

Corpus Callosum Damage Predicts Disability Progression and Cognitive Dysfunction in Primary-Progressive MS after Five Years

Benedetta Bodini,^{1*} Mara Cercignani,^{2,3} Zhaleh Khaleeli,¹ David H. Miller,² Maria Ron,² Sophie Penny,² Alan J. Thompson,¹ and Olga Ciccarelli¹

¹Department of Brain Repair and Rehabilitation, University College London Institute of Neurology, Queen Square, London, United Kingdom

²Department of Neuroinflammation, University College London Institute of Neurology, Queen Square, London, United Kingdom

³Neuroimaging Laboratory, IRCCS Santa Lucia Foundation, Rome, Italy

Abstract: We aim to identify specific areas of white matter (WM) and grey matter (GM), which predict disability progression and cognitive dysfunction after five years in patients with primary-progressive multiple sclerosis (PPMS). Thirty-two patients with early PPMS were assessed at baseline and after five years on the Expanded Disability Status Scale (EDSS), and EDSS step-changes were calculated. At year five, a subgroup of 25 patients and 31 healthy controls underwent a neuropsychological assessment. Baseline imaging consisted of dual-echo (proton density and T2-weighted), T1-weighted volumetric, and diffusion tensor imaging. Fractional anisotropy (FA) maps were created, and fed into tract-based spatial statistics. To compensate for the potential bias introduced by WM lesions, the T1 volumes underwent a lesion-filling procedure before entering a voxel-based morphometry protocol. To investigate whether FA and GM volume predicted EDSS step-changes over five years and neuropsychological tests scores at five years, voxelwise linear regression analyses were performed. Lower FA in the splenium of the corpus callosum (CC) predicted a greater progression of disability over the follow-up. Lower FA along the entire CC predicted worse verbal memory, attention and speed of information processing, and executive function at five years. GM baseline volume did not predict any clinical variable. Our findings highlight the importance of damage to the interhemispheric callosal pathways in determining physical and cognitive disability in PPMS. Disruption of these pathways, which interconnect motor and cognitive networks between the two hemispheres, may result in a disconnection syndrome that contributes to long-term physical and cognitive disability. *Hum Brain Mapp* 34:1163–1172, 2013. © 2012 Wiley Periodicals, Inc.

Key words: primary-progressive multiple sclerosis; tract-based spatial statistics; voxel-based morphometry; lesion filling; neuropsychological assessment

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: MS Society of Great Britain and Northern Ireland, and Department of Health's NIHR Biomedical Research Centres.

*Correspondence to: Dr. Benedetta Bodini, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, Queen

Square, London WC1N 3BG, UK. E-mail: b.bodini@ion.ucl.ac.uk

Received for publication 8 February 2011; Revised 7 September 2011; Accepted 29 September 2011

DOI: 10.1002/hbm.21499

Published online 13 February 2012 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

Quantitative magnetic resonance imaging (MRI) has the potential to provide prognostic indicators of clinical disability and cognitive dysfunction for patients with primary-progressive multiple sclerosis (PPMS), whose disease course can vary widely [Cottrell et al., 1999; Tremlett et al., 2005]. Over recent years, diffusion tensor imaging (DTI) [Basser et al., 1994] and volumetric MRI techniques have been employed to establish the clinical relevance of in-vivo brain white matter (WM) and grey matter (GM) abnormalities [Bakshi et al., 2008]. In particular, fractional anisotropy (FA), a DTI-derived measure [Basser et al., 1994], has been extensively used to quantify the integrity of WM tracts in MS; a reduction of this measure implies demyelination and axonal loss, and, despite being a measure which is pathologically non-specific, it has been shown to correlate significantly with cognitive scores in MS [Roosendaal et al., 2009] and predict motor impairment in relapsing-remitting MS [Kern et al., 2010]. With regard to GM, the most commonly used measure is GM volume, which, if reduced, reflects irreversible tissue loss (or atrophy) [Miller et al., 2002], and has been demonstrated to proceed relentlessly throughout the course of MS [De Stefano et al., 2010] and be clinically relevant [Fisher et al., 2008].

Several methods for detecting patient-control differences based on MRI parameters are available. Tract-based spatial statistics (TBSS) [Smith et al., 2006] and voxel-based morphometry (VBM) [Ashburner and Friston, 2000] are able to identify, respectively, WM regions of reduced FA and areas of GM atrophy in patients when compared with healthy controls. The advantage of these two techniques is that they localize specific brain regions of WM damage and GM volume loss in patients compared with healthy controls by searching across the whole brain, without the need to generate an a priori hypothesis. In a recent cross-sectional study of patients with early PPMS (i.e., patients who were studied within 5 years of symptom onset), we combined these two techniques, and reported a reduced FA in extensive WM regions including the whole corpus callosum (CC) and the cortico-spinal tracts [Bodini et al., 2009] in patients compared with healthy controls, and a diffuse reduction in patients GM volume, in areas such as the sensory-motor cortex bilaterally and the right superior temporal gyrus. In a subsequent longitudinal study in the same patient population, we employed magnetization transfer imaging histogram and volumetric analyses of WM and GM, and reported that WM lesion load is the most important predictor of subsequent cognitive dysfunction, but whole GM MTR also contributed [Penny et al., 2010]. Despite evidence of damage in both tissue compartments, and of their role in the accumulation of disability, it is unknown whether the factors responsible for clinical progression in PPMS are the whole (diffuse) GM tissue pathology or abnormalities in specific GM and WM regions, above and beyond the contribution provided by WM

TABLE I. Patients' clinical, radiological, and demographic characteristics at baseline

| | |
|--|-------------------|
| Number | 32 |
| Age, mean (SD) | 44.5 years (10.3) |
| Gender, female/male | 13/19 |
| Disease duration, mean (SD) | 3.3 years (0.9) |
| EDSS, median (range) | 4.5 (1.5–6.5) |
| T2 lesion load, ml, mean (SD) | 23.5 (17.3) |
| Years of education, ^a mean (SD) | 13 (2.9) |
| Premorbid IQ, ^a mean (SD) | 109.4 (9.6) |

^aValues calculated on the subgroup of 25 patients who underwent the neuropsychological testing at 5 years.

lesion load. This is an important question to address in order to understand the mechanisms underlying clinical deterioration, which is known to be quite variable in this type of MS.

Here, we hypothesize that specific location of grey and WM areas play a relevant role in predicting long-term physical and cognitive disability in patients with early PPMS. Therefore, we reanalyzed the same longitudinal dataset obtained in our unique cohort of early PPMS patients, and employed, for the first time, TBSS and VBM to identify specific areas of WM and GM which predict progression of physical disability over five years and cognitive dysfunction after five years.

METHODS

Study Design

Thirty-two patients with PPMS [Thompson et al., 2000] within five years of symptom onset (13 women, mean age 44.5 years, SD 10.3; see Table I for clinical, radiological, and demographic characteristics) underwent a whole brain imaging protocol, including diffusion sequences, and were clinically assessed on the Expanded Disability Status Scale (EDSS) [Kurtzke, 1983] at study entry. After five years, they were again assessed on the EDSS and invited to undergo an extensive neuropsychological assessment, which was also performed on a group of healthy controls.

Neuropsychological Assessment

Out of 32 patients recruited at baseline, 25 [11 women, mean age 51.3 years, (range 31–68), mean years of education 13.0 (range 10–19)] underwent an extensive neuropsychological assessment five years after entry into the study. Out of the seven patients who were not cognitively assessed, four male patients were too physically impaired (they were very dysarthric and ataxic) to provide valid results for a significant proportion of the tests, and therefore were excluded; one female and one male patient declined to take part in the neuropsychological tests. The remaining female patient was not contactable.

Thirty-one age- and gender-matched healthy controls [14 women, mean age 48.9 years, (range 30–65), mean years of education 13.4 (range 10–19)] also underwent the same neuropsychological tests. The exclusion criterion applied to all subjects was a history of other neurological or systemic illness, of impaired cognition, of psychiatric illness, of head injury resulting in loss of consciousness, or of alcohol or drug abuse.

The cognitive assessment included the following tests: (i) the National Adult Reading Test (NART) [Nelson, 1982], to estimate each patient's premorbid intellectual functioning; (ii) the immediate and delayed conditions of the story (SRT) and figure recall (FRT) subtests of the Adult Memory and Information Processing Battery [Coughlan and Hollows, 1985], to assess verbal and visual recall memory function; (iii) the Paced Auditory Serial Addition Test (3-second version; PASAT-3) [Cutter et al., 1999] and the Symbol Digit Modalities Test (SDMT) [Smith, 1982], to evaluate attention and speed of information processing; (iv) the Hayling Sentence Completion Task (HSCT) [Burgess and Shallice, 1997], a test of verbal response generation and inhibition, and the Brixton Spatial Anticipation Test [Burgess and Shallice, 1997], a spatial reasoning, rule detection and rule change task, to test executive functions.

Patients had not previously received any of these tests with the exception of the PASAT, which had been administered up to seven times in earlier follow-ups.

This study was approved by the Joint Medical Ethics Committee of the National Hospital for Neurology and Neurosurgery, London. Written and informed consent was obtained from all participants.

Processing of Clinical and Cognitive Data

Analysis was performed using Stata version 9.2 (Stata Corporation, College Station, Texas) and results with a $P < 0.05$ were considered significant.

Significant changes in EDSS between baseline and five years were assessed using the Wilcoxon signed ranks test. Since the EDSS is an ordinal (non-continuous scale), parametrical statistical methods cannot be used to analyze changes in the EDSS score, and, for this reason, the mean raw change in the EDSS score is considered to be an inappropriate statistical endpoint [Wingerchuk et al., 1997]. In addition, the meaning of the changes in EDSS is not equivalent throughout the scale. For example, the difference between EDSS 0.5 and 1.0 is 0.5, as well as the difference between EDSS 6.0 and 6.5 is 0.5; however, this 0.5 difference does not reflect the same extent of clinical deterioration. Finally, there is a variable mean duration at different EDSS levels [Weinshenker et al., 1991]. Therefore, to assess disability progression over the follow-up period, we followed the recommendation to calculate the EDSS step-change in each patient: one-step deterioration on the scale was defined as an increase of 1 if the baseline EDSS was

less than or equal to 5, or an increase of 0.5 if it was greater than 5. This approach is more sensitive in detecting significant clinical deterioration in the upper part of the scale (i.e., in more disabled patients), and has also been recommended for clinical trials [Wingerchuk et al., 1997].

As far as the neuropsychological scores are concerned, the raw scores of all the measures, except premorbid IQ, were converted to z-scores referenced to the control group, and these multiplied by -1 when appropriate, so that a lower score always indicated a poorer performance. A z-score ≤ -2 in a test was considered to be abnormal. Age, years of education, and premorbid IQ were compared between patients and healthy controls using *t*-tests; gender was compared between groups using a chi-square test. To investigate the difference in cognitive performance between patients and healthy controls, a multiple linear regression model was used, entering each cognitive measure, in turn, as dependent variable, and a binary group indicator, together with age, gender, years of education, and premorbid IQ, as covariates. Since our primary goal was to identify predictors of cognitive dysfunction, only the neuropsychological test scores that showed a significant difference between patients and healthy controls, or a trend towards a difference ($P < 0.08$), were entered in the subsequent steps of the analysis.

We have also tested for differences between male and female patients in EDSS scores at 5 years, and in all neuropsychological tests scores, using Mann-Whitney test and independent samples *t* tests, respectively.

Image Acquisition and Postprocessing

Patients were imaged at study entry using a 1.5T GE Signa scanner (General Electrics, Milwaukee, IL). MRI acquisition and protocol were as follows:

- a. Fast spin echo scan collecting proton-density-weighted (PD) and T2-weighted images [repetition time (TR) 2000 ms, echo times (TEs) 17/92 ms, field of view (FOV) 240×180 mm², matrix size 256×256 , 28 axial slices, 5-mm thickness]. Lesions were contoured on the PD images, with reference to the T2 images, using a semiautomated local contour thresholding technique [Plummer, 1992], and the total lesion load (LL) was calculated. In each patient, a lesion mask was created.
- b. Whole-brain, cardiac-gated, spin echo diffusion-weighted echo planar imaging sequence [FOV 240×240 mm², matrix size 96×96 (reconstructed to 128×128), image resolution $2.5 \times 2.5 \times 3$ mm (reconstructed to $1.9 \times 1.9 \times 3$ mm), TE 95 ms, TR 7 RRs, maximum b-factor 1000 smm⁻²; three series, each collecting 14 axial slices of 3-mm thickness, which were interleaved off-line; diffusion gradients were applied along 25 optimized directions, and three images with no diffusion weighting were also acquired]. After

correction for eddy-current induced distortions, the diffusion tensor was calculated on a voxel-by-voxel basis, and FA maps were generated using DTIfit, that is part of the FMRIB Software Library 4.1 software package (FSL, FMRIB Image Analysis Group, Oxford, UK) [Smith et al., 2004]. FA maps were fed into TBSS [Smith et al., 2006], to obtain a projection of all subjects' FA data onto a mean FA tract skeleton. We used the FSL FA atlas as a target for non-linear registration, while the mean FA of our patients was used to generate the mean skeleton, as recommended by the FSL website (www.fmrib.ox.ac.uk/fsl/tbss).

- c. Three-dimensional inversion-recovery fast spoiled gradient recall (3D FSPGR) T1-weighted (T1-w) sequence of the brain [FOV 300 × 225 mm, matrix size 256 × 160 (reconstructed to 256 × 256 for a final in plane resolution of 1.17 mm), TR 13.3 ms, TE 4.2 ms, inversion time 450 ms, 124 axial slices, 1.5-mm thickness].

The SPGR volumes were segmented and normalized to obtain GM, WM, and cerebro-spinal fluid volumes using SPM8 software (Wellcome Department of Cognitive Neurology, London, UK), according to the VBM protocol [Ashburner and Friston, 2005]. This protocol consists of an iterative combination of segmentations and normalizations and produced a GM probability map. As the presence of WM lesions is known to significantly affect brain segmentation [Nakamura and Fisher, 2009] and registration [Sdika and Pelletier, 2009], which are key steps in VBM, we applied the recently developed lesion-automated preprocessing (lesion-automated preprocessing, LEAP) [Chard et al., 2010] technique to the T1w images, before feeding them into the VBM procedure. Briefly, LEAP is an "in-painting" technique, which replaces lesional voxels with values derived from the intensity distribution within the WM outside visible lesions in the presegmentation phase [Chard et al., 2010]. This technique minimizes lesion-associated segmentation biases, which may significantly affect VBM results. As the smoothing kernel sensitizes the analysis to differences of comparable size as the kernel [Rosenfeld and Kak, 1982], GM images were modulated and smoothed using both 12-mm FWHM and 8-mm FWHM Gaussian kernels.

Investigation of Predictors of Physical Deterioration and Cognitive Dysfunction

White matter FA analysis

To investigate whether FA was associated with EDSS step-change and neuropsychological tests scores, a voxel-wise linear regression analysis was performed, adjusting for age, gender, years of education, and NART (the last two variables were used only when predicting neuropsychological scores). As mentioned above, only tests that identified abnormality in patients compared with healthy

controls were used. The analysis was based on permutation-based inference, and corrected for multiple comparisons using Threshold-Free Cluster Enhancement (TFCE; $P < 0.05$) [Smith and Nichols, 2009].

To investigate the effect of gender on significant findings, the voxel-wise linear regression analysis was repeated by modelling the interaction between gender (male or female) and each clinical and neuropsychological score, using permutation-based inference, and correcting for multiple comparisons using TFCE ($P < 0.05$).

To understand the contribution of WM lesions to the results, where significant associations were found between baseline FA and clinical/neuropsychological variable, the regression analysis was repeated by including, in turn, the total LL and the tract-specific LL, as additional covariate. This tract-specific LL was calculated by using the Johns Hopkins University WM tractography atlas, provided by FSL [Mori et al., 2005], which provides a probabilistic reconstruction of the main WM tracts. Each subject's T2-weighted scan was coregistered with the atlas, and the same transformation was applied to the lesion mask. Lesions belonging to a specific tract were identified to obtain the tract-specific LL.

Grey matter volumetric analysis

A voxel-wise multiple regression analysis was used to investigate the association between regional GM volume and EDSS step-change and neuropsychological tests, adjusting for the same variables as those used for TBSS, and correcting for multiple comparisons with family-wise error rate. P -values < 0.05 were considered to be significant.

RESULTS

Patients' Clinical and Neuropsychological Assessment

Patients showed progression of disability on the EDSS during the follow-up period [median EDSS at baseline = 4.5 (range 1.5–6.5), at 5 years = 6.4 (range 1.5–9), $P = 0.001$].

There were no significant differences between patients and healthy controls in age, gender, years of education, or premorbid IQ. The neuropsychological test results, the number of patients who failed on each test, and the comparison between patient and control groups in all tests are summarized in e-Table I (Supporting Information), and have been previously reported in detail [Penny et al., 2010]. Eighteen patients had abnormal scores on at least one test (six patients failed on one test only, six on two tests, and six on three or more tests). The neuropsychological tests on which patients performed significantly worse than healthy controls, and which were therefore retained for the subsequent part of the analysis, were the

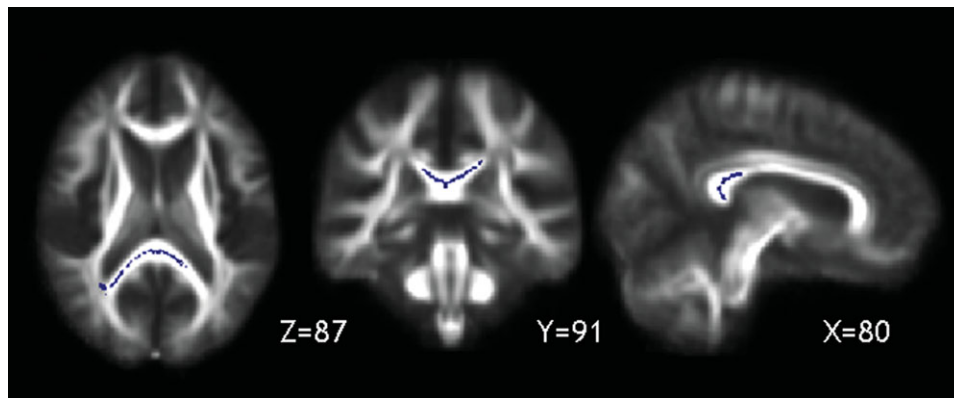


Figure 1.

The callosal regions along the TBSS skeleton, whose FA at study entry was associated with progression of physical disability over 5 years, are shown in blue and overlaid onto the patients' mean FA image ($P < 0.05$).

immediate and delayed story recall tests (SRT) and figure recall tests, the SDMT, and the HSCT.

There was no significant difference between male and female patients in EDSS scores at 5 years, and in all neuropsychological tests scores.

White Matter FA As Predictor

When looking for a significant association between FA across the whole skeleton and disability progression, a lower baseline FA in the splenium of the CC was found to be associated with greater progression of disability over five years, as measured by the EDSS step-change (MNI coordinates, $x = 83, y = 90, z = 92$, no. voxels = 8307, $P < 0.05$) (see Fig. 1).

With regard to the prediction of cognitive dysfunction, the strongest (and most consistent) association between lower FA at baseline and worse cognitive performance at

five years was found in the CC (see Table II for cluster sizes and coordinates). In particular: (i) lower FA in the splenium and genu of the CC was associated with worse immediate verbal memory function, as measured by the immediate SRT ($P = 0.003$; Fig. 2A); (ii) lower FA in the splenium, body, and genu of the CC (with extension into the right thalamic radiation) was associated with worse delayed verbal memory function, as measured by the delayed SRT ($P = 0.003$); (iii) lower FA in the body and splenium of the CC was associated with worse attention and speed of information processing, as measured by the SDMT ($P < 0.05$; Fig. 2B); (iv) lower FA in the genu and body of the CC (and in the bilateral corona radiata) was associated with worse executive functions, as measured by the HSCT ($P = 0.003$; Fig. 2C).

When less stringent thresholds for significance were used ($P < 0.05$), associations between FA and the

TABLE II. White matter regions showing a significant association between lower baseline FA and neuropsychological scores at 5 years

| Clinical and neuropsychological score at follow-up | No. of voxels per significant white matter cluster | MNI Atlas coordinates x, y, z | Significant regions | P value |
|--|--|---------------------------------|--|-------------|
| Immediate Story Recall Test | 6025 | 106, 82, 80 | Splenium of CC | $P = 0.003$ |
| | 1084 | 74, 171, 59 | Anterior part of genu of CC | |
| Delayed Story Recall Test | 4471 | 105, 84, 80 | Splenium of CC | $P = 0.003$ |
| | 2466 | 69, 173, 71 | Anterior part of right thalamic radiation and genu of CC | |
| Symbol Digit Modalities Test | 651 | 103, 112, 105 | Body of CC | $P = 0.05$ |
| | 2265 | 94, 104, 97 | Body of CC | |
| | 86 | 99, 84, 86 | Splenium of CC | |
| Hayling Sentence Completion Task | 1065 | 76, 161, 84 | Genu of CC | $P = 0.003$ |
| | 873 | 56, 80, 102 | Right posterior corona radiata | |
| | 456 | 107, 71, 121 | Left posterior corona radiata | |
| | 243 | 91, 135, 95 | Body of CC | |
| | 177 | 78, 122, 100 | Body of CC | |

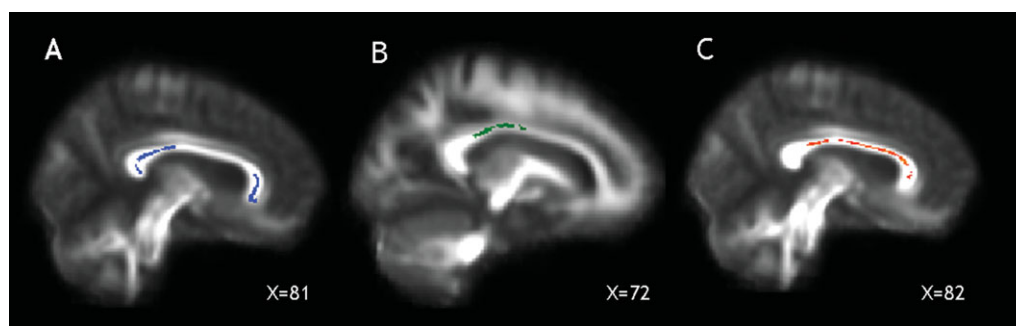


Figure 2.

The callosal areas, whose FA at study entry was associated with immediate verbal memory scores (as measured by SRT, $P < 0.003$), are shown in blue (A), those predicting attention and speed of information processing scores (SDMT, $P < 0.05$) at 5 years in green (B), and those whose FA was associated with executive function scores (as measured by HSCT, $P < 0.003$) are shown in red (C), and overlaid onto the patients' mean FA image.

immediate and delayed SRT and HSCT were also found in multiple areas of the skeleton, including the frontal and parietal WM (Fig. 1, Supporting Information) and the brainstem.

No significant gender-related differences in the EDSS and neuropsychological tests at five years were seen; no significant interactions between gender and EDSS or neuropsychological score were found.

Contribution of WM Lesions

When repeating the analyses adjusting for total LL, the association between lower FA in the genu and body of the CC and the executive functions (HSCT) at five years remained significant ($P < 0.05$), whilst the associations between FA and disability progression and the performance on immediate and delayed verbal memory tests (SRT) and attention and speed of information processing (SDMT) did not.

The mean volume of lesions localized in the CC (the region that showed the strongest association with clinical and neuropsychological variables) was 1.3 cc (SD1.47). When repeating the analyses adjusting for this CC lesion load, FA in the splenium of the CC no longer predicted disability progression, and lower FA in the splenium and body of the CC no longer predicted a poorer performance in attention and speed of information processing (SDMT). By contrast, the FA of the splenium of the CC at study entry remained significantly associated with a worse performance on immediate and delayed verbal memory tests (SRT; $P < 0.05$), and the FA of the body of the CC was associated with a poorer performance on a test of executive function (HSCT; $P < 0.05$) (Fig. 3) at five years.

Gray Matter Volume As Predictor

Baseline GM volume did not significantly predict EDSS step-change or any cognitive impairment at five years (with either 8-mm or 12-mm FWHM smoothing Kernel).

DISCUSSION

We applied a combination of TBSS and VBM (after a lesion-filling procedure) to the baseline FA and volumetric scans of patients with early PPMS. We found that the baseline FA of the CC, previously shown to be reduced in the same cohort of patients compared to that of healthy controls [Bodini et al., 2009], was significantly associated with long-term progression of physical disability, as measured by the EDSS step-change, and the occurrence of cognitive dysfunction at 5 years, in the domains of immediate and delayed verbal memory, executive function, and attention and speed of information processing, where patients performed significantly worse than controls. These findings were interpreted as indicative that lower CC FA at baseline was a predictor of both disability progression and cognitive dysfunction at five years. Future studies will



Figure 3.

The callosal area, whose FA at baseline significantly correlated with executive function scores (HSCT) at 5 years, after adjusting for corpus callosum T2 lesion load, is displayed in pink ($P < 0.05$) and overlaid onto the patients' mean FA image.

address the question whether CC FA is also relevant to the progression of cognitive impairment, which has been reported in MS [Langdon, 2011].

The CC ensures effective communication between motor and cognitive networks in the two hemispheres [Bogen et al., 1965]; in particular, the anterior parts of the CC (rostrum and genu) connect the orbitofrontal, lateral and medial frontal cortices, whereas the body and splenium connect parietal, temporal, and occipital homotopic regions [Abe et al., 2004]. The primary motor cortices, necessary for performing highly skilled motor movements, and several premotor areas, involved together with the supplementary motor areas in the initiation, planning, and regulation of movement, are connected between hemispheres through the CC [Chao et al., 2009]; the disconnection of these motor networks has been shown to result in a deterioration of motor function and upper limb function [Kern et al., 2010; Ozturk et al., 2010; Srikanth et al., 2010]. In our cohort, we found that lower FA in the splenium of the CC predicted the accumulation of motor disability over the follow-up period and, similarly, the raw EDSS at 5 years ($P = 0.02$; results not shown). The splenium of the CC is known to interconnect temporal and occipital cortices [Park et al., 2008], mainly underlying the transfer of auditory and visual information. However, the connections between its superior segment and Brodmann's area 5 [Chao et al., 2009], which is implicated in the earliest stages of sophisticated movement planning [Andersen and Buneo, 2002], could explain why its early damage predicts long-term disability progression in our patients.

Furthermore, the correlation we found between a lower baseline FA of the CC and a worse cognitive performance at 5 years, extends previous findings on the relevant role of the CC in ensuring effective interhemispheric cooperation, which is necessary to perform cognitive tasks that require the activation of cortical areas in both hemispheres. In particular, considerable evidence from functional neuroimaging studies showed that bilateral temporo-parietal cortices are implicated in performing episodic memory tests [Buckner and Wheeler, 2001; Rugg et al., 2002], and that complex tasks, such as those related to attention/speed of information processing, and executive function, require bilateral frontal and parietal involvement [Buchsbaum et al., 2005; Gazzaniga, 2005]. Evidence from several cross-sectional studies investigating different subject populations supports our findings [Jokinen et al., 2007; Voineskos et al., 2010]. For example, a reduction in the WM integrity of the splenium correlated with worse age-related memory and executive function in a sample of 53 healthy subjects [Voineskos et al., 2010], and atrophy of the CC fibers correlated with poor performance in tests assessing speed of mental processing, and attention and executive functions, in patients with age-related white-matter hyperintensities [Jokinen et al., 2007]. Cross-sectional studies have demonstrated that tissue damage localized in the CC plays a role in contributing to concurrent motor disability and cognitive dysfunction in different

phenotypes of MS [Kern et al., 2010; Mesaros et al., 2009; Ozturk et al., 2010]. We have now extended these results by demonstrating that CC damage has an important role in predicting long-term disability. The underlying mechanism may be the occurrence of a disconnection syndrome [He et al., 2009], whereby damage to structural connectivity [either through focal lesions or normal appearing white matter (NAWM) pathology] may, in turn, lead to an altered functional connectivity between GM regions and, ultimately, to clinical impairment [Dineen et al., 2009] and cognitive dysfunction [Dineen et al., 2009] in MS. The possibility of a disconnection syndrome is further supported by the findings that, when lower thresholds for significance were used for the analysis, the association between FA and cognitive performance on immediate and delayed verbal memory tests (SRT) and executive functions (HSCT) was found to extend to multiple regions of the skeleton, including frontal, temporal, and parietal WM, which are regions involved in memory [Buckner and Wheeler, 2001; Rugg et al., 2002] and executive function [Buchsbaum et al., 2005; Gazzaniga, 2005]. The finding of significant association between FA of the brainstem and executive function suggests that the disruption to brain stem-hemisphere interactions may further contribute to the disconnection syndrome, in agreement with reports of cognitive dysfunction in patients with isolated brain stem insult [Garrard et al., 2002].

When the association between FA and cognitive performance was corrected for total lesion load, FA in the genu and body of the CC remained significantly associated with performance on an executive function test (HSCT); similarly, FA of the body of the CC still predicted performance on the HSCT when adjusting for callosal lesion load (LL). Whilst the association between genu FA and executive function is expected to remain significant despite correction for the total LL, the association between more posterior portions of the CC and the performance on the HSCT is less obvious. The HSCT score is determined to some extent by general speed, and involves verbal processing and word generation. These cognitive functions require an effective interhemispheric connection between temporal and parietal cortical areas that are connected through more posterior regions of the CC. When adjusting for callosal LL, FA in the splenium of the CC predicted performance on immediate and delayed verbal memory tests (SRT), confirming that the reduced integrity of posterior callosal regions is linked to verbal memory dysfunction independently of local lesions. Overall, this persistence of significant correlations between FA and cognitive performance after corrections for LL supports the important role played by the normal-appearing WM (NAWM) damage in contributing to cognitive dysfunction [Mesaros et al., 2009] in MS; as previously hypothesised, the injury to the NAWM of the CC may result from a degeneration of axons transected in local T2 lesions and distal lesions whose related fibers cross the CC, through a mechanism of Wallerian degeneration [Coombs et al.,

2004; Dzedzic et al., 2010; Mesaros et al., 2009]. However, the contribution of GM injury and primary damage to axons (i.e., primary axonopathy) [Geurts et al., 2010] to axonal degeneration in the NAWM has yet to be conclusively established.

In our previous work on this same cohort, where we analyzed with a traditional approach the whole WM and GM matter histogram metrics and volumes, we found that whole brain GM MTR predicted the PASAT score, whereas whole brain GM and WM volume predicted performance on the Brixton test [Penny et al., 2010]. Here, the scores on both the PASAT and the Brixton have been excluded from the analysis, as they did not differ significantly between patients and healthy controls. Rather surprisingly, in the present study, we did not find any GM region that significantly predicted clinical outcome, when using a VBM approach. Since the extent of the areas being investigated is in principal unknown, we used two different smoothing kernels, but the results were consistently negative. The pathological correlates of GM atrophy are still unclear, and may reflect demyelination, loss of dendrites, axons, and synapses, possibly in combination. Even in early PPMS patients, GM atrophy has been identified in the precentral and postcentral gyri, and in the temporal cortex, and this has been reported to be relevant for concurrent clinical disability [Bodini et al., 2009; Khaleeli et al., 2007].

The lack of correlation between GM atrophy and disability progression is in contrast with the results of other studies, performed on patients with other forms of MS [Jasperse et al., 2007]. This could be a consequence of the regional distribution of atrophy in PPMS, which seems to differ from the pattern described in other forms of MS when using VBM [Ceccarelli et al., 2008]. This difference in regional distribution of atrophic GM areas could perhaps help to explain the lack of significant association between regional GM volumes at baseline and long-term clinical outcomes in our cohort of PPMS. Moreover, in RRMS, it has been reported a significant correlation between early changes in regional GM volume and clinical outcome [Jasperse et al., 2007]. It is possible that looking at regional changes in GM volume, rather than using baseline regional GM volumes only, as we did in this study, could prove more useful in finding significant correlations with long-term clinical outcomes. Finally, when investigating potential predictive measures of clinical outcome, MRI measures other than GM volume, for example those able to explore the microstructural damage occurring in the early stage of the disease, such as MTR, or sequences that allow detection of GM lesions, may provide significant results. In addition, it would be of a great interest to investigate whether there is a temporal relationship between WM abnormalities and GM changes, and whether this changes across brain regions and types of MS. From a technical point of view, in this work we used a novel approach to remove lesion-related segmentation bias before VBM. Preliminary results of the applications of this

method show that without region filling, the GM volume is in general overestimated (because the WM lesions have signal intensity similar to that of the cortex) [Chard et al., 2010]. Therefore, the application of this method should ensure that the segmented GM does not include the lesional tissue.

In conclusion, this study has provided insights into the mechanisms of progression in PPMS, by showing that early damage localized the CC is a relevant determinant of progression of motor disability and future cognitive dysfunction in this patient group. Disconnection of relevant brain regions is proposed as a potential mechanism of clinical progression in MS.

ACKNOWLEDGEMENTS

The authors are very grateful to all the patients for taking part in the study. This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

REFERENCES

- Abe O, Masutani Y, Aoki S, Yamasue H, Yamada H, Kasai K, Mori H, Hayashi N, Masumoto T, Ohtomo K (2004): Topography of the human corpus callosum using diffusion tensor tractography. *J Comput Assist Tomogr* 28:533–539.
- Andersen RA, Buneo CA (2002): Intentional maps in posterior parietal cortex. *Annu Rev Neurosci* 25:189–220.
- Ashburner J, Friston KJ (2000): Voxel-based morphometry—The methods. *Neuroimage* 11:805–821.
- Ashburner J, Friston KJ (2005): Unified segmentation. *Neuroimage* 26:839–851.
- Bakshi R, Thompson AJ, Rocca MA, Pelletier D, Dousset V, Barkhof F, Inglesse M, Guttman CR, Horsfield MA, Filippi M (2008): MRI in multiple sclerosis: Current status and future prospects. *Lancet Neurol* 7:615–625.
- Basser PJ, Mattiello J, LeBihan D (1994): Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 103:247–254.
- Bodini B, Khaleeli Z, Cercignani M, Miller DH, Thompson AJ, Ciccarelli O (2009): Exploring the relationship between white matter and gray matter damage in early primary progressive multiple sclerosis: An in vivo study with TBSS and VBM. *Hum Brain Mapp* 30:2852–2861.
- Bogen JE, Fisher ED, Vogel PJ (1965): Cerebral commissurotomy. A second case report. *JAMA* 194:1328–1329.
- Buchsbaum BR, Greer S, Chang WL, Berman KF (2005): Meta-analysis of neuroimaging studies of the Wisconsin card-sorting task and component processes. *Hum Brain Mapp* 25:35–45.
- Buckner RL, Wheeler ME (2001): The cognitive neuroscience of remembering. *Nat Rev Neurosci* 2:624–634.
- Burgess PW, Shallice T (1997): *The Hayling and Brixton Tests*. Bury St. Edmunds: Thames Valley Test Company.
- Ceccarelli A, Rocca MA, Pagani E, Colombo B, Martinelli V, Comi G, Filippi M (2008): A voxel-based morphometry study of grey matter loss in MS patients with different clinical phenotypes. *Neuroimage* 42:315–322.

- Chao YP, Cho KH, Yeh CH, Chou KH, Chen JH, Lin CP (2009): Probabilistic topography of human corpus callosum using cytoarchitectural parcellation and high angular resolution diffusion imaging tractography. *Hum Brain Mapp* 30: 3172–3187.
- Chard DT, Jackson JS, Miller DH, Wheeler-Kingshott CA (2010): Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *J Magn Reson Imaging* 32:223–228.
- Coombs BD, Best A, Brown MS, Miller DE, Corboy J, Baier M, Simon JH (2004): Multiple sclerosis pathology in the normal and abnormal appearing white matter of the corpus callosum by diffusion tensor imaging. *Mult Scler* 10:392–397.
- Cottrell DA, Kremenchutzky M, Rice GP, Koopman WJ, Hader W, Baskerville J, Ebers GC (1999): The natural history of multiple sclerosis: A geographically based study. V. The clinical features and natural history of primary progressive multiple sclerosis. *Brain* 122 (Part 4):625–639.
- Coughlan AK, Hollows AK (1985): The Adult Memory and Information Processing Battery. Leeds: St James's University Hospital, AK Coughlan.
- Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, Syndulko K, Weinshenker BG, Antel JP, Confavreux C, Ellison GW, Lublin F, Miller AE, Rao SM, Reingold S, Thompson A, Willoughby E (1999): Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 122 (Part 5):871–882.
- De Stefano N, Giorgio A, Battaglini M, Rovaris M, Sormani MP, Barkhof F, Korteweg T, Enzinger C, Fazekas F, Calabrese M, Dinacci D, Tedeschi G, Gass A, Montalban X, Rovira A, Thompson A, Comi G, Miller DH, Filippi M (2010): Assessing brain atrophy rates in a large population of untreated multiple sclerosis subtypes. *Neurology* 74:1868–1876.
- Dineen RA, Vilisaar J, Hlinka J, Bradshaw CM, Morgan PS, Constantinescu CS, Auer DP (2009): Disconnection as a mechanism for cognitive dysfunction in multiple sclerosis. *Brain* 132:239–249.
- Dziedzic T, Metz I, Dallenga T, Konig FB, Muller S, Stadelmann C, Bruck W (2010): Wallerian degeneration: A major component of early axonal pathology in multiple sclerosis. *Brain Pathol* 20:976–985.
- Fisher E, Lee JC, Nakamura K, Rudick RA (2008): Gray matter atrophy in multiple sclerosis: A longitudinal study. *Ann Neurol* 64:255–265.
- Garrard P, Bradshaw D, Jäger HR, Thompson AJ, Losseff N, Playford D (2002): Cognitive dysfunction after isolated brain stem insult. An underdiagnosed cause of long term morbidity. *J Neurol Neurosurg Psychiatry* 73:191–194.
- Gazzaniga MS (2005): Forty-five years of split-brain research and still going strong. *Nat Rev Neurosci* 6:653–659.
- Geurts JJ, Kooi EJ, Witte ME, van der Valk P (2010): Multiple sclerosis as an “inside-out” disease. *Ann Neurol* 68:767–768.
- He Y, Dagher A, Chen Z, Charil A, Zijdenbos A, Worsley K, Evans A (2009): Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. *Brain* 132:3366–3379.
- Jaspere B, Vrenken H, Sanz-Arigita E, de G, V, Smith SM, Polman CH, Barkhof F (2007): Regional brain atrophy development is related to specific aspects of clinical dysfunction in multiple sclerosis. *Neuroimage* 38:529–537.
- Jokinen H, Ryberg C, Kalska H, Ylikoski R, Rostrup E, Stegmann MB, Waldemar G, Madureira S, Ferro JM, van Straaten EC, Scheltens P, Barkhof F, Fazekas F, Schmidt R, Carlucci G, Pantoni L, Inzitari D, Erkinjuntti T (2007): Corpus callosum atrophy is associated with mental slowing and executive deficits in subjects with age-related white matter hyperintensities: The LADIS Study. *J Neurol Neurosurg Psychiatry* 78:491–496.
- Kern KC, Sarcona J, Montag M, Giesser BS, Sicotte NL (2010): Corpus callosal diffusivity predicts motor impairment in relapsing-remitting multiple sclerosis: A TBSS and tractography study. *Neuroimage* 55:1169–1177.
- Khaleeli Z, Cercignani M, Audoin B, Ciccarelli O, Miller DH, Thompson AJ (2007): Localized grey matter damage in early primary progressive multiple sclerosis contributes to disability. *Neuroimage* 37:253–261.
- Kurtzke JF (1983): Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology* 33:1444–1452.
- Langdon DW (2011): Cognition in multiple sclerosis. *Curr Opin Neurol* 24:244–249.
- Mesaros S, Rocca MA, Riccitelli G, Pagani E, Rovaris M, Caputo D, Ghezzi A, Capra R, Bertolotto A, Comi G, Filippi M (2009): Corpus callosum damage and cognitive dysfunction in benign MS. *Hum Brain Mapp* 30:2656–2666.
- Miller DH, Barkhof F, Frank JA, Parker GJ, Thompson AJ (2002): Measurement of atrophy in multiple sclerosis: Pathological basis, methodological aspects and clinical relevance. *Brain* 125:1676–1695.
- Mori S, Wakana S, Nagae-Poetscher LM, van Zijl PC (2005): MRI Atlas of Human White Matter. Elsevier: Amsterdam, The Netherlands.
- Nakamura K, Fisher E (2009): Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. *Neuroimage* 44:769–776.
- Nelson HE (1982): The National Adult Reading Test, Windsor: NFER-Nelson.
- Ozturk A, Smith SA, Gordon-Lipkin EM, Harrison DM, Shiee N, Pham DL, Caffo BS, Calabresi PA, Reich DS (2010): MRI of the corpus callosum in multiple sclerosis: Association with disability. *Mult Scler* 16:166–177.
- Park HJ, Kim JJ, Lee SK, Seok JH, Chun J, Kim DI, Lee JD (2008): Corpus callosal connection mapping using cortical gray matter parcellation and DT-MRI. *Hum Brain Mapp* 29:503–516.
- Penny S, Khaleeli Z, Cipolotti L, Thompson A, Ron M (2010): Early imaging predicts later cognitive impairment in primary progressive multiple sclerosis. *Neurology* 74:545–552.
- Plummer D (1992): DispImage: A display and analysis tool for medical images. *Rev Neuroradiol* 5:489–495.
- Roosendaal SD, Geurts JJ, Vrenken H, Hulst HE, Cover KS, Castellijn JA, Pouwels PJ, Barkhof F (2009): Regional DTI differences in multiple sclerosis patients. *Neuroimage* 44:1397–1403.
- Rosenfeld and Kak (1982): Digital Picture Processing 2, Academic Press: Orlando, FL, p. 42.
- Rugg MD, Otten LJ, Henson RN (2002): The neural basis of episodic memory: Evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci* 357:1097–1110.
- Sdika M, Pelletier D (2009): Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping. *Hum Brain Mapp* 30:1060–1067.
- Smith A (1982): Symbol Digit Modalities Test (SDMT) Manual (revised). Los Angeles: Western Psychological Services.
- Smith SM, Nichols TE (2009): Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44:83–98.
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De LM, Drobnjak I, Flitney

- DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De SN, Brady JM, Matthews PM (2004): Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 (Suppl 1):S208–S219.
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TE (2006): Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31:1487–1505.
- Srikanth V, Phan TG, Chen J, Beare R, Stapleton JM, Reutens DC (2010): The location of white matter lesions and gait—A voxel-based study. *Ann Neurol* 67:265–269.
- Thompson AJ, Montalban X, Barkhof F, Brochet B, Filippi M, Miller DH, Polman CH, Stevenson VL, McDonald WI (2000): Diagnostic criteria for primary progressive multiple sclerosis: A position paper. *Ann Neurol* 47:831–835.
- Tremlett H, Paty D, Devonshire V (2005): The natural history of primary progressive MS in British Columbia, Canada. *Neurology* 65:1919–1923.
- Voineskos AN, Rajji TK, Lobaugh NJ, Miranda D, Shenton ME, Kennedy JL, Pollock BG, Mulsant BH (2010): Age-related decline in white matter tract integrity and cognitive performance: A DTI tractography and structural equation modeling study. *Neurobiol Aging* [Epub ahead of print].
- Weinshenker BG, Rice GP, Noseworthy JH, Carriere W, Baskerville J, Ebers GC (1991): The natural history of multiple sclerosis: A geographically based study. IV. Applications to planning and interpretation of clinical therapeutic trials. *Brain* 114 (Part 2):1057–1067.
- Wingerchuk DM, Noseworthy JH, Weinshenker BG (1997): Clinical outcome measures and rating scales in multiple sclerosis trials. *Mayo Clin Proc* 72:1070–1079.