

I. Leng Tan
Ronald A. van Schijndel
Franz Fazekas
Massimo Filippi
Peter Freitag
David H. Miller
Tarek A. Yousry
Petra J. W. Pouwels
Herman J. Adèr
Frederik Barkhof

Image registration and subtraction to detect active T₂ lesions in MS: an interobserver study

Received: 27 June 2001
Received in revised form: 9 November 2001
Accepted: 14 November 2001

I. L. Tan · R. A. van Schijndel ·
P. J. W. Pouwels · F. Barkhof
MR Centre for MS research
Vrije Universiteit Medical Centre
P. O. Box 7057
1007 MB Amsterdam, The Netherlands

F. Fazekas
Department of Neurology
Karl-Franzens University
Neurologische Universitätsklinik
Department of Neurology
Auenbruggerplatz 22
8036 Graz, Austria

M. Filippi
Neuroimaging Research Unit
Department of Neuroscience
Scientific Institute Ospedale San Raffaele
Via Olgettina 60
20132 Milan, Italy

P. Freitag
Abt. Neuroradiologie
Kantonsspital/Universitäts Poliklinik
4031 Basel, Switzerland

D. H. Miller
Institute of Neurology
National Hospital for Neurology
and Neurosurgery Queen Square
London WC1N 3BG, UK

T. A. Yousry
University of Munich
Neurol. Klinikum Groshadern
Dept. of Neuroradiology
Marchioninstr. 15
81377 Munich, Germany

H. J. Adèr
Department of Clinical Epidemiology
and Biostatistics
Vrije Universiteit Medical Centre
Van der Boechorststraat 7
1081 BT Amsterdam, The Netherlands

F. Barkhof (✉)
Department of Radiology
Vrije Universiteit Medical Centre
P. O. Box 7057
1007 MB Amsterdam, The Netherlands
Tel.: +31-20/4440365
Fax: +31-20/4440397
E-Mail: f.barkhof@vumc.nl

■ **Abstract** Serial MRI studies are used to analyse change in multiple sclerosis (MS) lesion volume in clinical trials. As such an evaluation is very time consuming and subject to quantification errors, one might assess only the change in number or size of lesions using subtracted images. The advantage of subtracted images is that both new and/or enlarging and resolving and/or shrinking lesions can be evaluated, resulting in a more precise volume change than a net volume change. We studied the interobserver agreement in the detection of active MS lesions using paired dual-echo T₂-weighted spin-echo studies (3-mm slices) of 30 MS patients with a range of MS disease activity on MRI from treat-

ment trials. Using an automatic matching algorithm based on mutual information, the follow-up scan was registered to baseline, after which subtracted images were obtained. After a training session with formulation of guidelines, six observers identified new, enlarging, resolving and shrinking lesions on subtracted images. Weighted kappa (κ) values were calculated to assess interobserver agreement. Good agreement was found for new lesions (κ 0.69 ± 0.08), while moderate agreement was found for enlarging lesions (κ 0.52 ± 0.06). When new and enlarging lesions were combined, good agreement was found for “positive” activity (κ 0.71 ± 0.06). The interobserver agreement was poor for resolving lesions (κ 0.31 ± 0.07), and moderate for shrinking lesions (κ 0.53 ± 0.08). In conclusion, the use of subtracted images in the visual detection of new T₂ lesions resulted in a good level of interobserver agreement for “positive” disease activity. Subtraction of registered images is a reliable, time efficient method to assess disease progression in MS.

■ **Key words** MR imaging · brain · multiple sclerosis · registration · subtraction

Introduction

In phase III clinical multiple sclerosis (MS) trials, lesion volume as assessed by serial dual-echo T₂-weighted spin-echo (T₂SE) magnetic resonance imaging (MRI) series is used as a secondary outcome measure [2, 5, 15]. The evaluation of serial MRI data is quite time consuming, involving identification and quantification of lesions on at least two scans to detect changes in the order of 5–10% only. Such a small change in lesion volume can easily be biased by imperfect repositioning of scans, hampering the comparisons of images and/or introducing quantification error [4, 6, 8, 9, 20]. As the aim of the analysis is to detect change in MS disease burden, one might try to assess only change in the number or size of lesions directly from films. Visual detection of such lesions is more cost efficient, and allows the detection of unambiguous treatment effects [7, 16, 18]. Contrary to the detection of gadolinium enhancing lesions [3], the interobserver reliability of counting active T₂ lesions (i. e., new or enlarging) has proven to be low, even when incorporating consensus rules and training [17].

Since suboptimal scan repositioning has been thought to be a major source of the variability, the effect of image registration on interobserver agreement was assessed [21], by applying an automatic voxel based registration algorithm based on mutual information [10, 13, 14]. This matching criterion is subvoxel accurate, highly robust to noise and image inhomogeneity, and provides a completely automatic registration of pairs of images. Compared with the use of original images, the use of registered images improved interobserver agreement for new lesions significantly, but failed to do so for enlarging lesions [21].

Registered images can be used for subtraction purposes, thus further facilitating the analysis of changes between follow-up studies. In the analysis of MRI-monitored treatment trials of MS, this image assessment has not yet been used. To assess the value of subtracted images in the detection of changes in MS lesions from serial scans, an interobserver study preceded by guidelines formulation was set up. The subtracted images were treated as the primary carrier of information, using the registered images for reference only. The main focus was on “positive” change (i. e., new or enlarging lesions) since this is used as an outcome parameter in the analysis of MS treatment trials. In addition, we explored the inter-observer variation in detecting “negative” change (i. e., shrinking or resolving lesions) using this approach.

Subjects and methods

MR imaging

The study was performed using 30 MRI examinations from 26 patients (16 women and 10 men; mean age 38.6 years with range 21–49)

with clinically definite MS [19] of the relapsing-remitting (RR) or secondary progressive type (with an Expanded Disability Status Scale score [11] between 0.0 and 5.5), participating in clinical trials. MRI scans were obtained at 1.0 or 1.5 Tesla with particular care to achieve precise repositioning. Imaging started with a coronal scout image on which the falx cerebri could be well recognised. On this image, the midsagittal image was planned, revealing the midline structures. The line connecting the inferior base of the hypophysis and the fastigium of the fourth ventricle was used as angulation for the axial series. The Z-centre was placed on the inferior borders of the splenium of the corpus callosum. Axial imaging consisted of proton density and heavily T₂SE imaging (2000–3000/30–50 and 60–100/1 [TR/TE/excitations]) with a 1-mm in-plane resolution. To obtain contiguous 3-mm thick slices, 2 interleaved sets of 23 images with a 3-mm gap were combined.

A pair of T₂SE studies was selected per patient, according to good repositioning (slice difference of ≤1) without motion artefacts. The interval between the paired scans was on average 6 months (range 2–18). According to additional scans available (but not used further in this study), patients were selected in five strata according to the increase in number of enhancing lesions between the two scans; category I (n=4): no enhancing lesion, category II (n=6): 1–2 enhancing lesions, category III (n=10): 3–10 enhancing lesions, category IV (n=5): 11–20 enhancing lesions, category V (n=5): > 20 enhancing lesions. The number of active T₂-lesions in this period had not been assessed but was expected to follow that of the enhancing lesions, ensuring enough variability for a reliable statistical analysis.

For registration, an automatic voxel based registration algorithm was applied, using mutual information [10, 13] as matching criterion. Trilinear interpolation was used for both image interpolation and reslicing of data. The images of the follow-up scan were registered to baseline and subtracted images were obtained. Discarding the most upper and lower images of each series (which could be affected by the reslicing procedure), the central 40 images were printed on two 35 x 43 cm films using standard window/level setting.

Image analysis

Sets of films containing subtracted and registered images of each patient were presented in random order to six observers, all of whom had extensive experience in the assessment of serial scans in the context of MS clinical trials. First, a training session was held using eight additional sets of scans with a range of MRI activity, analysed in consensus by pairs of observers. The results were extensively discussed, followed by formulation of consensus rules. For the final readings, each observer was asked to identify new, enlarging, resolving and shrinking T₂ lesions directly on the subtracted images, using the registered images as a reference to check whether a lesion change suggested by the subtracted images was actually a genuine change. Lesions were marked on the most cranial image on which they first appeared. After completion of the image analysis, the marked lesions were classified according to their location as: (a) periventricular and (b) non-periventricular in the frontal/parietal/occipital regions, as (c) temporal or (d) infratentorial. To assess interobserver agreement, weighted kappa (κ) values and their standard deviations were calculated. To compare between κ values, a z-statistic was used. Kappa values below 0.4 indicate poor agreement, between 0.41 and 0.6 indicate moderate agreement, between 0.61 and 0.8 indicate good agreement, and more than 0.8 good agreement [1].

Since treatment trials will frequently use a combination of new and enlarging lesions, we also examined the level of agreement for “positive” (i. e., new and enlarging lesions) and for “negative” disease activity (i. e., resolving and shrinking lesions).

Results

Areas on original images that do not change in signal intensity on follow up, appear grey on subtracted images, while “positive” changes such as new or enlarging lesions will appear as bright areas against a grey background, while “negative” changes such as resolving or shrinking lesions appear as dark areas against a grey background. Fig. 1 shows the original baseline, registered follow-up and subtracted images of a patient of category V (i. e., > 20 enhancing lesions) at the level of the lateral ventricles. Compared to baseline, changes on the follow-up scan can be seen as both “positive” and “negative” changes in the right frontal area on the subtracted image (Fig. 1C). Slight reslicing artefacts may be seen, especially at the brain surface and grey/white matter borders. Inconstant pulsation artefacts and flow in vessels may also lead to subtraction errors.

The results of the training session were extensively discussed and consensus rules formulated. The following 4 types of lesions were defined; I: a new lesion was defined as a (non-artefactual) bright area visible against the grey background, with mandatory confirmation of its presence on the registered scan (compared to baseline). In case of suboptimal registration, e. g. due to motion of the patient between the interleaved series, any new lesion had to appear on two or more consecutive slices. II: The definition of an enlarging lesion was formulated according to its size. For lesions with a diameter larger than 5 mm on the original images, an overall increase by at least 100 % or an increase in size on at least

two consecutive slices on the registered images was required. For lesions with a diameter ≤ 5 mm both the above conditions had to occur. Change in signal intensity alone was not considered enough to define a lesion as enlarged; change in size was mandatory. III: A resolving lesion was a (non-artefactual) dark area visible against the grey background, with mandatory confirmation of its presence on the baseline scan (compared with the registered scan). IV: The definition of a shrinking lesion was an overall decrease by at least 50 % or a decrease in size on at least two consecutive slices on the registered scan for lesions with a diameter larger than 5 mm on the original images. For lesions with a diameter ≤ 5 mm, it was decided that shrinking can not be assessed reliably. Due to disturbing flow artefacts in the posterior fossa it was decided not to score enlarging or shrinking lesions in this region.

The six observers marked a total of 2191 lesions for 30 patients: 1605 new, 257 enlarging, 122 resolving and 207 shrinking lesions. An overview of the mean number of lesions per observer is shown in Fig. 2. Table 1 shows the mean number of lesions (standard deviation) together with the median (range) number of lesions. The κ values are shown for each type of lesion, and for both “positive” and “negative” activity. The κ value indicated good agreement for new lesions with 0.69 (standard deviation 0.08) and moderate agreement for enlarging lesions with a κ value of 0.52 (0.06). For “positive” activity (i. e. new or enlarging lesions), the κ value was 0.71 (0.06).

New lesions with only a moderately altered signal in-

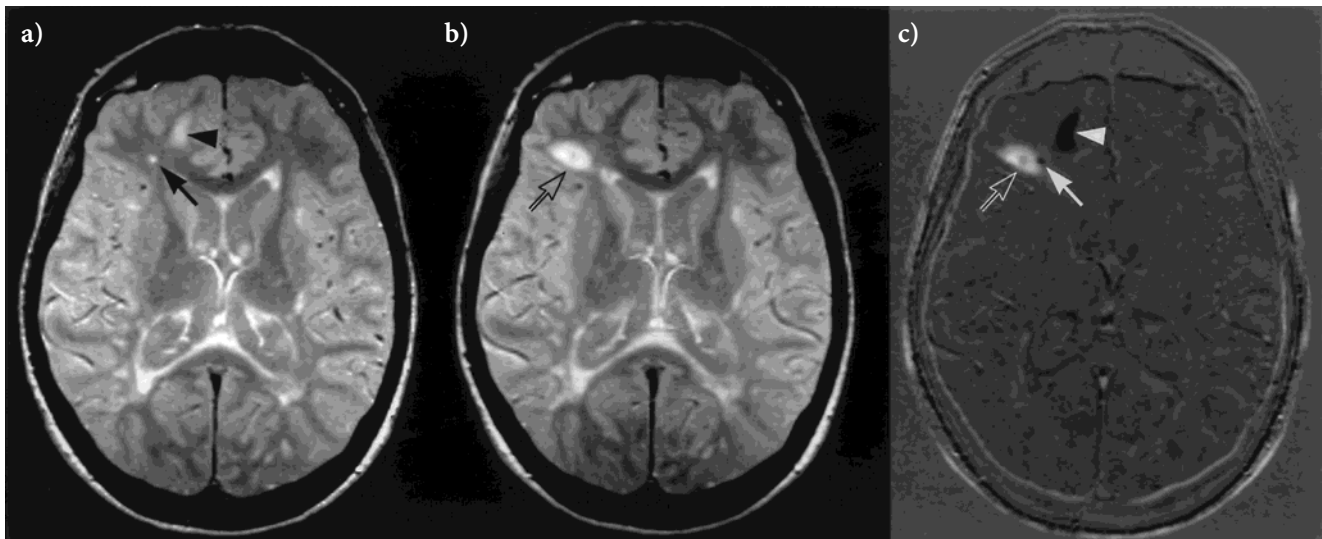


Fig. 1 Axial original baseline and registered follow-up proton density weighted MR images with subtracted image of the brain of an MS patient of category V at the level of the lateral ventricles. Compared to baseline (A), the small lesion in the right frontal region (arrow in A) seems to have enlarged into a large ovoid lesion on the registered image (open arrow in B). However, the subtracted image (C) demonstrates that the small lesion has clearly disappeared (“negative” change), while the ovoid lesion is new as it is shown as a bright area (“positive” change). The large disappearing subcortical lesion (arrowhead in A and C) is represented as a dark area due to “negative” change. Minor subtraction artefacts can be seen at the subcortical regions and near the brain surface.

tensity can be easily missed on subtracted images owing to their low contrast against the grey background as demonstrated in Fig. 3. This figure shows an original

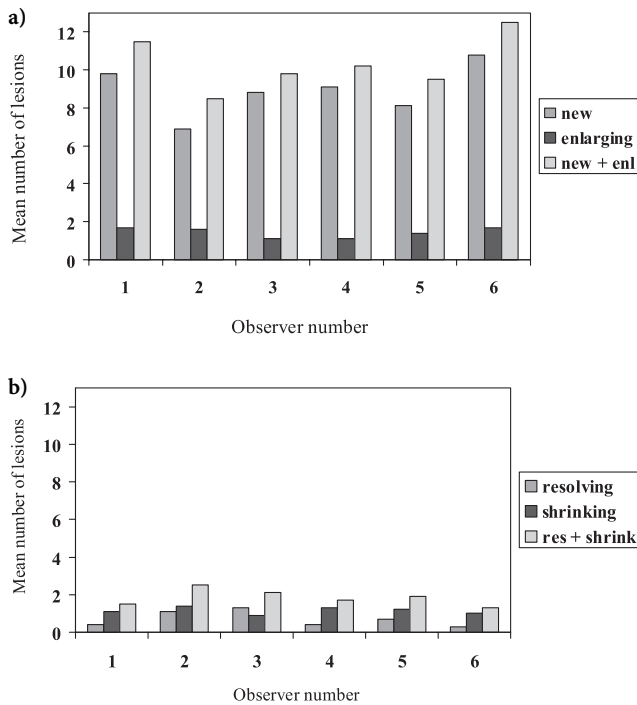


Fig. 2 (A) Mean number of new and enlarging lesions, and positive activity (new and enlarging lesions). (B) Mean number of resolving and shrinking lesions, and negative activity (of resolving and shrinking lesions).

baseline and registered follow-up image with the subtracted image at the level of the lateral ventricles. At least 10 lesions can be seen on the registered image. Compared with baseline, the bright new lesions in the right capsula interna and in the right occipital lobe were both classified as new lesions by all six observers. However, the new lesion in the splenium of the corpus callosum is less bright and, therefore, less conspicuous on both the follow-up and subtracted image. It was considered new by four observers but ignored by two.

The observers commented that dark areas were more difficult to interpret than changes resulting in bright areas. The interobserver agreement for resolving lesions was poor with a κ value of 0.31 (0.07), while for shrinking lesions it was moderate with κ 0.53 (0.08). For the combined measure of “negative” activity, the κ value was 0.50 (0.08).

When lesions were classified according to their location, the highest κ value was found in the non-periventricular region with a κ value of 0.73 (0.04) for “positive”

Table 1 Descriptive data and kappa values for visual detection of active T₂ lesions.

Type of lesion	Median (range)	Mean (sd)	κ (sd)
New	4.0 (0–36)	8.9 (9.6)	0.69 (0.08)
Enlarging	1.0 (0–11)	1.4 (2.0)	0.52 (0.06)
New and Enlarging	4.0 (0–37)	9.8 (10.4)	0.71 (0.06)
Resolving	0.0 (0–9)	0.7 (1.4)	0.31 (0.07)
Shrinking	0.0 (0–8)	1.1 (2.0)	0.53 (0.08)
Resolving and Shrinking	0.0 (0–17)	1.8 (2.8)	0.50 (0.08)

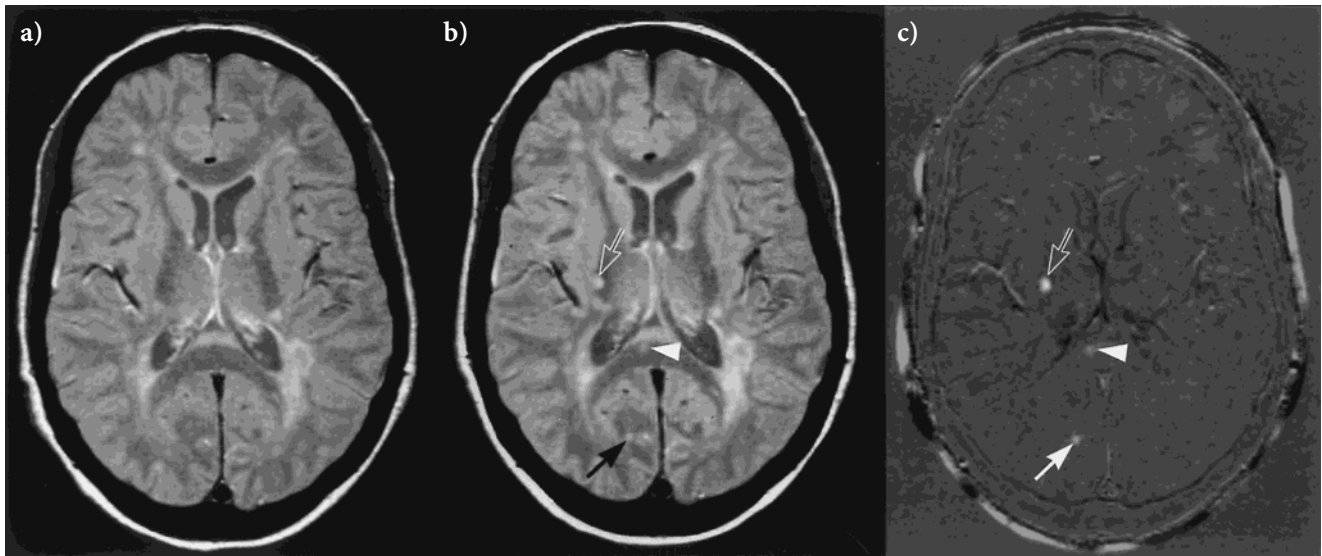


Fig. 3 Axial original baseline and registered follow-up proton density weighted MR images with subtracted image of the brain of another MS patient from category V at the level of the lateral ventricles. At least 10 lesions can be seen on the registered image (B). Compared to baseline (A), the bright new lesion in the right capsula interna (open arrow in B and C) and in the right occipital lobe (arrow in B and C) can be easily appreciated on the subtracted image and were both classified as new lesions by all six observers. In C, the new lesion in the splenium of the corpus callosum (arrowhead in B and C) is less bright and therefore less conspicuous on both the follow-up and subtracted image. It was considered new by four observers but ignored by two.

activity and 0.55 (0.08) for “negative” activity. The mean time for analysis of the scans was 8.4 minutes (standard deviation 2.5) and differed according to category of activity. The mean time was 2.5 minutes (0.3) for category I while the mean time was 15.3 minutes (11.1) for category V.

Discussion

The use of subtracted images for the visual detection of new T_2 lesions resulted in a good level of interobserver agreement with a κ value of 0.69 (0.08). The κ value for enlarging lesions was moderate, but when combined, the agreement for “positive” activity was again good. Since the separation of new and enlarging lesions may sometimes be ambiguous, a combined measure of “positive” changes seems to be a better indicator of disease activity than the analysis of new and enlarging lesions separately. Although difficult to compare, scan-rescan studies indicate that the difference in lesion volume between two scans can be as large as the expected average increase on annual scanning of 5–10% [6, 8, 20]. Direct identification using subtraction may not only prove to be more reliable; it is also more attractive, since only activity (i. e., change from baseline) is assessed. Finally, it may be more time efficient (although partly offset by the time needed for the registration procedure). Molyneux et al. [17] have reported the interobserver agreement for active lesions identification based on the visual analysis of non-registered T_2 SE images. They described the analysis of a data set of 16 RR MS patients by five observers, using baseline and month 9 studies. The images were analysed twice without and once with the use of consensus guidelines. Using guidelines, a mean number of new and enlarging lesions of about 6 and 1 was observed respectively, with κ values of 0.46 and 0.21. The κ values (before as well as after the use of consensus guidelines) were lower than those found in the present study (i. e., 0.69 and 0.52), in which a mean number of 8.9 new and 1.4 enlarging lesions was observed. Perfect agreement (i. e., a κ value of 1) is unlikely to be achieved even using subtraction when new lesions can have only a moderately altered signal intensity as demonstrated in Fig. 3. In an earlier interobserver study using registered images [21], the level of interobserver agreement remained poor for enlarging lesions. The use of subtraction resulted in moderate interobserver agreement; however, for reasons alluded to earlier, it is likely that the actual level of agreement was underestimated by the fact that enlarging and new lesions were misclassified rather than being undetected. In the analysis of T_2 lesion volumes in MS trials one looks at the net change in serial lesion volumes measurements. However, net changes reflect the volume of new or enlarging plus resolving or shrinking lesions. This inevitably leads to an underesti-

mation of the new lesion volume change. The assessment of resolving and shrinking lesions may be important in the follow up of MS disease activity. Using image registration in a data set of 19 RR MS patients, Lee et al. [12] demonstrated that over one year the volume of new lesions can be three times greater than the net lesion volume change and correlated better with the T_1 -weighted enhancing lesion volumes. Further, a trend was seen for patients with clinical progression to have a greater new T_2 lesion volume than those who did not. As shrinking and resolving lesions can be seen as dark areas on subtracted images, these lesions were also evaluated in the present study. These types of lesions had not been analysed before by the observers so new guidelines had to be formulated. The numbers of these lesions were quite low compared with those of lesions with “positive” activity. The κ values were poor to moderate, indicating that the guidelines for these new types of lesions have to be adjusted, more training, or other display methods (e. g. inverted grey scale, since the eye is more sensitive to bright versus grey than to dark versus grey) may be needed.

In conclusion, the use of subtracted images for the visual detection of new T_2 lesions resulted in a good level of interobserver agreement, with moderate agreement for enlarging lesions. As the level of interobserver agreement of new and enlarging lesions combined is good, the analysis of “positive” activity from subtraction images seems to be a reliable, time and cost-efficient method to assess disease progression in MS.

■ **Acknowledgements** This study was financially supported by MAGNIMS, a EC funded program (ERBCHRXCT940684) entitled “Development of optimal magnetic resonance techniques to monitor treatments for preventing disability in multiple sclerosis”.

Appendix

■ Consensus rules

Marking of new, enlarging, resolving and shrinking T_2 lesions on subtracted images.

■ In general

Conservative approach. Ignore both small hyper- and hypo-intense dots. The signal intensity of a lesion on the subtracted images should be clearly higher or lower than the surrounding grey background. Every lesion should be checked on the original and registered proton density scans to ensure that a lesion is a genuine lesion. For (sub)cortical and infratentorial lesions, the heavily T_2 -weighted images should be checked.

As lesions can change one slice position due to the registration procedure, adjacent slices on the registered

images should be checked to make sure that an apparent lesion is truly a lesion and not a “false positive” lesion.

In areas of artefacts (e. g., near the mastoid regions, the circle of Willis, periventricular and temporal regions, posterior fossa) lesions should be preferably seen on at least two consecutive slices. As the subtracted heavily T_2 -weighted images show more artefacts than the subtracted proton density images, care should be taken when using them in the analysis.

■ New lesions

Definition: a (non-artefactual) bright area that is visible against the grey background. Confirmation on the registered scan (compared with baseline) is mandatory. The form of the new lesion should match that of the same lesion on the registered image. In case of poor repositioning/registration: any new lesion must appear on two or more consecutive slices. When there is good repositioning/registration: a lesion can be visible on just a single slice. When there is a “bridge” connecting an existing lesion with an apparent new area with high signal intensity, consider the last as a new lesion. If not, consider the existing lesion enlarged (see below).

■ Enlarging lesions

Definition: a (non-artefactual) area of hyper-intensity that is visible against the grey background, and whose

form is (partly) congruent with the same lesion on the registered image. Change in signal intensity alone is not enough; change in size is mandatory.

- 1 A. for lesions with a diameter of $> 5\text{mm}$ \Rightarrow
 - a) overall increase by at least 100 % OR
 - b) increase in size on at least two consecutive slices.
- 1 B. for lesions with diameter $\leq 5\text{mm}$ \Rightarrow a) and b)
2. posterior fossa: no scoring of enlarging lesions.

■ Resolving lesions

Definition: a (non artifactual) dark area that is visible against the gray background. The form of the resolving lesion should match that of the original lesion on the baseline image.

■ Shrinking lesions

Definition: a (non-artefactual) dark area that is visible against the grey background and which form is (partly) congruent with the original lesion on the baseline image.

- 1 A. for lesions with a diameter of $> 5\text{mm}$ \Rightarrow
 - a) overall decrease by at least 50 % OR
 - b) decrease in size on at least two consecutive slices.
- 1 B. for lesions with diameter $\leq 5\text{mm}$ there is no marking of shrinking lesions. These lesions can only disappear.
2. posterior fossa: no scoring of shrinking lesions.

References

1. Altman DG (1991) Practical statistics for medical research. Chapman and Hall, London, pp 403–409
2. Barkhof F, Filippi M, Miller DH, Tofts P, Kappos L, Thompson AJ (1997) Strategies for optimizing MRI techniques aimed at monitoring disease activity in multiple sclerosis treatment trials. *J Neurol* 244:76–84
3. Barkhof F, Filippi M, van Waesberghe JH, et al. (1997) Improving interobserver variation in reporting gadolinium-enhanced MRI lesions in multiple sclerosis. *Neurology* 49:1682–1688
4. Capra R, Marciano N, Gasparotti R (1993) The effect of repositioning error on serial magnetic resonance imaging scans. *Arch Neurol* 50:570–571
5. Filippi M, Horsfield MA, Adèr HJ, et al. (1998) Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis. *Ann Neurol* 43:499–506
6. Filippi M, Marciano N, Capra R, et al. (1997) The effect of imprecise repositioning on lesion volume measurements in patients with multiple sclerosis. *Neurology* 49:274–276
7. Filippi M, Rovaris M, Sormani MP, et al. (1998) Intraobserver and interobserver variability in measuring changes in lesion volume on serial brain MR images in multiple sclerosis. *AJNR Am J Neuroradiol* 19:685–687
8. Gawne-Cain ML, Webb S, Tofts P, Miller DH (1996) Lesion volume measurement in multiple sclerosis: how important is accurate repositioning? *J Magn Reson Imaging* 6: 705–713
9. Goodkin DE, Vanderburg-Medendorp S, Ross J (1993) The effect of repositioning error on serial magnetic resonance imaging scans. *Arch Neurol* 50:569–570
10. Holden M, Hill DLG, Denton ERE, Jarosz JM, Cox TCS, Rohlfing T, Goodey J, Hawkes DJ (2000) Voxel similarity measures for 3D serial MR brain image registration. *IEEE Trans Med Imaging* 19:94–102
11. Kurtzke JF (1983) Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33:1444–1452
12. Lee MA, Smith S, Palace J, Matthews PM (1998) Defining multiple sclerosis disease activity using MRI T_2 -weighted difference imaging. *Brain* 121:2095–2102
13. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16:187–198
14. Maintz JBA, Viergever MA (1998) A survey of medical image registration. *Medical Image Anal.* 2:1–36
15. Miller DH, Albert PS, Barkhof F, et al. (1996) Guidelines for the use of magnetic resonance techniques in monitoring the treatment of multiple sclerosis. *Ann Neurol* 39:6–16

-
16. Miller DH, Molyneux PD, Barker GJ, et al. (1999) Effect of interferon- β 1b on magnetic resonance imaging outcomes in secondary progressive multiple sclerosis: results of a European multicenter, randomized, double-blind, placebo-controlled trial. *Ann Neurol* 46:850–859
 17. Molyneux PD, Miller DH, Filippi M, et al. (1999) Visual analysis of serial T₂-weighted MRI in multiple sclerosis: intra- and interobserver reproducibility. *Neuroradiology* 41:882–888
 18. Paty DW, Li DKB, the UBC MS/MRI Study Group and the IFNB Multiple Sclerosis Study Group (1993) Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. *Neurology* 43:662–667
 19. Poser CM, Paty DW, Scheinberg L, et al. (1983) New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 13:227–231
 20. Rovaris M, Gawne-Cain ML, Sormani MP, Miller DH, Filippi M (1998) The effect of repositioning on brain MRI lesions load in assessment in multiple sclerosis: reliability of subjective quality criteria. *J Neurol* 245:273–275
 21. Tan IL, van Schijndel RA, Fazekas F, et al. (2001) Improved interobserver agreement for visual detection of active T₂-lesions on serial MR scans in multiple sclerosis using image registration. *J Neurol* 248:789–794