Correlating MRI and clinical disease activity in multiple sclerosis:
Relevance of hypointense lesions on short-TR/short-TE (T1-weighted) spin-echo images

M.A.A. van Walderveen, MD; F. Barkhof, MD; O.R. Hommes, MD; C.H. Polman, MD; H. Tobi, MSc; S.T.F.M. Frequin, MD; and J. Valk, MD

Article abstract—Magnetic resonance imaging (MRI) is being used as an outcome criterion in therapeutic trials in multiple sclerosis (MS) on the assumption that it, as a sensitive marker of biologic disease activity, could serve as a surrogate marker of disability. We evaluated the relation between MRI findings and disability in a quantitative follow-up study of 48 MS patients. Median duration of follow-up was 24 months (range, 10 to 42 months). Computer-assisted volume measurements employing a seed-growing technique yielded a standard error of measurement of 0.275 cm². The median total area of the hyperintense lesions on the initial T2-weighted images was 8.4 cm². The median increase was 0.76 cm²/yr (9%). In a subgroup (n = 19) with short-TR/short-TE spin-echo (SE) images, we measured the hypointense lesion load. The median total area of the lesions at entry was 0.70 cm², with a median increase of 0.28 cm²/yr (40%). The total area of the hyperintense lesions on the initial T2-weighted images showed a weak correlation with the Expanded Disability Status Scale score at entry (Spearman rank correlation coefficient [SRCC] = 0.30; 0.02 < p < 0.05). The increase in disability showed a positive correlation (SRCC = 0.19) with the increase in hypointense lesion load on the T2-weighted SE images, but this correlation did not reach statistical significance (p = 0.09), probably because of lack of clinical progression. The number of active lesions detected by visual analysis of the T2-weighted SE images correlated significantly (SRCC = 0.40; 0.001 < p < 0.01) with the number of relapses during the interval between the initial and follow-up imaging examinations. The subgroup with short-TR/short-TE SE images (whose disease was clinically more active than the group as a whole) showed a significant correlation of increased disability with increase in hypointense lesion load (SRCC = 0.74; p < 0.002). Our results, and those from previous follow-up studies, suggest a positive correlation between MRI and clinical activity, but extended follow-up studies are needed to confirm the appropriateness of quantitative MRI as a surrogate marker of disability in MS. Short-TR/short-TE SE images seem to be more relevant than T2-weighted SE images in identifying the lesions that cause disability.

Magnetic resonance imaging (MRI) is sensitive in detecting brain abnormalities in multiple sclerosis (MS).1 T2-weighted spin-echo (SE) sequences are useful in displaying the overall lesion load, whereas gadolinium-enhanced short-TR/short-TE SE images identify the currently active lesions in MS.2 Previous studies have shown that many MS lesions are clinically silent. MRI-evident disease activity (new or enlarging lesions on the T2-weighted SE images and enhancing lesions on the short-TR/short-TE SE images) occurs more frequently than is detectable on clinical grounds.3-6 Histopathologic studies show that abnormalities seen with MRI represent MS lesions.7 The most sensitive biologic disease measure at present, MRI is increasingly used as an outcome measure in clinical trials to test the efficacy of new drugs,8 as was recently illustrated in the interferon beta-1b trial.9,10 Although the majority of new brain MRI lesions are clinically silent, the rationale for using MRI as an outcome criterion is that the more MRI lesions develop, the larger the increase in disability will be over time. However, the correlation between MRI lesions and clinical disability is not well understood. In follow-up studies, the number, size, and site of lesions have been used to seek correlation with the clinical status of the patients. Since computer-assisted techniques have become available, a more reliable and reproducible analysis of MRI lesion load is now possible. Minimal changes of individual lesions are hard to detect visually, whereas computer-assisted quantitation provides a more precise calculation of the size of lesions. One cross-
 sectional study showed a significant, although weak, correlation between the computed lesion volume and the Expanded Disability Status Scale (EDSS) score. A more recent quantitative follow-up study showed a significant correlation between MRI lesion load and EDSS score. In that study, a Spearman rank correlation coefficient of 0.62 between changes in lesion load on T2-weighted SE images and EDSS score was reported. Although the EDSS score is an ordinal one with several limitations, the follow-up time in the study of Filippi et al was long enough (5 years) to have allowed identification of a stronger relationship between the clinical and the MRI variables if such a relationship had existed.

One possible explanation for the lack of a closer relation between MRI activity and disability is that not all abnormalities seen on T2-weighted SE images are clinically significant. Inflammation, edema, gliosis, demyelination, and axonal loss are all represented as areas of high intensity on T2-weighted SE images. Short-TR/short-TE SE images sometimes show areas of hypointensity in a part of the lesions seen on T2-weighted SE images. These areas of hypointensity may represent the chronic MS lesion in which astrocyte growth and scarring, and therefore permanent neurologic damage, have occurred; they might be of greater clinical significance than the total lesion load seen on T2-weighted SE images.

We report a clinical and MRI follow-up study of 48 MS patients. The purpose of this study is to determine whether an increase in quantified lesion load correlates with an increase in disability, and whether such a correlation is stronger with lesions observed on short-TR/short-TE SE images. By addressing these issues, one can evaluate the value of MRI as a surrogate marker of disability.

Methods. A cohort of 48 patients (12 men, 36 women) was followed clinically and with MRI. The patients were classified at point of entry into the study. The cohort consisted of 32 patients with clinically definite MS, 13 with laboratory-supported definite MS, and three with clinically probable MS. The mean age was 33.2 years (range, 21 to 62 years), and the median disease duration was 4.0 years (mean, 5.4 years; range, 0 to 25 years). Thirty-three patients had relapsing-remitting MS and 15 had chronic progressive MS (one primary progressive, 14 secondary progressive). The EDSS score was estimated shortly before or after the MRI examination by a neurologist blinded to the MRI analysis. Baseline EDSS score was estimated on the basis of frequent prior EDSS scores for patients who suffered from a relapse at the time of imaging. (A relapse was defined as any clinical deterioration lasting more than 24 hours that was not related to a concurrent disease.) The number of relapses between the first and the second MRI examinations was documented. The median duration of follow-up was 24 months (range, 10 to 42 months). Patients were treated only with short courses of intravenous methylprednisolone when indicated; no other immunosuppressive therapy was given.

Imaging was performed on a 0.6-T machine (Technicare, Solon, OH) with a standard headcoil. The same imaging protocol was used for the first and the second MRI examinations. Identical slice repositioning was achieved by means of unenhanced images in two or three consecutive planes, thereby correcting for differences in the patient’s position according to internal landmarks (angulation along the line connecting the caudal border of the pituitary gland and the fastigium of the fourth ventricle; the Z center was aligned with the caudal border of the splenium of the corpus callosum). From the logical mid sagittal image, double oblique two-dimensional axial series were planned. T2-weighted SE images were obtained with 2,755/60, 120/2 (TR/TE/excitations). Nineteen slices with an in-plane resolution of 1.0 × 1.3 mm and a slice thickness of 5 mm were obtained with a gap of 1.25 mm, resulting in a Z range of 11.75 cm. Nineteen patients received contrast-enhanced short-TR/short-TE SE images on both occasions. Gadopentetate dimeglumine (Magnevist, Schering AG, Berlin) was administered intravenously as a bolus in a dose of 0.2 mmol/kg. Short-TR/short-TE SE images (two series of nine slices) were obtained with 450/28/4, starting 5 to 10 minutes after gadolinium injection.

The images were analyzed by observers blinded to the clinical data, using visual inspection and computer-assisted techniques; for both techniques, the intra-observer variability (precision) was assessed in five patients. As the follow-up images were analyzed shortly after one another, the images used for the intra-observer variability were also analyzed in a short time interval. On visual inspection, the number of new or enlarging lesions on the follow-up (T2-weighted) image (compared with the initial image) was assessed as a measure of MRI activity by two experienced readers in conference. The short-TR/short-TE SE images were not analyzed visually. Computer-assisted volume measurements were performed by a single observer on Sparc-2 and -10 workstations (Sun, Palo Alto, CA) using home-developed software. On the long-TR (T2-weighted) SE images, the surfaces of the hyperintense lesions were measured in comparison with the surrounding white matter (figure 1A). On the short-TR/short-TE SE images, obtained after administration of gadopentetate dimeglumine, we measured the surfaces of the hypointense lesions (figure 1B) and the surfaces of lesions that showed enhancement. A hypointense lesion was defined as any region visible on the short-TR/short-TE SE images corresponding to a region of high signal intensity on the T1-weighted images, with low signal intensity relative to the surrounding white matter. We mainly used a seed-growing method for calculating the surfaces of the lesions: after a seed point is positioned in a part of the lesion, the boundaries of this region are grown toward interactively set threshold intensities (figure 1, C and D). Once the thresholds have been established for the first lesion, these can usually be applied to the other lesions, thus speeding up the routine. If no satisfactory definition of subsequent lesions is obtained, the threshold levels can be adapted for each lesion. When this technique failed (eg, due to low contrast between the lesion and the surrounding tissue), a manual tracing technique was used (infrequently). Once the boundaries of the lesions had been established, the surface of the lesions was calculated per slice (figure 1, E and F). A distinction was made between supratentorial and infratentorial (cerebellum, mesencephalon, medulla oblongata, andpons) lesions. By multiplying the surfaces by the interlice distance (0.626 cm), it is possible to obtain an estimate of the actual volumes of the lesions.
The comparison results, with a significance level of 0.05 to correct for multiple comparisons, were calculated for the correlation coefficient (Spearman Rank correlation coefficient). The Spearman Rank correlation coefficient mainly distributed, medians instead of means are used for statistical analysis that was performed using the statistical package RMDP. As most of the data were not normally distributed, medians instead of means are used for the statistical analysis. The Spearman Rank correlation coefficient.
Table. Intraobserver variation for computer-assisted measurement (brain area in cm²) and for visual analysis (number of new or enlarging lesions on the T₂-weighted images)

<table>
<thead>
<tr>
<th></th>
<th>Pt 1</th>
<th>Pt 2</th>
<th>Pt 3</th>
<th>Pt 4</th>
<th>Pt 5</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T₂-weighted images</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation 1</td>
<td>9.43</td>
<td>6.26</td>
<td>31.44</td>
<td>45.69</td>
<td>30.34</td>
<td>24.63</td>
</tr>
<tr>
<td>Observation 2</td>
<td>9.00</td>
<td>5.88</td>
<td>30.63</td>
<td>44.36</td>
<td>29.79</td>
<td>23.93</td>
</tr>
<tr>
<td>Difference</td>
<td>0.43</td>
<td>0.38</td>
<td>0.81</td>
<td>1.33</td>
<td>0.65</td>
<td>0.70</td>
</tr>
<tr>
<td>Percentage difference</td>
<td>4.5</td>
<td>6.1</td>
<td>2.6</td>
<td>2.9</td>
<td>1.8</td>
<td>3.58*</td>
</tr>
<tr>
<td><strong>Short-TR/short-TE SE images</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation 1</td>
<td>0.68</td>
<td>0.70</td>
<td>10.37</td>
<td>6.88</td>
<td>17.67</td>
<td>7.26</td>
</tr>
<tr>
<td>Observation 2</td>
<td>0.61</td>
<td>0.72</td>
<td>9.90</td>
<td>6.31</td>
<td>17.32</td>
<td>6.97</td>
</tr>
<tr>
<td>Difference</td>
<td>0.07</td>
<td>0.02</td>
<td>0.47</td>
<td>0.57</td>
<td>0.55</td>
<td>0.29</td>
</tr>
<tr>
<td>Percentage difference</td>
<td>10.3</td>
<td>2.9</td>
<td>4.5</td>
<td>8.3</td>
<td>2.0</td>
<td>5.6e</td>
</tr>
<tr>
<td><strong>Visual analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observation 1</td>
<td>5</td>
<td>2</td>
<td>21</td>
<td>8</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Observation 2</td>
<td>5</td>
<td>2</td>
<td>36</td>
<td>14</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Difference</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Percentage difference</td>
<td>0</td>
<td>0</td>
<td>67</td>
<td>75</td>
<td>25</td>
<td>33*</td>
</tr>
</tbody>
</table>

Difference = observation 1 minus observation 2. Absolute values are given. Percentage difference = difference/observation 1 for each of the patients.

* Mean percentage difference = sum of the percentage differences of the 5 patients/number of patients.

Results. The median EDSS score was 2.0 (mean, 2.2) at the time of the initial scan and 2.0 (mean, 2.5) at the time of the second scan. Forty-one patients showed an increase in EDSS score of less than one full point during the follow-up period; seven patients showed an increase in EDSS score equal to or greater than one full point during follow-up. (Forty-six patients showed an increase in EDSS score of less than one full point a year; two patients showed an increase in EDSS score equal to or greater than one full point a year.) Although the increase in disability was significantly larger than zero (Wilcoxon signed rank test, p = 0.003), the progression in this group was small in clinical terms. The median number of relapses was 2.0 (range, 0 to 6) during the interval between the first and second MRI examinations (mean, 2.1).

It typically took 30 to 45 minutes to analyze a set of T₂-weighted SE images (19 slices per patient), depending on the extent of abnormalities present and the experience of the observer. For the short-TR/short-TE SE images, about 30 minutes are needed, the comparison with the T₂-weighted images being the time-consuming factor. The standard error of measurement was 0.275 cm² for the T₂-weighted images. In general, the intra-observer variation seems dependent on the extent of the abnormalities (table), being smaller for the larger lesion loads present on the T₂-weighted images (3.6%) than for the markedly smaller lesion loads on the short-TR/short-TE SE images (5.6%). By contrast, the intra-observer variation of the visual analysis (table) was more marked in cases with a high number of active lesions.

The median number of new lesions was 4.0 (range, 0 to 52), and the median number of enlarged lesions was 0.0 (range, 0 to 17). A median total of 4.5 active lesions (range, 0 to 69) was seen on the second image. The 48 patients initially showed a median total lesion load of 8.4 cm² (range, 0.6 to 73.4) and a median total lesion volume of 5.3 cm³. On the follow-up image, the median total lesion load had increased to 11.4 cm² (range, 1.5 to 91.6) and the median total lesion volume to 7.1 cm³. The yearly increase in total lesion load on T₂-weighted images ranged from −1.9 to 9.3 cm², with a median of 0.76 cm² (9%), which was highly significant (Wilcoxon signed rank test, p ≤ 0.00001).

We measured hypointense lesions on the short-TR/short-TE SE images in the 19 patients with such images. We calculated an initial median lesion load of 0.70 cm² (range, 0 to 18.2) and a median lesion volume of 0.44 cm³ on the initial image. At follow-up, the median lesion load was 1.3 cm² (range, 0 to 24.8) and the median lesion volume 0.6 cm³. The annual increase in hypointense lesion load ranged from −0.7 to 3.3 cm², with a median of 0.28 cm² (40%), which was statistically significant (Wilcoxon signed rank test, p = 0.01).

EDSS score at the time of the first image correlated weakly with the measured lesion volume on the T₂-weighted SE (SRCC = 0.30; 0.02 < p < 0.05) but not on the short-TE/short-TR images. The disease duration correlated only weakly (SRCC = 0.30; p = 0.04) with the computed initial supratentorial lesion volume on the T₂-weighted SE images. The number of active lesions correlated significantly with the number of relapses (SRCC = 0.40; 0.001 < p < 0.01) but not with the difference in EDSS score over the interval period (SRCC = 0.18; p > 0.1). We then divided the patients into an MRI active group (number of active lesions > 0) and an MRI stable group (number of active lesions = 0), but we did not find a significant difference in increase in disability.
For the T2-weighted SE images, the increase in EDSS score correlated positively, but not significantly, with any measured lesion volume (supratentorially, SRCC = 0.19; infratentorially, SRCC = 0.21). The relation between the increase in EDSS score and the increase in total lesion load on T2-weighted images is shown in figure 2 (SRCC = 0.19; NS). The correlation between T2 lesion load and disability did not differ between patients with a relapsing-remitting disease course and those with a secondary progressive form of MS. The number of relapses during the follow-up period did not correlate significantly with the increase in T2 lesion load (SRCC = 0.22; NS).

For the short-TR/short-TE SE images, the difference in EDSS score between the first and second images showed a strong correlation (SRCC = 0.74; \(p < 0.002\)) with the increase in hypointense lesion load (figure 3). For the subgroup of 19 patients with these images, we checked the correlation coefficient between the increase in EDSS score and the increase in lesion load on the T2-weighted SE images; it was higher (SRCC = 0.34), but not significantly (0.1 < \(p < 0.2\)), than the correlation coefficient for the whole group (SRCC = 0.19), illustrating that the subgroup was biased toward more active disease. The increase in hypointense lesion load did not correlate with the number of relapses in the follow-up period (SRCC = -0.29; NS).

To establish the predictive value of lesion load at entry, we correlated the initial lesion load with changes in lesion load and disability at follow-up. The initial lesion load on T2-weighted images correlated (SRCC = 0.48; \(p < 0.001\)), as expected, with the increase in lesion load (figure 4) and also correlated with the increase in EDSS score (SRCC = 0.38; 0.001 < \(p < 0.01\)). The increase in EDSS score correlated only weakly with the measured hypointense lesion load on the initial short-TR/short-TE SE images (SRCC = 0.42; 0.05 < \(p < 0.1\)); the initial lesion load correlated weakly (SRCC = 0.41; 0.05 < \(p < 0.1\)) with the increase in lesion load of these lesions.

**Discussion.** The increase in MRI lesion load, measured on the T2-weighted images, correlated weakly, but not significantly, with the increase in disability as measured by EDSS score (figure 2). Filippi et al found a correlation coefficient of 0.62 between lesion load on T2-weighted images and EDSS score in their follow-up study in patients with clinically isolated syndromes suggestive of MS,\(^{12}\) which supports the use of MRI as a surrogate marker of disability in treatment trials. That we did not detect a relation is probably related to our follow-up period being relatively short and to the group of MS patients we studied being relatively stable clinically (only seven patients showed an increase in EDSS score equal to or greater than one full point during follow-up). Although one might ex-
pect a correlation with the infratentorial region, the increase in infratentorial lesion load did not correlate with the increase in EDSS score. This could be due to the difficulty in measuring infratentorial lesion volumes: the lesions in this area are very small, there are disturbing flow artifacts from the third and fourth ventricles, and there is relatively poor contrast of these lesions (perhaps due to the cerebellum not being in the center of the coil). Another possible explanation is that the EDSS score mainly represents spinal disease activity, although preliminary results from spinal imaging studies also do not explain changes in disability (Dr. A.J. Thompson, personal communication).

We found that MRI lesion load at entry predicts the increase in disability and the increase in lesion load at follow-up. This is in agreement with the study of Filippi et al., which found that MRI lesion load at presentation in patients with clinically isolated syndromes is a strong predictor of the degree of disability and the increase in lesion load 5 years later. When MRI is used as a selection criterion for treatment trials, patients with a high initial lesion load are more likely to develop progressive disability (figure 4).

Intra-observer variability for quantitation is dependent on the extent of abnormalities present on the images (table), being lower (ie, quantitation is more accurate) in cases of extended abnormalities (in which it is difficult to analyze differences in size visually and in which the intra-observer variability for the visual analysis is high), as in patients with secondary progressive disease. The intra-observer variation for lesion quantitation falls well below the median increase per year. This kind of quantitation seems to be an appropriate and objective method by which to evaluate MRI lesion load and MRI disease activity over a long period (ie, more than 1 year). Visual analysis of MRIs could be more useful for monitoring disease activity in short-term studies (less than 6 months), especially when monthly gadolinium-enhanced images are used.

In agreement with the quantitative measurement, there was no relation between the number of active lesions (new or enlarging) on visual analysis and the increase in EDSS score in the interval between imaging examinations, although this has been reported in other studies. This might be due to the difficulty in analyzing differences in size, especially when the lesion load is high and confluent lesions are present. Also, we did not analyze the number of lesions shrinking or "disappearing" (beyond the scanner resolution). The number of active (new or enlarging) lesions did, however, correlate significantly with the number of relapses in the interval period. Usually, sequential (monthly) contrast-enhanced MRI is used to monitor disease activity in MS, especially to evaluate treatment effects in short-term (phase II) studies. Our results show that active lesions on T2-weighted images might also be useful as an outcome measure in phase III studies over longer periods (and in larger samples), in which it is difficult to perform frequent (monthly) imaging.

The most salient finding in our follow-up study is the strong correlation between the difference in EDSS score and the hypointense lesion load on short-TR/short-TE SE images (figure 3). The lesions visible on the T2-weighted SE images represent both active and inactive MS plaques, and it is not possible to differentiate between early lesions (edema, inflammation, and mild demyelination) and chronic lesions (severe demyelination, axonal loss, and gliosis). Some MS lesions visible on T2-weighted SE images can also been seen on short-TR/short-TE SE images as regions with hypointensity and a more or less distinct boundary. Little is known about the value of these hypointense lesions. One study underlines the importance of these lesions in differentiating between MS and subcortical arteriosclerotic encephalopathy (SAE); hypointense lesions on the short-TR/short-TE SE images are uncommon in SAE, even though high-signal lesions are visible on T2-weighted SE images. SAE lesions mainly represent mild demyelination but little gliosis or axonal loss. Hypointense lesions on short-TR/short-TE SE images may represent chronic MS plaques, in which astrocyte growth and scarring appear.

What then causes the low signal on these short-TR/short-TE SE images? It could be due to changes in T1 relaxation time or to changes in other MRI variables. Barnes et al. showed that prolongation of T1 relaxation time is present in experimental glosis, in contrast to edema, where prolongation of T1 relaxation time is almost twice that of T1 relaxation time. Hiehle et al. analyzed unenhanced T1-weighted images of MS patients with magnetic transfer imaging (MTI). They divided the MS lesions visible on the T1-weighted SE images into either isointense or hypointense lesions visible on the T1-weighted images. The MTI ratio of the hypointense lesions appeared to be decreased in comparison with that of the isointense lesions. Hiehle et al concluded that the hypointense plaques represented the most demyelinated lesions. In hypointense short-TR/short-TE SE lesions, permanent neurologic damage has probably occurred, and such lesions may correlate better with clinical disability than do T2-weighted SE lesions. Indeed, a recent study showed an inverse correlation of average MTI ratio with EDSS score (SRCC = -0.44), this correlation being stronger than that between total lesion area on T2-weighted images and EDSS score (SRCC = 0.33).

In conclusion, computer-assisted volume measurement can be useful in monitoring disease activity in MS. The intra-observer variability for this method is small compared with that for visual analysis of MRIs. We showed that an increase in MRI lesion load correlates positively with an increase in disability, while the number of active lesions correlated significantly with the number of relapses. Extended follow-up studies with larger...
samples over longer periods (in patients with greater increase in disability) are needed to assess the definite role of quantitative MRI as a surrogate marker of disability in MS treatment trials. Short-TR/short-TE SE images seem to be more specific than T2-weighted SE images in identifying the lesions that cause disability.

Acknowledgments

We thank Ton Schweigmann for his technical support, Corrado Barrati for his help in analyzing the images, and Luc Truyen for his useful comments. We also thank Frank Hoogenraad for developing the software for the computer program and for his technical support.

References

10. Paty DW, Li DKB, the UBC MS/MRI Study Group, the IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. II. MRI analysis results of a multicenter, randomized, double-blind, placebo-controlled trial. Neurology 1993;43:662-667.