Association between MRI parameters and the MS severity scale: a 12 year follow-up study


Submitted
Abstract

Objective: Several MRI parameters are known to be associated with short-term outcome in multiple sclerosis (MS) patients. The goal of this study was to identify brain MRI parameters that are associated with overall disease severity.

Methods: In 46 MS patients, short interval MRI data were obtained at baseline and at first follow-up (FU) after 3.3 years. Overall disease severity was expressed using the MS Severity Score (MSSS) based on assessments at last FU (on average after 12.2 years). MRI parameters included both lesion loads and atrophy measures. Average and change parameters were calculated for MRI parameters and subsequently used as independent variables in regression models. MSSS was the dependent variable.

Results: FU was obtained in 43/46 patients (94%) and median expanded disability status scale (EDSS) score increased significantly from 2.5 to 4.0. At last FU median MSSS was 4.3 (range 2.2 to 6.9). In univariate analyses, both change and crosssectional T1-hypointense lesion lesion load (BHLL) and ventricular atrophy measures were associated with MSSS. A multiple regression model included the change parameter of BHLL. This model explained 20% of variance in MSSS, which increased to 34% when type of disease (relapsing remitting or secondary progressive), age and sex were entered additionally.

Conclusion: MRI measures of axonal loss are associated with higher overall disease severity in MS patients.
Introduction

In patients with multiple sclerosis (MS), prediction of long-term outcome in terms of clinical disability is important for both patients and clinicians. Natural history studies have identified clinical parameters that predict disability at long-term follow-up (FU).\textsuperscript{1-8} Although these clinical parameters do have modest prognostic value for the development of clinical disability later in the disease, the highly variable disease course and outcome as measured by the expanded disability status scale (EDSS)\textsuperscript{9} still remain largely unpredictable. Compared to data known from average disease progression in large groups of patients, models using clinical parameters refine the prediction of the time to reach EDSS 6 for the individual patient by only 30-40%.\textsuperscript{6,10} Substantial variation remains unexplained by those models, stressing the need for additional (paraclinical) parameters associated with progression of disability. Magnetic resonance imaging (MRI) may be valuable for this purpose: it is widely used as a diagnostic tool for MS and plays an important role in the most recent diagnostic criteria.\textsuperscript{11}

In most patients, typical MS lesions are seen on T2-weighted images and MRI helps to predict conversion to clinically definite MS (CDMS).\textsuperscript{12} Studies that follow patients with a clinically isolated syndrome (CIS) from onset describe significant associations between lesion loads and disability at FU.\textsuperscript{7,8,13-15} In the more advanced disease stage of CDMS, both cross-sectional and longitudinal studies show moderate associations between lesion loads on T2 or T1-weighted images and EDSS.\textsuperscript{16,21,22,24,25-28} Compared to lesion loads, brain atrophy seems to be associated more strongly with (future) disability.\textsuperscript{25,29-32} There is only limited data on the clinico-radiological association in the long-term, since only moderate duration of FU is available for most studies–this relationship is therefore largely unknown.\textsuperscript{13,33} A long-term FU study on patients that presented with CIS found that associations between change in EDSS and change in T2 lesion load weaken over time.\textsuperscript{13} Concern rises when translating these results to CDMS patients: are the already weak associations between MRI parameters and disability still valid at long-term FU?

The current study presents the 12-year clinical FU of a cohort from which two previous clinical and MRI observations have been reported previously.\textsuperscript{27} Our aim was to develop a multi-parametric model with MRI parameters obtained during
the initial short FU period that are most strongly associated with overall disease severity as expressed by the MS Severity Score (MSSS) based on assessments at long-term FU.

**Methods**

**Patients**
The initial cohort included 46 patients (28 women) with CDMS according to the Poser criteria. Baseline observations were performed between 1989 and 1991. After a median period of 40 months, a second observation (first FU) was performed according to the same protocol. Clinical data including age at onset, disease duration at entry, disability (EDSS) at entry and at first FU (at the time of the second MRI), type of disease course: relapsing-remitting (RR) or SP were available. Data on the use of disease modifying therapy (DMT) beyond the first FU visit were not systematically available–at the time of the first FU visit DMT were not available in the Netherlands. For long-term clinical observation (last FU), patients were contacted and hospital charts were reviewed for disability on the basis of EDSS. For 10 patients no recent hospital charts were available and EDSS was obtained by phone using a standard questionnaire. To classify overall disease severity, we used the MSSS, in which EDSS-levels are adjusted for disease duration. The institutional medical ethics committee approved the study and informed consent was obtained from every patient.

**MRI**
All patients had two scans: at baseline and at first FU. Scans were performed on a 0.6T machine (Technicare, Solon, Ohio, USA) using an identical protocol. Accurate repositioning was performed using internal landmarks. Axial T2 and T1-weighted spin-echo images (2755/60,120/2 and 450/28/4 repetition time/echo time/excitations)) were obtained. Each series consisted of 19 slices with a slice thickness of 5 mm (1.25 mm gap), covering the whole brain. The T1-weighted spin-echo sequence at baseline was performed before and 5 to 10 minutes after the administration of 0.1 mmol/kg gadolinium-DTPA. At first FU no gadolinium was administered.
Gadolinium-enhancing lesion loads (GdLL, only baseline), hypointense T1 lesion loads (BHLL) and hyperintense T2 lesion loads (T2LL) were quantified by a single reader using an in-house developed, semi-automated thresholding technique. Baseline and first FU scans were evaluated during the same session. Lesion loads have been reported previously. Additionally, we determined parenchymal and ventricular volumes on T1-weighted images. Intracranial volumes were determined on corresponding T2-weighted images. To control for differences in skull size, brain parenchymal fraction (BPF) and ventricular fraction (VF) were calculated by dividing parenchymal (BPF) and ventricular (VF) volumes by the intracranial volume.

**Statistical analysis, descriptive**

Because most data were not normally distributed, medians and interquartile range (IQR) were used to describe the data. The skewed distribution also necessitated the use of non-parametrical statistical tests: Mann Whitney U test was used to evaluate differences in MRI parameters between subgroups when lowest and highest quartiles of MSSS at last FU are compared. Longitudinal MRI data were normalized for duration of FU between first and second MRI. To describe correlations between (changes in) MRI and clinical parameters, Spearman rank correlation coefficients were used.

**Statistical analysis, multiple regression model**

To find the MRI parameter(s) that are most strongly associated with disease severity at last FU, a model was composed using a forward selection approach. MSSS as determined at last FU was the dependent variable. Because of the relatively small group size, summary parameters were used to optimize model building. This approach decreases the number of independent parameters and enables the use of information at baseline, first FU and change parameters without loss of statistical power. The effects of high correlations between baseline and first FU parameters are also avoided this way. Average values were calculated for: T2LL, BHLL, BPF and VF (baseline value + first FU value / disease duration at baseline (years) + disease duration at first FU (years)). Furthermore, a change parameter was calculated for T2LL, BHLL, BPF and VF (first FU value–baseline value / interval (years) between baseline and first FU). Performance of the model is reported by
percentage of explained variance. All statistical procedures were performed using SPSS 12.0 for Windows.

Results

Clinical parameters
During long term FU, two patients died due to causes not related to MS and were excluded from the analysis. One patient was lost to FU and had to be excluded. Characteristics at baseline and first FU were not different between excluded and reported patients. Data on the remaining 43 patients (94% adherence rate) are presented here. At baseline, median age was 38 years (Table 3.10).

Table 3.10. Clinical characteristics

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Group (n=43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration interval baseline–follow-up 1 (y)</td>
<td>3.3 (3.2 ; 3.6)</td>
</tr>
<tr>
<td>Duration interval follow-up 1–follow-up 2 (y)</td>
<td>8.6 (6.9 ; 10.3)</td>
</tr>
<tr>
<td>Duration interval baseline - follow-up 2 (y)</td>
<td>12.2 (10.4 ; 13.8)</td>
</tr>
<tr>
<td>Type MS at baseline (RR,SP)</td>
<td>27 ; 16</td>
</tr>
<tr>
<td>Age at baseline (y)</td>
<td>38 (33 ; 44)</td>
</tr>
<tr>
<td>Disease duration at baseline (y)</td>
<td>5.0 (2.4 ; 7.3)</td>
</tr>
<tr>
<td>Sex (female; male)</td>
<td>27 ; 16</td>
</tr>
<tr>
<td>EDSS at baseline</td>
<td>2.5 (2.0 ; 3.5)</td>
</tr>
<tr>
<td>EDSS at follow-up 1</td>
<td>2.5 (1.5 ; 5.0)</td>
</tr>
<tr>
<td>EDSS at follow-up 2</td>
<td>4.0 (3.5 ; 6.5)</td>
</tr>
<tr>
<td>MSSS (based on data from follow-up 2)</td>
<td>4.3 (2.2 ; 6.9)</td>
</tr>
</tbody>
</table>

RR = relapsing remitting; SP = secondary progressive; MSSS = MS Severity Score.

Twenty-seven patients had RRMS, the other 16 SPMS at interception. Median EDSS score at baseline and at the first FU was 2.5, and increased significantly to 4.3 (IQR 2.3–6.9) after an average of 12.2 years of FU, with death due to MS (EDSS 10) occurring in 2 patients. Patients with SPMS had significant higher MSSS compared to RRMS patients: 7.0 (IQR 3.3; 9.8) compared to 3.4 (1.8 ; 5.2), \( P=0.008 \).
MRI parameters

All patients had hyperintense lesions consistent with white matter MS lesions on the T2-weighted images. T2LL at baseline was highly variable (IQR 4.5; 32.4 ml) (Table 3.11).

Table 3.11. MRI characteristics

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Group (n=43)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2LL baseline (ml)</td>
<td>14.1 (4.5; 32.4)</td>
</tr>
<tr>
<td>BHLL baseline (ml)</td>
<td>3.4 (0.8; 11.2)</td>
</tr>
<tr>
<td>Enhancement (yes/no)</td>
<td>18/25</td>
</tr>
<tr>
<td>BPF baseline</td>
<td>0.830 (0.802; 0.865)</td>
</tr>
<tr>
<td>VF baseline</td>
<td>0.0226 (0.0159; 0.0327)</td>
</tr>
<tr>
<td>T2LL follow-up 1 (ml)</td>
<td>22.2 (7.6; 46.7)</td>
</tr>
<tr>
<td>BHLL follow-up 1 (ml)</td>
<td>6.4 (1.3; 15.6)</td>
</tr>
<tr>
<td>BPF follow-up 1</td>
<td>0.820 (0.789; 0.862)</td>
</tr>
<tr>
<td>VF follow-up 1</td>
<td>0.0224 (0.0153; 0.0331)</td>
</tr>
<tr>
<td>Change in T2LL/y (ml/y)</td>
<td>1.4 (0.6; 2.8)</td>
</tr>
<tr>
<td>Change in BHLL/y (ml/y)</td>
<td>0.5 (0.1; 1.3)</td>
</tr>
<tr>
<td>Change in BPF/y (* 100)</td>
<td>-0.343 (-0.923; 0.000)</td>
</tr>
<tr>
<td>Change in VF/y (*100)</td>
<td>-0.0028 (-0.1461; 0.0073)</td>
</tr>
</tbody>
</table>

T2LL = hyperintense T2 lesion load; BHLL = hypointense T1 lesion load; BPF = brain parenchymal fraction; VF = ventricular fraction.

At baseline, a strong correlation between BHLL and T2LL was found ($r=0.83$, $P<0.001$). At first FU, median T2LL increased to 22.2 ml. In all but two patients an accumulation of T2LL was observed, at a median rate of 1.4 ml/year (Table 3.11). BHLL had increased to 6.4 ml, at a median rate of 0.5 ml/year. The change in BHLL was not correlated with BHLL at baseline. However, changes in T2LL were moderately associated with T2LL at baseline ($r=0.53$, $P<.001$). BHLL was associated with VF at both baseline ($r=0.63$, $P<.001$) and first FU ($r=0.74$, $P<.001$). Median BPF had decreased at first FU from 0.830 to 0.820. VF increased most in patients with the highest quartile MSSS (from 0.0267 to 0.0357); in this group the baseline and FU VF was also statistically significantly higher compared to patients with the lowest quartile MSSS: 0.0185 compared to 0.0267 ($P=0.034$) and at first FU 0.0162 compared to 0.0357 ($P=0.012$). Both baseline and first FU lesion loads
(T2LL and BHLL) were not statistically significantly different between the group with highest quartile MSSS and the lowest MSSSS quartile group. Accumulation of BHLL was higher in the group with highest quartile MSSS compared to the lowest MSSS quartile group (median 1.2 ml/year compared to 0.2 ml/year, \( P = 0.043 \)). No statistically significant differences in accumulation of T2LL were found between the two groups.

**Table 3.12. Linear regression: univariate analysis**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Beta</th>
<th>P value</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2LL baseline (ml)</td>
<td>0.26</td>
<td>0.092</td>
<td>0.05</td>
</tr>
<tr>
<td>BHLL baseline (ml)</td>
<td>0.40</td>
<td>0.008</td>
<td>0.14</td>
</tr>
<tr>
<td>BPF baseline</td>
<td>-0.18</td>
<td>0.275</td>
<td>0.03</td>
</tr>
<tr>
<td>VF baseline</td>
<td>0.31</td>
<td>0.063</td>
<td>0.07</td>
</tr>
<tr>
<td>T2LL follow-up 1 (ml)</td>
<td>0.29</td>
<td>0.064</td>
<td>0.06</td>
</tr>
<tr>
<td>BHLL follow-up 1 (ml)</td>
<td>0.44</td>
<td>0.004</td>
<td>0.17</td>
</tr>
<tr>
<td>BPF follow-up 1</td>
<td>-0.21</td>
<td>0.211</td>
<td>0.02</td>
</tr>
<tr>
<td>VF follow-up 1</td>
<td>0.42</td>
<td>0.010</td>
<td>0.15</td>
</tr>
<tr>
<td>Change in T2LL/y (ml/y)</td>
<td>0.26</td>
<td>0.087</td>
<td>0.05</td>
</tr>
<tr>
<td>Change in BHLL/y (ml/y)</td>
<td>0.44</td>
<td>0.003</td>
<td>0.18</td>
</tr>
<tr>
<td>Change in BPF/y</td>
<td>0.02</td>
<td>0.920</td>
<td>0.00</td>
</tr>
<tr>
<td>Change in VF/y</td>
<td>0.35</td>
<td>0.033</td>
<td>0.10</td>
</tr>
</tbody>
</table>

T2LL = hyperintense T2 lesion load; BHLL = hypointense T1 lesion load; BPF = brain parenchymal fraction; VF = ventricular fraction. Adjusted \( R^2 \) = explained fraction of variance in MSSS at last follow-up.

**Linear regression models**

In the univariate analysis, both change and crossectional lesion (BHLL) and atrophy (VF) measures were associated with MSSS (Table 3.12).

To create a model with the MRI parameters that are most strongly associated with disease severity represented by MSSS score, the following summary parameters were used in a forward selection procedure; GdLL (baseline only), average and change of T2LL, BHLL, BPF and VF. The final MRI model contained the change parameter of BHLL only. This model was uncorrected for clinical parameters and explained 20% of the variance in MSSS. After inclusion of type of disease course (RR or SP), age and sex in the model, the change parameter of BHLL survived in the forward analysis. The percentage of explained variance in MSSS increased to 34%.
Discussion

Studies concerning associations between changes in lesion loads and concomitant disability changes or changes in disability at long-term FU are sparse in MS. Most of these studies included patients directly from onset and have reported moderate associations between (change in) MRI and disability.\textsuperscript{8,13,15} A study on these associations where patients were followed from onset of the disease showed associations to decrease in strength over time.\textsuperscript{13} A possible explanation could be the noise introduced due to the large number of patients lost to FU. On the other hand, a recent study reported the increasing predictive value of conventional MRI measures for clinical outcome during (13 years) FU.\textsuperscript{33}

The current study included MS patients with varying disease duration that were followed-up for a median of 12 years, enabling classification of disease severity (MSSS) in a reliable way. Despite being small, our study had the benefit of having a low number of patients that were lost to FU (6%). Main outcome is that in a group of CDMS patients with varying disease duration, MRI parameters reflecting axonal loss (BHLL and ventricular widening) are associated with the higher disease severity at long-term FU. Change in BHLL was included in the final multivariate logistic regression model. Recently, stratification of patients was proposed using measures similar to those included in our final model: using black hole fraction as measure for focal brain tissue destruction. That model also included atrophy as measure for diffuse brain tissue destruction, a measure that was associated with MSSS in our study as well.\textsuperscript{39} Their longitudinal observations showed these stratified groups to be distinct at baseline and FU and thus not to represent different stages of the disease.

Brain atrophy is present from the earliest stages of MS. Besides cross-sectional associations between atrophy and clinical status, as measured by the EDSS or scales of cognitive function, there are numerous short-term FU studies that show these associations may also exist longitudinally.\textsuperscript{26,30-32,40-44} Previous studies extended these findings with FU durations of 6 and 8 years.\textsuperscript{25,29} and recently a 13 year FU study was published.\textsuperscript{33} No atrophy measures (independently) made it into our final model but the univariate analysis shows these measures still remain important during long-term FU and confirms the unfavourable impact of atrophy on disease progression.
BHLL is thought to represent focal brain tissue destruction and axonal loss more specifically than T2LL. Therefore, it is not surprising that associations with clinical status improve in most studies when BHLL is used rather than T2LL. This applies to cross-sectional and short-term longitudinal studies. Our observations suggest that the impact of BHLL on disease progression is also valid during long-term FU. Several studies reported on the associations between BHLL and (central) brain atrophy measures like VF in the various stages of the disease. It is thought that destructive T1-hypointense lesions in the periventricular regions lead to local axonal loss and decreasing volume of brain tissue with subsequently enlarging ventricles. In some studies BHLL and disease severity were not associated. This may depend on the disease duration, overall disease severity and prevalence of T1-hypointense lesions in the studied population: the present study included patients with well established MS with relatively high BHLL.

A limitation of (longitudinal) studies on differences in the progression of disability between groups of patients is the lack of reliable classifying methods. The EDSS is the most widely used method to classify MS related disability. Several limitations, including the lack of correction for disease duration and its non-linearity have been reported. Regarding these limitations, we did not use the annualized change in EDSS to determine the progression of disability during FU. The retrospective character of the study with EDSS measured at only three time-points also precluded the use of the exact time to reach EDSS 4 (or other values of the EDSS). Instead, we used the MSSS to classify disease severity: the MSSS provides a tool to compare the severity of the disease course between patients and is based on EDSS score adjusted for disease duration.

Weaknesses of the study can be divided in both MRI (performed at only 0.6T, impossibility to use more sophisticated brain volume change methods rather than the more ‘noisy’ BPF and VF measures) and clinical limitations (partial use of EDSS by phone, small group size and limited information on the use of DMT and relapse rate). Although use of DMT could not be included in the analysis, this probably has not affected our results: DMT is known to effect the progression on MRI but our MRI measurements were obtained before DMT became available. Although some patients started using DMT between the first and last FU, this probably did not effect our results either since, until now, no major effects of DMT on long-term disability have been reported but several studies indicate that some benefit may be expected.
In conclusion the findings of our study support the idea that MRI measures of axonal loss are associated with clinical status at long-term FU.

Acknowledgements

A. Minneboo is supported by a grant (98-348 MS) of the Dutch MS Research Foundation. The MS Center of the VU University Medical Center is supported by the Dutch MS Research Foundation and VU University Medical Center. We thank drs. Hoogervorst and Kalkers for their work on obtaining the EDSS-by-phone.

References


Chapter 3.4


