Quantitative assessment of MRI lesion load in multiple sclerosis
A comparison of conventional spin-echo with fast fluid-attenuated inversion recovery

M. Filippi, 1 T. Yousry, 3 C. Baratti, 2 M. A. Horsfield, 5 S. Mammi, 1 C. Becker, 3 R. Voltz, 4 S. Spuler, 4 A. Campi, 2 M. F. Reiser 3 and G. Comi 1

Departments of 1 Neurology and 2 Neuroradiology, Scientific Institute Ospedale San Raffaele, Milan, Italy, the Departments of 3 Radiology and 4 Neurology, Klinikum Grosshadern, Munich, Germany and the 5 Department of Medical Physics, University of Leicester, UK

Summary
In this study, we compared a fast fluid-attenuated inversion recovery (fast-FLAIR) sequence to conventional spin-echo (CSE) in the evaluation of brain MRI lesion loads of seven patients with clinically definite multiple sclerosis. Interleaved CSE (3000/20, 5 mm contiguous axial slices) and fast-FLAIR (9000/150/2200, 5 mm contiguous axial slices) sequences were performed on a 1.0 T machine. Lesions were counted consensually by two observers and then segmented independently by two other observers using a local thresholding technique, with subsequent manual editing in the case of poorly defined lesions. Four hundred and two lesions were detected in at least one of the two sequences: 128 were seen only on fast-FLAIR, 17 only on CSE. Forty-one lesions were larger on fast-FLAIR, while no lesion was considered larger on CSE. The numbers of periventricular (P = 0.05), cortical/subcortical (P = 0.02) and discrete (P = 0.05) lesions detected using fast-FLAIR were higher than those detected using CSE. The median lesion load was 7185 mm$^3$ on CSE and 8418 mm$^3$ on the fast-FLAIR, the average being 18% (range = 11.6–29%) higher on the fast-FLAIR images. Lesion contrast ratio was higher for lesions on the fast-FLAIR than on the CSE sequence (P < 0.0001). The percentages of poorly defined lesions which needed manual editing after the local thresholding technique was applied and the total time needed for the measurements were lower (P < 0.001) when fast-FLAIR images were used compared with CSE. This resulted in a reduced inter-rater coefficient of variation in measuring lesion volumes. Our data indicate that fast-FLAIR sequences are more sensitive than CSE in detecting multiple sclerosis lesion burden and that fast-FLAIR is a promising technique for natural history studies and clinical trials in multiple sclerosis.

Keywords: multiple sclerosis; MRI; lesion volumes; conventional spin-echo; fast fluid-attenuated inversion recovery

Abbreviations: CR = contrast ratio; CSE = conventional spin-echo; fast-FLAIR = fast fluid-attenuated inversion recovery

Introduction
The authors of several recent studies have proposed the use of FLAIR sequences in the evaluation of many neurological conditions (De Coene et al., 1992; Hajnal et al., 1992; White et al., 1992; Bergin et al., 1993; De Coene et al., 1993; Schwieso et al., 1993; Thomas et al., 1993; Bradley et al., 1994; Litt and Moore, 1994; Rydberg et al., 1994; Thorpe et al., 1994; Baratti et al., 1995; Hashemi et al., 1995; Takanashi et al., 1995), including multiple sclerosis (De Coene et al., 1992; White et al., 1992; Bergin et al., 1993; De Coene et al., 1993; Bradley et al., 1994; Litt and Moore, 1994; Rydberg et al., 1994; Thorpe et al., 1994; Baratti et al., 1995; Hashemi et al., 1995; Takanashi et al., 1995). Fluid attenuated inversion recovery sequences produce heavily T2-weighted images with suppression of the CSF signal by
combining a long inversion time inversion recovery sequence with long echo-time. The main limitation of these sequences (i.e. the very long acquisition time) has recently been partly overcome by the concomitant use of a rapid acquisition with relaxation enhancement (Henning et al., 1986) imaging (fast-FLAIR).

Multiple sclerosis lesions are characterized by heterogeneous pathological features, all resulting in increased water mobility and consequently in an increase of T₂ (McDonald et al., 1992). Hence T₂-weighted images are traditionally used for diagnosing multiple sclerosis. In contrast, for monitoring the evolution of the disease, it is necessary to obtain reliable and reproducible measures of the total lesion load; this can be done using lesion segmentation techniques based on signal intensity threshold (Pannizzo et al., 1992, Wicks et al., 1992; Kapouleas et al., 1993; Paty et al., 1994; Filippi et al., 1995, f). In this respect, T₂-weighted images are not wholly suitable, since the hyperintense CSF signal would be above the threshold. At present, mixed proton-density, T₁- and T₂-weighting, giving partial suppression of the CSF, is commonly used for quantitative assessment of lesion load in multiple sclerosis (Filippi et al., 1995).

Our aims in this study were to evaluate conspicuity of multiple sclerosis lesions as detected by CSE and fast-FLAIR and to compare the number of lesions and the total lesion loads detected with these two sequences when using a local thresholding technique to segment the lesions.

**Patients and methods**

**Patients**

Seven patients with clinically definite multiple sclerosis (Poser et al., 1983), recruited from the multiple sclerosis population attending the Department of Neurology, Klinikum Grosshadern, University of Munich, Germany, entered the study. The mean age was 42 years (SD = 14), mean duration of the disease was 5 years (range 3–30) and mean expanded disability status scale (Kurtzke, 1983) score was 2.0 (range 0.0–5.0). Four patients had relapsing–remitting multiple sclerosis and three secondary progressive multiple sclerosis. All patients were clinically stable at the time of the MRI examination (no clinical relapses for at least the preceding 6 months) and had not taken immunosuppressive or immunomodulating treatments (including steroids) for at least 6 months. At the time MRI was performed, each patient was assessed neurologically, and disability was assessed using the expanded disability status scale. Written informed consent was obtained from all the patients before inclusion in the study.

**MRI**

Interleaved CSE (TR = 3000, echo-time = 20, 18 axial slices 5 mm thickness; 256x256 image matrix, field of view = 200 mm) and fast-FLAIR (TR = 9000, echo-time = 150, inversion time = 2200, 18 axial slices 5 mm thickness; 256x256 image matrix, field of view = 200 mm; echo-train length = 15) sequences were performed using a 1.0 T machine in the same session without moving the patient from the scanner. The total acquisition time for CSE was 6 min 54 s and for fast-FLAIR was 5 min 6 s. Patients were positioned according to guidelines established by a European Community Committee for multiple sclerosis (Miller et al., 1991).

**Quantification of the MRI abnormalities**

Conventional spin-echo and FLAIR scans were compared consecutively in a random order and then side-by-side by two of us (M.F. and T.Y.) by agreement. Lesions on each sequence were counted and marked on the hardcopies. Only hyperintense areas which were considered lesions by both the raters with high confidence were counted as lesions. The site of each lesion was also recorded: those within the brainstem or cerebellum were referred to as posterior fossa, those abutting the lateral ventricles as periventricular, those in or immediately adjacent to cerebral cortex as cortical/subcortical and supratentorial lesions away from the ventricles or cortex as discrete. The lesions present on both the scans, those present on only one of the two scans and those larger on one of the two scans were recorded.

The lesion load was assessed quantitatively by estimating the total volume of the white matter lesions using a semi-automated local thresholding technique. The software used was the ‘just/image’ library (University of North Carolina, Chapel Hill, North Carolina, USA) and Dispir image display software (Mr David Plummer, University College, London, UK) running on a computer workstation (Sun Microsystems, Mountain View, Calif., USA). The measurements were performed using a mouse-controlled cursor on the computer display by clicking on the perimeter of the lesions. The computer program first examines the image in a region close to where the mouse was clicked to find the strongest local intensity gradient, which it considers to be the edge of the lesion. Then the lesion is outlined by following a contour of iso-intensity from this initial edge point, thus defining the lesion as a region of the image where the signal intensity is locally above the signal intensity at the initial edge position. This sometimes gave poor results because other structures (such as abutting grey matter) adjacent to the lesion were as bright, leading to the contour moving away from the lesion outline. When this happened, the lesions were outlined manually by the rater, by moving the cursor in order to define the boundary of the lesion. Each outline was stored on computer disk before automatic computation of the lesion volume. The lesion volume was calculated simply as the lesion area multiplied by the slice thickness. The total time needed for the measurements and the amount of manual editing were recorded for each sequence in all the patients.

The lesion contrast, normalized to the background tissue intensity, was calculated as a measure of the conspicuity of
the lesion against a background of both CSF and normal appearing white matter. The contrast ratio (CR) was calculated for CSF (CR-CSF) by subtracting the mean signal intensity of the lesion from the mean signal intensity (weighted in proportion to the number of pixels in the area) in two areas of CSF, and dividing by the mean signal intensity in CSF: CR-CSF = (signal lesion–signal CSF)/signal CSF. A similar ratio was calculated for normal appearing white matter (CR-normal appearing white matter). The areas chosen in the CSF and the normal appearing white matter were usually in the same slices in which the lesions were present. When this was not possible, the slices closest to the lesions in which CSF and normal appearing white matter were present were evaluated. In order to avoid partial volume effects, CRs were calculated only for well-defined lesions visible on at least two adjacent slices. The same lesions were assessed in both CSE and fast-FLAIR images.

All the MRI measurements were performed independently by two observers (C.B. and S.M.) and the mean values of these two measurements for each parameter entered the analysis.

Statistical analysis

Differences in numbers of lesions seen and lesion load using the two sequences were studied using the Wilcoxon Rank Signed test, since these data were not normally distributed. Differences in lesion CR, percentages of poorly defined lesions which needed manual editing and time needed to perform the measurements were studied using the two-tailed Student’s t test for paired data. The inter-observer agreement in measuring lesion volumes from both sequences was calculated as previously indicated (Bland and Altman, 1986; Filippi et al., 1995e). The inter-observer variability for each technique was assessed by calculating the coefficient of variation. The coefficient of variation was computed as the ratio of SEMs of the two lesion volume measurements (Zar, 1984).

Results

Four hundred and two lesions (26, posterior fossa; 108, periventricular; 184, cortical/subcortical; 84, discrete) were seen using the two sequences. Two hundred and sixteen lesions (8, posterior fossa; 7, periventricular; 71, cortical/subcortical; 58, discrete) were seen and were considered to have similar sizes on both the sequences. Another 41 lesions (0, posterior fossa; 16, periventricular; 18, cortical/subcortical; 7, discrete) were also seen on both the sequences but were considered larger on fast-FLAIR images, while the reverse was never found. Finally, 17 lesions (15, posterior fossa; 1, periventricular; 0, cortical/subcortical; 1, discrete) were seen only on CSE and 128 (3, posterior fossa; 12, periventricular; 95, cortical/subcortical; 18, discrete) were seen only on fast-FLAIR (Fig. 1). In Table 1, the medians and the ranges of the lesion numbers seen in the patients are presented for the four anatomical locations. Significantly higher numbers of periventricular (P = 0.05), cortical/subcortical (P = 0.02) (Fig. 2) and discrete (P = 0.05) lesions were seen on fast-FLAIR.

The median lesion load was 7185 mm$^3$ (range = 3675–26961 mm$^3$) on CSE images and 8418 mm$^3$ (range = 4740–30116 mm$^3$) on the fast-FLAIR (using the Wilcoxon Rank Signed test, difference = 28; P = 0.02). Thus, the average increase was 18% (range = 11.6–29%) for the latter. The lesion volumes measured using fast-FLAIR images were higher than those obtained using CSE in all cases for both observers (Table 2). The mean inter-observer agreement in lesion volume measurements was 94.3% (range = 89.5–97%) for CSE and 97.6% (range = 97–99.8%) for the fast-FLAIR. The coefficient of variance was 4.1% for CSE and 2.8% for fast-FLAIR.

Lesion conspicuity was studied in 59 lesions. Mean CR-CSF was 0.45 (SD = 0.1) for lesions on CSE images and 0.75 (SD = 0.11) for those on fast-FLAIR images (P < 0.0001). Mean CR-normal appearing white matter was 0.24 (SD = 0.1) for lesions on CSE images and 1.18 (SD = 0.25) for those on fast-FLAIR images (P < 0.0001).

The mean percentages (± standard deviation) of poorly defined regions of interest with the local thresholding technique (which needed manual editing after the contouring technique was applied) were 33 (±7) for CSE sequences and 25 (±8) for the fast-FLAIR sequences (P < 0.001). The total times needed for evaluation of lesion load using CSE (mean ± SD = 73 ± 10 min) and fast-FLAIR (mean ± SD = 64 ± 5 min) images was significantly different (P < 0.001).

Discussion

In several pathological studies (Brownell and Hughes, 1962; Lumdsen, 1970; Revesz et al., 1994) it has already been demonstrated that cortical/subcortical lesions are a frequent finding in the brains of patients with multiple sclerosis. However, using CSE, such lesions normally go undetected either because they are masked by the hyperintense CSF signal, or because their signal intensity is similar to that of grey matter. With FLAIR, the CSF suppression can reveal these cortical/subcortical lesions, and the very long echo-time enhances the lesion/grey matter contrast (Fig. 2). In addition, the rapid acquisition with relaxation enhancement sequence (and variants such as fast-FLAIR sequences) suffers from the unusual property that the brightness of structures in the image depends on their size, not just on the intrinsic molecular properties (Constable and Gore, 1992). Because the central part of the data matrix is acquired at the nominal echo time, which in our case is relatively long, the effect is to enhance signal in the outer parts of the data matrix. This emphasizes small features and edges, and may account for some of the improved conspicuity of the smaller, cortical/subcortical lesions. The greater T2 weighting also improves lesion/normal appearing white matter contrast leading to better lesion conspicuity throughout the white matter.
Fig. 1 (A) Axial CSE image; (B) axial fast-FLAIR image at the same anatomical level. The visible white matter abnormalities are much more extensive in B. In detail, four cortical/subcortical lesions are clearly seen in the frontal and parietal lobes of the fast-FLAIR scan (arrows), while the presence of two of them can be only suspected (right hemisphere) and the other two are not visible at all (left hemisphere) in the CSE image.

Table 1 Median numbers (and ranges) of lesions seen on CSE and fast-FLAIR sequences for the four anatomical locations

<table>
<thead>
<tr>
<th></th>
<th>CSE</th>
<th>Fast-FLAIR</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
<td>1 (0-16)</td>
<td>0 (0-2)</td>
<td>n.s.</td>
</tr>
<tr>
<td>PV</td>
<td>13 (9-18)</td>
<td>15 (10-21)</td>
<td>0.05</td>
</tr>
<tr>
<td>C</td>
<td>11 (7-23)</td>
<td>25 (15-38)</td>
<td>0.02</td>
</tr>
<tr>
<td>D</td>
<td>9 (4-15)</td>
<td>13 (6-17)</td>
<td>0.05</td>
</tr>
<tr>
<td>All</td>
<td>37 (28-59)</td>
<td>59 (36-66)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

n.s. = not significant; PF = posterior fossa; PV = periventricular; C = cortical; D = discrete. Statistical analysis: Wilcoxon Rank Signed test.

(Fig. 2), whilst in the periventricular regions, CSF suppression markedly facilitates the delineation of the boundary between the lesion and the ventricle. The opposite seems to happen in the posterior fossa, where CSE detected 15 lesions not seen with the fast-FLAIR sequence. A possible explanation for the poor sensitivity of fast-FLAIR in this region is that the relatively long echo time used might be responsible for a loss of contrast in lesions where T2 is not greatly elevated. However, the poor sensitivity of fast-FLAIR sequences in detecting posterior fossa lesions is of concern, since lesions in this site are thought to contribute particularly to disability (Filippi et al., 1995b).

In this study, we found that the lesion volumes detected using the fast-FLAIR sequence are on average 18% higher than those detected using CSE, which, at present, is most commonly used to assess lesion volume in multiple sclerosis (Filippi et al., 1995f). Some of this was due to extra cortical/subcortical lesions seen in the fast-FLAIR scans (Fig. 1). However, cortical/subcortical lesions tend to be smaller than those in white matter, and clearly these are not the only contribution to the increase in total lesion load. More white matter lesions were also evident in the fast-FLAIR images, probably due to the improved contrast which makes small lesions, which may go undetected because of partial volume effects (Plante and Turkstra, 1991), visible. The contribution of these small lesions has been investigated in two previous studies (Barbosa et al., 1994; Filippi et al., 1995c). Barbosa et al. (1994), using pixel-by-pixel relaxation time mapping, found that small lesions of 1–2 pixels in size, which normally go undetected and are considered part of the normal appearing white matter, account for an important fraction of the total lesion load in multiple sclerosis patients. Filippi et al. (1995c) found that the measured lesion load increases significantly as the image resolution in the slice direction is reduced, at least in part, because of the improved detection of smaller lesions. In addition, some lesions both in the cortex and in white matter appeared larger on fast-FLAIR scans. There is evidence from Filippi et al. (1995a) that there are abnormalities, detected by magnetization transfer imaging,
Lesion volumes measurements in MS using CSE and FLAIR

Fig. 2 (A) Axial CSE image; (B) axial fast-FLAIR image at the same anatomical level. Although diffuse hyperintensities are more evident in the posterior white matter of the central semiovalia in B, lesions are characterized by better contrast. In addition, a subcortical lesion in the left frontal region is present in B but it is not visible in A.

Table 2 Mean lesion volumes (mm$^3$) present on CSE and fast-FLAIR sequences*

<table>
<thead>
<tr>
<th>Patient</th>
<th>CSE</th>
<th>Fast-FLAIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26961</td>
<td>30116</td>
</tr>
<tr>
<td>2</td>
<td>7826</td>
<td>8735</td>
</tr>
<tr>
<td>3</td>
<td>5590</td>
<td>6885</td>
</tr>
<tr>
<td>4</td>
<td>7185</td>
<td>8418</td>
</tr>
<tr>
<td>5</td>
<td>13320</td>
<td>15632</td>
</tr>
<tr>
<td>6</td>
<td>6445</td>
<td>7625</td>
</tr>
<tr>
<td>7</td>
<td>3675</td>
<td>4740</td>
</tr>
</tbody>
</table>

*The values presented are the mean lesion volumes obtained from the measurements of two observers.

in the normal appearing white matter adjacent to visible lesions of more disabled patients, and it would seem that the improved contrast of the fast-FLAIR sequence shows that some lesions extend beyond the boundary which would be perceived on a CSE scan.

These data, resulting from different approaches, suggest that the lesion load missed using conventional MRI techniques is important and that the missing lesion load might be biased towards different variables, such as a particular lesion size, lesion load (Filippi et al., 1995c) or relaxation characteristic of the lesions due to their pathological heterogeneity (McDonald et al., 1992). Further improvements in the sensitivity of multiple sclerosis lesion load measurement might be expected by combining the fast-FLAIR sequence with thin slice acquisition (Bradley et al., 1994; Hashemi et al., 1995). Hashemi et al. (1995) have indeed demonstrated abnormalities consistent with multiple sclerosis on 2 mm thick fast-FLAIR images in 43% of their patients considered to have no abnormalities on 5 mm CSE images. However, it remains to be established whether fast-FLAIR sequences are also more sensitive than CSE in detecting lesion load change over time, which, at present, is used as secondary end-point measure in phase III clinical trials.

The better CR of lesions in fast-FLAIR sequences, with consequent clearer definition of their edges and reduced human interaction lead to improved inter-observer reproducibilities in lesion volume measurements. The time required for manual editing was indeed reduced with fast-FLAIR, even though the total lesion volumes and numbers of lesions were greater. Manual editing was performed mainly for central slices of fast-FLAIR images in two patients, because of the presence of diffuse areas with high signal intensities in periventricular regions, a feature which is also seen in FLAIR scans of normal subjects (Thorpe et al., 1994). On the other hand, more manual editing was necessary for all the patients when the CSE sequence was used, especially for lesions located in the posterior fossa, temporal lobe and subcortical areas. The improved lesion conspicuity seen with fast-FLAIR should make any technique for lesion...
segmentation based on pixel intensities more robust. Artifacts are sometimes seen on fast-FLAIR images due to CSF flow bringing fresh spins into the image slice. These have proved difficult to eliminate entirely (Ryderberg et al., 1994), but it is normally easy to distinguish them from pathologically based hyperintensities (Fig. 2). Nevertheless, they may pose a problem when attempting to automate lesion detection and quantification.

The time required to evaluate the scans influences reproducibility since if multiple observers are employed because of time constraints, the reproducibility will also be reduced (Filippi et al., 1995e). The good reproducibility of the results of fast-FLAIR means that it is potentially very important for future multiple sclerosis clinical trials since it will become the technique of choice and hence more patients will be studied using it.

In several studies it has been found that the correlation between changes on CSE sequences and the development of disability in multiple sclerosis is not as strict as might be expected (McDonald et al., 1994; Miller, 1994; Filippi et al., 1995d). The extent of abnormalities in the spinal cord and/or the pathological heterogeneity of multiple sclerosis lesions are two possible explanations for this discrepancy, for which there is emerging evidence (Miller, 1994). In addition, inaccuracies in measuring lesion volumes should also be considered. The results of the present study suggest that fast-FLAIR sequences may improve clinical/MRI correlations for two reasons. First, detection of cortical/subcortical lesions is clearly important, since patients with multiple sclerosis have characteristically high prevalence of cognitive deficits (Peyser et al., 1990). Secondly, the improved overall sensitivity of fast-FLAIR in detecting lesions, and particularly in refining the extent of the lesions, may reveal, in more detail, the changes in lesion load over time. Indeed, it is conceivable that, by using fast-FLAIR instead of CSE, more subtle changes in lesion load will be detected, improving both our understanding of the disease and our ability to monitor it.

Acknowledgements
This work derives, in part, from the collaboration made possible by the EC funded (ERBCHRXCT 940684) European Magnetic Resonance Network in Multiple Sclerosis (MAGNIMS). The Dispimage display program was written and provided by Mr D. Plummer (Department of Medical Physics, University College, London, UK).

References


Received December 20, 1995. Accepted February 19, 1996