is not a significant predictor for PASAT in PPMS 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, 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, and in our study T2 lesion volume was an independent predictor of NGMV. 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. Unfortunately, our data did not allow a distinction between the effect of T1 hypointense and T2 lesions on GM volume.

GM volume explained physical disability as measured by the EDSS and cognitive impairment, measured by the PASAT, better than WM volume or T2 lesion volume. Associations between GM volume and EDSS scores have been reported by several previous studies. The PASAT measures working memory and sustained attention, and is often used as a measure of cognitive impairment in MS. Our finding that GM volume is the strongest predictor of both physical disability and cognitive impairment emphasizes the clinical relevance of GM damage. This has also been recognized in recent longitudinal studies investigating atrophy as well as cortical lesions.

Now that it is becoming increasingly clear that MS involves not only 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 accrues faster with progressing disease, 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 the 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 UCSF when compared with those of patients from the other two sites. A 3 T MR system was used at UCSF and 1.5 T MR systems were used in the other two centres; this field strength might have contributed to the observed differences. Unfortunately, MR sequences to specifically investigate cortical lesions, such as double inversion-recovery (DIR), have not been acquired in our study.

### Table 4. (Multiple) linear regression for PASAT, adjusted for centre

<table>
<thead>
<tr>
<th>Linear regression</th>
<th>Multiple linear regression&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Multiple linear regression&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>p-value</td>
</tr>
<tr>
<td>Age</td>
<td>-0.22</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>Disease duration</td>
<td>-0.26</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LogT2LV</td>
<td>-0.15</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LogT1LV</td>
<td>-0.19</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NWMV</td>
<td>0.19</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NGMV</td>
<td>0.28</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>LogT2LV is initially used as the single lesion variable because of the collinearity problem, but is not in the final model because $p > 0.1$.

<sup>b</sup>LogT1LV is used as the single lesion variable in this model. When centre was the only independent variable, $R^2$ was 0.01. The adjusted $R^2$ of both final models was 0.11.

strength difference may be responsible. The trend within centres was the same and the between-centre differences were limited as much as possible by correcting all models in our study for centre effects; nevertheless these effects are still likely to have been of influence. Some studies have reported that frontal and temporal cortical areas may exhibit more severe pathology than others.\textsuperscript{40–42} 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;\textsuperscript{16} therefore it may 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,\textsuperscript{43} such as the thalamus, and mixed WM–GM structures such as the hippocampus, 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 adjust for lesion misclassification. In the automated segmentation process some of the WM lesions may have been misclassified as 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 SPMS, our results are unlikely to be false-positive due to this phenomenon.

In conclusion, our study has shown that, in a large group of MS patients, GM pathology in (secondary) progressive MS is clearly dominant of WM volume loss. In addition, we have shown that GM volume explains clinical disability better than WM volume, and is itself better predicted by T2 or T1 hypointense lesion volume than WM volume. These findings are important for our understanding of MS and for future clinical trial design.

Acknowledgements
The authors would like to thank all people involved in the GeneMSa consortium.

Funding
GlaxoSmithKline sponsored data acquisition of the multinational GeneMSa study, but were not involved in data interpretation for the current research.

Conflict of interest statement
None declared.

References
14. Miller DH, Barkhof F, Frank JA, Parker GJ and Thompson AJ. Measurement of atrophy in multiple...


