White matter integrity improves in multiple sclerosis patients receiving natalizumab treatment

Oliver T Wiebenga, MD,[1,2] Menno M Schoonheim, MS,[2] Hanneke E Hulst, MS,[2] Gijsbert JA Nagtegaal, MS,†[1,2] Eva MM Strijbis, MD,[3] Martijn D Steenwijk, MS,[1] Chris H Polman, PhD,[3] Petra JW Pouwels, PhD,[4] Frederik Barkhof, PhD,[1] Jeroen JG Geurts, PhD[2]

Departments of Radiology and Nuclear Medicine[1], Anatomy & Neuroscience[2], Neurology[3], Physics & Medical Technology[4]

Neuroscience Campus Amsterdam and VU University Medical Center, Amsterdam, the Netherlands

Corresponding author: Oliver T Wiebenga, VU University Medical Center, Department of Radiology, PO Box 7057, 1007 MB, Amsterdam, The Netherlands

Email: o.wiebenga@vumc.nl

Abstract: (max 250)	249
Body (not including abstract, references, figure legends)	3.266
Number of figures:	3
Number of color figures	1
Number of tables:	4
# of references:	42

O.T. Wiebenga was sponsored by the Dutch MS Research Foundation, grant number 09-358d. MR imaging was sponsored by Biogen Idec.

ABSTRACT

Objective: To investigate the effect of natalizumab treatment on the extent and severity of white matter (WM) damage over time, compared to standard disease-modifying-treatment (DMT) in multiple sclerosis (MS).

Methods: The study included 22 MS patients starting natalizumab at baseline, 17 matched MS patients continuing standard DMT and 12 matched healthy controls. Imaging included diffusion tensor imaging (DTI) and conventional MRI sequences at baseline, month 6 and month 12. Tract-Based-Spatial-Statistics (TBSS) was performed at each time-point to compare diffusion measures between patients and controls, investigating the extent and severity of WM damage. Subjects also underwent comprehensive neuropsychological investigation.

Results: Natalizumab patients showed an extent of abnormalities in 56.8% of the WM voxels at baseline, which was reduced to 47.2% at month 12. Severity of FA-damage was expressed as a Z-score (compared to controls). At baseline, patients starting natalizumab showed reduced FA in the WM (Z= -0.67), which reduced over time (Z=-0.59 at month 12; p=0.04). In standard DMT patients, 41.4% of the WM voxels at baseline was abnormal in terms of FA, which was 39.1% at month 12. From baseline to month 12, severity of FA-damage did not change. In all patients, the severity of FA damage was correlated with overall cognitive performance (R =0.604, p=<0.001; at baseline).

Interpretation: Both extent and severity of WM abnormalities improve significantly under natalizumab treatment. No improvement of WM damage was

observed in the patients on standard DMT. This may have important implications for restriction of physical and cognitive impairment in the future.

INTRODUCTION

Multiple Sclerosis (MS) is an inflammatory demyelinating disease of the central nervous system, characterized by focal damage and atrophy of the white[1] (WM) and grey matter[2] (GM). Physical and cognitive dysfunction start early on in the disease[3], and strongly impact quality of life.[4] Conventional MRI measures, such as WM lesion loads, show only modest correlations with patient functioning and progression. Recent advanced magnetic resonance imaging (MRI) such as diffusion tensor imaging (DTI) have shown better correlations with clinical outcome measures, partly due to the demonstration of subtle abnormalities in the normal appearing white matter (NAWM).[5-7]

Natalizumab is a relatively new disease modifying treatment option. It has been shown to have a strong anti-inflammatory effect, impacting the formation of new WM lesions.[8, 9] Additionally, it has also been shown to affect atrophy[10, 11] and clinical measures such as the number of relapses and progression of disability.[8, 9] The large clinical effect of natalizumab can only partly be explained by its effect on lesion formation, given the poor relationship between lesion load and EDSS scores.[12] Its effect might rather be understood in terms of restriction of more subtle damage in the (NA)WM, as measured with more advanced MRI techniques.

This study investigated the effect of natalizumab on the extent and severity of WM damage as measured with DTI, as well as relations with patient functioning.

A single center, prospective and observational design was used, following patients starting with natalizumab treatment for twelve months over three visits. Patients were compared with patients on standard disease-modifying-treatment (DMT) and healthy controls.

MATERIAL AND METHODS

Study design The study is a prospective, single-centre, observational, longitudinal healthy control and patient controlled study with three time-points; baseline, month 6 and month 12 (Figure 1). The study population of consisted of three groups: 22 relapsing-remitting MS (RRMS) patients starting natalizumab treatment, 17 RRMS patients following and continuing standard DMT (interferons or glatiramer acetate) and 12 healthy controls. Subjects continuing standard DMT were matched according to the levels of patients starting natalizumab for age, sex, disability (Expanded Disability Status Scale - EDSS[13]) and duration of current standard DMT therapy. The healthy controls were age- and sex-matched according to the levels of the patients starting natalizumab at baseline.

Inclusion criteria of the MS patients consisted of a diagnosis of clinical definite MS[14] and an age between 18 and 65 years. All subjects were anamnestically screened for exclusion criteria, i.e., presence or history of psychiatric or neurological disease (for patients: neurological disease other than MS), presence of contra-indications for MRI and presence or history of alcohol or drug abuse. The study protocol was approved by the institutional ethics review board of the VU University Medical Center and informed consent was obtained from the participants before their first examination.

No serious or unanticipated adverse events attributed to MS-medication developed in the patients groups.

Natalizumab patients started natalizumab treatment (300 mg i.v. once every four weeks) close to the baseline visit. Patients were included right from natalizumab treatment onset if possible. At the baseline measurement, six patients were completely treatment-naïve for natalizumab, 13 patients had received one infusion, and three patients had received two infusions. Patients from this group were recruited and treated in the outpatient clinic of the department of Neurology of the VU University Medical Center Amsterdam. Patients that were in the screening phase (approximately two weeks) for eligibility of natalizumab were approached for participation in the study. Patients were only approached once the decision to start natalizumab treatment was already made. The treating neurologist could withdraw patients from treatment if they did not show evidence of therapeutic benefit. Patients starting natalizumab were screened according to the inclusion criteria used at our centre, including at least one prior periods of standard DMT with break-through disease or rapidly evolving active RRMS. All patients continued natalizumab treatment for the duration of the study.

Standard DMT patients were already receiving standard DMT as their regular medical treatment, and continued this treatment at baseline, where 11 patients were continuing interferons (IFN-ß-1a/b - dose and route of administration dependent on type) and 6 patients were continuing glatiramer acetate (20mg s.c. once daily). After the baseline visit, one patient from this group discontinued standard DMT because of radiologically and clinically stable disease in

combination with suffered side-effects (necrosis and scarring of skin at injection sites). Another patient discontinued standard DMT one month before the month 12 visit due to conversion to SPMS. Both patients remained in the study.

MR imaging

All imaging was performed on a 1.5T whole-body scanner (Siemens Sonata, Erlangen, Germany) using an eight-channel phased-array head coil. Structural imaging sequences included a 3D-T1 weighted magnetization prepared rapid acquisition gradient-echo (MPRAGE) sequence (TR 2.700ms, TE 5ms, TI 950ms; 176 sagittal slices with 1.3 mm section thickness, 1.3x1.3 mm^2 in-plane resolution) for brain volume measurements and an axial turbo spin-echo proton density PD/T2-weighted sequence (TR 3.130 ms, TE 24 and 85ms, 46 contiguous 3 mm slices, 1x1 mm^2 in-plane) for WM lesion detection. Diffusion-weighted echo-planar images (TR 8.500 ms, TE 86 ms and isotropic resolution, 2x2x2 mm) were acquired with 60 volumes with noncollinear diffusion gradients (b value of 700 sec/mm2) and 10 volumes without directional weighting.

Brain and lesion volumes T2-hyperintense WM lesions were quantified using an automated segmentation method.[15] Normalized whole brain (NBV), WM (NWMV) and GM (NWGMV) volumes were calculated using the 3DT1 MPRAGE images and SIENAX[16] (part of FSL, version 5.02 www.fsl.fmrib.ox.ac.uk) and corrected for head size. Brain volumes were calculated after lesion filling, using

an automated lesion-filling technique (LEAP; Lesion Automated Preprocessing).[17]

Extent and severity of WM integrity damage

Fractional Anisotropy (FA) images were created using FSL, by fitting a tensor model to the raw diffusion data after motion- and eddy current correction. Fractional anisotropy (FA) and mean diffusivity (MD) were derived for each voxel. All subject's FA and MD data were then aligned into a common space. Next, the mean FA image was created and thresholded at 0.2 to create a mean FA skeleton. Each subject's aligned FA and MD data was then projected onto this skeleton and the resulting data were fed into tract-based spatial statistics (TBSS).[18]

At each time point, cross-sectional group differences in FA and MD of the mean WM skeleton were analyzed using randomize (500 permutations) using an FWE-corrected threshold of $p < 0.05$ and correcting for age and sex.

Following the TBSS pipeline, extent and severity of damage was calculated as follows:

1. The *extent* of damage was calculated per group, by calculating the percentage of significantly abnormal ($p < 0.05$) voxels within the WM skeleton for the diffusion parameters FA and MD, compared to the healthy controls.

2. The *severity* of damage was calculated individually, by converting the diffusion measures FA and MD to Z scores, based on the mean and standard deviations of healthy control voxels. This was done for each voxel by subtracting the mean value for that voxel of the control group, and dividing this by the standard deviation for that voxel of the control group. A single whole-skeleton mean

Z-score was calculated for the diffusion parameters FA and MD, indicating the severity of damage across the entire WM skeleton per subject.

Severities of axial- (AD) and radial diffusivity (RD) were also calculated at baseline and month 12.

Neuropsychological evaluation

All subjects underwent an elaborate neuropsychological assessment on the day of scanning. The cognitive domains most frequently affected in MS were investigated using certain tests from Rao's Brief Repeatable Battery for Neurological disease (BRB-N)[19], such as the Symbol Digit Modalities Test for information processing speed, the Spatial Recall Task (SPART 10/36) for visuospatial memory, and the World List Generation (WLG) test for verbal fluency. The PASAT (Paced Auditory Serial Addition Test) was excluded because of learning effects.[20] Additionally, the Verbal Learning and Memory Task (VLGT; Dutch equivalent of the California Verbal Learning Test) for verbal memory, the Stroop color-word test for attention and inhibition, the digit span forward plus backward for working memory and the Delis-Kaplan Executive Function System Trail Making Test (TMT) for executive functioning were administered. To reduce training effects in this longitudinal study, parallel versions were used for the different subtests where available (Symbol Digit Modalities Test, SPART 10/36, VLGT).

The (sub)test raw baseline scores were converted into Z-scores to uniformly quantify the deviation compared to (sub)test scores of the healthy controls. An average (overall) cognition Z-score was also calculated.

To correct for subtle learning effects in the longitudinal analysis the raw (sub-)test scores were converted to Reliable Change Indices (RCI). With this method, the significance of the change on an individual test score is based on the difference between baseline and retest scores for the normative subject sample.[21] The following formula: ((X2-X1) – (mean[hc2] – mean[hc1])) / SED was used. Where X is the test score of a cognitive test of one subject and mean[hc] was the mean of the healthy controls, whereas 1 and 2 were the different time points on which cognitive testing was performed. The standard deviation of the mean delta score of the healthy controls was used as the standard error of the difference score (SED).

An RCI score above zero indicates a performance that improves more than can be expected by the normal learning curve of healthy controls, whereas an RCI score below zero indicates a performance worse than the learning effect of HC.

Questionnaires

Fatigue and symptoms of anxiety and depression were measured using the Checklist Individual Strength (CIS)[22] and Hospital -Anxiety and -Depression Scale (HADS-A and -D)[23] questionnaires, respectively.

Statistical analysis

Statistical analyses were performed with SPSS for Windows version 20.0. When the variables were normally distributed, a multivariate GLM was used with age, sex and education included as covariates. Longitudinal analyses were performed using paired t-tests comparing baseline with month 6 and baseline with month 12, Non-parametric analysis was done using Kruskall-Wallis and (post-hoc) Mann-Whitney. All analyses were Bonferroni-corrected and $p<0.05$ was considered statistically significant.

RESULTS

Baseline descriptives

As shown in table 1, the three groups did not differ significantly on age and sex. Compared to the healthy controls, MS patients had more depressive symptoms (natalizumab p =0.004, standard DMT p=0.007) and fatigue (natalizumab p=0.006, standard DMT p=0.001) at baseline and also a slightly lower level of education (natalizumab p =0.008, standard DMT p=0.014). There were no differences between the MS groups for any of these measures, nor for EDSS, disease duration and duration of prior standard DMT.

Clinical disability

The median EDSS of the natalizumab patients remained stable at 3.0 for all time-points (p=1.0, see table 2). The depression score (HADS-D), however, was significantly reduced in this group: from 4.5 at baseline to 2.0 at month 12 (p=0.024).

The median EDSS of the patients continuing standard DMT significantly increased from 2.5 at baseline to 3.0 at month 12 (p=0.012). The depression score did not significantly change in this group (see table 1, p=0.69).

Cognition

At baseline both patient groups had lower Z-scores for the Symbol Digit Modalities Test (natalizumab p=0.001, standard DMT p<0.001) and the World List Generation Professions (natalizumab p=0.003, standard DMT p=0.002)

(Table 1) compared to healthy controls. The average cognition Z-score was significantly lower for both patient groups (natalizumab p<0.001, standard DMT p=0.005) compared to the healthy controls.

No significant cognitive changes were seen over time for any of the groups (see Table 4), apart from the immediate recall of the VLGT, where only patients with natalizumab declined over time (p=0.024), compared to a relatively stable score in patients with standard DMT.

Volumetric MRI measures

At baseline, the patient groups did not differ regarding T2 lesion volumes (see table 1, p=0.281). Over the period of 12 months, the median T2-lesion volume remained stable in the natalizumab patients (p= 1.0). In the patients continuing standard DMT the T2-lesion volume significantly increased from a median of 4.9 ml at baseline to 5.4 ml at month 12 (p=0.024).

Both patient groups had lower NWMV compared to healthy controls (natalizumab p=0.044, Standard DMT p=0.024) at baseline. Only the standard DMT patients showed lower NBV compared to controls (p=0.022). No significant differences were found between the patient groups at baseline for any brain volume metric. Over 12 months, a reduction in NWMV was seen in both patient groups (natalizumab p=0.024, Standard DMT p=0.012), while NBV was only reduced in patients on Standard DMT (p<0.001).

DTI: *extent* of WM damage

Natalizumab patients had reduced FA in 56.8% of the investigated WM skeleton voxels at baseline, including major WM bundles such as the corpus callosum, the capsula externa, forceps major, corticospinal tract, corona radiata, superior longitudinal fasciculus among other regions (see Figure 2a). At month 6, this number remained stable at 55.6%, but was reduced to 47.2% at month 12. Improvements of FA at month 12 were seen throughout the WM skeleton.

Mean diffusivity (MD) values fluctuated around a stable point over the three time points (54.3% to 48.1% to 55.7% at baseline, month 6 and month 12, respectively; see Figure 2b.) Increased MD was generally seen in WM where the FA was also decreased.

Patients continuing standard DMT had reduced FA in 41.4% of the skeleton at baseline, and 47.2% and 39.1% at months 6 and 12. Affected areas overlapped with the regions affected in the natalizumab group, but were less extensively involved. Increased MD was seen in 33.7%, 32.0% and 36.1% of the skeleton at baseline, month 6 and month 12, respectively.

DTI: *severity* of WM damage

In the natalizumab patients, the severity of FA-damage at baseline was Z= -0.67, and remained stable at month six at Z= -0.68. At month 12, however, the severity was significantly reduced to Z= -0.59 (p=0.04 compared to baseline, see Table 3). In the patients continuing standard DMT the severity remained stable at all time points, ranging from -0.64 to -0.72 to -0.67, with no significant change over time. There was no significant difference in severity of FA-damage between the

patient groups (p=1.0). The severity of MD abnormalities is shown in Table 3; no significant change over time was seen for this diffusivity measure in any group.

No significant changes in AD- and RD- severity of WM damage was found in any group between baseline and month 12. The severity of AD-damage in natalizumab patients was 0.36 at baseline and 0.43 at month 12 (p=0.68), and for the standard DMT patients 0.26 at both time points (p=1.0). The severity of RD-damage was 0.94 at both time points (p=1.0) where it was 0.82 at baseline and 0.87 at month 12 (p=0.85) in Standard DMT patients.

Relating *severity* of WM damage to cognition

In the patient group, a significant correlation was found between average cognition and the FA severity Z-score ($r = 0.604$, $p<0.001$ at baseline and $r= 0.525$, $p<0.001$ at month 12; see Figure 3), indicating that patients with more severe WM damage have more cognitive impairment.

DISCUSSION

In this study, the clinical and radiological evolution of MS patients starting natalizumab treatment was monitored over a period of 12 months and was compared to MS patients continuing standard DMT and to healthy controls. At baseline, patients starting natalizumab showed a higher extent of FA damage in WM, as measured with DTI than standard DMT patients (56.8 vs 41.4%). However there was no baseline difference in the severity of damage, nor in lesion volume or EDSS.

Lesion volumes remain constant in natalizumab-treated patients, while both extent and severity of WM damage both improve over time. Depressive symptoms improved in natalizumab patients only. Lesion volumes, and EDSS scores, both significantly worsened in standard DMT patients while cognition remained largely stable in both groups. Decrease of NWMV was seen in both patient groups over 12 months.

Previous in-vivo studies have shown FA reductions in focal lesions in MS, as well as in the NAWM.[6, 24] A post-mortem histopathological DTI study of MS brain slices has confirmed the strong correlation of FA with myelin as well as axons in both WM lesions and NAWM.[25] Several other studies have shown clinically relevant abnormalities in the NAWM in MS patients, using DTI[26] as well as magnetization transfer ratio (MTR)[27] and MR spectroscopy.[27] In our study, lesion volumes remain constant in natalizumab-treated patients, while the extent and severity of WM damage decreases over time.

The strong anti-inflammatory effect of natalizumab[28] may lead to a more advantageous milieu for axonal repair and remyelination. Looking at specific diffusivities potentially pinpoints signs of improved axonal integrity (a reduction in axial diffusivity, AD) [29, 30] versus remyelination (a reduction in radial diffusivity, RD).[31, 32] Unfortunately we were unable to find significant changes in AD and RD, indicating that subtle changes in either of these parameters probably led to measurable changes in FA, but were too small to pick up individually in the current sample. Future studies are required to investigate to which extent changes were within lesional tissue and/or the NAWM as both were included in our TBSS analysis. It would also be very interesting to monitor diffusivity changes of individual lesions under standard DMT and natalizumab treatment over time.

Previous studies have shown the strong clinical effect of natalizumab, as in a reduced risk of sustained progression was notably found after two years of treatment in the two registration studies.[8, 9] In our study, EDSS scores in the natalizumab group remained stable, whereas depressive symptoms improved. The clinical changes could be the direct result of the observed reduction in DTI abnormalities, as for example a previous study found a link between DTI-abnormalities and depression,[33] but needs further investigation in future studies.

Our study did not show direct effect of natalizumab on cognitive performance, as both patient groups remained largely stable. There are few studies that have investigated the effect of natalizumab on cognitive performance. Those studies that have looked at this relation, and suggest an improvement in cognition in a short time frame, are hampered by a restricted sample size, lack of adequate

control groups, and most did not include imaging parameters.[34-40] Both patient groups generally displayed normal learning curves in their cognitive scores over a period of twelve months, comparable to those of the healthy controls (i.e. an RCI around 0, see Table 4). We did not find significant improvements of any cognitive measure in any patient group. In fact, the natalizumab group significantly worsened on verbal memory, perhaps due to their more severe disease status to begin with. If we were to look at raw cognitive scores of natalizumab-treated patients alone they would show a pseudo-increase in cognitive performance over time, stressing the need for control scores for future longitudinal studies looking at cognition in MS.

Despite the remarkable and significant effect of natalizumab on FA-values after 12 months of treatment, we did not see an effect on cognitive functioning. With longer follow-up times it could be expected that natalizumab-induced improvements and/or stabilizations in brain tissue could protect cognitive function in these patients. This is underlined by the strong correlation between the severity of FA abnormalities and overall cognition found in our data, as well as in other previous studies,[5, 7, 41] and the relatively slow rate of cognitive decline as seen in other studies.[42]

In conclusion, this non-randomised observational study showed a reduction in both extent and severity of WM damage in MS patients after twelve months of natalizumab treatment. Additionally, a stabilization of lesion load and physical disability was seen in these patients, combined with improvements on depressive symptoms. In contrast, standard DMT patients did not show reduced WM

damage, had an increase in both lesion volume and physical disability. Future studies with longer follow-up times are now required to investigate possible effects of natalizumab on cognition, given the strong correlation of DTI metrics and cognitive dysfunction in MS.

Reference List

1. Compston,A. & Coles,A. Multiple sclerosis. *Lancet* **372**, 1502-1517 (2008).

2. Geurts,J.J., Calabrese,M., Fisher,E., & Rudick,R.A. Measurement and clinical effect of grey matter pathology in multiple sclerosis. *Lancet Neurol.* **11**, 1082-1092 (2012).

3. Schulz,D., Kopp,B., Kunkel,A., & Faiss,J.H. Cognition in the early stage of multiple sclerosis. *Journal of Neurology* **253**, 1002-1010 (2006).

4. Chiaravalloti,N.D. & DeLuca,J. Cognitive impairment in multiple sclerosis. *Lancet Neurology* **7**, 1139-1151 (2008).

5. Hulst,H.E. *et al.* Cognitive impairment in MS: impact of white matter integrity, gray matter volume, and lesions. *Neurology* **80**, 1025-1032 (2013).

6. Roosendaal,S.D. *et al.* Regional DTI differences in multiple sclerosis patients. *Neuroimage.* **44**, 1397-1403 (2009).

7. Schoonheim,M.M. *et al.* Sex-specific extent and severity of white matter damage in multiple sclerosis: Implications for cognitive decline. *Hum. Brain Mapp.*(2013).

8. Polman,C.H. *et al.* A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *New England Journal of Medicine* **354**, 899-910 (2006).

9. Rudick,R.A. *et al.* Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. *New England Journal of Medicine* **354**, 911-923 (2006).

10. Miller,D.H. *et al.* MRI outcomes in a placebo-controlled trial of natalizumab in relapsing MS. *Neurology* **68**, 1390-1401 (2007).

11. Radue,E.W. *et al.* Natalizumab plus interferon beta-1a reduces lesion formation in relapsing multiple sclerosis. *J. Neurol. Sci.* **292**, 28-35 (2010).

12. Barkhof,F. The clinico-radiological paradox in multiple sclerosis revisited. *Curr. Opin. Neurol.* **15**, 239-245 (2002).

13. Kurtzke,J.F. Rating Neurologic Impairment in Multiple-Sclerosis - An Expanded Disability Status Scale (Edss). *Neurology* **33**, 1444-1452 (1983).

14. Polman,C.H. *et al.* Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann. Neurol.* **58**, 840-846 (2005).

15. Steenwijk,M.D. *et al.* Accurate white matter lesion segmentation by k nearest neighbor classification with tissue type priors (kNN-TTPs). *NeuroImage: Clinical* **3**, 462-469 (2013).

16. Smith,S.M. *et al.* Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage.* **17**, 479-489 (2002).

17. Chard,D.T., Jackson,J.S., Miller,D.H., & Wheeler-Kingshott,C.A. Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *J. Magn Reson. Imaging* **32**, 223-228 (2010).

18. Smith,S.M. *et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage.* **31**, 1487-1505 (2006).

19. Rao,S.M., Leo,G.J., Bernardin,L., & Unverzagt,F. Cognitive Dysfunction in Multiple-Sclerosis .1. Frequency, Patterns, and Prediction. *Neurology* **41**, 685-691 (1991).

20. Bever,C.T., Jr., Grattan,L., Panitch,H.S., & Johnson,K.P. The Brief Repeatable Battery of Neuropsychological Tests for Multiple Sclerosis: a preliminary serial study. *Mult. Scler.* **1**, 165-169 (1995).

21. Walker,L.A.S., Mendella,P.D., Stewart,A., Freedman,M.S., & Smith,A.M. Meaningful change in cognition in multiple sclerosis: method matters. *Can. J. Neurol. Sci.* **38**, 282-288 (2011).

22. Vercoulen,J.H. *et al.* Dimensional assessment of chronic fatigue syndrome. *J. Psychosom. Res.* **38**, 383-392 (1994).

23. Zigmond,A.S. & Snaith,R.P. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* **67**, 361-370 (1983).

24. Bammer,R. *et al.* Magnetic resonance diffusion tensor imaging for characterizing diffuse and focal white matter abnormalities in multiple sclerosis. *Magn Reson. Med.* **44**, 583-591 (2000).

25. Schmierer,K. *et al.* Diffusion tensor imaging of post mortem multiple sclerosis brain. *Neuroimage.* **35**, 467-477 (2007).

26. Liu,Y. *et al.* Whole brain white matter changes revealed by multiple diffusion metrics in multiple sclerosis: a TBSS study. *Eur. J. Radiol.* **81**, 2826-2832 (2012).

27. Bellmann-Strobl,J. *et al.* MR spectroscopy (MRS) and magnetisation transfer imaging (MTI), lesion load and clinical scores in early relapsing remitting multiple sclerosis: a combined cross-sectional and longitudinal study. *Eur. Radiol.* **19**, 2066-2074 (2009).

28. Yednock,T.A. *et al.* Prevention of Experimental Autoimmune Encephalomyelitis by Antibodies Against Alpha-4-Beta-1 Integrin. *Nature* **356**, 63-66 (1992).

29. Budde,M.D., Xie,M., Cross,A.H., & Song,S.K. Axial diffusivity is the primary correlate of axonal injury in the experimental autoimmune encephalomyelitis spinal cord: a quantitative pixelwise analysis. *J. Neurosci.* **29**, 2805-2813 (2009).

30. Budde,M.D. *et al.* Toward accurate diagnosis of white matter pathology using diffusion tensor imaging. *Magn Reson. Med.* **57**, 688-695 (2007).

31. Song,S.K. *et al.* Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage.* **17**, 1429-1436 (2002).

32. Song,S.K. *et al.* Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage.* **26**, 132-140 (2005).

33. Akbar,N. *et al.* Diffusion tensor imaging abnormalities in cognitively impaired multiple sclerosis patients. *Can. J. Neurol. Sci.* **37**, 608-614 (2010).

34. Iaffaldano,P. *et al.* Impact of natalizumab on cognitive performances and fatigue in relapsing multiple sclerosis: a prospective, open-label, two years observational study. *PLoS. One.* **7**, e35843 (2012).

35. Lang,C., Reiss,C., & Maurer,M. Natalizumab may improve cognition and mood in multiple sclerosis. *Eur. Neurol.* **67**, 162-166 (2012).

36. Mattioli,F., Stampatori,C., Bellomi,F., & Capra,R. Natalizumab efficacy on cognitive impairment in MS. *Neurol. Sci.* **31 Suppl 3**, 321-323 (2011).

37. Mattioli,F., Stampatori,C., & Capra,R. The effect of natalizumab on cognitive function in patients with relapsing-remitting multiple sclerosis: preliminary results of a 1-year follow-up study. *Neurol. Sci.* **32**, 83-88 (2011).

38. Portaccio,E. *et al.* Natalizumab may reduce cognitive changes and brain atrophy rate in relapsing-remitting multiple sclerosis: a prospective, non-randomized pilot study. *Eur. J. Neurol.* **20**, 986-990 (2013).

39. Stephenson,J.J. *et al.* Impact of natalizumab on patient-reported outcomes in multiple sclerosis: a longitudinal study. *Health Qual. Life Outcomes.* **10**, 155 (2012).

40. Svenningsson,A. *et al.* Natalizumab treatment reduces fatigue in multiple sclerosis. Results from the TYNERGY trial; a study in the real life setting. *PLoS. One.* **8**, e58643 (2013).

41. Yu,H.J. *et al.* Multiple white matter tract abnormalities underlie cognitive impairment in RRMS. *Neuroimage.* **59**, 3713-3722 (2012).

42. Weinstein,A. *et al.* Neuropsychologic status in multiple sclerosis after treatment with glatiramer. *Arch. Neurol.* **56**, 319-324 (1999).

TABLES

Table 1. Baseline variables of MS patients and controls (mean ± SD).

	natalizumab patients (n=22)	**standard DMT patients (n=17)**	**healthy controls (n=12)**	**p-value**
MS therapy	Starting natalizumab at baseline	Continuing standard DMT	n/a	
Descriptives				
Age (years)	37.2 ± 8.8	38.2 ± 5.0	35.1 ± 5.3	0.492
Sex (male/female) [a]	9/13	8/9	3/9	0.322
Education [b]	6.0 (4-7)	6.0 (5-7)	7.0 (5-7)	0.016 [d]
HADS-A [b]	6.0 (1-13)	6.6 (1-14)	4.8 (1-12)	0.369
HADS-D [b]	4.5 (0-18)	5.0 (0-12)	1.8 (0-10)	0.008 [d]
CIS-20 [b]	68.5 (14-125)	86 (31-114)	42.5 (17-85)	0.002 [d]
MS characteristics				
EDSS [ab]	3.0 (1.5-6.5)	2.5 (1.0-6.5)	n/a	0.615
Disease duration since onset (y)	8.3 ± 6.2	9.1 ± 5.2	n/a	0.662
Prior standard DMT	2.9 ± 3.1	4.5 ± 4.0	n/a	0.169

duration at baseline (y)			
Volumes			
NGMV (L)	0.75 ± 0.04	0.73 ± 0.06	0.77 ± 0.04	0.134
NWMV (L)	0.69 ± 0.04	0.69 ± 0.04	0.73 ± 0.03	0.020 [d]
NBV (L)	1.44 ± 0.06	1.42 ± 0.08	1.50 ± 0.06	0.025 [d]
T2 lesion volume (ml) [c]	6.2	(2.4-14.9)	4.9	(2.5-12.0)	n/a	0.281
Cognition Z scores			
Symbol Digit Modalities Test	-2.17 ± 0.92	-1.71 ± 1.13	0.00 ± 1.00	<0.001 [d]
SPART-total recall	-1.90 ± 1.94	-0.82 ± 1.83	0.00 ± 1.00	0.021 [f]
SPART-delayed recall	-1.42 ± 1.57	-0.80 ± 1.93	0.00 ± 1.00	0.076
VLGT - immediate recall	-1.33 ± 1.84	-1.66 ± 1.53	0.00 ± 1.00	0.052
VLGT - short term free recall	-1.66 ± 1.89	-1.40 ± 1.64	0.00 ± 1.00	0.068
VLGT - long term free recall	-1.65 ± 1.86	-1.48 ± 1.72	0.00 ± 1.00	0.070
WLG - Animals	-0.97 ± 0.73	-0.75 ± 0.81	0.00 ± 1.00	0.019 [f]
WLG - Professions	-1.69 ± 1.15	-1.79 ± 1.58	0.00 ± 1.00	0.001 [d]

WLG - 4 letter "M" words	-0.81 ± 1.0	-0.85 ± 0.94	0.00 ± 1.00	0.072
Digit Span – forward	-0.62 ± 1.06	-0.29 ± 0.98	0.00 ± 1.00	0.670
Digit Span – backward	-0.96 ± 1.13	-0.05 ± 1.36	0.00 ± 1.00	0.067
TMT - Letter Number Switching	-2.33 ± 2.45	-1.58 ± 2.08	0.00 ± 1.00	0.059
Stroop – interference	-0.28 ± 1.44	-0.63 ± 1.25	0.00 ± 1.00	0.697
Average Cognition	-1.61 ± 0.73	-1.22 ± 1.01	0.00 ± 1.00	<0.001 [d]

Abbreviations: EDSS = Expanded Disability Status Scale, Standard DMT = Standard Disease Modifying-therapy (interferons or glatiramer acetate), HADS = Hospital Anxiety and Depression Scale, A = Anxiety, D = Depression, CIS-20 = Checklist for Individual Strength questionnaire, y = years, NBV = Normalized Brain Volume, NGMV = normalized Grey Matter Volume, NWMV = Normalized White Matter Volume, SPART = Spatial Recall Task, VLGT = Verbal Memory and Learning Task, WLG = World List Generation, TMT = Trail Making Test

When normally distributed a multivariate GLM was used with age, sex and education included as covariates, GLM main effect p-values are shown. Non-parametric testing was performed using Kruskall Wallis and post-hoc Mann-Whitney. P-values < 0.05 are considered significant.

[a] Chi-Square test, [b] indicates median and range, [c] indicates median and interquartile range, [d] significant in both patient groups compared to healthy controls, [e] only significant between Standard DMT patients and healthy controls, [f] only significant between natalizumab patients and healthy controls

Table 2. Clinical scales of baseline and follow-up (median and range)

Time point	Baseline		Month 6		Month 12		p-value baseline-month 6	p-value baseline-month 12
natalizumab patients (n=22)								
EDSS	3.0	(1.5-6.5)	3.0	(1.5-6.5)	3.0	(1.0-6.5)	1.0	1.0
HADS-A	6.0	(1-13)	5.0	(0-15)	6.0	(0-15)	1.0	1.0
HADS-D	4.5	(0-18)	3.5	(0-20)	2.0	(0-17)	1.0	0.024
CIS-20	68.5	(14-125)	64.0	(9-125)	62	(10-117)	1.0	0.072
standard DMT patients (n=17)								
EDSS	2.5	(1.0-6.5)	3.0	(1.0-6.5)	3.0	(1.5-7.0)	0.640	0.012
HADS-A	7.0	(1-14)	6.0	(1-11)	5.0	(1-11)	1.0	1.0
HADS-D	5.0	(0-12)	4.0	(1-10)	3.0	(0-14)	0.690	1.0
CIS-20	86.0	(31-114)	85.0	(28-102)	81.0	(16-105)	0.060	0.354
healthy controls (n=12)								
HADS-A	3.5	(1-12)	3.5	(0-16)	2.5	(0-9)	1.0	1.0
HADS-D	1.0	(0-10)	0.0	(0-9)	0.0	(0-5)	1.0	1.0
CIS-20	42.5	(17-85)	36.0	(16-100)	52.0	(17-82)	1.0	1.0

Abbreviations: EDSS = Expanded Disability Status Scale, standard DMT = standard Disease Modifying-therapy (interferons or glatiramer acetate), HADS = Hospital Anxiety and Depression Scale, A = Anxiety, D = Depression, CIS-20 = Checklist for Individual Strength questionnaire

EDSS, HADS and CIS were tested with the related samples wilcoxon signed rank test Bonferroni-corrected

Table 2. A decrease in HADS-D scores in patients treated with natalizumab and an increase in EDSS in standard DMT patients between baseline and month 12 is shown.

Table 3[O.T.1]**. DTI and volumetric variables of baseline and follow-up (mean ± SD)**

	Baseline	**Month 6**	**Month 12**	p base -m6
natalizumab patients (n=22)				
FA-severity (Z-score)	-0.67	-0.68	-0.59	1.0
MD-severity (Z-score)	0.89	0.84	0.93	1.0
NGMV (L)	0.75 ± 0.04	0.75 ± 0.05	0.75 ± 0.05	1.0
NWMV (L)	0.69 ± 0.04	0.70 ± 0.04	0.67 ± 0.03	1.0
NBV (L)	1.44 ± 0.06	1.45 ± 0.06	1.42 ± 0.06	1.0
T2 lesion vol. (ml)[a]	6.2 (2.4-14.9)	6.9 (2.2-13.6)	6.9 (2.5-14.9)	0.132
standard DMT patients (n=17)				

FA-severity (Z-score)	-0.64		-0.72		-0.67		1.0
MD-severity (Z-score)	0.74		0.72		0.78		1.0
NGMV (L)	0.73	± 0.06	0.73	± 0.05	0.74	± 0.05	0.318
NWMV (L)	0.69	± 0.04	0.70	± 0.03	0.67	± 0.03	1.0
NBV (L)	1.42	± 0.08	1.42	± 0.06	1.41	± 0.07	1.0
T2 lesion vol. (ml)[a]	4.9	(0.25-12.0)	5.1	(0.26-11.9)	5.4	(0.28-13.3)	1.0
healthy controls (n=12)							
NGMV (L)	0.77	± 0.04	0.77	± 0.04	0.78	± 0.04	1.0
NWMV (L)	0.73	± 0.03	0.73	± 0.03	0.71	± 0.05	1.0
NBV (L)	1.50	± 0.06	1.50	± 0.06	1.49	± 0.07	1.0

Abbreviations: FA = Fractional Anisotropy, MD = Mean Diffusivity, NGMV = normalized Grey Matter Volume, NAWM = Normalized White Matter Volume, NBV = Normalized Brain Volume,

[a] indicates median and interquartile range, *significant; Bonferroni-corrected

Table 4. Reliable Change Index (RCI) (mean ± sd) of natalizumab patients and standard DMT patients between baseline and month 12

RCI Baseline - Month 12	natalizumab patients	standard DMT patients	p-value
Symbol Digit Modalities Test	-0.02 ± 0.68	-0.04 ± 0.77	0.993
SPART-total	0.18 ± 1.46	-0.49 ± 1.42	0.432
SPART-delayed	0.54 ± 1.69	0.06 ± 1.99	0.629
VLGT - immediate recall	-0.67 ± 0.88	0.20 ± 1.03	0.024*
VLGT- short term free recall	-0.15 ± 1.01	0.06 ± 0.75	0.575
VLGT - long term free recall	-0.21 ± 1.36	0.47 ± 0.84	0.211
WLG - Animals	0.15 ± 0.90	-0.40 ± 0.81	0.185
WLG - Professions	-0.19 ± 0.88	-0.04 ± 1.17	0.835
WLG – 4 letter "M" words	-0.61 ± 0.88	-0.36 ± 1.08	0.129
Digit Span - forward	-0.01 ± 0.87	0.06 ± 0.85	0.842
Digit Span - backward	-0.31 ± 0.89	-0.34 ± 1.42	0.572
TMT - Letter Number Switching	0.41 ± 3.23	0.59 ± 2.91	0.922
Stroop - interference	-1.00 ± 2.14	-0.78 ± 1.63	0.595

Abbreviations: standard DMT = standard Disease Modifying-therapy (interferons or glatiramer acetate), SPART = Spatial Recall Task, VLGT = Verbal Memory and Learning Task, WLG = World List Generation, TMT = Trail Making Test, * indicates significant difference between the patient groups

LEGENDS

Figure 1. Study design

Figure 2a. Spatial extent of lowered FA (in blue) of white matter voxels in the TBSS skeleton (green) (X = 77, Y = 109, Z = 77) of natalizumab patients (upper three rows) and standard DMT patients (lower three rows). Patients starting natalizumab had reduced FA in 56.8% ($p<0.05$) of the WM skeleton at baseline, was 55.6% at month 6 and was 47.2% at month 12. Patients continuing standard DMT had reduced FA in 41.4% of the WM skeleton at baseline, was 47.2% at month 6 and was 39.1% at month 12.

Figure 2b. Spatial extent of increased MD (in red) in white matter voxels in the TBSS skeleton (green) (X = 101, Y = 109, Z = 77) of natalizumab patients (upper three rows) and standard DMT patients (lower three rows). Patients starting natalizumab had increased MD in 54.3% ($p<0.05$) of the WM skeleton at baseline, was 48.1% at month 6 and was 55.7% at month 12. Patients continuing standard DMT had increased MD in 33.7% of the WM skeleton at baseline, was 32.0% at month 6 and was 36.1% at month 12.

Figure 3. Correlation between average cognition Z-score and average skeleton Z-score of MS patients at baseline and at month 12

FIGURES

Figure 1.

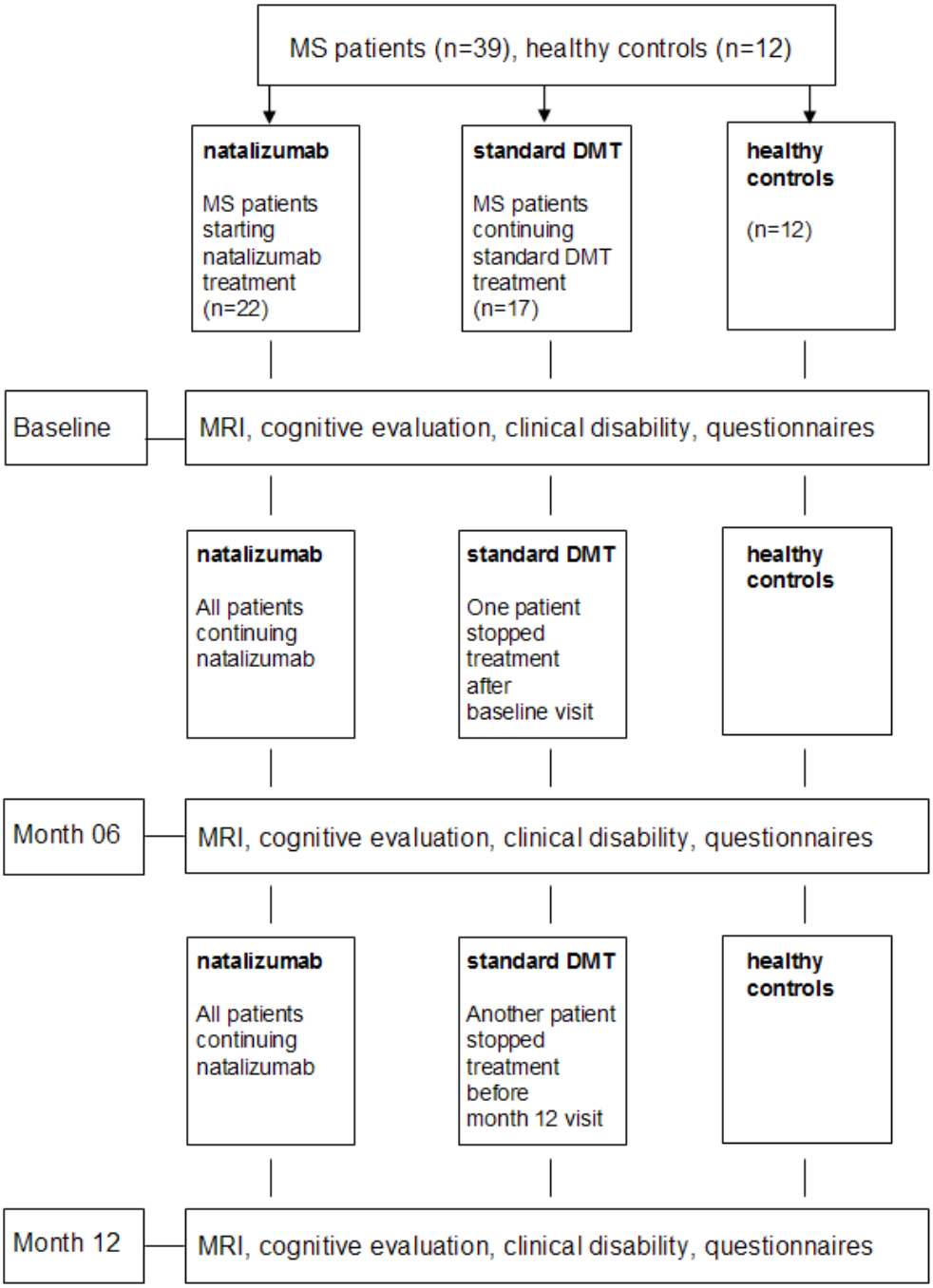


Figure 2a.

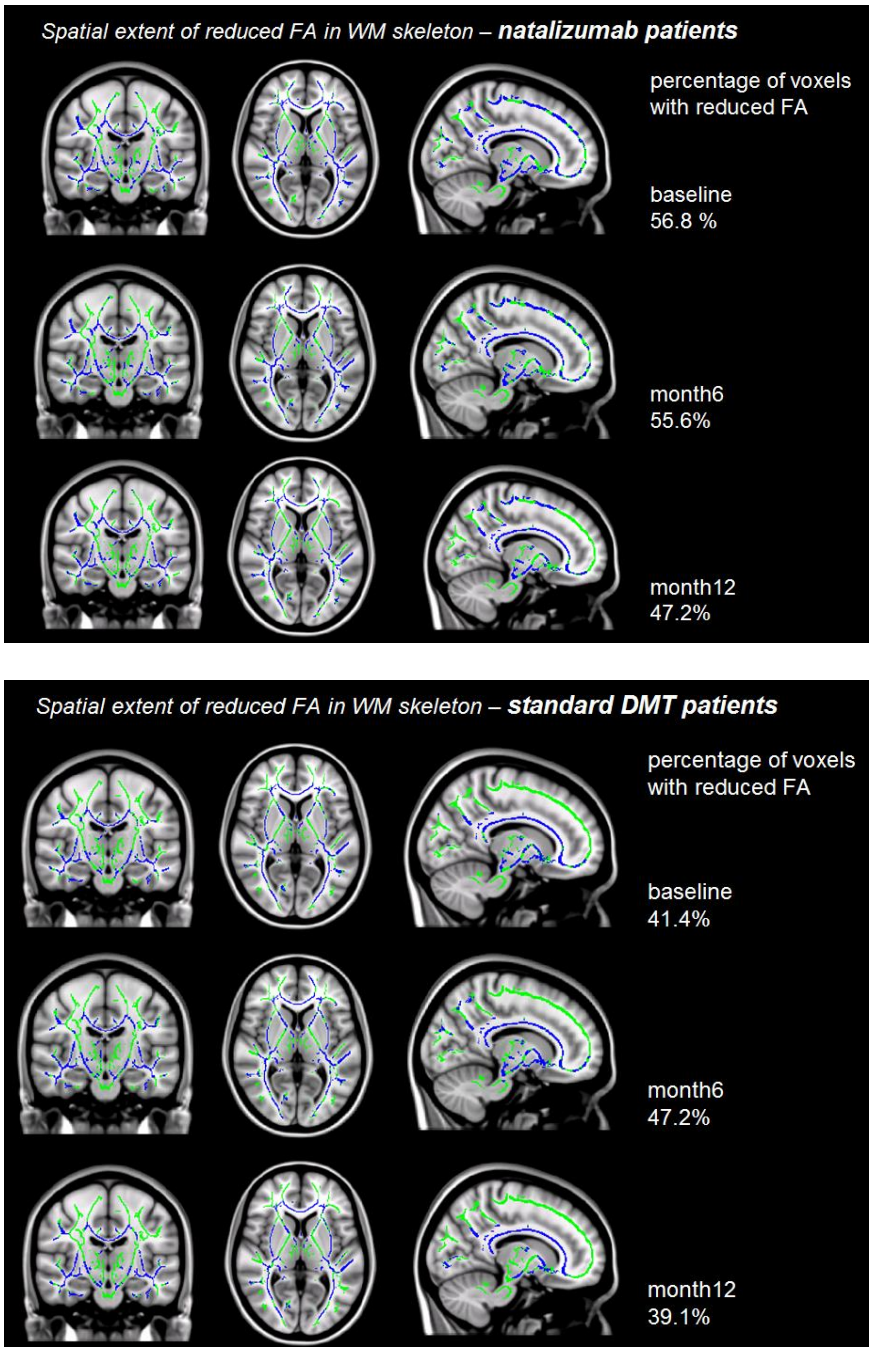
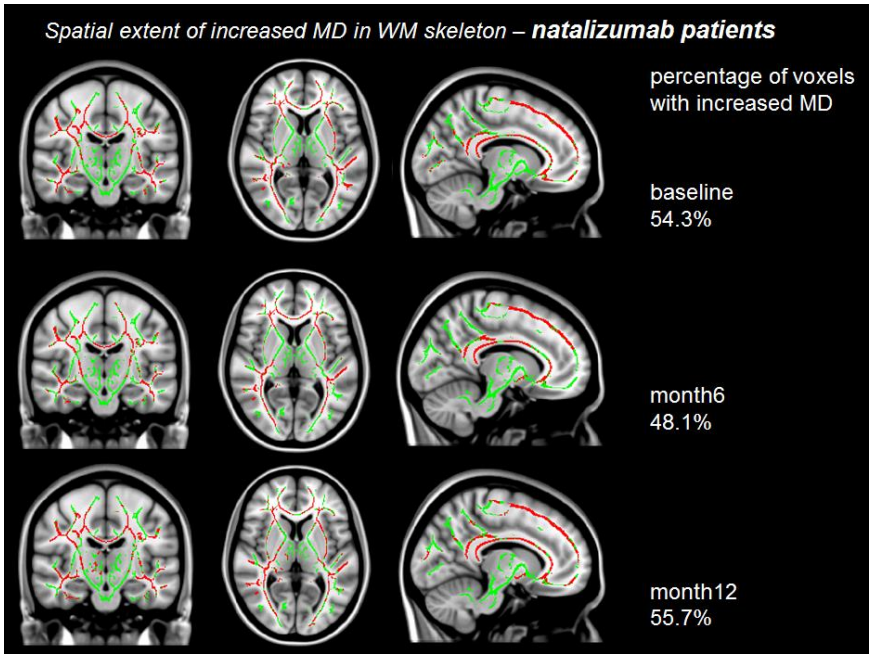

Spatial extent of reduced FA in WM skeleton – natalizumab patients
percentage of voxels with reduced FA
baseline 56.8 %
month6 55.6%
month12 47.2%
Spatial extent of reduced FA in WM skeleton – standard DMT patients
percentage of voxels with reduced FA
baseline 41.4%
month6 47.2%
month12 39.1%

Figure 2b.


Spatial extent of increased MD in WM skeleton – natalizumab patients
percentage of voxels with increased MD
baseline 54.3%
month6 48.1%
month12 55.7%

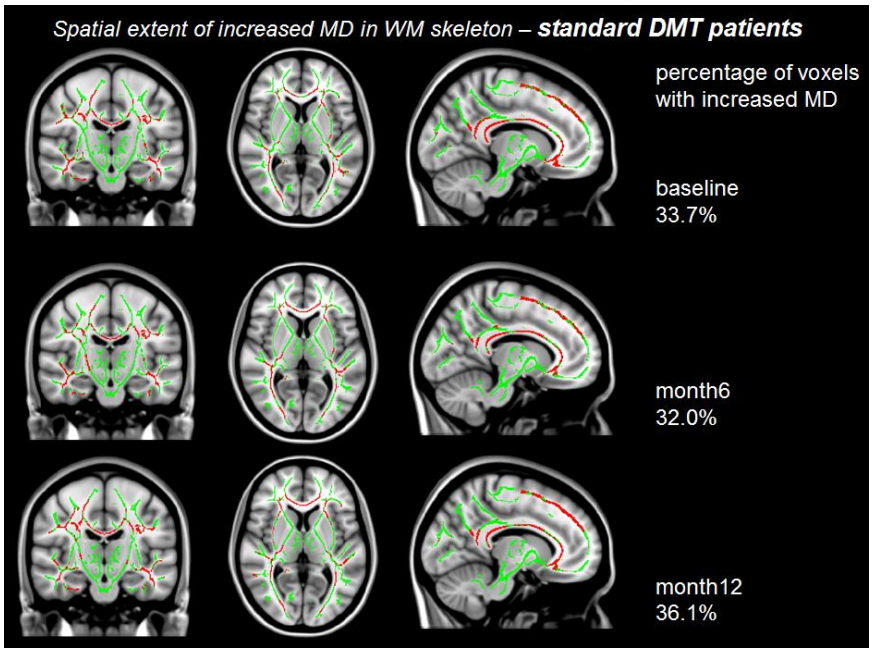
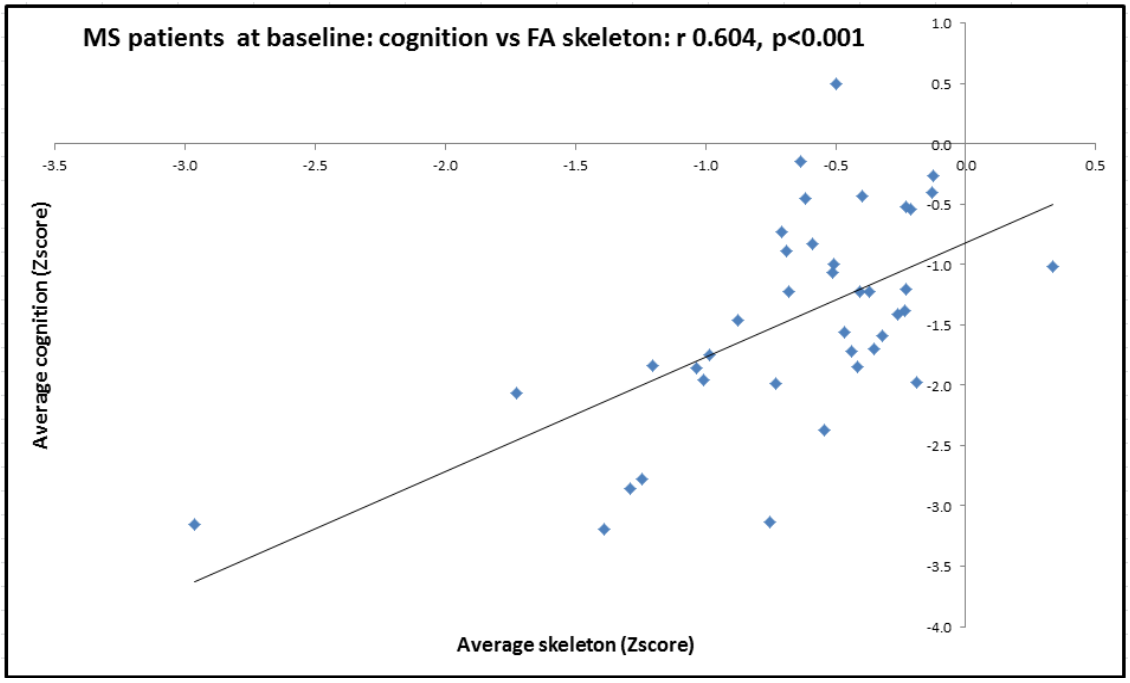
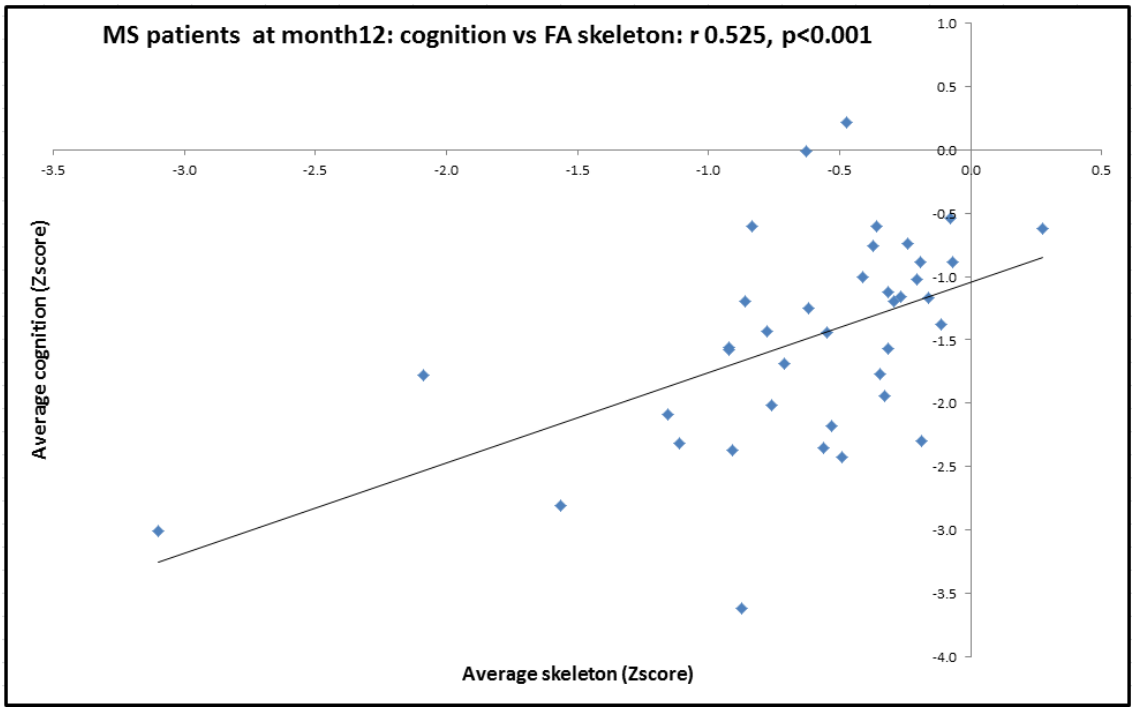

Spatial extent of increased MD in WM skeleton – standard DMT patients
percentage of voxels with increased MD
baseline 33.7%
month6 32.0%
month12 36.1%

Figure 3.


MS patients at baseline: cognition vs FA skeleton: r 0.604, p<0.001
Average cognition (Zscore)
Average skeleton (Zscore)
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-3.5
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5


MS patients at month12: cognition vs FA skeleton: r 0.525, p<0.001
Average cognition (Zscore)
Average skeleton (Zscore)
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-3.5
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5