White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI

Mark Mühlau1, Dorothea Buck1, Annette Förschler2, Christine C Boucard1, Milan Arsic1, Paul Schmidt1,3, Christian Gaser4,5, Achim Berthele1, Muna Hoshi1, Angela Jochim1, Helena Kronsbein1, Claus Zimmer2, Bernhard Hemmer1 and Rüdiger Ilg1

Abstract
Background: In MS, the relationship between lesions within cerebral white matter (WM) and atrophy within deep gray matter (GM) is unclear.
Objective: To investigate the spatial relationship between WM lesions and deep GM atrophy.
Methods: We performed a cross-sectional structural magnetic resonance imaging (MRI) study (3 Tesla) in 249 patients with clinically-isolated syndrome or relapsing–remitting MS (Expanded Disability Status Scale score: median, 1.0; range, 0–4) and in 49 healthy controls. Preprocessing of T1-weighted and fluid-attenuated T2-weighted images resulted in normalized GM images and WM lesion probability maps. We performed two voxel-wise analyses: 1. We localized GM atrophy and confirmed that it is most pronounced within deep GM; 2. We searched for a spatial relationship between WM lesions and deep GM atrophy; to this end we analyzed WM lesion probability maps by voxel-wise multiple regression, including four variables derived from maxima of regional deep GM atrophy (caudate and pulvinar, each left and right).
Results: Atrophy of each deep GM region was explained by ipsilateral WM lesion probability, in the area most densely connected to the respective deep GM region.
Conclusion: We demonstrated that WM lesions and deep GM atrophy are spatially related. Our results are best compatible with the hypothesis that WM lesions contribute to deep GM atrophy through axonal pathology.

Keywords
Multiple sclerosis, relapsing–remitting multiple sclerosis, clinically-isolated syndrome, magnetic resonance imaging, voxel-based morphometry, gray-matter atrophy, white-matter lesions, brain damage

Date received: 4th October 2012; accepted: 18th January 2013

Introduction
Demyelinating lesions within cerebral white matter (WM), as visible in T2-weighted (w) magnetic resonance images (MRIs) are the most striking finding in multiple sclerosis (MS) and they have become the most important paraclinical criterion for its diagnosis. Intriguingly, cerebral gray matter (GM) changes are also observed in MS, even in the early stages. Yet there is an inconsistency between the histopathological and MRI studies: In histopathological studies, MS pathology (i.e. lymphocyte-driven inflammation, demyelination, axonal transection and remyelination) is reported to occur within the cortical GM primarily, and to occur less frequently or in later stages in deep GM, although mere measurements of neuronal counts yield considerable neuronal loss in the thalamus. In contrast, MRI studies applying voxel-based morphometry (VBM) thus analyzing
both cortical and deep GM, mostly detect a marked accentuation of atrophy within deep GM, particularly in patients with clinically-isolated syndrome (CIS), a state of high risk for conversion to clinically-definitive MS. Consequently, deep GM atrophy, as detectable with MRI even in early stages of MS, is postulated not to reflect direct damage from MS pathology, but to result from the disconnection via axonal transection within WM lesions (WMLs), leading to degeneration along the axonal projections. This assumption is further justified by the finding that GM atrophy of the lateral geniculate body correlates with lesion probability within the optic radiation. Moreover, WMLs not only occur more frequently within than outside of the tracts connecting to the thalamus, but WMLs also correlate with thalamic volume. On the other hand, others demonstrate a temporal, but not a spatial, association of WML occurrence and GM loss or do not find any spatial relationships between WMLs and GM atrophy.

Aiming at defining the role of WMLs for deep GM atrophy, our goals were to confirm that there was pronounced atrophy within the deep GM and to study the spatial relationship of WMLs and deep GM atrophy in a large cohort of patients with early MS. We reasoned as follows: If deep GM atrophy resulted from disconnection, GM atrophy of one region would be explained by WMLs within tracts connecting to it. This would also imply the occurrence of anatomically-plausible distance effects; that is, that WMLs cause GM atrophy within distant, but connected, regions. Alternatively, GM atrophy and WMLs could reflect the spread of inflammation from WMLs to bordering GM, or represent concurrent events that would be independent across space.

Materials and methods

Patients and healthy controls

This study was approved by the local ethics committee. As our aim was to investigate the spatial relationship between WMLs and GM atrophy in CIS and early MS, we specified the following inclusion criteria: an Expanded Disability Status Scale (EDSS) score between 0 and 4; relapsing–remitting MS (RRMS) or CIS; CIS was defined as WML probability maps (LPMs). The resulting GM and WM images were generated one binary lesion image of the WMLs present, per subject. Next, lesion voxels were removed from the T1w images and filled with values approximating normal WM intensity, as proposed recently, so that the GM content of the resulting T1w images could be analyzed next, across the whole brain by VBM, without a relevant impact of WMLs. The refilled T1w images were normalized and segmented into GM and WM by VBM8, which has proved an effective preprocessing pipeline, including high-dimensional warping ("DARTEL"). The resulting GM and WM images were modulated, accounting only for nonlinear volume changes. Finally, the GM images were smoothed with a Gaussian kernel of 6 mm, in order not to miss the changes of small structures. Binary lesion images were normalized by applying the deformation field, derived from the refilled T1 images. Normalized lesion images were blurred by a Gaussian kernel of 12 mm, as was recently proposed. These normalized and blurred lesion images are referred to as WML probability maps (LPMs).

Global volumes

For the calculation of total WM lesion volume (LV), voxel values of unsmoothed normalized lesion maps were summed up. Global volumes of GM and WM were calculated by summing up the voxel values of the unsmoothed modulated GM and WM images. To calculate global values of deep and cortical GM, values were separately summed up within a mask of deep and cortical GM. To create the cortical mask, we removed deep GM from the GM template 6 of VBMS, using commercial software (Amira 5.3.3, Visage Imaging), then binarized the remainder (threshold, 0.1) and added a 3 mm margin, by using the following steps:

1. Image pre-processing

We used SPM8 software (http://www.fil.ion.ucl.ac.uk/spm) and its toolbox VBM8 (http://dbm.neuro.uni-jena.de/vbm8). Recently, we extended VBM8 by Lesion Segmentation Tool (LST; http://www.applied-statistics.de/lst.html). We used the default options of SPM8, VBM8 and LST, unless indicated otherwise. After co-registration, T1w and FLAIR images were used for the segmentation of T2-hyperintense WMLs by LST. Operating in original ("native") space, this procedure generated one binary lesion image of the WMLs present, per subject. Next, lesion voxels were removed from the T1w images and filled with values approximating normal WM intensity, as proposed recently, so that the GM content of the resulting T1w images could be analyzed next, across the whole brain by VBM, without a relevant impact of WMLs. The refilled T1w images were normalized and segmented into GM and WM by VBM8, which has proved an effective preprocessing pipeline, including high-dimensional warping ("DARTEL"). The resulting GM and WM images were modulated, accounting only for nonlinear volume changes. Finally, the GM images were smoothed with a Gaussian kernel of 6 mm, in order not to miss the changes of small structures. Binary lesion images were normalized by applying the deformation field, derived from the refilled T1 images. Normalized lesion images were blurred by a Gaussian kernel of 12 mm, as was recently proposed. These normalized and blurred lesion images are referred to as WML probability maps (LPMs).
freely available software (FMRIB Software Library, www.fmrib.ox.ac.uk/fsl/). We used the Wake Forest University atlas (http://fmri.wfubmc.edu/software/PickAtlas) to create the deep GM mask: we chose the binary masks of the caudate nucleus, putamen and thalamus; dilated each of the three (dilation value, 1) and created a single mask. Global volumes were analyzed by the use of standard software (IBM SPSS Statistics, version 19). For group comparisons, we used $t$-tests. We also analyzed the effect of the global volumes (WM LV, WM, cortical and deep GM) on EDSS, by one ordinal logistic regression model. As explaining variables, we included the four volumes, as well as age and sex.

Voxel-wise data analyses

**Statistical models, thresholds and atlases.** SPM8 enables univariate MRI data analysis based on a voxel-wise general linear model (GLM), in which each voxel value constitutes the response (dependent) variable. For each explanatory (independent) variable (e.g. behavioral data, age, sex or other imaging parameters), a $\beta$ value is estimated, to which weights can be attributed. Within this framework, standard statistical tests can be realized, if they are special cases of the GLM (e.g. $t$-test and multiple linear regression). Of note, results are corrected for variance explained by nuisance variables.\(^{24}\)

To account for multiple statistical tests across voxels, we applied a height threshold of 0.05 family-wise error corrected (voxel level) as implemented in SPM8.\(^{25}\) Only once, this threshold did not reveal significant voxels; so we relaxed the height threshold (voxel level) to 0.001 uncorrected, in this case, which will be clearly indicated. Finally, results were attributed to anatomical regions by atlases, as implemented in the FMRIB Software Library.

**Spatial distribution of gray-matter atrophy.** We performed two voxel-wise analyses:

1. To compare the patients with HCs, we applied the $t$-test extended by the nuisance variables of age and sex, as both variables exert a substantial influence on GM.\(^ {22}\)
2. To correlate GM atrophy with WM LV, we only analyzed the patients by one multiple linear regression model, which included total WM LV as a variable of interest that was weighted minus 1 (decreasing GM, with increasing WM LV). As nuisance variables, we included age, sex, immunomodulatory therapy (no, 0; yes, 1) and corticosteroid administration in the previous 30 days (no, 0; yes, 1), to covary out any potential influence of these parameters.

**White-matter lesion probability to explain regional deep gray-matter atrophy.** In preparation, voxel values of four regional

---

**Table 1.** Demographic features and global imaging parameters of the study population.

<table>
<thead>
<tr>
<th></th>
<th>Healthy controls</th>
<th>Patients</th>
<th>Decline (%)</th>
<th>$P$ value (group comparison)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>49</td>
<td>249</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>36.4 ± 13.0 (20–59)</td>
<td>36.8 ± 10.7 (18–68)</td>
<td>n/a</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>Sex (male/female)</strong></td>
<td>16/33 (33/67%)</td>
<td>87/162 (35/65%)</td>
<td>n/a</td>
<td>n.s.</td>
</tr>
<tr>
<td><strong>CIS/RRMS</strong></td>
<td>n/a</td>
<td>81/168</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>EDSS</strong></td>
<td>n.d.</td>
<td>1.4±1.0 (0–4.0)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Corticosteroids in the previous 30 days</strong></td>
<td>0 (29%)</td>
<td>71 (29%)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Immunomodulatory therapy(^ a)</strong></td>
<td>0</td>
<td>108 (43%)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Deep gray matter (ml)</strong></td>
<td>29.2 ± 3.28 (21.5–36.0)</td>
<td>27.2 ± 4.17 (14.1–36.3)</td>
<td>6.6</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Cortical gray matter (ml)</strong></td>
<td>807 ± 51.6 (672–890)</td>
<td>784 ± 52.3 (629–918)</td>
<td>2.9</td>
<td>0.005</td>
</tr>
<tr>
<td><strong>White matter (ml)</strong></td>
<td>684 ± 41.9 (591–774)</td>
<td>670 ± 45.6 (508–852)</td>
<td>1.9</td>
<td>0.056</td>
</tr>
<tr>
<td><strong>Lesion volume (ml)</strong></td>
<td>0.17 ± 0.45 (0–2.3)</td>
<td>5.7 ± 8.9 (0.02–50)</td>
<td>n/a</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

CIS: clinically-isolated syndrome; n/a: not applicable; n.d.: not determined; n.s.: not significant ($P$ value > 0.1); RRMS: relapsing–remitting multiple sclerosis. Mean values ± standard deviations as well as ranges (minimum–maximum) are given. 

\(^{a}\)Therapies (n patients): intramuscular interferon beta-1a (37), subcutaneous interferon beta-1b (23), subcutaneous interferon beta-1a (12), glatiramer acetate (16), natalizumab (18), azathioprine (1), intravenous immunoglobulin (1).
maxima of deep GM atrophy (derived from the correlation analysis with total WM LV) were extracted (caudate and pulvinar, each left and right) by the plot function of SPM8, which adjusts voxel values for the nuisance variables. Next, we subjected all four variables to one voxel-wise multiple linear regression model, in order to identify region-specific WML probability. Six contrasts were calculated. Again, the variables of interest were weighted minus 1. In the first four contrasts, only one variable was weighted (caudate and pulvinar, each left and right). Of note, WM areas identified this way specifically explain GM atrophy of the weighted region, because the model controls for variance explained by the other three GM regions, so that these can be interpreted as control regions. In contrasts numbered 5 and 6, we searched for distance effects, by weighting two variables. Of note, when weighting two variables minus 1, voxels are only identified if the sum of the respective two β values is significantly lower than 0. Hence, two scenarios are possible:

1. If there were no lesion location explaining GM atrophy of both weighted regions, we would expect to

Figure 1. Regional gray-matter atrophy, explained by total white-matter lesion volume. Regions of decreasing gray matter with increasing total white-matter lesion volume are shown. Axial slices are projected onto the SPM template. The smoothed GM images were subjected to a multiple linear regression model, with the variable of interest of total WM LV (weight, minus 1) and the nuisance variables of age, sex, immunomodulatory therapy and corticosteroid administration in the previous 30 days (weights: −1, 0, 0, 0, 0). The height threshold (voxel level) was set to a P < 0.05, family-wise error corrected. Only voxels are shown that show a gray matter probability of more than 0.3, according to the tissue probability map of SPM8. Increasing significance is color-coded according to the bar in the right lower corner. MNI coordinates are indicated in the left upper corners. The right side of the image is indicated by the letter “R” (neurological convention). GM: gray matter; LV: lesion volume; MNI: Montreal Neurological Institute; R: right; SPM: Statistical Parametric Maps; WM: white matter.
identify about the voxels identified after separate weighting of one or the other variable (as described for the first four contrasts).

2. In case of distant effects, circumscribed WML probability may explain GM atrophy of both regions, resulting in identification of further regions.

In contrast number 5, we weighted both caudate nuclei. As both regions are connected to the contralateral hemisphere via anterior transcallosal fibers, we expected WML probability within the anterior portion of the corpus callosum, to explain GM atrophy of both caudate nuclei. In contrast 6, we weighted both pulvinar regions. As both regions are connected to the contralateral hemisphere via posterior transcallosal fibers, we expected WML probability within the posterior portion of the corpus callosum to explain GM atrophy of both pulvinar regions. Finally, we repeated estimation of the multiple linear regression model after inclusion of the nuisance variables of age, sex, CIS or RRMS, immunomodulatory therapy and previous corticosteroid administration, to covary out the potential effects of these variables.

Results

Global volumes and spatial distribution of gray-matter atrophy

Global volumes of GM were significantly lower in the patients, as compared to HCs. Atrophy was more pronounced in the deep, rather than in cortical GM. Global WM showed a trend towards decline (Table 1). The ordinal logistic regression model revealed a significant effect on the EDSS score only for WM LV (P value, < 0.002) and cortical GM (P value, 0.013), but not for deep GM volume (P value, 0.914) nor WM volume (P value, 0.312). Analyzing the spatial distribution of GM atrophy, we primarily identified areas of deep GM. As expected, the results derived from the group comparison (not shown) and from the correlation analysis with total WM LV restricted to the patient group (Figure 1, Table 2) greatly resembled each other.

White-matter lesion probability explaining regional deep gray-matter atrophy

Generation of the four variables of regional deep GM atrophy, subjected to the regression analyses of WM LPMs, is illustrated in Figure 2(a) and the corresponding peak voxels are given in Table 2. Results derived from the voxel-wise multiple regression analysis on region-specific WML probability are shown in Figure 2(b), in which different contrasts are arranged in analogy to a contingency table, so that weighted variables are indicated by row and column of each field. When weighting only one variable, we detected ipsilateral WML probability in the area known to be most densely connected to the respective region, namely the frontal lobe when weighting one caudate and the posterior parietal lobe when weighting one pulvinar. In one case (right caudate), we detected WML probability only after relaxing the significance threshold to 0.001, uncorrected. Weighting both caudate nuclei resulted in WML probability primarily in frontal WM, including one local maximum in the anterior portion of the corpus callosum. Weighting both pulvinar regions revealed wide-spread WML probability, primarily in parietal and occipital WM, including one local maximum in the posterior portion of the corpus callosum. Extension of the GLM by the nuisance variables of age, sex, CIS or RRMS disease, immunomodulatory therapy, and corticosteroid administration in the previous 30 days did not change the results in a meaningful way.
Aiming at the role of WMLs in deep GM atrophy, we investigated the spatial relationship of WMLs and deep GM atrophy in early MS. First, we analyzed global volumes. As was demonstrated in previous studies,\textsuperscript{7,8} deep GM showed more atrophy than cortical GM. WM LV showed the strongest effect on the subject’s EDSS score. In addition, only cortical GM showed a significant effect on the EDSS score, although atrophy was less pronounced here compared to

**Discussion**

Aiming at the role of WMLs in deep GM atrophy, we investigated the spatial relationship of WMLs and deep GM atrophy in early MS. First, we analyzed global volumes. As was demonstrated in previous studies,\textsuperscript{7,8} deep GM showed more atrophy than cortical GM. WM LV showed the strongest effect on the subject’s EDSS score. In addition, only cortical GM showed a significant effect on the EDSS score, although atrophy was less pronounced here compared to
deep GM. In other words, only cortical GM atrophy contributes to disability, independent of WM LV. Next, we performed voxel-wise analyses of GM atrophy by both group comparison and correlation with total WM LV within the patient group. In accordance with our analysis of global volume, we found widespread deep GM atrophy. This finding confirmed recent MRI studies applying VBM in CIS patients that were comparable with our cohort, with regard to their level of disability (mean EDSS in both studies, 1.0).

In particular, we aimed to relate the spatial distribution of WMLs and GM atrophy. We focused on both caudate and pulvinar regions, since GM atrophy was most pronounced here. As the caudate, and the pulvinar regions are densely interconnected to several cortical regions, no atlas-based masks of WM tracts connecting these areas are available; therefore, we followed a merely data-driven approach, by correlating WM LPMs with the values of regional GM atrophy. To this end, we extracted the maxima of regional atrophy from the caudate and pulvinar region (each left and right) and subjected all four variables to a single linear multiple regression, in order to identify region-specific WML probability. Our approach must be regarded very conservative as all results demonstrating region-specific WM probability were corrected for the overall atrophy within deep GM, as estimated by at least another two regions. Therefore, it is justified that once, we relaxed the statistical threshold to 0.001 uncorrected, which can still be regarded as conservative. Moreover, we repeated estimation of the multiple linear regression model after inclusion of the nuisance variables of age, sex, CIS or RRMS, immunomodulatory therapy and previous corticosteroid administration, although we rated a differential influence of these variables to the four regions under investigation as very unlikely. As expected, we obtained virtually the same results after inclusion of these additional nuisance variables in the model. Each time, region-specific WML probability was localized within the ipsilateral area that is most densely connected to the respective deep GM region. The caudate is most densely connected to the frontal cortex, and the pulvinar to the posterior parietal cortex, which accorded well with our findings of the most significant WML probability in these regions. Furthermore, weighting two variables revealed additional regions of significant WML probability. WML probability within the anterior portion of the corpus callosum explained GM atrophy of both caudate nuclei and WM probability within the posterior portion explained GM atrophy of both pulvinar regions. These findings indicate anatomically plausible distance effects, because both the caudate and pulvinar regions are connected to the contralateral hemisphere by transcallosal fibers. Hence, our results demonstrated not only a spatial correlation of WMLs and deep GM atrophy, but also anatomically plausible distance effects, which can best be explained by axonal transection within WMLs and subsequent degeneration along the axonal projections.

This conclusion may at first seem contradictory to studies demonstrating deep GM atrophy to explain cognitive decline better than WMLs, yet deep GM atrophy as detectable at T1w MRI may still be the better parameter for assessing disease progression. Inherent to its secondary nature, measurement of deep GM atrophy may quantify WMLs with regard to both destructive power (i.e. axonal transection) and eloquence of location. On the other hand, the T2w signal of WMLs covers only part of MS pathology, as lesions are not specified with regard to active inflammation, demyelination, axonal transection nor remyelination.

We acknowledge the limitations of our study. We cannot rule out that deep GM atrophy in part results from inflammation that spreads from WMLs to bordering deep GM. Furthermore, our cross-sectional study does not indicate the direction of the relationship between WMLs and deep GM atrophy; thus, our data do not rule out the possibility that deep GM pathology causes WMLs, which was also previously considered by others. Yet this notion is incompatible with the commonly held view that active inflammation provides for the formation of WMLs. We will also refrain from extending our conclusions on deep GM to cortical GM. Our finding that cortical, but not deep, GM contributes to disability independent of WM LV, suggests there may be different relationships of cortical and deep GM to WMLs. Most importantly, other MRI sequences, such as double-inversion recovery and histopathological studies have clearly demonstrated direct damage of cortical GM from MS pathology, even in the early stages.

In conclusion, we demonstrated an anatomically plausible spatial relationship between WMLs and deep GM atrophy, including distance effects. Hence, our data are compatible with the hypothesis that WMLs contribute to deep GM atrophy most likely through axonal transection, with subsequent degeneration along axonal projections.

Conflict of interest
The authors declare that there are no conflicts of interest.

Disclosures
Mark Mühlau has received research support from Merck Serono, and travel expenses for attending meetings from Bayer and Merck Serono. Dorothea Buck has received honoraria for lecturing, travel expenses for attending meetings and financial support for research from Bayer, Biogen Idec, Merck Serono, Novartis, Sanofi Aventis and Teva Neuroscience. Milan Arsic received research support from Merck Serono. Achim Berthele is a consultant for Biogen Idec, Bayer and Merck Serono; he has received research support from Bayer; he has received honoraria for lecturing from Biogen Idec, Bayer, Merck Serono, Teva Neuroscience and Novartis; and he has received Investigator fees for Phase II–IV clinical studies from Biogen Idec, Novartis, Merck Serono and Galapagos. Bernhard Hemmer served on scientific advisory boards for Roche, Biogen Idec, Bayer, Novartis, Merck Serono and Metanomics; he is a consultant for Roche, Biogen Idec, Bayer, Novartis, Merck Serono and Metanomics; he received research support from Novartis, Merck Serono, Biogen Idec and Bayer; he received honoraria for
lecturing from Teva Neuroscience, BiogenIdec, Bayer, Novartis and Merck Serono; and he has received travel expenses for attending meetings from Roche, Teva Neuroscience, BiogenIdec, Bayer, Novartis and Merck Serono.

**Funding**

This work was supported by the German Ministry for Education and Research (BMBF) to B.H., M.M. (German Competence Network Multiple Sclerosis, KKNMS; 01GI0917 and 01GI1307B), and C.G. (grant 01EV07099) as well as Merck Serono (to MA).

**References**