Brain tissue volume changes in relapsing-remitting multiple sclerosis: correlation with lesion load

Mario Quarantelli, Andrea Ciarmiello, Vincenzo Brescia Morra, Giuseppe Orefice, Michele Larobina, Roberta Lanzillo, Vittorio Schiavone, Elena Salvatore, Bruno Alfano, and Arturo Brunetti

Biostructure and Bioimaging Institute, National Council for Research, Naples, Italy
Diagnostic Imaging Department, Nuclear Medicine, IRCCS Fondazione G. Pascale, Naples, Italy
Department of Neurological Sciences, University Federico II, Naples, Italy
Department of Diagnostic Imaging, University Federico II, Naples, Italy

Received 19 April 2002; revised 26 August 2002; accepted 13 September 2002

Abstract

The aim of this study was to simultaneously measure in vivo volumes of gray matter (GM), normal white matter (WM), abnormal white matter (aWM), and cerebro-spinal fluid (CSF), and to assess their relationship in 50 patients with relapsing-remitting multiple sclerosis (RR-MS) (age range, 21–59; mean EDSS, 2.5; mean disease duration, 9.9 years), using an unsupervised multiparametric segmentation procedure applied to brain MR studies. Tissue volumes were normalized to total intracranial volume providing corresponding fractional volumes (fGM, faWM, fWM, and fCSF), subsequently corrected for aWM-related segmentation inaccuracies and adjusted to mean patients’ age according to age-related changes measured in 54 normal volunteers (NV) (age range 16–70). In MS patients aWM was 23.8 ± 29.8 ml (range 0.4–138.8). A significant decrease in fGM was present in MS patients as compared to NV (49.5 ± 3.2% vs 53.3 ± 2.1%; P < 0.0001), with a corresponding increase in fCSF (13.0 ± 3.8% vs 9.1 ± 2.4%; P < 0.0001). No difference could be detected between the two groups for fWM (37.5 ± 2.6% vs 37.6 ± 2.2%). faWM correlated inversely with fGM (R = 0.434, P < 0.001 at regression analysis), and directly with fCSF (R = 0.473, P < 0.001), but not with fWM. There was a significant correlation between disease duration and EDSS, while no relationship was found between EDSS or disease duration and fractional volumes. Brain atrophy in RR-MS is mainly related to GM loss, which correlates with faWM. Both measures do not appear to significantly affect EDSS, which correlates to disease duration.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Relapsing-remitting multiple sclerosis; MRI; Brain atrophy; Segmentation; Brain volume

Introduction

Brain atrophy is a common finding in multiple sclerosis (MS) (Dawson, 1916; Loizou, et al., 1982). Progression of brain atrophy, above normal age-related brain volume changes, has been detected in MS using linear measures of width of the ventricles and/or volumetric measures of the brain/intracranial volume (ICV) ratio (Adams and Kozoil, 2000; Simon et al., 1999; Rudick et al., 1999; Losseff et al., 1996), over a period of time as short as 1 year.

Tissue damage in MS is not limited to the white matter (WM) as from inflammation axonal transection ensues (Ferguson et al., 1997; Trapp et al., 1998), and this process may extend by Wallerian degeneration, indirectly demonstrated by N-acetylaspartate reduction in MR spectroscopy studies (Davie et al., 1994, 1995; Rooney et al., 1997), possibly contributing to gray matter (GM) loss.

Accordingly, both GM and WM atrophy can be postulated, but whether they are simultaneously present and sizeable is currently debated, as to date only two studies have assessed volume changes in these tissues separately in re-
lapses-remitting MS (RR-MS), with conflicting results. In one case (Ge et al., 2001) tissue loss appeared limited to WM, albeit a significant correlation between GM and abnormal white matter volume (aWM) was found, while a second study (Chard et al., 2002) found a decrease in both GM and WM fractions, although lesion load (LL) correlated only with GM loss.

The aim of our study was to simultaneously measure in vivo, in a group of patients with RR-MS, aWM and normal brain tissue volumes, assessing their relationship and the possible association with disability, using an unsupervised multiparametric segmentation procedure based on a relaxometric approach (Alfano et al., 2000).

Materials and methods

Patients

Fifty MRI studies (18 males; 32 females; mean age, 35.6 ± 8.7 years; age range 21–59) from patients with clinically definite MS according to Poser criteria (Poser et al., 1983) with a relapsing-remitting course (Lublin and Reingold, 1996) were analyzed. All patients had been previously treated only by brief courses of steroids during clinical exacerbations.

Expanded Disability Status Scale (EDSS) testing (Kurtzke et al., 1983) was performed within 11.1 ± 9.9 days (mean ± standard deviation; range, 0–38 days) from the MRI. For comparison, MRI studies from 54 normal volunteers (NV) spanning over a wider age range (mean age, 38.5 ± 13.1 year; age range, 16–70; 22 males, 32 females) were used. Exclusion criteria for NV were evidence of cardiovascular, metabolic, neurological, and psychiatric impairment, as well as abnormal MRI examination.

Both NV and MS patients agreed to participate in the study by signing a written informed consent, and the ethical committees of participating Institutions previously approved the protocol.

MR studies

MRR protocol included two interleaved sets of 15 slices covering the entire brain obtained at 1.5 T (Magnetom, Siemens, Erlangen, Germany), sampling the brain at a total covering the entire brain obtained at 1.5 T (Magnetom, Siemens, Erlangen, Germany), sampling the brain at a total of 30 levels. Each of the two sets was composed of two conventional spin-echo sequences, generating 15 T1w (600/15 TR/TE) and N(H)/T2w (2200/15–90 TR/TE1–TE2) images (25 cm FOV, 256 × 256 acquisition matrix, 4-mm-thick axial slices).

Relaxation rate calculation

The segmentation procedure used in the present work is primarily based on the relaxation rates of the voxels for tissue assignment. For each level, from the T1w, T2w, and N(H) triplet of images, relaxation rate (R1 and R2) and N(H) maps were calculated, as described in detail elsewhere (Alfano et al., 1995) and summarized hereinafter.

To calculate the R2 maps, pixel values are calculated from the two N(H) and T2w images according to the following equation,

\[
\frac{S_2}{S_1} = e^{-\frac{TE_2 - R_2}{R_1}}
\]

where TE1 and TE2 are the echo time of the first and second echo, respectively, and S1 and S2 are the corresponding signal intensities.

To calculate the R1 map pixel values, which cannot be obtained analytically, the program preliminarily generates a look-up table of the signal ratios expected for R1 values ranging from 0 to 2.55 s⁻¹ at the TR and TE of the T1w and N(H) images. For each voxel the signal intensity ratio is then looked up in the table, and the corresponding R1 is assigned. Finally, N(H) maps are calculated by correcting N(H) images according to T2 decay on a pixel-by-pixel basis.

Segmentation procedure

R1 and R2 and N(H) maps were then segmented into GM, WM, aWM, and cerebrospinal fluid (CSF) with a fully automated multispectral procedure (Alfano et al., 1997, 2000), which provides for each level the corresponding binarily segmented images (i.e., each intracranial voxel is labeled as belonging univocally to GM, WM, aWM, or CSF).

Briefly, the program takes advantage of the known position of the brain tissue clusters (both normal tissue and aWM) in the multiparametric R1, R2, N(H) space to identify WM, GM, and CSF, as well as aWM voxels. To achieve this, a set of regions of interest (ROI) defined a priori in the multiparametric space is used to preliminarily assign the voxels accordingly.

Segmentation is then performed through the following steps:

(a) realignment across slices of the whole brain R1 histograms to the central slice aiming at correcting the R1 maps for residual RF inhomogeneities (R1 maps are inherently insensitive to RF inhomogeneities); (b) adjustment of the GM/WM R1 cutoff at 45% of the limits of the realigned R1 histogram of the whole brain. The GM/WM R1 cutoff is the only ROI coordinate tailored on the single study (all the other ROI limits are fixed); (c) elimination, using an iterative erosion algorithm, of the voxels belonging to extrameningeal tissues with relaxation rates overlapping those of intracranial tissues (i.e., nasal mucosa and intraocular fluids); (d) assignment of the voxels located outside the predefined ROIs to the prevalent tissue among the surrounding voxels; (e) assignment of the clusters of voxels falling into the aWM ROI (potential lesions) either to aWM or to another tissue (i.e., the tissue with the most similar relaxometric features) by applying a function that takes into account the
shape and size of the potential lesions as well as the composition of surrounding tissues (i.e., only potential lesions that are roundish and surrounded mainly by WM are assigned to aWM).

Assessment of the reproducibility of the segmentation procedure has shown an inter-study standard deviation of 2.5 ml in repeat aWM measurement in MS patients (Alfano et al., 2000), and of 1.1% in repeat normal tissue fractional volumes measurements in NV studies (Alfano et al., 1998), while comparison of the results of the algorithm versus manual definition of plaques in 16 MS studies provided R = 0.993 at linear regression analysis and 87.3% sensitivity (Alfano et al., 2000).

The segmentation program was written using FORTRAN, and MRI data were analyzed on a DEC-Alpha workstation (Digital Equipment, Maynard, MA) running VMS, requiring an average processing time of 2 min/study (including calculation of the relaxation rate maps, automated segmentation, and calculation of fractional volumes).

Validation of normal brain tissue segmentation in MS

The segmentation technique relies mainly on calculated relaxation rates of normal brain tissues, which are fixed for each tissue and relatively independent on sequence parameters, to individuate GM, WM, CSF, and MS plaques.

The only parameter tailored on each study is the R1 cutoff between GM and WM, which is calculated at a fixed percentage between the two tissue clusters (see step b of the segmentation procedure). However, a critical issue in the present work was the possible bias introduced in the segmentation of normal brain tissues by the presence of aWM (with R1 values comparable to GM), which could in theory modify the definition of the R1 threshold between GM and WM clusters. A validation of the accuracy of normal brain tissue volumes in the presence of aWM was therefore included in the present work, to preliminarily verify the effect of the presence of aWM upon GM, WM, and CSF volume measurements.

A set of simulated MS studies was thus generated by substituting in 4 NV studies variable amounts of WM with plaques selected from 4 MS studies with mid to high lesion load. The voxel clusters identified in the MS studies by the segmentation program as aWM were substituted to randomly selected areas segmented as WM in the 4 NV studies. The procedure was iteratively replicated after a 1-voxel 2D erosion of the plaques to generate 37 simulated MS studies with progressively decreasing LL ranging from 230.2 ml to 0 (43 ± 58.7 ml; mean ± standard deviation).

The relationship between aWM presence and the accuracy in fractional normal tissue volumes and the resulting correction factors (see data normalization section) were assessed by linear regression analysis. The software for generation of simulated MS studies was written using Interactive Data Language (IDL, Research Systems, Inc.; Boulder, CO).

Data normalization

For subsequent analysis, in order to assess WM changes, WM and aWM were summed for each study providing total WM volume (tWM). GM, aWM, tWM, and CSF were then divided by ICV (sum of GM, tWM, and CSF), thus providing corresponding fractional results (fGM, faWM, fWM, and fCSF). Before further analysis, fractional volumes underwent a two-step correction, in order to take into account first the effect of aWM on the accuracy of the segmentation, and then the age effect.

In simulated MS studies, where the faWM was 2.8 ± 3.8% (mean ± standard deviation, range 0–14.3%), analysis of the accuracy of normal tissue volumes (fGM, fWM, and fCSF) divided by the corresponding volumes measured...
In the original NV studies prior to aWM addition, there was an underestimation of fGM (the measured value being decreased by \(-0.38\%\) for each aWM unit, Fig. 1A), and an overestimation of fWM \((+0.44\%\); Fig. 1B) and of fCSF \((+0.23\%\) per aWM unit; Fig. 1C). For subsequent analysis, fGM, fWM and fCSF values were thus corrected accordingly.

In the ensuing step, age-related changes, as assessed in the NV data set, were taken into account. Aging-related changes in fractional volumes were preliminarily analyzed independently for males and females to assess age effects using linear regression analysis, which failed to demonstrate a significant difference between the rates of age-related decline of fGM and increase of fCSF in males and females (Fig. 2A).

fWM did not show any age-related change in both genders, and there was no significant difference on the Student’s t test between fWM in females and males (Fig. 2B). NV data were thus pooled for subsequent definition of fGM and fCSF age-related changes. For comparison between NV and MS patients, fGM and fCSF were then corrected for age-related changes by adjusting them to the mean age of patient studies (35.6 years), according to the corresponding rates of yearly decline (for fGM, 0.013\% year\(^{-1}\)) or of yearly increase (for fCSF, 0.015\% year\(^{-1}\)), as measured in the data base of NV. fWM was left uncorrected as no statistically significant relationship with age was found in NV.

**Statistical analysis**

Differences in fractional age-adjusted volumes between NV and MS patients were tested using Student’s t test. To test the hypothesis of relationship between faWM and changes in normal brain tissue volumes, linear regression analysis was performed (Zar, 1984). To test the hypothesis of relationship between disease duration, faWM, and frac-
tional brain tissue volumes versus the clinical status as assessed by EDSS, Spearman’s rank correlation coefficient ($S_r$) was used (Zar, 1984). A significance level of 0.01 was chosen. Statistical analysis was performed using Excel (Microsoft Corporation, Redmond, WA).

**Results**

Mean aWM in NV was 0.32 ± 0.44 ml (mean ± standard deviation; range, 0–1.7 ml), which appeared, at visual inspection, essentially related to noise (i.e., to the presence of scattered voxels with relaxometric properties equal to aWM. Mean aWM in MS studies was 23.8 ± 29.8 ml (mean ± standard deviation; range, 0.4–138.8 ml), corresponding to a faWM of 1.7 ± 2.2% (range, 0.04–11.1%).

In Fig. 2 fGM, fWM, and fCSF values are plotted versus age for both NV and MS patients. fGM and fCSF age-related changes can be appreciated in NV, as well as a reduction in fGM in MS patients, with a corresponding increase in fCSF. fWM values do not significantly differ between NV and MS patients.

Comparison by Student’s t-test between NV and patient’s fractional volumes, after correction of fGM and fWM for aWM and adjustment of fGM and fCSF to the mean patient age, confirmed a significant decrease in fGM (49.5 ± 3.2% vs 53.3 ± 2.1% in MS patients and NV, respectively; $P < 0.0001$), with a corresponding increase in fCSF (13.0 ± 3.8% vs 9.1 ± 2.4%; $P < 0.0001$), while no significant difference emerged between fWM values of NV (37.5 ± 2.6%) and MS patients (37.6 ± 2.2%). A significant negative correlation was found for fGM vs faWM, with a corresponding increase in fCSF (for fGM $R = -0.434$, $P < 0.001$ at regression analysis; for fCSF $R = 0.473$, $P < 0.001$; Fig. 3). No significant correlation was found between fWM and faWM.

In MS patients, mean EDSS was 2.5 ± 0.7 (mean ± standard deviation; range, 1–4), and mean disease duration was 9.9 ± 5.8 years (mean ± standard deviation; range, 1–33). When testing for a relationship between EDSS or disease duration vs imaging data (faWM, age-corrected fGM and fCSF, fWM), no relationship could be found. There was a significant correlation between disease duration and EDSS, Spearman’s rank correlation being $S_r = 0.438$ ($P < 0.005$).

**Discussion**

We found a significant correlation, in RR-MS patients, between T2 lesion load and GM loss, balanced by CSF increase, as measured with a fully automated method, based on the postprocessing of standard clinical MRI sequences.

Previous studies, not assessing separately GM and WM, found both RR- and chronic progressive-MS patients presenting aWM correlation with brain parenchyma fractional volume decrease (Miki et al., 1999b; Rudick et al., 1999) or with ventricular enlargement and corpus callosum thinning. (Losseff et al., 1996; Liu et al., 1999) while at longitudinal analysis a lack of correlation in RR- or secondary progressive-MS (SP-MS) has been reported (Ge et al., 2000b).

The concept of brain atrophy as a potential comprehensive marker of disease status has then been advocated by groups who have found a relationship between brain/ICV ratio and EDSS in MS (Losseff et al., 1996; Miki et al., 1999b; Rudick et al., 1999; Ge et al., 2000b). We did not find such a correlation between EDSS and brain volumetry, possibly because we restricted our analysis to a clinically homogeneous group of RR-MS patients, a category of MS patients where indeed heterogeneous results have been found. Although others found a relationship between brain/ICV and EDSS in RR-MS (Rudick et al., 1999; Miki et al., 1999b), longitudinal analysis disclosed a lack of correlation between EDSS and brain/ICV ratio in RR-MS (opposite to SP-MS) (Ge et al., 2000b), this result being confirmed when analyzing separately GM and WM in RR-MS (Ge et al., 2001).

In agreement with previous works, we did not find a significant correlation between aWM and EDSS (Miki et al., 1999a,b; Filippi et al., 1995), or aWM and disease duration (Miki et al., 1999b).

Unlike our findings, Ge et al. (2001) did not find a reduction in fGM in RR-MS patients, although a significant negative correlation between fGM and aWM was found. Our patient cohort, however, comprises a larger range of aWM (up to 138.8 ml with more than 40% of the patients with aWM greater than 20 ml, as compared to their aWM lower than 60 ml, with only 10% of the patients above 20 ml). This difference in lesion load between the two groups of patients, given the correlation between aWM and fGM (a finding common to both studies), may have partly hindered in their case the detectability of a significant mean fGM reduction in MS patients, a finding also reported by Chard et al. (2002).

Unlike other studies (Ge et al., 2001; Chard et al., 2002) we did not find a reduction in fWM once LL (plaque volume) was added to WM, reduction in WM being an obvious finding if plaque volume is not considered. It is somewhat surprising that a residual decrease in WM, caused at least by pre- and postlesion axonal loss, was not found even in patients with high aWM. A possible speculative explanation of this finding is that the axonal loss is somewhat compensated, in terms of volume, by gliosis and edema, two phenomena which occur in areas segmented as plaques, and then added to the fWM compartment in our data analysis. Also, plaques located in GM structures, which should account for at least 9% of the aWM (Brownell and Hughes, 1962; Peterson et al., 2001), are pooled in the fWM volume when summing plaques with WM in determining fWM volume, providing a potential source of overestimation of the fWM in MS. Other segmentation-related issues should also be considered, which may potentially affect the


GM/WM segmentation, hampering the detection of WM loss.

It is, for example, of note that modifications in tissues outside aWM (e.g., in the so-called normal appearing white matter, NAWM, for which both T2 and T1 changes have been indeed reported) may interphere with the segmentation of normal brain tissues. While T2 changes in NAWM (Whittall et al., 2002) represent less of a problem for GM/WM differentiation, which is essentially obtained in the R1 domain, T1 changes in NAWM appear more relevant to the segmentation process.

NAWM T1 change figures have recently been published (Vaithiaanatar et al., 2002), the two modifications in MS patients therein reported being a less than 8% increase in the NAWM T1 histogram (i.e., a reduction of the R1, a potential source of WM overestimation/GM underestimation) and a 68% increase in its standard deviation (a stronger source of WM underestimation/GM overestimation). The net consequence of these two changes on our segmentation technique, if any, would be an overestimation of GM and underestimation of WM, mainly due to the predominant effect of the widening of the R1 distribution of WM. However, our results show an opposite trend. Therefore, this effect cannot explain the observed findings.

Regarding the correction for aWM-related bias in the segmentation of normal brain tissue, a question can be raised regarding the specificity of the correction (i.e., if the specific features of the aWM of the single patient can introduce different degrees of segmentation inaccuracies depending on its mean relaxometric properties). This would be in fact theoretically possible if the relaxometric features of the MS lesions were largely distinct from patient to patient. However, since MS plaque characteristics can be extremely variable in the same patient ranging from mildly decreased R2 (T2 hyperintense areas) to markedly low R1 (CSF-like “black holes” in T1), this heterogeneity appears to be more frequently intra- rather than interpatient. It should also be reminded here that the effect of aWM on GM/WM segmentation was clearly inferior to the observed GM and CSF volume changes measured by the segmentation procedure, which proved to be robust with respect to normal brain tissue volume measures in the presence of aWM (Fig. 1).

In conclusion, we have found that brain atrophy in RR-MS mainly relates to GM loss, which in turn correlates with lesion load, global WM volume appearing unaffected by the pathology once MS plaque volume is taken into account. LL and GM loss, as well as corresponding CSF increase, do not appear to significantly affect the clinical status as assessed by EDSS, which correlates instead with disease duration. However, a definitive confirmation of these findings may only come from pathologic examinations, and in the absence of such a “gold standard” the hypothesis for volumetric compensation of WM loss can only be speculative at this stage. Further studies are needed especially to clarify the impact of these changes on the whole clinical status, particularly aiming at defining the relationship between these measures and the cognitive impairment sometimes present in these patients.

Acknowledgments

Part of this work has been carried out within PVEOut, an EC cofinanced R&T project (5th Framework Programme, Contract QLG3-CT2000-594). We gratefully acknowledge support by the “Progetto Neuroscienze” of CNR/MIUR, and by Ministero della Sanità (Progetto Alzheimer, Attività di ricerca finalizzata D. Lgs. 502/92 e D. Lgs. 229/99).

References


