Magnetic resonance imaging in monitoring the treatment of multiple sclerosis: Concerted Action Guidelines

D H Miller, F Barkhof, I Berry, L Kappos, G Scotti, A J Thompson

Abstract
Serial gadolinium enhanced MRI of the brain detects much clinically silent disease activity in early relapsing-remitting and secondary progressive multiple sclerosis (MS), and thus has an important role in monitoring the effects of therapy. Based on the proceedings of a recent Commission of the European Communities (CEC) workshop and a review of the literature, guidelines are presented for using MRI to monitor treatment trials in MS. The guidelines consider: A) MRI system and techniques; B) patient selection; C) trial design; D) analysis of results. Priorities for future research are also indicated.

The traditional measure of therapeutic outcome in multiple sclerosis (MS) is clinical disability. Because fixed disability usually increases slowly (despite frequent short term fluctuations), clinical trials are of long duration (usually at least two years). Clinical assessment is necessarily subjective, and considerable differences occur between experienced neurologists when quantifying disability.

MRI of the brain readily detects lesions in multiple sclerosis (MS). Most of the MRI lesions are asymptomatic, and serial studies have revealed new lesions five to 10 times more frequently than clinical relapses. Thus serial brain MRI, when added to clinical assessment, should allow more sensitive and objective monitoring of therapy.

There are many issues which need to be considered when formulating the “optimum” MRI monitoring protocol. As part of the programme of Concerted Action on Multiple Sclerosis organised under the auspices of the Commission of the European Communities (CEC), a workshop on “The monitoring of MS by magnetic resonance” was held in London, United Kingdom on 9–10 November 1990. Following the meeting, a small group was nominated to prepare a set of guidelines for the use of MRI in treatment trials. We present these guidelines, which are based primarily on the conference proceedings but include a review of the literature. The guidelines are likely to change as new developments arise in this rapidly changing field. Nevertheless, it is hoped that they will provide a basis for multicentre studies and maximise the information which can currently be derived from MRI, when monitoring the effects of new treatment.

A) MRI system
I Strength of magnet
High resolution is needed for accurate monitoring of MS since many plaques are small such as 2–5 mm diameter. At low fields (for example, 0.15T) a slice thickness of 10 mm is usually required to obtain satisfactory images in a reasonable time, whereas at 0.5T and above, good images are obtained with 5 mm thick slices.

Patients with secondary progressive MS have been studied serially at 0.15T and 0.5T. At the lower field with 10 mm thick slices, 10/25 (40%) new lesions disappeared completely on follow up scans, whereas only 7/109 (6%) new lesions disappeared at 0.5T using 5 mm thick slices. We therefore recommend that studies are performed on systems with a field strength of 0.5T or more. Serial examination of individual patients should be undertaken on the same scanner, as there is always some variation in image quality between scanners.

2 Repositioning
To detect changes on follow up scans with reliability, accurate repositioning of the patient is essential. The history of stereotaxy reminds us that if we want to be very precise, only fixation devices and stereotactic methods can be used. An alternative that will be widely available in the future will be the use of computer programs of volumetric display and reconstruction and orientation of slices.

At present, the most suitable methods are ones which orientate around certain well visualised anatomical landmarks. Whichever method is chosen, we recommend that it is the same for all centres involved in a multicentre trial. The following method has been used successfully in some of our units: the patient is placed in a comfortable manner within the head coil and moved to the scanning position within the magnet bore. A low resolution coronal pilot scan is performed, followed by a high resolution sagittal T1-weighted scan to demonstrate the anterior and posterior commissures, which serve as the landmarks for the oblique orientation of the axial plane (parallel to the bicommissural line). Height adjustment is achieved by placing the middle slice over the inter-commissural line. Motion artefact can be reduced by using a head fixation device such as a velcro band.

3 Slice orientation
Axial imaging has been used in virtually all serial studies in MS. One study sug-
gested that additional infratentorial lesions may be detected on sagittal images, but there were a number of differences between the axial and sagittal protocols which favoured the latter. If all other parameters are constant, there is probably little difference in lesion detection with varying slice orientation. We recommend using an axial profile as this is the most widely accepted.

4 Spatial resolution
There is inevitably a trade off between optimal resolution and length of imaging time. With this caveat in mind we make a number of recommendations
It is desirable to obtain the thinnest slice which provides an adequate signal-to-noise ratio in a reasonable period of time. For multicentre studies, with systems anticipated to have field strengths of 0.5 to 1.5T, we recommend a uniform 5 mm slice thickness. It is essential to have a minimal interslice gap (less than 25% of the slice thickness). We recommend using a 25-6 cm field of view and a 256 x 256 matrix (providing 1 mm² pixels) with 2 excitations, and a quadrature head receiver coil if available.

5 Imaging sequences for unenhanced MRI
T2-weighted spin echo (SE) sequences are more sensitive than T1-weighted ones for the demonstration of lesions in MS. A moderately T2-weighted sequence, where the CSF signal is slightly less intense than that of normal white matter, is optimal for demonstrating periventricular lesions. A more heavily T2-weighted sequence, in which lesion intensity is greater than that of grey matter, but in which CSF intensity is not excessive, is optimal for demonstrating sub-cortical lesions. Berry (workshop) detected more lesions when both moderate and more heavily T2-weighted sequences were used.

Although there are a number of faster techniques for obtaining T2-weighted images, none produces as good a signal-to-noise ratio as standard T2-weighted SE imaging. We therefore recommend the latter approach, with simultaneous acquisition of a moderate and a more heavily T2-weighted sequence.

6 Gadolinium enhancement
In chronic relapsing experimental allergic encephalomyelitis, an animal model of immune mediated demyelinating disease resembling MS in some important respects, gadolinium-(Gd)-DTPA enhancement (which occurs in lesions with an abnormal blood-brain barrier) is invariably associated with inflammation. Gd-DTPA enhancement is common in MS lesions, where it also indicates the presence of inflammation.

Enhancement occurs in almost all new lesions which appear during serial monthly scans of patients with relapsing-remitting or secondary progressive MS, the few which do not enhance are invariably small (Thompson, London). The enhancing phase of new lesions is transient, having disappeared within four to six weeks in about 2/3 of lesions.

Three observations emphasise the value of gadolinium enhancement in monitoring MS. First, enhancement occurs in some older lesions - there is no other change in their appearance on T2-weighted unenhanced MR1. Thompson (London) reported that enhancement in such lesions increased the overall detection of active lesions by 10%. Secondly, enhancing lesions have been observed in the cortex and subcortical white matter when they are not visible on unenhanced T2-weighted MR1—detection of lesions in these regions is poor using standard T2-weighted sequences because of the longer T2 of normal cortex. Thirdly, Gd-DTPA enhancement frequently increases the conspicuousness of small new lesions and thereby confirms their sometimes equivocal appearance on a T2-weighted scan.

T1-weighted SE sequences (for example, SE500/20 at 0.5T) are more sensitive than inversion recovery sequences in detecting enhancing MS lesions, and should be used. Scanning should start at least five minutes after injection of Gd-DTPA, as most lesions display maximum enhancement five to 30 minutes post injection. Images should be obtained at the same time post injection on all follow up scans.

Direct comparison of different doses of Gd-DTPA has not been studied for MS. With brain tumours, there is satisfactory enhancement with 0.1 mmol/kg, little additional enhancement with 0.2 mmol/kg, but appreciably less with 0.05 mmol/kg. We therefore recommend a dose of 0.1 mmol/kg, which should be injected over three minutes.

Minor side effects from Gd-DTPA, such as nausea and headache, are infrequently described. Extravascular injection produces local pain, and leads to tissue reaction in rats; we therefore recommend that injection is made either under direct observation outside of the scanner, or via a secure intravenous line with the patient in situ in the scanner—the latter approach has the advantage of identical positioning for the pre- and post-contrast sequences.

Anaphylactoid reactions have been described with Gd-DTPA, although with great rarity—probably of the order of one in 100,000—and no fatal reactions have occurred. Appropriate resuscitation facilities should be available, but the currently reported low incidence of serious reactions does not preclude using Gd-DTPA in monitoring treatment trials in MS. Gd-DTPA should not be used in pregnancy.

Other gadolinium containing contrast agents (for example, Gd-DOTA, Gd(DP-DO3A)) probably have a similar efficacy as Gd-DTPA in terms of enhancement capability, although they have not been evaluated as extensively. It remains to be seen whether they will have even less frequent side effects.

To conclude, gadolinium enhancement should be used in monitoring treatment in
patients with relapsing-remitting or secondary progressive MS (we do not recommend enhancement in patients with primary progressive MS—see section B). It increases the yield of active inflammatory lesions, improves their clarity, and is sufficiently well tolerated to allow serial administration.

B) Patient selection

The clinical course of MS exhibits a number of distinctive patterns. In selecting patients for MRI monitoring, it is necessary to consider four main clinical subgroups which are defined as follows:

1. Early relapsing-remitting MS History of relapses and remissions, without progressive deterioration; the illness is of less than 10 years duration.

2. Benign MS Relapsing-remitting MS, minimal or no disability (usually a Kurtzke EDSS rating < 3) after at least 10 years disease duration.

3. Secondary progressive MS After an initially relapsing-remitting course, progressive deterioration for at least six months, with or without superimposed relapses.

4. Primary progressive MS Progressive deterioration from the outset, without any relapses or remissions.

Only five to 10% of MS patients have the primary progressive course. The remainder have an initial relapsing-remitting course, of whom about two-thirds subsequently develop secondary progression, while most of the remainder have a benign course.

There are important differences between some of these subgroups on brain MRI. Patients with primary progressive MS have lesions which are fewer in number and smaller in size than those with secondary progressive MS. Patients with secondary progressive MS have more confluent periventricular lesions than those with benign MS, and may also have more infratentorial lesions. In contrast to other clinical subgroups, Gd-DTPA enhancement is rarely seen in primary progressive MS, even in new lesions.

Thompson (London) reported serial MRI studies at two to four week intervals in the four clinical subgroups. On follow up scans, lesions were classified as active if they were new, or had enlarged, or displayed new enhancement without changing in size. He found that over a six month period the mean number of active lesions per patient was 20 per year in both the secondary progressive and early relapsing-remitting groups (seven times the clinical relapse rate), 8 per year in the benign group, and 3 per year in the primary progressive group. In secondary progressive and early relapsing-remitting MS, about 80% of new lesions showed Gd-DTPA enhancement; in benign MS, only 33% of new lesions enhanced, and in primary progressive MS, only 5% enhanced. The new non-enhancing lesions in the latter two groups were invariably small (limited to a single 5 mm slice). Kappos (workshop) noted similar differences between the clinical subgroups.

From these findings, we conclude that for MRI monitoring trials, it is important to separate patients with early relapsing-remitting or secondary progressive MS from those with primary progressive MS. It is, however, reasonable to combine patients with early relapsing-remitting and secondary progressive MS into a single group. Because of the much lower frequency of enhancing lesions, we recommend that gadolinium is not used in monitoring patients with primary progressive MS. Patients with benign MS are, by definition, rarely an appropriate group to study in a treatment trial, but should their inclusion be felt appropriate, we recommend that they are studied as a separate group.

C) Design of trial for early relapsing-remitting and secondary progressive MS

1. Introduction MRI monitoring will be most useful in patients with early relapsing-remitting and secondary progressive MS, since they have the most asymptomatic disease activity. The following study design refers exclusively to these two patient groups. There are three general caveats to consider in all treatment trials using MRI monitoring:

a) Clinical monitoring The problems encountered when measuring therapeutic efficacy by clinical criteria alone are reviewed elsewhere. Nevertheless, since therapeutic efficacy must ultimately be determined clinically, regular clinical assessments measuring disability and recording relapses should always be performed in parallel with MRI.

There are other reasons why monitoring brain MRI alone is inadvisable. 1) There is a poor correlation of brain MRI changes and clinical disability, in part because much of the disability arises from spinal cord lesions; 2) Myelin breakdown products lead to a non-specific increase in T2 and are not separately visible on standard sequences—thus MRI does not selectively identify demyelination which produces much symptomatology in MS. It is conceivable that one treatment might suppress demyelination (thus producing clinical stability), but not inflammation (which produces continuing gadolinium enhancement on MRI). Conversely, another treatment may suppress inflammation (and hence stop gado- linium enhancement), but not demyelination (thus clinical deterioration continues).

b) Spinal cord/optic nerves Compared with the brain, MRI of the spinal cord and optic nerves is more time consuming, and currently detects fewer lesions and little asymptomatic disease. At present, we do not recommend imaging of these regions in treatment trials designed to modify the long term course of MS—it may be useful when studying the effects of treatment on specific symptomatic lesions in these regions, for example, in optic neuritis.

c) Should the study be controlled or blind? Because of the considerable variations in frequency of MRI activity between and within
patients over time (see section C3), open and uncontrolled studies are of limited value in determining therapeutic effect. Patients should be prospectively randomised and controlled. While the double-blind procedure is necessary for clinical monitoring, the patient does not need to be “blind” if MRI monitoring alone (which we do not recommend) is being performed. It is advisable to record any corticosteroid treatment given during the study, since high dose intravenous methylprednisolone can produce transient suppression of Gd-DTPA enhancement during treatment, and may shorten the subsequent duration of enhancement.

2 Frequency of scanning and duration of study
With fortnightly serial unenhanced MRI at 0-15T in relapsing-remitting MS, Paty (Vancouver) found that 36% of active lesions (that is, new or enlarging) were seen for less than four weeks. However, given the low sensitivity for small lesions at this field strength, it is possible that subtle changes which may have been apparent with monthly scans on a higher field system were missed. Thompson (London) performed fortnightly studies at 0-5T, and found that while Gd-DTPA enhancement persisted for less than four weeks in 25% of new lesions, such lesions invariably persisted on the unenhanced scan even though they were small, that is, they would have been detected as new lesions on the monthly unenhanced MRI. There is, of course, a much lower yield of enhancing lesions when the scanning interval is greater than one month.

Thus at 0-5T and above, fortnightly Gd-DTPA enhanced MRI detects few additional active lesions compared with monthly scanning. Our patients have complied well with monthly examinations up to six months, and we consider that this is the most suitable interval for studies of such a duration.

In deciding the duration of a study, it is necessary to consider how much MRI activity will occur in a given period. Over a six month period, Thompson (London) identified a total of 172 active lesions (mean = 10; range 0–69) in 17 patients with early relapsing-remitting or secondary progressive MS (table). In the first three months, 4/17 patients had no active lesions, while only two had none after six months. These results suggest sufficient data for analysis should be available after six months, and that this would therefore be an appropriate study period. It will be necessary to monitor for longer than six months if a treatment is expected to take longer to be effective. The degree of compliance with monthly examinations over a year or more is uncertain. Further work is needed to compare the efficacy and tolerance of different scanning frequencies (for example monthly, three monthly, and six monthly) in such long term studies.

3 Patient number
Since MRI detects changes more frequently than clinical methods, it seems likely that its use as a monitor will allow therapeutic efficacy to be established with smaller numbers of patients in a relatively short period of time. We have made some case number calculations for a placebo controlled study in which monthly gadolinium enhanced MRI is performed for six months. The calculations assume that Thompson’s 17 patients are representative of placebo–treated controls. A number of different outcomes could be measured; two are now presented as examples.

a) The proportion of patients developing active lesions
Eighty eight per cent of Thompson’s patients developed one or more active lesions during the six months. Twenty one patients in both the active and placebo-treated groups would be needed to show a 50% reduction in this proportion (Chi-squared test; a error = 5%, b error = 10%). Eleven patients would be needed in each group to show a 70% reduction, and six would be needed to show a 90% effect.

b) The mean number of active lesions/patient
In Thompson’s patients, the number of active lesions/patient is not normally distributed (table). We cannot therefore calculate a sample size based on differences in the mean. The log of the number of active lesions/patient is approximately normally distributed. If we use the mean of the log to measure outcome, 150 patients would be needed in each group to show a 50% reduction (2 sample t test, a error = 5%, b error = 10%). Fifty patients would be needed in each group to demonstrate a 70% reduction, and 14 to show a 90% effect.

A third outcome which might be measured is the total number of active lesions that develop in both active and placebo–treated groups, but such data is non-parametric, and formulas do not exist for calculating case numbers. A fourth outcome would be the change in total lesion area/volume, but this will probably be unsuitable for a study duration of only six months (see section D3).

The case numbers calculated are not intended to be used as definitive guidelines for treatment trials, based as they are on serial data from only 17 untreated patients; clearly, further natural history studies are needed to

---

Table. Number of active* lesions over six months in early relapsing-remitting and secondary progressive MS

<table>
<thead>
<tr>
<th>Disease type</th>
<th>First 3 months</th>
<th>Second 3 months</th>
<th>Total (6 months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 RR</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2 RR</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3 RR</td>
<td>7</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>4 RR</td>
<td>9</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>5 RR</td>
<td>9</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>6 SP</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7 SP</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8 SP</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9 SP</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10 SP</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>11 SP</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>12 SP</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>13 SP</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>14 SP</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>15 SP</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>16 SP</td>
<td>11</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>17 SP</td>
<td>31</td>
<td>38</td>
<td>69</td>
</tr>
</tbody>
</table>

*includes all new, enlarging or old-enhancing lesions.
RP = early relapsing-remitting MS.
SP = secondary progressive MS.
Magnetic resonance imaging in monitoring the treatment of multiple sclerosis: Concerted Action Guidelines

consolidate this data. Nevertheless, two observations can be made. First, whatever outcome is measured, a highly effective treatment (90%) is likely to produce significant results with small patient numbers. Secondly, when measuring outcome as the mean number of active lesions, substantial patient numbers will be needed to detect a moderate (50%) treatment effect, due to the very large inter-patient variation in the number of active MRI lesions (table).

If over time there was less variation in MRI activity within the same patient than there was between patients, it should be possible to demonstrate treatment effect with smaller numbers, by using patients as their own control, that is, having a period of MRI monitoring before the start of treatment. In Thompson's patients, although there appears to be some consistency within patients over time, using patients as their own control did not result in a reduction in sample size.

4 Summary
In a randomised controlled study, a useful treatment effect should be observed after six months of monthly gadolinium enhanced scanning, whereas clinical monitoring generally requires at least two years of follow up. This early indication of efficacy should be obtained from a small number of patients when the treatment is highly effective. Because the amount of MRI activity varies considerably both between and within patients over time, more substantial numbers may be required to demonstrate less marked but nevertheless useful treatment effects; in this context, multicentre studies will be essential, which underlies the need for a common protocol and database, such as those being developed by Concerted Action on Multiple Sclerosis under the auspices of the CEC. Such short duration MRI studies will influence the decision to proceed with longer term clinical trials.

D) Analysis of results
1 Quality control
a) Hard copies of images. All images should be stored on hard copies. Qualitative analysis of scans will be helped by keeping the images reasonably large. We suggest that there are not more than 12 images on each 30 x 40 cm film. There should be a geometrical marker on hard copies.

b) Other forms of image storage. All images should be stored on tape or disc. Electronic storage is required for volume analysis.

c) Geometrical uniformity. For quantitative analysis, it is important to maintain geometrical uniformity of images between centres and within the same centre over time. The use of a single phantom containing "lesions" would facilitate standardisation between the centres (du Boulay, London), and it is proposed to develop such a phantom for studies in Europe.

2 Qualitative analysis of scans
At each follow up scan, the number of new, enlarging and enhancing lesions should be recorded. In the case of multicentre studies, this analysis should be performed at a single, common centre, by two experienced neuroradiologists.

The minimal size requirement for a lesion is 3 pixels. On unenhanced T2-weighted scans, a lesion should be visible on both echos, although it may be more conspicuous on one. Enhancement should only be recorded when a lesion displays hyperintensity compared with surrounding brain parenchyma on a T1-weighted sequence (T1-weighted pre-contrast images are not necessary).

3 Quantitative analysis of scans
Lesion area/volume measurement
Crude estimates of lesion size can be obtained by measuring its largest diameter. Semi-automated measurement of lesion size is obtained by an observer using a computer assisted tracking method. The intra-rater variability of the technique, however, varies from five to 20% (Paty, Vancouver; Berry, Toulouse; Wicks, London), depending on the degree of experience of the observer. Inter-rater variability is even higher, and the technique is extremely time consuming. More fully automated techniques have been developed, but none is completely satisfactory; for example, thresholds set to exclude normal grey matter will miss white matter lesions with a T2 between that of cortex and white matter.

Consideration of the natural history of MS suggests limitations in using lesion area/volume measurements to monitor MS in the short term. First, new lesions wax and wane in size, and thus the total lesion area/volume fluctuates with time, rather than increasing in a steady fashion. Secondly, given the lengthy duration of MS, the increases in stable lesion area/volume which occur over a brief study period (such as, six months) will often fall within the margin of error of the measurements.

Thus measurement of total lesion area/volume is inappropriate for short term studies, though it may be useful in the long term, for example, over a period of two or more years. In such long term treatment trials, lesion area/volume should be measured yearly from a T2-weighted brain MRI, by a single, experienced observer. To minimise the fluctuations due to rapidly changing acute lesions, we recommend performing two scans one month apart at each yearly follow up, recording the smallest area/volume of the two.

Relaxation time measurement
In MS, the T1 and T2 relaxation times of acute/enhancing lesions display a wide range which overlap with the range of values seen in chronic/non-enhancing lesions and do not therefore differentiate the age or activity of lesions. The normal appearing white matter in some patients has significant elevations in T1 and T2 but the implications of such a finding for prognosis and monitoring are uncertain. We do not recommend the measurement of relaxation times either in lesions or normal
appearing white matter in current treatment protocols.

E) Areas for future research

Further work is needed to determine the best method of monitoring MS by MRI in long term studies (over two years or more), and the value of MRI, in addition to clinical monitoring, in studies of such duration. A study is required to determine whether very frequent scanning (for example, at twice weekly intervals) detects additional disease activity—it is possible for example that enhancement may be very transient (a few days only) in some lesions, particularly in patients with primary progressive MS.

It is important to develop more sensitive and rapid methods for demonstrating spinal cord lesions, since these are a major source of disability in MS. It is also important to develop methods for monitoring demyelination and axonal loss, as these processes probably account for much of the permanent disability in MS. Proton magnetic resonance spectroscopy shows potential for detecting both processes, 43 while diffusion imaging 44 and T2-magnetisation decay curves 45 may provide an indication of axonal loss.

We thank Dr P Sakaris for statistical advice, and Dr P Rudge and Professor W J McDonald for their helpful comments.