DOI: 10.1093/brain/awg232 Advanced Access publication August 21, 2003 *Brain* (2003), **126,** 2323–2332

Conventional and magnetization transfer MRI predictors of clinical multiple sclerosis evolution: a medium-term follow-up study

Marco Rovaris,[1] Federica Agosta,[1] Maria Pia Sormani,[1] Matilde Inglese,[1] Vittorio Martinelli,[2] Giancarlo Comi[2] and Massimo Filippi[1]

[1]*Neuroimaging Research Unit and* [2]*Clinical Trial Unit, Department of Neuroscience, Scientific Institute and University Ospedale San Raffaele, Milan, Italy*

Correspondence to: Dr Massimo Filippi, Neuroimaging Research Unit, Department of Neuroscience, Scientific Institute and University Ospedale San Raffaele, via Olgettina, 60, 20132 Milan, Italy
E-mail: filippi.massimo@hsr.it.

Summary

The correlation between conventional MRI lesion load accumulation and multiple sclerosis clinical evolution is only modest. The assessment of brain parenchymal volume and of its changes over time may provide adjunctive MRI markers reflecting the more disabling aspects of multiple sclerosis pathology. Magnetization transfer (MT) MRI is sensitive to 'occult' multiple sclerosis-related brain damage and might also contribute to overcome the clinical/MRI paradox. In this study, we assessed the value of conventional and MT MRI-derived metrics in predicting the medium-term clinical evolution of patients with different multiple sclerosis phenotypes. We studied 73 patients, with relapsing–remitting multiple sclerosis (n = 34), secondary progressive multiple sclerosis (n = 19) and clinically isolated syndromes suggestive of multiple sclerosis (n = 20), and 16 healthy subjects. Brain dual-echo, T_1-weighted (only in patients) and MT MRI scans were obtained at baseline and after 12 months. T_2-hyperintense and T_1-hypointense lesion volumes, normalized brain volume and average lesion MT ratio (MTR) were measured. MTR histograms from the whole brain tissue were also obtained. Clinical multiple sclerosis evolution and neurological disability were re-assessed in all patients after a median follow-up of 4.5 years. A multivariate analysis was performed to establish which clinical and MRI-derived variables were significant predictors of neurological deterioration at the end of the study period. At the end of follow-up, 34 patients showed significant neurological deterioration. The final multivariable model included average brain MTR percentage change after one year [P = 0.02, odds ratio (OR) = 0.86] and baseline T_2-hyperintense lesion volume (P = 0.04, OR=1.04) as independent predictors of medium-term disability accumulation (r^2 = 0.23). In this cohort of patients, abnormal values of average brain MTR changes showed a relatively high specificity (76.9%) and positive predictive value (59.1%) for Expanded Disability Status Scale score deterioration in individual cases. In patients with multiple sclerosis, a comprehensive estimation of the short-term changes of both conventional and MT MRI-detectable lesion burden might provide useful prognostic information for the medium-term clinical disease evolution.

Keywords: multiple sclerosis; MRI; magnetization transfer MRI; clinical evolution

Abbreviations: BET = brain extraction tool; CIS = clinically isolated syndrome; CI = confidence interval; EDSS = Expanded Disability Status Scale; MT = magnetization transfer; MTR = magnetization transfer ratio; NABT = normal-appearing brain tissue; NBV = normalized brain volume; OR = odds ratio; PBVC = percentage brain volume change; RRMS = relapsing–remitting multiple sclerosis; SPMS = secondary progressive multiple sclerosis; TLV = total lesion volume.

Introduction

A major issue in the management of multiple sclerosis is the identification of factors that might help in predicting the individual patient's outcome. Even though several epidemiological studies have generated multiparametric models including various clinical predictors (Weinshenker *et al*., 1991), the results are still sub-optimal in terms of accuracy,

mainly because of the wide inter- and intra-patient variability as regards the clinical severity and rate of progression of multiple sclerosis.

The advent of MRI, with its exquisite sensitivity in detecting the presence of multiple sclerosis lesions (Ormerod *et al*., 1987), has markedly improved the confidence in making a diagnosis of multiple sclerosis. Serial T_2-weighted and post-gadolinium T_1-weighted MRI scans are also able to reveal disease activity and lesion burden accumulation over time, with greater sensitivity than the clinical assessment of relapses and disability progression (Miller *et al*., 1993). Nevertheless, in patients with established multiple sclerosis, the correlation between clinical and MRI findings is only modest (Kappos *et al*., 1999; Molyneux *et al*., 2001; Rovaris *et al*., 2003). Only in patients at presentation with a clinically isolated syndrome (CIS) suggestive of multiple sclerosis does the burden of T_2-hyperintense lesions at disease onset seems to be a robust predictor of subsequent evolution to definite multiple sclerosis (Brex *et al*., 2002). However, even in these patients the increase of lesion load in the following years is only moderately correlated to the long-term accumulation of neurological disability (Brex *et al*., 2002). In the last few years, there has been considerable interest in measuring brain volume decrease (atrophy) as an adjunctive MRI marker reflecting the more disabling aspects of multiple sclerosis pathology (Miller *et al*., 2002). Cross-sectional and longitudinal studies have shown moderate correlations between brain atrophy and multiple sclerosis clinical aspects (Fisher *et al*., 2002; Kalkers *et al*., 2002; Miller *et al*., 2002). However, important issues regarding the mechanisms underlying the development of multiple sclerosis-related brain atrophy have not yet been fully addressed and the assessment of atrophy seems to be less sensitive to disease changes than other MRI-derived outcomes (Miller *et al*., 2002).

Several reasons have been advocated to explain the discrepancy between clinical and MRI evolution in multiple sclerosis patients (Filippi and Grossman, 2002). These include the limitations of conventional MRI in terms of inability to characterize the pathological heterogeneity of multiple sclerosis lesions and to quantify the severity of tissue damage occurring in the normal-appearing brain tissue (NABT). The use of quantitative magnetic resonance-based techniques might, at least partially, overcome the aforementioned limitations and provide an *in vivo* accurate estimate of the disabling aspects of multiple sclerosis pathology (Filippi and Grossman, 2002). Among these techniques, magnetization transfer (MT) MRI has become one of the most extensively applied to the assessment of multiple sclerosis (Rovaris and Filippi, 2003). A low magnetization transfer ratio (MTR) indicates a reduced capacity of the macromolecules in the CNS to exchange magnetization with the surrounding water molecules, reflecting damage to myelin or to the axonal membrane. The most compelling among much evidence indicating that markedly decreased MTR values correspond to areas where severe tissue loss has occurred is the strong correlation of MTR values from multiple sclerosis lesions and normal-appearing white matter with the percentage of residual axons and the degree of demyelination found in a post-mortem study of patients with multiple sclerosis (van Waesberghe *et al*., 1999).

By means of histogram-based analysis (van Buchem *et al*., 1996), MT MRI data can be post-processed to obtain metrics which enable the assessment and grading of overall brain damage in multiple sclerosis patients. Metrics from whole brain MTR histograms have been found to be moderately to strongly correlated with multiple sclerosis-related physical (Filippi *et al*., 1999*a*; Kalkers *et al*., 2001) or cognitive (Rovaris *et al*., 1998; van Buchem *et al*., 1998) disability in cross-sectional studies. In addition, whole brain MTR histogram-derived quantities seem to be sensitive to longitudinal changes over relatively short periods of time (Filippi *et al*., 2000) in patients with relapsing-remitting multiple sclerosis (RRMS) and secondary progressive multiple sclerosis (SPMS).

To our knowledge, the potential of MT MRI to provide paraclinical predictors of multiple sclerosis clinical outcome has been investigated only in a recent, preliminary study (Santos *et al*., 2002) of a small group of RRMS and SPMS patients. In the present study, we assessed conventional and MT MRI-derived metrics and their changes over a short period of time as predictors of medium-term clinical evolution in patients with different multiple sclerosis phenotypes.

Material and methods

Patients

Patients included had either suffered from clinically definite multiple sclerosis (Poser *et al*., 1983) for at least 2 years, with a RR or SP (Lublin *et al*., 1996) disease phenotype, or from CIS suggestive of multiple sclerosis, with the first clinical attack in the preceding 3 months and at least four focal abnormalities on T_2-weighted scans. All RRMS, SPMS and CIS patients who had taken part in a short-term follow-up study (Filippi *et al*., 2000) were asked to participate into the present one. None of the patients had taken immunosuppressive or immunomodulating treatments for at least 12 months prior to entry into the study. In addition, they had neither relapses nor steroid treatment during the 3 months preceding both the baseline and the follow-up scanning sessions. At the time MRI scans were performed, patients were assessed neurologically by a single physician, who was unaware of the MRI results, and disability was measured using the Expanded Disability Status Scale (EDSS) (Kurtzke *et al*., 1983). Neurological assessment and EDSS rating were repeated for a third time after a median follow-up period of 4.5 years (range 3.2–5.9 years). At follow-up evaluations, patients were considered clinically worsened if they had an EDSS score increase $\geqslant$1.0 when the baseline EDSS was <6.0, or an EDSS score increase $\geqslant$0.5 when the baseline EDSS was $\geqslant$6.0. EDSS changes had always to be confirmed by a second visit

after a 3 month, relapse-free interval. The occurrence of changes of the clinical disease phenotype [defined as evolution to clinically definite multiple sclerosis for CIS patients and entrance into the progressive phase for RRMS patients) (Poser *et al*., 1983; Lublin *et al*., 1996) was also assessed. During the study period, the occurrence of clinical multiple sclerosis relapses was assessed by the same physician. Sex- and age-matched controls with no previous history of neurological diseases and with a normal neurological examination were also studied. Approval from the San Raffaele Ethics Committee and written informed consent from all the subjects were obtained before study initiation.

Image acquisition

Brain MRI scans were obtained at study entry and after 12 months (±10 days) using the same 1.5 Tesla scanner (Magnetom SP63, Siemens, Erlangen, Germany) on a regular course of maintenance. During each session, the following scans were performed without moving the patient from the gantry: dual-echo conventional spin echo [TR (repetition time) = 2400, TE (echo time) = 30/80]; T_1-weighted conventional spin echo (TR = 768, TE = 15) (only in patients); 2D gradient echo (TR = 600, TE = 12, $\alpha = 20°$) with and without a saturation pulse. The radio frequency saturation pulse was 1.5 kHz below the water frequency, with a Gaussian envelope of duration of 16.4 ms, a bandwidth of 250 Hz and an amplitude of 3.4×10^{-6} Tesla. For the dual-echo and T_1-weighted scans, 24 contiguous interleaved axial slices were acquired with 5 mm slice thickness, 256 × 256 matrix and 250 × 250 mm field of view. MT MRI scans were obtained using the same acquisition parameters, except for the number of slices, which was 20. The set of slices for the MT images was positioned to obtain the same central 20 slices as for the dual-echo and T_1-weighted scans. The slices were positioned to run parallel to a line that joins the most infero-anterior and infero-posterior parts of the corpus callosum (Miller *et al*., 1991). At follow-up, patients were carefully repositioned following published guidelines (Miller *et al*., 1991).

Image review and analysis

Lesions were first identified by agreement by two experienced observers, without knowing to whom the scans belonged and the scan order of acquisition, on the hard copies of the first echo of the dual-echo scans, and on the T_1-weighted scans; the second echo of the dual-echo scans was always used to increase confidence in lesion identification. For T_1-weighted scans, only areas with a signal intensity between that of the grey matter and that of the CSF and with corresponding lesions on both echoes of the dual-echo images were considered as hypointense lesions. Total lesion volume (TLV) measurements were then performed by a single observer, again without knowing to whom the scans belonged and the order of scan acquisition, using a semi-automated segmentation technique (Rovaris *et al*., 1997).

Using T_1-weighted images, both longitudinal (two time-points) percentage brain volume change (PBVC) and cross-sectional (single time-point) normalized brain volume (NBV) were estimated. PBVC was estimated using SIENA (Structural Image Evaluation, using Normalization, of Atrophy) (Smith *et al*., 2001) and NBV was estimated using SIENAX (Siena Cross-Sectional) (Smith *et al*., 2002). In both methods, the first stage is the extraction of the brain from each input MRI using BET (Brain Extraction Tool, Oxford Centre for Functional MRI of the Brain, Oxford University, UK). The original images are then registered to a canonical image in a standardized space. In the longitudinal method, to estimate changes between the images, SIENA finds all brain surface edge points using tissue-type segmentation, and then correlates differentiated 1D perpendicular profiles taken around the position of these points in both images. Brain atrophy is quantified by taking the mean perpendicular edge motion over all edge points and converting this into PBVC. The normalizing factor in the conversion is found by a self-calibration step, which involves finding estimated atrophy on an artificially scaled version of one image with respect to itself. Accuracy in measuring PBVC was found to be ~0.2% (Smith *et al*., 2001). In SIENAX, a similar registration process is applied, but instead of a second time point image, standard space average brain and skull images are used. The estimate of brain tissue volume for a subject is then multiplied by the normalization factor to yield the NBV. Reproducibility tests have resulted in a mean standard error across a group of normal subjects of ~0.5% (Smith *et al*., 2002). From each patient, the NBV at baseline and the PBVC between baseline and 12 month scans were obtained.

MTR histograms were derived from the whole brain parenchyma (van Buchem *et al*., 1996). First, the two sets of gradient echo images were co-registered using a technique based on mutual information (Studholme *et al*., 1997). From the two sets of co-registered images, MTR maps were derived on a pixel-by-pixel basis. The resulting MTR maps were co-registered with the corresponding proton density-weighted images (Studholme *et al*., 1997)., Extra-cranial tissues were then removed. To reduce any further image noise, all pixels with MTR <10% were also excluded. For all the histograms, the peak height and the average MTR were analysed. On the co-registered MTR maps, we also calculated the average lesion MTR, as described elsewhere (Filippi *et al*., 1999).

Statistical analysis

Baseline and follow-up TLV and MT MRI-derived metrics were compared using a Student *t*-test for paired data. To compare TLV, NBV, PBVC and MT MRI metrics between groups, a one-way ANOVA (analysis of variance) was used and three comparisons were decided *a priori* (*a priori* contrasts): healthy subjects versus all patients; CIS versus

Table 1 *Demographic and clinical characteristics of patients with different clinical phenotypes*

	CIS	RRMS	SPMS
Number of patients	20	34	19
Male/female	9/11	9/25	4/15
Mean age (SD) (years)	28.2 (4.9)	32.7 (8.4)	40.5 (10.6)
Median disease duration (range) (years)	–	7 (2–25)	8 (3–23)
Median EDSS at baseline (range)	0.0 (0.0–1.5)	2.5 (1.0–5.5)	5.5 (3.5–6.5)
Median EDSS at one-year follow-up (range)	1.0 (0.0–2.5)	2.0 (1.0–7.0)	6.0 (3.5–6.5)
Median EDSS at final follow-up (range)	1.0 (0.0–3.0)	3.0 (1.0–7.0)	6.5 (3.5–8.0)
EDSS worsening at final follow-up	7 (35%)	13 (38%)	14 (74%)
Mean number of on-study relapses (range)	1.8 (0–8)	3.6 (0–10)	1.7 (0–7)
On-study immunomodulatory treatment	10 (50%)	22 (65%)	8 (42%)

Table 2 *Conventional MRI findings at study entry and after a 1-year follow-up*

Variable		All patients	CIS (n = 20)	RRMS (n = 34)	SPMS (n = 19)
T_2 TLV – baseline (ml)	Mean (SD)	15.9 (14.9)	6.5 (3.4)	17.6 (16.0)	22.7 (15.9)
	Range	1.0–72.9	1.8–15.5	1.0–72.9	6.3–70.8
T_2 TLV – follow-up (ml)	Mean (SD)	16.4 (15.6)	9.2 (4.0)	15.8 (15.8)	24.9 (19.0)
	Range	2.0–80.2	4.4–21.8	2.0–80.2	5.5–75.1
	P*	0.38	<0.0001	0.04	0.10
T_2 TLV – % change	Median	+6.6	+42.1	+16.1	+6.1
	Range	–53, +202	+1, +150	–53, +202	–32, +45
T_1 TLV – baseline (ml)	Mean (SD)	2.6 (5.1)	0.2 (0.2)	1.5 (2.1)	8.2 (8.2)
	Range	0.0–22.3	0.0–1.0	0.0–8.4	0.1–22.3
T_1 TLV – follow-up (ml)	Mean (SD)	4.0 (6.4)	1.2 (0.8)	3.1 (3.6)	9.9 (10.4)
	Range	0.1–36.9	0.3–3.2	0.1–14.0	1.0–36.9
	P*	<0.0001	<0.0001	0.004	0.20
T_1 TLV – % change	Median	+90.1	+508	+61.4	+23.2
	Range	–95, +3384	+191, +1513	–95, +1057	–61, +3384
NBV – baseline (ml)	Mean (SD)	1356.0 (108.8)	1414.5 (75.3)	1317.6 (121.1)	1349.7 (92.6)
	Range	1030.3–1508.0	1192.4–1508.0	1030.3–1499.5	1111.8–1506.9
PBVC (1-year follow-up versus baseline NBV) (%)	Median (SD)	–0.7 (1.2)	–0.3 (0.6)	–0.7 (0.9)	–1.4 (2.0)
	Range	–6.6, +2.2	–1.3, +1.1	–3.4, +0.9	–6.6, +2.2

*Pairwise comparisons versus baseline values (see the text for details about statistical analysis methods).

RRMS; RRMS versus SPMS. The number of *a priori* contrasts was determined by the available degrees of freedom and their nature was decided based on clinical relevance. A univariate logistic regression model adjusted for follow-up duration was used to screen the clinical and MRI variables as independent predictors of the probability of having an EDSS deterioration at final follow-up. Those variables with a P value <0.20 entered a multivariate analysis where the presence or absence of EDSS deterioration was the dependent variable. To investigate the potential role of brain MTR histogram metrics as predictors of multiple sclerosis evolution in individual cases, we classified patients' values for each of these quantities as abnormal when they were ⩾2 SD below the corresponding mean values obtained from the cohort of healthy subjects. We then calculated the sensitivity, specificity, positive predictive value and accuracy for each of these quantities in predicting the presence of EDSS deterioration at final follow-up assessment. When calculating the statistical significance of all these tests, no correction for multiple comparisons was performed due to the exploratory nature of this study.

Results

Seventy-three patients (51 females and 22 males; CIS = 20, RRMS = 34, SPMS = 19) and 16 controls (11 females, five males, mean age 30.9 years, range 19–39 years) were studied. Of the original patients' cohort (Filippi *et al*., 2000), only five RRMS patients did not participate into this study, because of unwillingness (three patients) or inability (two patients) to

Table 3 *MT MRI findings at study entry and after a 1-year follow-up*

Variable		All patients	CIS (n = 20)	RRMS (n = 34)	SPMS (n = 19)
Average lesion MTR –baseline (%)	Mean (SD)	42.4 (3.2)	43.0 (3.1)	42.2 (3.3)	42.0 (3.1)
	Range	32.8–49.4	32.8–48.3	34.6–47.3	37.0–49.4
Average lesion MTR – follow-up (%)	Mean (SD)	41.2 (2.9)	42.7 (2.1)	41.4 (2.8)	39.0 (2.6)
	Range	33.3–46.9	37.4–46.1	34.9–46.9	33.3–44.6
	P*	0.001	0.67	0.10	<0.0001
Average lesion MTR – % change	Median	–2.2	–2.2	–0.6	–5.8
	Range	–19.0, +29.2	–10.0, +29.2	–15.1, +7.2	–19.0, +3.4
Average brain MTR –baseline (%)	Mean (SD)	45.4 (2.1)	46.3 (1.2)	44.8 (2.2)	45.6 (2.5)
	Range	39.0–51.6	43.5–48.1	39.0–50.4	43.3–51.6
Average brain MTR – follow-up (%)	Mean (SD)	43.6 (1.7)	44.4 (1.3)	43.7 (1.8)	42.5 (1.4)
	Range	38.8–47.2	41.4–46.7	38.8–47.2	40.2–46.0
	P*	<0.0001	<0.0001	0.006	<0.0001
Average brain MTR – % change	Median	–3.7	–3.9	–2.6	–6.4
	Range	–16.0, +7.2	–10.0, –1.2	–15.1, +7.2	–16.0, +4.2
Histogram peak height – baseline	Mean (SD)	101.5 (19.0)	113.2 (14.0)	97.2 (18.3)	96.8 (20.2)
	Range	63.0–137.6	88.6–137.6	63.0–136.9	66.7–130.5
Histogram peak height – follow-up	Mean (SD)	99.5 (17.3)	114.9 (11.7)	96.1 (15.4)	89.3 (15.1)
	Range	55.2–139.8	95.7–139.8	55.2–120.1	56.4–117.9
	P*	0.17	0.56	0.53	0.04
Histogram peak height – % change	Median	–2.4	+1.5	–2.5	–6.7
	Range	–30.0, +40.2	–21.1, +40.0	–23.1, +29.2	–30.0, +18.2

*Pairwise comparisons versus baseline values (see the text for details about statistical analysis methods).

attend the scheduled visits. Mean patients' age was 33.5 (range 19–53) years and median EDSS scores at baseline and final follow-up were 2.5 and 3.0 (range 0.0–8.0). Clinical characteristics for the three patient subgroups are reported in Table 1. At a retrospective evaluation of study entry data, all CIS patients were found to fulfil the criteria of McDonald and colleagues (McDonald *et al*., 2001) for paraclinical evidence of spatial disease dissemination, based upon either MRI findings alone (n = 16) or the combination of MRI and CSF findings (n = 4). During the follow-up period, 12 CIS patients developed clinically definite multiple sclerosis and 5 RRMS patients entered a SP course. A total of 173 relapses were experienced by 58 patients (12 CIS, 33 RRMS and 13 SPMS). All relapses were treated with intravenous methylprednisolone (1 g per day for 3–5 days without tapering). At final follow-up, 34 patients (47%) were considered clinically worsened. During the study, 40 patients (55%) started immunomodulatory treatment with either interferon β-1b (10 patients 875 μg weekly) or interferon β-1a (30 patients 66 μg weekly). Twenty-five of these patients started their treatment during the first year of follow-up, 1–6 months before undergoing the second MRI scan. The mean duration of treatment during the study period was 36.0 (range 10–45) months.

No abnormalities were detected on dual-echo brain MRI scans from controls at baseline and one-year follow-up. The mean values (SD) of brain MTR histogram parameters obtained from controls at baseline and follow-up were 47.5% (1.1) and 47.4% (1.6) for average MTR, 116.9 (10.5) and 116.8 (17.7) for histogram peak height. There were no significant differences between baseline and follow-up values for any of these quantities.

Tables 2 and 3 report the values of brain TLV, NBV and MT MRI quantities for CIS, RRMS and SPMS patients, at study entry and 1-year follow-up. A significant increase was observed for T_2-hyperintense TLV in CIS patients and for T_1-hypointense TLV in CIS and RRMS patients. Significant decreases were observed for: average lesion MTR in SPMS patients; for average brain MTR in CIS, RRMS and SPMS patients; and for histogram peak height in SPMS patients. All the MTR histogram-derived quantities resulted to be lower in multiple sclerosis patients than in healthy controls (P values <0.001). Mean T_2 and T_1 TLV at baseline and the respective changes were significantly higher in RRMS than in CIS patients (P = 0.004 and P = 0.007 for baseline, P < 0.001 for one-year changes); average brain MTR (at baseline) and histogram peak height (at baseline and follow-up) were significantly lower in RRMS than in CIS patients (P values: 0.008, 0.001 and <0.001, respectively). Mean T_1 TLV at baseline and follow-up were significantly higher in SPMS than in RRMS patients (P values: <0.001 and 0.002); the observed 1-year decreases of average lesion MTR and average brain MTR were significantly greater in SPMS than in RRMS patients (P values: 0.002 and 0.003). Baseline NBV was significantly smaller in RRMS than in CIS patients (P = 0.002). PBVC after 1 year did not differ significantly either between CIS and RRMS or between RRMS and SPMS patients (P values: 0.31 and 0.07).

Table 4 reports the results of the univariate logistic regression analysis. Multiple sclerosis clinical phenotype, disease duration, baseline T_2 and T_1 TLV, average brain and lesion MTR percentage changes entered the multivariate analysis. The final multivariable model included average brain MTR percentage change after one year [odds ratio

Table 4 *Univariate logistic regression analysis of the predictive value of clinical and MRI-derived quantities for patients' EDSS worsening at medium-term follow-up (dependent variable)*

Independent variable	Odds ratio (95% CI)	*P* value
Patients' age	1.03 (0.97–1.08)	0.22
Disease duration*	1.08 (0.99–1.18)	0.08
Multiple sclerosis clinical phenotype (SPMS versus CIS + RRMS)*	4.79 (1.51–15.19)	0.008
On-study disease-modifying treatment	0.48 (0.14–1.64)	0.24
On-study relapse rate	0.98 (0.81–1.18)	0.82
Baseline T_2 TLV*	1.05 (1.01–1.09)	0.03
Baseline T_1 TLV*	1.10 (0.98–1.23)	0.12
T_2 TLV percentage change	0.89 (0.33–2.40)	0.81
T_1 TLV percentage change	0.99 (0.89–1.09)	0.83
Baseline NBV	1.00 (1.00–1.01)	0.69
PBVC	0.98 (0.65–1.48)	0.91
Baseline average brain MTR	1.01 (0.99–1.04)	0.29
Baseline histogram peak height	0.99 (0.96–1.01)	0.38
Baseline average lesion MTR	1.00 (0.99–1.02)	0.68
Average brain MTR percentage change*	0.85 (0.75–0.97)	0.01
Histogram peak height percentage change	1.61 (0.04–67.84)	0.80
Average lesion MTR percentage change*	0.93 (0.85–1.00)	0.06

*Variables entering the multivariate analysis. Percentage changes were for 1-year follow-up versus baseline scans. See the text for further details.

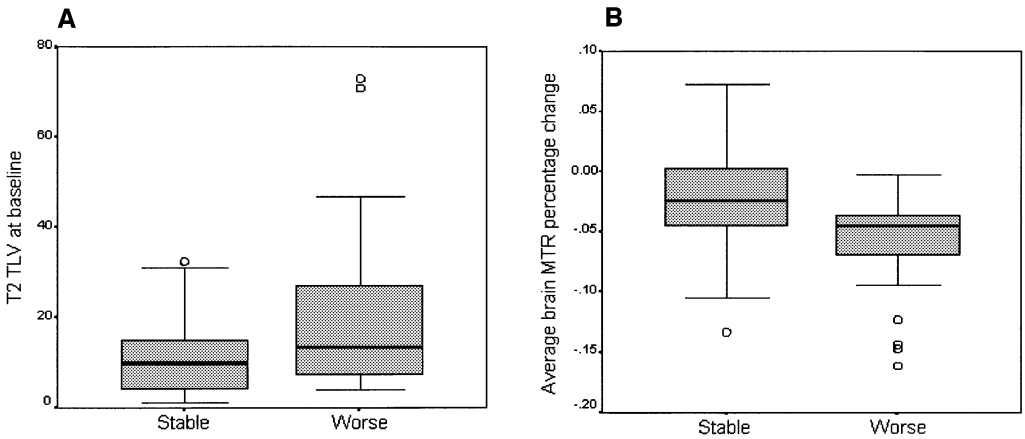

Fig. 1 Plots of T_2 TLV at study entry (**A**) and average brain MTR percentage changes after one year (**B**) in patients who were clinically stable versus worsened at study end (see the text for definitions). Plots are based on the median, quartiles and extreme values. The box represents the interquartile range, which contains 50% of individual patients' values. The whiskers are lines that extend from the box to the highest and lowest values, excluding outliers. A line across the box indicates the median value.

(OR): 0.86, 95% confidence interval (CI): 0.76–0.98; $P = 0.02$] and baseline T_2 TLV (OR: 1.04, 95% CI: 1.00–1.09; $P = 0.04$) as independent predictors of medium-term EDSS deterioration (Nagelkerke $r^2 = 0.23$) (Fig. 1). These results did not change when patients with CIS were not considered for the analysis (data not shown).

Average brain MTR and histogram peak height at baseline were abnormal in 40 (54.8%) and 29 (39.7%) patients, respectively. Average brain MTR and histogram peak height percentage changes after 1 year were abnormal in 22 (30.1%) and five (6.8%) patients, respectively. Average brain MTR percentage change after 1 year showed the best trade-off between sensitivity (44.1%) and specificity (76.9%) in predicting EDSS deterioration at follow-up, leading to a positive predictive value of 59.1% and to an accuracy of 58.0%. Histogram peak height at baseline showed a positive predictive value of 51.7% and an accuracy of 54.7% in predicting medium-term EDSS worsening. The positive predictive value and the accuracy in predicting EDSS deterioration at final follow-up were always lower than 50% for average brain MTR at baseline and histogram peak height percentage changes after 1 year.

Discussion

Our study included both patients with RRMS or SPMS and those at presentation with CIS suggestive of multiple sclerosis. CIS patients had an increased risk of developing multiple sclerosis, considering that, at a retrospective analysis, all of them showed paraclinical evidence of spatial disease dissemination since the first clinical attack (McDonald *et al*., 2001). During the study period, 60% of them had a second clinical attack and were, therefore, diagnosed as having clinically definite multiple sclerosis, but a greater percentage of them (i.e. 80%) had MRI evidence of temporal disease dissemination at 1-year follow-up scans (data not shown). Therefore, the inclusion of CIS patients makes our sample representative of a cohort of multiple sclerosis patients at different disease stages. The relatively modest predictive value of disease duration on medium-term clinical deterioration seen in this patients cohort might be, at least partially, explained by the variable durations of the presymptomatic phase of the disease (Mastronardo *et al*., 1999). Actually, the relatively high MRI lesion burden of CIS patients might be a reflection of a longer disease duration than that elapsed between the first clinical episode and the time of study entry. Such an underestimation of disease duration might have occurred in ~30% of our patients and this might have reduced the prognostic value of disease duration on medium-term clinical evolution of the entire cohort. Interestingly, SPMS patients had a significantly increased likelihood of medium-term disability accumulation compared with those with CIS and RRMS, but the disease phenotype did not enter the final multivariable model as a predictor of EDSS deterioration. That this model included only average brain MTR percentage change and baseline T_2 TLV over one year indicates that MRI outcomes can provide prognostic information independently of the clinical characteristics of the multiple sclerosis patients. However, we are aware of the limitations of the EDSS scale for the detection of significant clinical worsening in multiple sclerosis patients. In the last few years, additional clinical scales, such as the multiple sclerosis Functional Composite Score (Rudick *et al*., 2001), have been proposed to improve the reliability and sensitivity to changes of clinical assessment. However, these scales were not available when the present study was initiated. As a consequence, their contribution to the definition of clinical outcomes of prognostic value for multiple sclerosis evolution needs to be explored by future studies.

About 60% of our patients started a disease-modifying therapy during the study and were being treated for an average period of 3 years; however, the presence and duration of treatment did not affect the final clinical outcome. Since the choice of treatment was made by the referring neurologists on an individual patient basis, the apparent lack of any treatment effect on multiple sclerosis evolution might greatly depend upon a selection bias. Untreated patients were most probably those with less disease activity and absent or mild clinical deterioration over the observation period. However, these results are consistent with those of several parallel group, placebo-controlled trials (IFNB Multiple Sclerosis Study Group, 1995; Johnson *et al*., 1995; Jacobs *et al*., 1996; PRISMS Study Group, 1998) and with those of a recent meta-analysis of interferon β trials in RRMS (Filippini *et al*., 2003), where the efficacy of immunomodulating treatments on the accumulation of multiple sclerosis-related disability was found to be absent or only modest.

Over a 1-year period, we observed a significant increase of T_2 TLV in CIS patients only, whereas changes in T_1 TLV were found to be significant in both CIS and RRMS patients. Conventional MRI-visible lesion load remained stable in patients with SPMS. On the contrary, average brain MTR changes after 1 year were significant in all disease phenotypes, and average lesion MTR showed a significant decrease in SPMS patients only. These findings are consistent with previous reports (Filippi *et al*., 2000; Inglese *et al*., 2003); they confirm that the sensitivity of T_2- and T_1-weighted MRI to multiple sclerosis-related changes is highest in the early phases of the disease and tends to decrease when patients enter the late, progressive stage. In this latter stage, measures reflecting the evolution of tissue damage of pre-existing T_2-visible lesions and NABT are more sensitive than T_2 lesion load accrual to multiple sclerosis changes which are likely to lead to the accumulation of irreversible disability. Admittedly, MRI results may have been partially influenced by the administration of interferon treatments, which are known to prevent the accumulation of T_2-visible lesion burden. However, this was the case only for about one third of our patients, i.e. for those who started interferon therapy during the first year of the study. In addition, this treatment does not seem to alter the evolution of brain MTR histogram-derived findings either in RRMS (Richert *et al*., 1998) or in SPMS (Inglese *et al*., 2003).

Although baseline NBV was significantly smaller in patients with more advanced (RR) or disabling (SP) forms of multiple sclerosis than in those with CIS, and correlated with patients' EDSS (*r* value: 0.40, data not shown), no significant differences between disease phenotypes were found in terms of PBVC after 1 year, nor there was any relationship between baseline NBV or 1-year PBVC and medium-term clinical worsening. Our results are consistent with those from a recent longitudinal study (Kalkers *et al*., 2002) of patients with heterogeneous multiple sclerosis phenotypes, where the rate of development of brain atrophy was found to be largely independent of the disease course and other clinical characteristics, including EDSS changes over time. On the other hand, a medium-term follow-up study of a large cohort of RRMS patients (Fisher *et al*., 2002) has shown that atrophy rate during an initial, 2-year observation period was the most significant MRI predictor of disability 6 years later. Although the assessment of brain atrophy is a relatively straightforward and robust approach, with limited intra-observer and inter-scanner variabilities (Miller *et al*., 2002), the use of different post-processing methodologies might, at least partially, explain the discrepant results. Our data

indicate that the measurement of an 'end-stage' phenomenon (i.e. brain parenchymal volume decrease) does not seem to be the optimal strategy to provide reliable prognostic information on the clinical evolution of multiple sclerosis over a medium-term period.

The most intriguing and novel part of this study is that a 1-year MRI follow-up can provide prognostic information on the medium-term clinical outcome of multiple sclerosis patients with varying disease phenotypes. The strongest predictive value can be achieved by combining T_2 TLV at study entry and observed average brain MTR changes after 1 year, which explain ~23% of the variance in EDSS worsening. However, we did not find any significant relationship between baseline MT MRI findings and patients' clinical evolution. Moreover, the final multivariable model included both baseline T_2 lesion load and short-term brain MTR changes as independent predictors of medium-term clinical worsening. The discrepancy between these results and those of Santos *et al*. (2002) are likely to be explained by the different sample size and clinical characteristics of the patients. Our group of patients is larger and more representative of multiple sclerosis at different stages, including patients at the earliest clinical phase of the disease. Our findings indicate that both the overall, MRI-visible disease burden and the subsequent accumulation of 'occult' brain tissue damage over a 1-year period significantly influence the medium-term disability worsening in multiple sclerosis patients. This also suggests that a 'multiparametric' MRI assessment of multiple sclerosis patients may represent a desirable approach to obtain reliable paraclinical outcomes of prognosis (Mainero *et al*., 2001). In this latter study, a model including T_1 lesion load, average brain tissue diffusivity and brain N-acetylaspartate/creatine ratio was found to explain ~50% of the observed EDSS variance in a cross-sectional evaluation of 23 multiple sclerosis patients. Changes of T_2 lesion load mainly reflect inflammatory-based multiple sclerosis activity, but are not sensitive to progressive tissue damage of chronic lesions and NABT. This might be the reason why the correlation between T_2 lesion load and clinical evolution is higher in the early stages of multiple sclerosis. On the other hand, metrics derived from whole brain MTR histograms reflect both the accumulation of MRI-visible disease burden and the severity of tissue disruption within established lesions and NABT (van Buchem *et al*., 1996; Rovaris and Filippi, 2003). The combination of these different information strengthens the correlation between MRI and clinical findings in patients with various clinical phenotypes of multiple sclerosis, since, at any stage of the disease, they provide a comprehensive *in vivo* estimation of both the 'inflammatory' and the 'neurodegenerative' components of the disease pathology (Martino *et al*., 2002).

At present, MRI is being widely used to monitor the evolution of multiple sclerosis damage both in natural history studies and in trials. On the contrary, the utility of MRI as an adjunctive, paraclinical tool for the management of individual multiple sclerosis patients is still debatable. We attempted to address this issue by assessing the predictive value of MT MRI findings for the clinical evolution of individual patients of this cohort. The accuracy of abnormal values of brain MTR histogram-derived metrics in predicting the medium-term clinical worsening of our patients was found to be suboptimal. Admittedly, this does not constitute a proper validation of MT MRI as a prognostic surrogate for multiple sclerosis clinical evolution, also considering that the high percentage of patients who did not present significant EDSS worsening at medium-term follow-up surely bears on the specificity of MT MRI findings by increasing the number of false positives. However, these findings support the notion that, whereas neither conventional nor MT MRI findings can be used as stand-alone predictors of multiple sclerosis evolution, only the integration of clinical and MRI data will probably enable us to define the prognosis of multiple sclerosis patients earlier and more accurately than clinical assessment alone would permit. Prospective, multiparametric MRI studies with longer follow-up periods and larger cohorts of patients than those of the present study and with clinical assessment based on more reliable scales than EDSS are warranted.

References

Brex PA, Ciccarelli O, O'Riordan JI, Sailer M, Thompson AJ, Miller DH. A longitudinal study of abnormalities on MRI and disability from multiple sclerosis. New Engl J Med 2002; 346: 158–64.

Filippi M, Grossman RI. MRI techniques to monitor multiple sclerosis evolution: the present and the future. Neurology 2002; 58: 1147–53.

Filippi M, Iannucci G, Tortorella C, Minicucci L, Horsfield MA, Colombo B, et al. Comparison of multiple sclerosis clinical phenotypes using conventional and magnetization transfer MRI. Neurology 1999a; 52: 588–94.

Filippi M, Inglese M, Rovaris M, Sormani MP, Horsfield MA, Iannucci G, et al. Magnetization transfer imaging to monitor the evolution of multiple sclerosis. A 1-year follow-up study. Neurology 2000; 55: 940–6.

Filippini G, Munari L, Incorvaia B, Ebers GC, Polman C, D'Amico R, et al. Interferons in relapsing remitting multiple sclerosis: a systematic review. Lancet 2003; 361: 545–52.

Fisher E, Rudick RA, Simon JH, Cutter G, Baier M, Lee JC, et al. Eight-year follow-up study of brain atrophy in patients with multiple sclerosis. Neurology 2002; 59: 1412–20.

IFNB Multiple Sclerosis Study Group and the University of British Columbia Multiple Sclerosis/MRI Analysis Group. Interferon β-1b in the treatment of multiple sclerosis: final outcome of the randomized, controlled trial. Neurology 1995; 45: 1277–85.

Inglese M, van Waesberghe JHTM, Rovaris M, Beckmann K, Barkhof F, Hahn D, et al. The effect of interferon β-1b on quantities derived from MT MRI in secondary progressive multiple sclerosis. Neurology 2003; 60: 853–60.

Jacobs DL, Cookfair DL, Rudick RA, Herndon RM, Richert JR, Salazar AM, et al. Intramuscular interferon beta-1a for disease

progression in relapsing multiple sclerosis. Ann Neurol 1996; 39: 285–94.

Johnson KP, Brooks BR, Cohen JA, Ford CC, Goldstein J, Lisak RP, et al. Copolymer 1 reduces relapse rate and improves disability in relapsing-remitting multiple sclerosis: results of a phase III multicenter, double- blind, placebo-controlled trial. Neurology 1995; 45: 1268–76.

Kalkers NF, Hintzen RQ, van Waesberghe JH, Lazeron RHC, van Schijndel RA, Ader HJ, et al. Magnetization transfer histogram parameters reflect all dimensions of multiple sclerosis pathology, including atrophy. J Neurol Sci 2001; 184: 155–62.

Kalkers NF, Ameziane N, Bot JCJ, Minneboo A, Polman CH, Barkhof F. Longitudinal brain volume measurement in multiple sclerosis. Rate of brain atrophy is independent of the disease subtype. Arch Neurol 2002; 59: 1572–6.

Kappos L, Moeri D, Radue EW, Schoetzau A, Schweikert K, Barkhof F, et al. Predictive value of gadolinium-enhanced magentic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. Lancet 1999; 353: 964–9.

Kurtzke JF. Rating neurologic in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983; 33: 1444–52.

Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology 1996; 46: 907–11.

Mainero C, De Stefano N, Iannucci G, Sormani MP, Guidi L, Federico A, et al. Correlates of multiple sclerosis disability assessed in vivo using aggregates of MR quantities. Neurology 2001; 56: 1331–4.

Martino G, Adorini L, Rieckmann P, Hillert J, Kallmann B, Comi G, et al. Inflammation in multiple sclerosis: the good, the bad, and the complex. Lancet Neurol 2002; 1: 499–509.

Mastronardo G, Rocca MA, Iannucci G, Pereira C, Filippi M. A longitudinal MR study of the presymptomatic phase in a patient with clinically definite multiple sclerosis. AJNR Am J Neuroradiol 1999; 20: 1268–72.

McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. Ann Neurol 2001; 50: 121–7.

Miller DH, Barkhof F, Berry I, Kappos L, Scotti G, Thompson AJ. Magnetic resonance imaging in monitoring the treatment of multiple sclerosis: concerted action guidelines. J Neurol Neurosurg Psychiatry 1991; 54: 683–8.

Miller DH, Barkhof F, Nauta JJP. Gadolinium enhancement increases the sensitivity of MRI in detecting disease activity in multiple sclerosis. Brain 1993; 116: 1077–94.

Miller DH, Barkhof F, Frank JA, Parker GJM, Thompson AJ. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain 2002; 125: 1676–95.

Molyneux PD, Barker GJ, Barkhof F, Beckmann K, Dahlke F, Filippi M, et al. Clinical-MRI correlations in a European trial of interferon beta-1b in secondary progressive multiple sclerosis. Neurology 2001; 57: 2191–7.

Ormerod IEC, Miller DH, McDonald WI, du Boulay EPGH, Rudge P, Kendall BE, et al. The role of NMR imaging in the assessment of multiple sclerosis and isolated neurological lesions: a quantitative study. Brain 1987;110: 1579–16.

Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983; 13: 227–31.

PRISMS (Prevention of Relapses and Disability by Interferon β-1a Subcutaneously in Multiple Sclerosis) Study Group. Randomised double-blind placebo-controlled study of interferon β-1a in relapsing/remitting multiple sclerosis. Lancet 1998; 352: 1498–504.

Richert ND, Ostuni JL, Bash CN, Duyn JH, McFarland HF, Frank JA. Serial whole-brain magnetization transfer imaging in patients with relapsing-remitting multiple sclerosis at baseline and during treatment with interferon β-1b. AJNR Am J Neuroradiol 1998; 19: 1705–13.

Rovaris M, Filippi M. Magnetization transfer imaging. In: Filippi M, Comi G, editors. New frontiers of MR-based techniques in multiple sclerosis. Milan: Springer-Verlag Italia; 2003. p. 11–32.

Rovaris M, Filippi M, Calori G, Rodegher M, Campi A, Colombo B, et al. Intra-observer reproducibility in measuring new putative MR markers of demyelination and axonal loss in multiple sclerosis: a comparison with conventional T_2-weighted images. J Neurol 1997; 244: 266–70.

Rovaris M, Filippi M, Falautano M, Minicucci L, Rocca MA, Martinelli V, et al. Relation between MR abnormalities and patterns of cognitive impairment in multiple sclerosis. Neurology 1998; 50: 1601–8.

Rovaris M, Comi G, Ladkani D, Wolinsky JS, Filippi M. Short-term correlations between clinical and MR imaging findings in relapsing-remitting multiple sclerosis. AJNR Am J Neuroradiol 2003; 24: 75–81.

Rudick RA, Cutter G, Baier M, Fisher E, Dougherty D, Weinstock-Guttman B, et al. Use of the multiple sclerosis functional composite to predict disability in relapsing multiple sclerosis. Neurology 2001; 56: 1324–30.

Santos AC, Narayanan S, de Stefano N, Tartaglia MC, Francis SJ, Arnaoutelis R, et al. Magnetization transfer can predict clinical evolution in patients with multiple sclerosis. J Neurol 2002; 249: 662–8.

Smith SM, De Stefano N, Jenkinson M, Matthews PM. Normalized accurate measurement of longitudinal brain change. J Comput Assist Tomogr 2001; 25: 466–75.

Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 2002; 17: 479–89.

Studholme C, Hill DLG, Hawkes DJ. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multi-resolution optimization of voxel similarity measures. Med Phys 1997; 24: 25–35.

vanBuchem MA, McGowan JC, Kolson DL, Polansky M, Grossman RI. Quantitative volumetric magnetization transfer

analysis in multiple sclerosis: estimation of macroscopic and microscopic disease burden. Magn Reson Med 1996; 36: 632–36.

vanBuchem MA, Grossman RI, Armstrong C, Polansky M, Miki Y, Heyning FH, et al. Correlation of volumetric magnetization transfer imaging with clinical data in multiple sclerosis. Neurology 1998; 50: 1609–17.

vanWaesberghe JH, Kamphorst W, De Groot C, van Walderveen MA, Castelijns A, Ravid R, et al. Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. Ann Neurol 1999; 46: 747–54.

Weinshenker BG, Rice GP, Noseworthy JH, Carriere W, Baskerville J, Ebers GC. The natural history of multiple sclerosis: a geographically based study. 3. Multivariate analysis of predictive factors and models of outcome. Brain 1991; 114: 1045–56.

Received February 14, 2003. Revised March 21, 2003.
Second revision March 28, 2003. Accepted May 12, 2003