Evidence for early neurodegeneration in the cervical cord of patients with primary progressive multiple sclerosis

K. Abdel-Aziz [a,b,*], T. Schneider [a,c], B.S. Solanky [a,c], M.C. Yiannakas [a,c], D.R. Altmann [a,d], C.A.M Wheeler-Kingshott [a,c], A.L. Peters [e], B.L. Day [e], A.J. Thompson [a,b,f], O. Ciccarelli [a,b,f].

[a] NMR Research Unit, UCL Institute of Neurology, London, UK.

[b] Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, London, UK.

[c] Department of Neuroinflammation, UCL Institute of Neurology, London, UK.

[d] Medical Statistics Department, London School of Hygiene and Tropical Medicine, London, UK

[e] Sobell Department, UCL Institute of Neurology, London, UK.

[f] National Institute of Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre (BRC)

[*] Corresponding author and reprint request to:

Khaled Abdel-Aziz
NMR Research Unit,
Queen Square MS Centre
Department of Brain Repair and Rehabilitation,
UCL Institute of Neurology,
Queen Square House,
Queen Square,
London, WC1N 3BG
U.K
Email: k.abdel-aziz@ucl.ac.uk

Tel: +44 (0)8451555000 ext. 724307

Fax: +44 (0)207 278 5616

Grant Support: The NMR Research Unit is supported by the UK MS Society. This study has been supported by the UK MS Society (Award Ref No: 984). TS is supported by the EPSRC (grant reference EP/I027084/1). 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.

ABSTRACT

Spinal neurodegeneration is an important determinant of disability progression in patients with primary progressive multiple sclerosis (PPMS). Advanced imaging techniques, such as single-voxel ^{1}H-MR spectroscopy (MRS) and q-space imaging (QSI), have increased pathological specificity for neurodegeneration, but are challenging to implement in the spinal cord and have yet to be applied in early PPMS. By combining these imaging techniques with new clinical measures, which reflect spinal cord pathology more closely than conventional clinical tests, we explored the potential for spinal MRS and QSI to detect early spinal neurodegeneration that may be responsible for clinical disability.

Data from We recruited 23 21 PPMS patients within six years of disease onset and 26 24 healthy controls were analysed. Patients were clinically assessed on grip strength, vibration perception thresholds (VPT) and postural stability, in addition to the Expanded Disability Status Scale, 9-hole peg test, timed 25-foot walk test, Multiple Sclerosis Walking Scale-12, and Modified Ashworth Scale (MAS). All subjects underwent MRS and QSI of the cervical cord and conventional brain and spinal MRI at 3T. Multivariate analyses and multiple regression models were used to assess the differences in imaging measures between groups and the relationship between MRI measures and clinical scores, correcting for age, gender, spinal cord cross-sectional area, brain T2 lesion volume, and brain white matter and grey matter volume fractions.

Although patients did not show significant cord atrophy when compared with healthy controls, they had significantly lower total N-acetyl-aspartate (tNAA) (mean 4.01 versus 5.31 mmol/L, P=0.020) and Glutamate-Glutamine (Glx) (mean 4.65 versus 5.93 mmol/L, P=0.043) than controls. Patients showed an increase in QSI-derived indices of perpendicular

diffusivity in both the whole cord and major columns compared with controls ($P<0.05$ for all indices). Lower tNAA was associated with higher disability, as assessed by EDSS (Coefficient= -0.41, $0.01<P<0.05$), MAS (Coefficient= -3.78, $0.01<P<0.05$), VPT (Coefficient= -4.37, $P=0.021$) and postural sway ($P<0.001$). Lower Glx predicted increased postural sway ($P=0.017$). Increased perpendicular diffusivity in the whole cord and columns was associated with increased scores on the MAS, VPT and postural sway ($P<0.05$ in all cases).

These imaging findings indicate reduced structural integrity of neurons, demyelination, and abnormalities in the glutamatergic pathways in the cervical cord of early PPMS, in the absence of extensive spinal cord atrophy. The observed relationship between imaging measures and disability suggests that early spinal neurodegeneration may underlie clinical impairment, and should be targeted in future clinical trials with neuroprotective agents to prevent the development of progressive disability.

INTRODUCTION

The clinical phenotype of primary progressive multiple sclerosis (PPMS) is characterised by sustained disability progression from disease onset and is typically associated with severe locomotor disability (Thompson *et al.*, 2000), with a median time to DSS 6 (walking with a cane) of between 6 to 8.5 years (Runmarker and Andersen, 1993; Cottrell *et al.*, 1999; Confavreux *et al.*, 2000). The rate of disability progression is highly variable, but occurs more quickly early in the disease course and reflects, in part, neuroaxonal loss and neuronal dysfunction in the spinal cord (Bjartmar *et al.*, 2000). There would be great value in developing and applying imaging markers of neurodegenerative processes to the spinal cord in early PPMS in order to improve our understanding of the early pathological events that occur in the injury pathway responsible for clinical disability. This step is considered to be crucial in the translational pathway that aims to validate biomarkers that predict clinical outcomes and treatment response in clinical trials in PPMS (Fox *et al.*, 2012).

Advanced quantitative MRI (qMRI) has been applied in the brain in early PPMS and has improved our understanding of the mechanisms leading to tissue damage, beyond that associated with macroscopic T2 lesions (Wheeler-Kingshott *et al.*, 2014). Measures provided by diffusion tensor imaging (DTI) and ^{1}H-MR spectroscopy (^{1}H-MRS), have been shown to correlate with disability (Ramio-Torrenta *et al.*, 2006; Sastre-Garriga *et al.*, 2005; Bodini *et al.*, 2013), and to predict progression (Khaleeli *et al.*, 2007; Khaleeli *et al.*, 2008). Applying similar techniques to the spinal cord has been technically challenging (Wheeler-Kingshott *et al.*, 2014). However, recent developments have led to applications of advanced qMRI in the spinal cord in relapsing-remitting multiple sclerosis (RRMS) and have provided insights into

underlying spinal tissue pathology (Ciccarelli *et al.*, 2007; Farrell *et al.*, 2008; Marliani *et al.*, 2010; Ciccarelli *et al.*, 2013; Kearney *et al.*, 2014).

One of the most promising qMRI techniques is high b-value Q-space imaging (QSI), a model free diffusion weighted imaging (DWI) technique (Callaghan *et al.*, 1988). QSI is thought to be highly specific for axonal injury (Assaf *et al.*, 2005) and has shown better sensitivity for detecting pathophysiological changes within lesions and normal appearing white matter (NAWM), compared to DTI in the brains of patients with MS (Assaf *et al.*, 2002). A small pilot study in relapse-onset MS demonstrated the feasibility of using high b-value QSI in the spinal cord with improved detection of abnormal diffusion compared with the conventional DWI acquisition and analysis (Farrell *et al.*, 2008).

Spinal cord ^{1}H-MRS is used to quantify metabolites which reflect specific pathological processes, and can complement structural imagingComplementary to structural MRI, is metabolic imaging, such as spinal cord ^{1}H-MRS, which is used to quantify metabolites that are markers of specific pathological processes (Ciccarelli *et al.*, 2014). Commonly quantified metabolites in the spinal cord include: total N-acetyl-aspartate (tNAA), a marker of neuroaxonal integrity and metabolic function (Moffett *et al.*, 2007), Myo-inositol (Ins), a marker of astrocytic activation and proliferation (Brand *et al.*, 1993), and total Choline (tCho), which reflects changes in steady state levels of membrane phospholipids released during myelin breakdown (Henning *et al.*, 2008; Marliani *et al.*, 2010). More recently, our group developed a new protocol capable of quantifying glutamate-glutamine (Glx), a marker of neuronal integrity and neurotransmitter pool, in the spinal cord (Solanky *et al.*, 2013). Although there have been a few spinal cord MRS studies in patients with RRMS and neuromyelitis optica (NMO), which have consistently shown neuronal loss and metabolic

dysfunction, as reflected by reduced concentration in tNAA in the cervical cord of patients compared to controls (Marliani *et al.*, 2007; Ciccarelli *et al.*, 2007; Ciccarelli *et al.*, 2013), to date none have included patients with PPMS.

Besides the need to utilise more pathologically specific *in vivo* spinal cord imaging techniques, there is also a need to incorporate objective clinical measures, which are more sensitive to changes in clinical functions mediated by spinal pathways than conventional clinical tests, such as the Expanded Disability Severity Scale (EDSS) (Kurtzke, 1983). Measures such as postural stability, vibration perception thresholds (VPT) and dynamometry are more responsive to small clinical changes due to damage in the spinal cord than the EDSS, and have been shown to increase the sensitivity for detecting correlations between MRI abnormalities in the spinal cord and disability (Zackowski *et al.*, 2009; Oh *et al.*, 2013).

In the current study we have used a combination of MRS and QSI to investigate changes in the cervical cord which underlie disability in patients with early PPMS, to test two hypotheses; i) MRS and QSI demonstrate early neurodegeneration in the upper cervical cord in patients with PPMS before the occurrence of spinal cord atrophy; ii) in patients, there is a relationship between MRS and QSI measures and disability, as reflected by newer spinal-cord specific clinical scores, alongside standard MS clinical scales, suggesting that early spinal cord neurodegeneration is linked with clinical impairment in PPMS.

MATERIALS AND METHODS

Study participants

We prospectively recruited patients with a diagnosis of PPMS (Polman *et al.*, 2005), aged between 18 – 65 years, within six years from disease onset, as well as, age and gender matched healthy controls. On the day of the MRI, patients were clinically assessed. All subjects provided written, informed consent prior to taking part in the study which was approved by our local research ethics committee.

Clinical Assessments

All patients were assessed using conventional clinical scales, including the EDSS (Kurtzke, 1983), 9-Hole Peg Test (HPT) (Goodkin *et al.*, 1988) and Timed 25-foot Walk Test (TWT) (Cutter *et al.*, 1999). For the purpose of statistical analysis, the average of two trials of the TWT and the average of four trials of the HPT (averaged as reciprocals of the mean times from two trials for each hand) (Fischer *et al.*, 1999) were calculated. We also used the Multiple Sclerosis Walking Scale-12 (MSWS-12) (Hobart *et al.*, 2003), and the Modified Ashworth Scale (MAS) (Bohannon and Smith, 1987). The MAS values from 16 muscle groups in the upper and lower limbs were converted from a 0–4 scale (which includes a value of 1+ between scores of 1 and 2) to a 0 –5 scale; the resulting values were summated to obtain an overall score ranging from 0 to 80 (Stein *et al.*, 2007).

Clinical scales with the potential to be sensitive to spinal cord pathways injury were also applied, including the mean grip strength from both upper limbs, using the Jamar hydraulic dynamometer (Sammons Preston Incorporated, Bolingbrook, IL, USA) (Svens and Lee, 2005), and the vibration perception thresholds (VPTs), which were measured from all four

limbs at the lateral malleoli and ulna styloid processes using the biosthesiometer (Bio-Medical Instrument Company, Newbury, Ohio). Mean VPTs were calculated and used in the analysis. Finally, postural stability was assessed using a modified version of a recently described protocol for quantifying stance instability (Bunn *et al.*, 2013). Subjects were asked to stand relaxed and still, facing a blank wall at a distance of 1 metre, in a well-lit room, for 40 second-long trials. Three trials under each of the four conditions were recorded, consisting of 2 stance widths (inter-malleolar distance of 32 cm and 4 cm) under 2 visual conditions (eyes either open or closed). Body sway was measured using a 3-D orientation sensor (MTx: Xsens, Enschede, NL), which was fixed to the skin, just below the C7 spinous process. The device measured the instantaneous angular position of the trunk in the anteroposterior (pitch) and mediolateral (roll) planes and was sampled at 100Hz. Summary measures were made on these signals using custom scripts written in Matlab (The Mathworks, Natick, MA USA). The raw data were low-pass filtered at 10Hz using a zero-phase, 5th order Butterworth filter. The amount of angular motion was then calculated separately for the roll and pitch body sway data and from the combined motion given by square root (pitch-motion^2 + roll-motion^2), termed total sway. All three signals were summarised by summing the sample-to-sample absolute change in signal and then dividing by the duration of the trial to yield average angular speeds of body sway reported in degrees/second. The mean of the three trials per condition were used for statistical analysis. An index of exacerbation of sway on eye closure was obtained from the Romberg quotient calculated as sway eyes closed/sway eyes open at both stance widths.

Spinal cord and brain MRI Protocol

All scans were performed using a 3T Achieva system (Philips Medical Systems, Best, Netherlands). To reduce motion artefacts during scanning and improve image quality, an MR compatible cervical collar was worn by all volunteers (Yiannakas *et al.*, 2012).

Using the manufacturer's 16-channel neurovascular coil (Phillips Healthcare Systems), single voxel MRS was performed using a recently optimised protocol (Solanky *et al.*, 2013). Conventional turbo spin-echo sequences (TSE) were used to acquire structural images for radiological reading and to guide guiding voxel placement. T2w images were acquired in the coronal plane [parameters: TR = 4000 ms; TE = 100ms; FOV= 160 x 250 mm^2; voxel size = 0.6 x 0.6 x 3.0 mm^3; NEX = 2; 13 contiguous slices; scan time =1:36 minutes] and PD/T2w images were acquired in the sagittal plane using a dual echo TSE [parameters: TR = 4000 ms; TE = 15/80ms; FOV= 256 x 160 mm^2; echo train length (ETL) = 12; voxel size = 1.0 x 1.0 x 3.0 mm^3; NEX = 2; 12 contiguous slices; scan time = 5:44 minutes]. For spectroscopy, volumes of interest (VOIs) with dimensions of approximately 5.4 x 7.76 x 55mm^3 (2.3 ml) were prescribed using the reference images and centred on the C2/3 intervertebral disc (**Figure 1)**. The dimensions of the VOI were adjusted in the anterior-posterior (AP) direction dependent on the size of each volunteers spinal cord (Ciccarelli *et al.*, 2007; Marliani *et al.*, 2010). MRS data was acquired using a point resolved spectroscopy (PRESS) localisation sequence [parameters: TE = 30ms; 376 averages with triggered, first order iterative shimming, multiply optimised insensitive suppression train (MOIST) water suppression, 4 outer volume suppression (OVS) slabs in the AP and rostrocaudal directions and cardiac gating (TR = 3RR ≈ 3000 ms) using a peripheral pulse unit (350ms delay), scan time = 19:42 minutes].

For cord mean cross-sectional area (CSA) measurements and confirmation of lesion location, the cervical cord was imaged in the axial plane, perpendicular to the longitudinal axis of the cord with the imaging volume centred on the C2/3 intervertebral disc, using a fat-suppressed 3D slab-selective fast field echo (FFE) sequence [parameters: TR = 23 ms; TE = 5ms; flip angle α = 7°; FOV= 240 x 180 mm^2; voxel size = 0.5 x 0.5 x 5 mm^3; NEX = 8; 11 axial contiguous slices; scan time = 15:58 minutes]. In order to match the position and orientation of the volumetric scan to the spectroscopy voxel, the prescription values used for the MRS acquisition were copied and manually entered by the operator when setting up the 3D-FFE scan.

Using the manufacturer's 32-channel receive head coil (Philips Medical Systems, Best, Netherlands), each subject underwent a cardiac gated DWI acquisition [parameters: voxel size=1×1×5 mm^3 (interpolated in k-space to a 0.5 x 0.5mm^2 in-plane resolution) FOV = 64 × 64mm^2; TR = 9RR, TE = 129ms] performed with the volume centred on the C2/C3 disc to ensure similar coverage as the spectroscopy voxel; 12 axial contiguous slices covering a 60mm length of the cervical cord, typically giving coverage of the C1-3 spinal segments (**Figure 1**). The 32 channel head coil was used because it gave superior SNR during QSI sequence optimisation experiments (Schneider *et al.*, 2011). A ZOOM sequence was used with outer-volume suppression to minimise artefacts (Wilm *et al.*, 2007). Thirty DWI volumes with equally spaced **q**-values (Farrell *et al.*, 2008) (Schneider and Wheeler-Kingshott, 2014) and two non-diffusion weighted (b0) volumes were acquired with diffusion weighting in two perpendicular (x and y) and one parallel (z) direction relative to the main axis of the spinal cord [parameters: diffusion pulse duration δ=11.4 ms, diffusion time Δ=75ms, gradient strength G linearly increased in 31 steps from 0 to 87.5mT/m in x and y direction and 62mT/m in z direction; scan time = 22:28 minutes].

To achieve the maximum possible gradient strength on the scanner, we exploited the combination of parallel gradient amplifiers in the scanner, which each generate a maximum diffusion gradient strength of 62mT/m along the major axis of the scanner bore. Assuming axial symmetry of the axons along the long axis of the spinal cord, we applied gradient amplifiers in two orthogonal directions that maximise gradient strength perpendicular to axis of the spinal cord (**Supplementary figure 1**). This allowed us to generate a guaranteed maximum gradient strength of $\sqrt{2}$*62mT/m in the xy direction. In the z direction we use a maximum gradient of 62 mT/m. q-value were the same in xy and z direction, but the increase in gradient strength allowed us to use a smaller the gradient pulse duration of 11.4ms in xy direction (16ms in z). The full protocol is given in **supplementary table 1**.

For calculation of brain T2 lesion volumes, PD/T2 weighted images were acquired using a dual-echo TSE sequence [parameters: TR = 3500 ms; TE = 15/85 ms; flip angle α = 90°; FOV= 240 x 180 mm^2; voxel size = 1 x 1 x 3 mm^3; NEX = 1; 50 axial contiguous slices; scan time = 4:01 minutes]. For calculation of brain tissue volumes a 3D T1-weighted magnetisation-prepared gradient-echo sequence was used [TR = 6.9 ms; TE = 3.1 ms; TI = 824 ms; flip angle α = 8°; FOV= 256 x 256 mm^2; voxel size = 1 x 1 x 1 mm^3; NEX = 1; 180 sagittal contiguous slices; scan time = 6:31 minutes].

Imaging post-processing

Spinal cord metabolite quantification

Metabolite concentrations were quantified using the user-independent LCModel (version 6.3) package (Provencher, 1993) and a set of basis spectra, comprising seventeen metabolites

including the macromolecules, simulated using GAMMA (Smith *et al.*, 1994) as previously described (Solanky *et al.*, 2013). NAA (N-acetyl-aspartate) + NAAG (N-acetylaspartyl glutamate), (hereafter tNAA), tCho (Choline + phosphocholine), tCr (creatine + phosphocreatine), Ins and Glx concentrations were quantified using the unsuppressed water signal obtained from the same voxel as a reference (Gasparovic *et al.*, 2006) and formed the focus of our analysis. Corrections for T2 values were not performed because the TE used is relatively short, compared to the T2 relaxation times of the metabolites under study (Wansapura *et al.*, 1999; Edden *et al.*, 2007) and, therefore, it is expected that changes in T2 would be negligible. Measuring T2 values for each metabolite would not have been possible in a patient cohort within clinically feasible scan times

The signal-to-noise ratio (SNR) and full width of half maximum (FWHM) of the tNAA peak provided by LCModel were used to assess spectral quality and Cramér–Rao Lower Bounds (CRLB) values of <20% for tNAA, tCr, tCho and Ins and <30% for Glx were used to confirm the reliability of the spectral fit (Provencher, 2014). Poor quality spectra were excluded from the analysis. Criteria for exclusion were poor water suppression or FWHM > 0.13 with SNR < 3.

Spinal cord cross sectional area measurement

Image segmentation and CSA measurements were performed using the 3D-FFE dataset in Jim 6.0 Software (Xinapse systems, Northants, England). Three contiguous 5mm axial slices, centred on the C2/3 disc were segmented using the active surface model method (Horsfield *et al.*, 2010). The mean cross-sectional area of these three slices was then calculated.

Spinal cord QSI and ROI analysis

QSI indices from the q-space analysis, which characterise the diffusion properties of water, are derived from the displacement probability density function (dPDF), which is the average probability of a spin moving a certain distance during a given diffusion time. At a given diffusion time, a tall, narrow dPDF suggests a low diffusion constant and/or restricted diffusion, whereas a low, broad dPDF suggests a high diffusion constant and/or more unrestricted diffusion (Farrell *et al.*, 2008).

The two perpendicular diffusion directions were averaged (xy) to increase the signal-to-noise ratio. The measurements were then linearly regridded to be equidistant in q-space and the diffusion dPDF was computed using inverse Fast Fourier Transformation. To increase the resolution of the dPDF, the signal was extrapolated in q-space to a maximum q=200mm^{-1} by fitting a bi-exponential decay curve to the DWI data (Farrell *et al.*, 2008). The dPDF was computed from the extrapolated DWI data on a voxel-by-voxel basis using the inverse Fast Fourier Transformation**. Supplementary figure 2** illustrates the processing pipeline.

Data was corrected for motion using reg_aladin from the NiftyReg toolkit (Ourselin *et al.*, 2000). Registration was performed between the interleaved b=0 acquisitions of the xy and z protocol using the first b=0 of the xy protocol as reference. The estimated registration was then applied to the intermediate DWI images. The quality of the motion correction was assessed in each subject and mis-registered slices/subjects were excluded from the study.

Voxel-wise maps of the full width at half maximum (FWHM), which represents the width of the dPDF, and the zero displacement probability (P0), representing the height of the dPDF, were computed for xy and z. Conventional ADC maps were also derived from the low b-

value part of the decay curve ($b < 1100 s/mm^2$) for both xy and z directions, using a constrained non-linear least squares fitting algorithm (Farrell *et al.*, 2008).

To assess region specific differences in QSI indices, the whole 60mm length of upper cervical spinal cord was first extracted from CSF and other tissue types, and four regions of interest (ROIs) were created using the ROI tool in JIM 6.0 and positioned using the b0 images for each axial slice for orientation. ADC and QSI indices were measured from ROIs in the anterior, right lateral, left lateral and posterior columns, as well as whole cord (**Supplementary figure 3**). No statistical differences were found between QSI indices from the right and left lateral columns; therefore for ease of analysis, a mean value from both columns was calculated for each of the QSI indices.

Brain T2 Lesion volumes and grey matter and white matter volume fractions

Brain T2-lesion volume (T2LV) was calculated by outlining lesions on T2-weighted MRI scans using a semi-automated edge finding tool (JIM v. 6.0) by a single observer (KA). Total lesion volume was recorded in mLs for each subject.

To avoid segmentation errors due to white matter (WM) lesions, an automated lesion-filling technique was employed (Chard *et al.*, 2010). Lesion masks were created based on 3D-T1 weighted sequences. The lesion-filled images were segmented into WM, grey matter (GM) and cerebrospinal fluid (CSF), using the 'new segment' option in SPM8 (statistical parametric mapping; Wellcome Trust Centre for Neuroimaging, University College London (UCL) Institute of Neurology, London). Segmentations were reviewed to exclude errors. WM

and GM fractional (WMF and GMF) volumes, relative to total intracranial volume (the sum of GM, WM and CSF volumes), were calculated.

Statistical analysis

Analyses were performed in Stata 13.1 (Stata Corporation, College Station, Texas, USA). Adjusted differences between patients and controls were obtained by multiple regression of the relevant imaging measure on a subject type indicator, with age, gender and CSA as covariates. This analysis was then repeated to evaluate the adjusted difference between controls and patients with and without spinal cord lesions within the C1-3 region of interest. In the case of CSA, group differences were obtained with a multiple regression model co-varying for age and gender.

In patients, univariable associations between metabolites and whole cord QSI metrics were examined with Pearson correlations. Associations between spinal cord imaging measures and EDSS, MAS, HPT and VPTs were examined with multiple regression of the clinical variable on the spinal cord imaging measure as predictor, with the following potential confounders as covariates: age, gender, mean cord area, brain T2 lesion volume, GMF and WMF; because of the large number of covariates, these were entered singly into the model, and the unadjusted association is only reported where it was not materially affected by entering any of these covariates. Where regression residuals showed signs of non-normality (e.g. for EDSS), the non-parametric bias corrected and accelerated bootstrap was used (1000 - 5000 replicates, depending on the p-value resolution required), and then, if more precise determination was too computer intensive, the P-value was reported as a range.

For associations between spinal cord measures and measures of postural stability, multivariate regression was used because of the highly related nature of the clinical measures: by performing joint tests of association, the danger of spurious significant results was minimised by reporting associations only where the joint test was significant; where the joint test is not significant, there is no global evidence for any of the individual associations tested, in which case these are not reported as significant even when individually $P<0.05$. The multivariate associations were carried out with potential confounders entered as described above.

RESULTS

Participant demographics and characteristics

Twenty-three patients with early PPMS and 26 healthy controls were recruited. One patient was unable to tolerate the scan and was therefore excluded; a second patient's scans were excluded from the final analysis due to severe motion-related image degradation. Two control subjects were also excluded from the study due to the detection of unexpected pathology on structural spinal cord imaging. Therefore data from 21 patients and 24 age and gender-matched healthy controls were included in the final analysis (**Table 1**). Patients had short disease duration and mild to moderate levels of disability; further details on patient characteristics, disability and conventional brain MRI are summarised in **Table 1.** Conventional MRI of the cervical cord identified cervical cord lesions in 18 out of 21 patients (see conventional MRI findings presented patient-by-patients with age and disease duration in **supplementary Table 12)**. Of the 18 patients with cervical cord lesions, 142 patients had lesions within the lower C1 to upper C4 C3 segments covered by the MRS and QSI volumes.

Spectroscopy quality indicators

Typical post-processed spectra are shown in **Figure 2**. The FWHM and SNR estimated by LCModel (reported as mean ± SD) were 0.11 ± 0.03 ppm and 4.4 ± 1.4 respectively. Mean CRLBs for each metabolites were; tNAA (8%), tCr (11%), tCho (10%), Ins (10%) and Glx (21%). The reproducibility of MRS measurements achieved with this protocol have previously been reported (Solanky *et al.*, 2013).

Differences in spinal cord measures between patients and controls

There was no significant difference in CSA between patients and controls, after adjusting for age and gender (P = 0.092). Patients had lower spinal tNAA and Glx concentrations than healthy controls, after correction for age, gender and CSA, and this was most marked in patients with spinal cord lesions within the spectroscopic volume (**Table 2** and **Figure 2**). Ins concentrations were borderline significantly higher in patients than healthy controls but were significantly elevated in patients with a C1-C3 lesion (**Table 2**).

Patients had significantly higher perpendicular diffusivity (indicating increased movement of water perpendicular to the main cord axis, as reflected by increased ADCxy and FWHMxy and reduced P0xy), in the whole cord and the anterior, posterior and lateral columns, and a significant increase in parallel diffusivity (ADCz) confined to the posterior columns when compared with controls, after adjusting for age, gender and CSA (**Table 2**, **Table 3** and **Figure 2**). Perpendicular diffusivity derived from QSI indices (FWHMxy and P0), but not ADCxy was also significantly higher in patients with normal appearing spinal tissue compared with healthy controls **(Table 2)**.

Univariable analysis of spinal cord metabolite concentrations and QSI metrics in patients

In patients, spinal cord tNAA concentration was negatively correlated with whole cord ADCxy (r = - 0.581, p = 0.011) and FWHMxy (r = - 0.636, p = 0.005) and positively correlated with whole cord P0xy (r = 0.646, p = 0.004) (**Supplementary Figure 4**). Other spinal metabolite concentrations did not correlate significantly with each other, or with cord QSI indices.

Associations between whole cord imaging measures and clinical disability

In patients, following adjustment for age, gender, CSA, brain T2 lesion volume, GMF and WMF, a significant association was seen between lower spinal tNAA concentrations and increased global and spinal-cord specific disability measures (as reflected by higher EDSS, MAS, VPT and postural sway, respectively) (**Tables 4 and 5**). Lower spinal Glx and higher Ins were both also independently associated with increased postural instability **(Table 5)**. When looking at the relationship between QSI measures and disability, increased QSI-derived perpendicular diffusivity was associated with higher spasticity (MAS), higher VPT, and increased postural instability **(Tables 4 and 5)**.

Associations between column-specific QSI indices and clinical disability

Following adjustment for age, gender, CSA, brain T2 lesion volume, GMF and WMF, increased perpendicular diffusivity within the major spinal columns was associated with increased disability. In particular, increased spasticity was independently associated with perpendicular diffusivity in all the columns (in particular, lower P0xy in the anterior, lateral and posterior columns, increased FWHMxy in the anterior and lateral columns, and increased ADCxy in the lateral columns). Reduced vibration sensation was independently associated with increased perpendicular diffusivity (reduced P0xy and increased FWHMxy and ADCxy) in the anterior, lateral and posterior columns. Instability in the roll plane was independently associated with increased perpendicular diffusivity (reduced P0xy) in the posterior column, while instability in the pitch plane was independently associated with increased perpendicular diffusivity (increased ADCxy) in the anterior column. A summary of associations is presented in **Table 6**.

Discussion

In this study, we have demonstrated lower concentrations of tNAA and Glx in the upper cervical cord of patients with early PPMS compared to controls, which suggest the presence of neurodegeneration, including neuronal loss and/or metabolic dysfunction, and changes in the glutamatergic pathway. The increased QSI-derived perpendicular diffusivity (increased FWHMxy and decreased P0xy) in patients compared with controls further confirms the occurrence of reduced neuronal integrity, possibly with demyelination. Significant associations between spinal cord tNAA, Glx and QSI-derived perpendicular diffusivity and newer measures of clinical disability, such as postural stability and VPT, suggest that these imaging measures reflect abnormalities that contribute to clinical impairment. Thus, the evidence for early neurodegeneration in the spinal cord, in the absence of extensive spinal cord atrophy, and its link with clinical impairment, provide insights into the pathological events that occur in PPMS and indicate that this should become a target for therapeutic intervention. qMRI measures will be further developed and validated as useful biomarkers of disease progression and treatment response in clinical trials.

Differences in metabolite concentrations and QSI measures between patients and controls

The lower tNAA concentrations in the spinal cord of PPMS patients when compared with healthy controls are consistent with metabolite abnormalities in the brain, where tNAA is lower in cortical grey matter and NAWM in early PPMS compared with controls (Sastre-Garriga *et al.*, 2005). In addition, our findings are qualitatively similar to those seen in acute (Ciccarelli *et al.*, 2007; Henning *et al.*, 2008; Ciccarelli *et al.*, 2010) and chronic (Marliani *et al.*, 2010; Ciccarelli *et al.*, 2013) spinal cord lesions in RRMS. The majority of the early

PPMS patients included in the present study (N=1412) had a lesion (or part of a lesion) within the spectroscopic voxel and in these patients, spinal tNAA concentrations were lower than patients without a lesion., suggesting that lesional tissue abnormalities may have contributed to the observed tNAA changes. There were too few subjects in the study to detect a statistically significant difference in tNAA concentrations between patients with and without spinal lesions within the spectroscopic voxel: we estimated that the sample size required to detect a difference between those two groups with 80% power (alpha 0.05) using the spectroscopy protocol described in this study would be 168 subjects per group; this finding suggests that tNAA concentration is the lowest in the lesional tissue of the spinal cord, but may be reduced, although less extensively, in the normal-appearing white matter when compared with the healthy tissue, which is similarly to what has been previously been demonstrated in the brain (Caramanos *et al.*, 2005).

Glx, which represents the sum of Glu and its precursor Gln, was also significantly lower in patients than controls., These changes were most significant in patients with spinal cord lesions within the spectroscopic voxel and likely reflecting changes in the spinal glutamatergic pathway. . Glu makes up the majority of the Glx signal (Baker *et al.*, 2008), and is predominantly found in the synaptic terminals, with relatively little present in the extracellular compartment and glial cells (Kaiser *et al.*, 2005; Muhlert *et al.*, 2014). It is therefore possible that lower spinal Glx could in part, be explained by neuro-axonal degeneration. In the brains of patients with early PPMS, Glx is reduced in the cortical grey matter, but not the NAWM (Sastre-Garriga *et al.*, 2005). Similarly, in patients with clinically stable RRMS, Glu and Glx are both reduced in grey matter regions (Muhlert *et al.*, 2014). Together these results would suggest that the reductions in Glx reflect reduced synaptic density in the grey matter secondary to neuronal loss. We found that Glx did not correlate well with tNAA, which suggests that impairment of glutamatergic metabolism as well as

neuroaxonal loss may occur in early PPMS, and these metabolites reflect different aspects of the underlying tissue changes.

During relapses, Glu is elevated in acute cerebral white matter lesions (Srinivasan *et al.*, 2005) and elevated Glu is associated with accelerated neuronal loss on long-term follow-up (Baranzini *et al.*, 2010). It has been suggested that activated inflammatory cells are the source of transient excesses in Glu (Piani *et al.*, 1991), and as inflammation ebbs, Glu concentrations return to normal (Stover *et al.*, 1997; Gurwitz and Kloog, 1998).

Using QSI we measured diffusivity, parallel and perpendicular to the long axis of the spinal cord and found significantly higher perpendicular diffusivity in patients compared with controls. The changes in the dPDF shape (Figure 2) in our patient group are likely to reflect a breakdown in myelin and axonal membranes which both act as microstructural barriers to perpendicular diffusion (Beaulieu, 2002) and correspond to what would be expected based on findings from murine and canine models of dysmyelination and axonal loss (Biton *et al.*, 2006; Farrell *et al.*, 2010; Wu *et al.*, 2011; Anaby *et al.*, 2013), as well as to what has previously been reported in patients with relapse-onset MS (Assaf *et al.*, 2002; Farrell *et al.*, 2008). The differences in QSI measures between patients and controls are also in agreement with the tNAA and Glx changes detected and provide corroborating evidence for early neurodegeneration in the cervical cord. Importantly, in patients with normal appearing spinal tissue within the diffusion imaging volume, FWHMxy and P0xy remained significantly different to controls, whereas ADCxy did not, suggesting that QSI indices are more sensitive to microstructural injury than conventional ADC measures.

In addition to the differences in tNAA, Glx and QSI-derived perpendicular diffusivity between groups, we found that patients had higher spinal Ins levels than controls but this finding did not reach statistical significance. However, wWe calculated that for Ins, to have

80% power to detect a patient vs control difference of the size observed (which is about two thirds of a standard deviation) at 5% significance, 40 subjects per group would be necessary. Patients with a spinal cord lesion within the spectroscopic voxel did have significantly elevated Ins concentrations, Therefore, the difference in Ins between groupswhich may is likely to reflect genuine metabolic differences and suggests that astrocytic proliferation and activation (or gliosis) occurs in the spinal cord lesions in early PPMS. Previous studies have suggested that gliosis is an early pathological process in MS, and gliosis may be an important mechanism of disease progression (Ciccarelli *et al.*, 2014). Our results suggest this process is more active in lesional than non-lesional tissue.

With regard to tCho, which is a marker of inflammation and membrane turnover (Henning *et al.*, 2008; Marliani *et al.*, 2010), since the observed differences between groups is less than 15% of the SD, it would take hundreds of subjects per arm to detect such a small difference, suggesting that this metabolite is unlikely to be useful for distinguishing patients from healthy controls in future studies.

In our study, spinal CSA, a measure of tissue loss, which is often used as an imaging surrogate of axonal loss and has started to be used in MS clinical trials (Kearney *et al.*, 2013), was not significantly different between patients and controls despite CSA measurements being performed on a sequence with high in-plane resolution. Earlier studies with larger sample sizes (Bieniek *et al.*, 2006) and those which included patients with longer disease duration (Losseff *et al.*, 1996), demonstrated significant cord atrophy in PPMS. Based on CSA measures from our cohort of patients and controls, we estimate that the sample size required to detect significant differences in CSA in early PPMS, using the method described in this study with 80% power (alpha = 0.05), is 68 subjects per group. This would suggest that much smaller sample sizes are required to detect group differences early in the disease course with newer qMRI measures that reflect neurodegenerative processes other than

atrophy alone. We cannot exclude that alternative image segmentation methods, such as the edge detection and partial volume correction method proposed by Tench *et al* (Tench *et al.*, 2005) may have enabled detection of significant cord atrophy in this patient group and this merits further study in future. In order to validate these new measures for clinical trials, it is important to test whether these qMRI measures and metabolite concentrations are sensitive to changes occurring over time and predict clinical outcome at follow-up.

Association between spinal cord metabolites and diffusion indices

The modifications to gradients and pulse lengths necessary to perform QSI on clinical scanners have the effect of exaggerating the contribution of slow diffusing water to QSI metrics (Assaf *et al.*, 2002), consequently, diffusion of intra-axonal water is highly represented (Assaf and Cohen, 2000; Assaf *et al.*, 2000; Assaf *et al.*, 2005) and it has therefore been suggested that QSI metrics make useful markers of axonal integrity. Interestingly, in our study, spinal tNAA concentration which reflects axonal integrity correlated more strongly with QSI derived indices of perpendicular diffusivity than ADC suggesting these indices are more indicative of axonal integrity.

Associations between whole cord imaging measurements and clinical disability

Using new clinical scales, which reflect disability in functions mediated by spinal cord pathways, we have extended previous findings of significant associations between tNAA and neurological disability, as measured by the EDSS, in the spinal cord of RRMS patients (Ciccarelli *et al.*, 2007; Blamire *et al.*, 2007), by demonstrating that significant associations exist in patients with early PPMS and that Glx levels are associated with postural stability.

We found that, in patients, higher Ins concentrations were associated with poor postural stability, suggesting that spinal cord gliosis may be a process of clinical importance in early PPMS. This is in agreement with spinal cord MRS studies in RRMS, which have shown an increased Ins concentration in patients than controls (Marliani *et al.*, 2010) and a relationship between higher Ins and higher EDSS scores (Ciccarelli *et al.*, 2007).

In agreement with the MRS results, we found that increased whole cord QSI-derived perpendicular diffusivity, which reflects increased movement of water in the direction perpendicular to the main axis of the cord, as a consequence of reduced neuronal integrity and/or demyelination, is independently associated with increased spasticity, VPTs and postural instability. Our findings extend on those from an earlier pilot study which found a significant increase in QSI-derived perpendicular diffusivity within spinal cord lesions in patients with relapse-onset MS compared to healthy controls (Farrell *et al.*, 2008), and suggest that whole cord QSI reflects clinically meaningful pathological changes in the spinal cord.

Associations between column-specific diffusion indices and disability

We found several significant associations which were expected based on *a priori* knowledge of the neurological function of tracts running in specific spinal cord columns. Specifically, increased QSI-derived perpendicular diffusivity within the anterior and lateral columns, where the corticospinal tracts are located, independently predicted spasticity. Instability in the roll plane and diminished vibration sense were predicted by increased perpendicular

diffusivity in the posterior columns, where afferent sensory tracts conveying vibration sense and proprioception run. It is interesting that this effect emerges with the feet wider apart, when the body is normally more stable. It has previously been suggested that the increased stability with increasing stance width, in part, arises from hip proprioceptors being increasingly able to signal lateral sway, because of the mechanical linkage between hips and ankles (Day *et al.*, 1993), which may be degraded where there is posterior column pathology. When we examined the association between the imaging measures and postural stability, we found that higher ADCxy in the anterior column was associated with increased instability in the pitch plane, which implies that pitch plane abnormalities are predominantly linked to pathology of the tracts running in the anterior columns that mediate motor organisation or coordination. The coordination of joints is probably more demanding in the sagittal (pitch) plane, since there are more degrees of freedom due to independent action of leg joints. In contrast, in the frontal (roll) plane, the knees cannot contribute much to instability, while the ankle and hips are no longer independent (Day *et al.*, 1993).

For associations between imaging and clinical measures, we did not adjust for multiple comparisons since we were investigating a number of different hypotheses, and in such contexts correction can be inappropriate (Rothman, 1990; Perneger, 1998); nevertheless, as always there is a danger of spurious significant results, and p-values close to 0.05 should be interpreted with caution, and regarded as hypothesis-generating, to be examined in future studies.

Limitations and future directions

Although we have used state-of-the-art spinal cord sequences, there are a number of limitations of the current study that future work could try to address. Using a clinical scanner, our MRS protocol reliably quantified Glx (Glu + Gln) in the spinal cord for the first time in an MS patient group. Strategies for separating Glu and Gln at 3T such as TE-averaged PRESS have been developed and used in the brain (Hurd *et al.*, 2004; Hancu, 2009), but they may not be feasible in the spinal cord, using a 3T scanner as much larger voxel sizes would be needed. Future technical developments may make it possible to directly measure Glu with no Gln overlap in the spinal cord, which would allow a more specific evaluation of the role of Glu in MS pathophysiology in the spinal cord.

In addition, the smaller gradients and longer gradient pulses needed to perform QSI on a clinical scanner have the effect of narrowing the dPDF produced by q-space analysis, possibly leading to an under-estimate of the FWHM. It has been proposed that these should be considered as apparent values (Assaf *et al.*, 2005; Farrell *et al.*, 2008). Therefore direct comparison with previously published studies should be made with care, and only after taking into account differences in gradient settings.

It was beyond the scope of the current study to establish whether the QSI indices used in this study are more sensitive to spinal microstructural changes, than the more established DTI-derived indices such as fractional anisotropy (FA), radial (RD) and axial (AD) diffusivity. AnSome attempts to address this question hasve been made in the past. In 2002, Assaf *et al* examined 13 patients with MS using DTI and q-space imaging and demonstrated greater sensitivity of q-space metrics at detecting abnormalities in the normal appearing white matter and lesional brain tissue compared with FA (Assaf *et al.*, 2002). This finding was reproduced in a later study from the same group in 2005 (Assaf *et al.*, 2005), when they also showed that

q-space displacement values correlated strongly to NAA/Cr ratios suggesting they are highly specific for axonal loss.

A future longitudinal extension of the current study will investigate whether QSI and MRS measures are predictive of disability and cord atrophy at 1 year and 3 years. We will also examine whether the predictive accuracy can be improved by combing metabolic and structural metrics into a parametric model. Application of these new imaging techniques to patients with other MS subtypes is also required. This information may help to stratify patients for treatments and clinical trials on the basis of their spinal cord pathology and predicted clinical course. Further work is still needed to establish the relationship between QSI derived indices from the lateral columns and lateralised disability and to assess how closely longitudinal changes in imaging measures reflect clinical change in order to validate the use of these advanced spinal cord imaging protocols to provide potential imaging biomarkers for future clinical trials of neuroprotective agents.

References

Anaby, D., Duncan, I. D., Smith, C. M. &Cohen, Y. (2013). q-Space diffusion MRI (QSI) of the disease progression in the spinal cords of the Long Evans shaker: diffusion time and apparent anisotropy. *NMR Biomed* 26(12): 1879-1886.

Assaf, Y., Ben-Bashat, D., Chapman, J., Peled, S., Biton, I. E., Kafri, M., Segev, Y., Hendler, T., Korczyn, A. D., Graif, M. &Cohen, Y. (2002). High b-value q-space analyzed diffusion-weighted MRI: Application to multiple sclerosis. *Magnetic Resonance in Medicine* 47(1): 115-126.

Assaf, Y., Chapman, J., Ben-Bashat, D., Hendler, T., Segev, Y., Korczyn, A. D., Graif, M. &Cohen, Y. (2005). White matter changes in multiple sclerosis: correlation of q-space diffusion MRI and 1H MRS. *Magn Reson Imaging* 23(6): 703-710.

Assaf, Y. &Cohen, Y. (2000). Assignment of the water slow-diffusing component in the central nervous system using q-space diffusion MRS: Implications for fiber tract imaging. *Magnetic Resonance in Medicine* 43(2): 191-199.

Assaf, Y., Mayk, A. &Cohen, Y. (2000). Displacement imaging of spinal cord using q-space diffusion-weighted MRI. *Magnetic Resonance in Medicine* 44(5): 713-722.

Baker, E. H., Basso, G., Barker, P. B., Smith, M. A., Bonekamp, D. &Horska, A. (2008). Regional apparent metabolite concentrations in young adult brain measured by (1)H MR spectroscopy at 3 Tesla. *J Magn Reson Imaging* 27(3): 489-499.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed* 15(7-8): 435-455.

Bieniek, M., Altmann, D. R., Davies, G. R., Ingle, G. T., Rashid, W., Sastre-Garriga, J., Thompson, A. J. &Miller, D. H. (2006). Cord atrophy separates early primary progressive and relapsing remitting multiple sclerosis. *J Neurol Neurosurg Psychiatry* 77(9): 1036-1039.

Biton, I. E., Duncan, I. D. &Cohen, Y. (2006). High b-value q-space diffusion MRI in myelin-deficient rat spinal cords. *Magn Reson Imaging* 24(2): 161-166.

Bjartmar, C., Kidd, G., Mork, S., Rudick, R. &Trapp, B. D. (2000). Neurological disability correlates with spinal cord axonal loss and reduced N-acetyl aspartate in chronic multiple sclerosis patients. *Ann Neurol* 48(6): 893-901.

Blamire, A. M., Cader, S., Lee, M., Palace, J. &Matthews, P. M. (2007). Axonal damage in the spinal cord of multiple sclerosis patients detected by magnetic resonance spectroscopy. *Magnetic Resonance in Medicine* 58(5): 880-885.

Bodini, B., Cercignani, M., Toosy, A., Stefano, N. D., Miller, D. H., Thompson, A. J. &Ciccarelli, O. (2013). A novel approach with "skeletonised MTR" measures tract-specific microstructural changes in early primary-progressive MS. *Hum Brain Mapp*.

Bohannon, R. W. &Smith, M. B. (1987). Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther* 67(2): 206-207.

Brand, A., Richter-Landsberg, C. &Leibfritz, D. (1993). Multinuclear NMR studies on the energy metabolism of glial and neuronal cells. *Dev Neurosci* 15(3-5): 289-298.

Bunn, L. M., Marsden, J. F., Giunti, P. &Day, B. L. (2013). Stance instability in spinocerebellar ataxia type 6. *Mov Disord* 28(4): 510-516.

Callaghan, P. T., Eccles, C. D. &Xia, Y. (1988). NMR microscopy of dynamic displacements: k-space and q-space imaging. *Journal of Physics E: Scientific Instruments* 21(8): 820.

Caramanos, Z., Narayanan, S. &Arnold, D. L. (2005). *1H-MRS quantification of tNA and tCr in patients with multiple sclerosis: a meta-analytic review.*

Chard, D. T., Jackson, J. S., Miller, D. H. &Wheeler-Kingshott, C. A. (2010). Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. *J Magn Reson Imaging* 32(1): 223-228.

Ciccarelli, O., Altmann, D. R., McLean, M. A., Wheeler-Kingshott, C. A., Wimpey, K., Miller, D. H. &Thompson, A. J. (2010). Spinal cord repair in MS: does mitochondrial metabolism play a role? *Neurology* 74(9): 721-727.

Ciccarelli, O., Barkhof, F., Bodini, B., De Stefano, N., Golay, X., Nicolay, K., Pelletier, D., Pouwels, P. J., Smith, S. A., Wheeler-Kingshott, C. A., Stankoff, B., Yousry, T. &Miller, D. H. (2014). Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. *Lancet Neurol* 13(8): 807-822.

Ciccarelli, O., Thomas, D., De Vita, E., Wheeler-Kingshott, C., Kachramanoglou, C., Kapoor, R., Leary, S., Matthews, L., Palace, J., Chard, D., Miller, D., Toosy, A. &Thompson, A. (2013). Low myo-inositol indicating astrocytic damage in a case series of NMO. *Ann Neurol*.

Ciccarelli, O., Wheeler-Kingshott, C. A., McLean, M. A., Cercignani, M., Wimpey, K., Miller, D. H. &Thompson, A. J. (2007). Spinal cord spectroscopy and diffusion-based tractography to assess acute disability in multiple sclerosis. *Brain* 130(Pt 8): 2220-2231.

Confavreux, C., Vukusic, S., Moreau, T. &Adeleine, P. (2000). Relapses and progression of disability in multiple sclerosis. *N Engl J Med* 343(20): 1430-1438.

Cottrell, D. A., Kremenchutzky, M., Rice, G. P., Koopman, W. J., Hader, W., Baskerville, J. &Ebers, G. C. (1999). The natural history of multiple sclerosis: a geographically based study. 5. The clinical features and natural history of primary progressive multiple sclerosis. *Brain* 122 (Pt 4): 625-639.

Cutter, G. R., Baier, M. L., Rudick, R. A., Cookfair, D. L., Fischer, J. S., Petkau, J., Syndulko, K., Weinshenker, B. G., Antel, J. P., Confavreux, C., Ellison, G. W., Lublin, F., Miller, A. E., Rao, S. M., Reingold, S., Thompson, A. &Willoughby, E. (1999). Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 122 (Pt 5): 871-882.

Day, B. L., Steiger, M. J., Thompson, P. D. &Marsden, C. D. (1993). Effect of vision and stance width on human body motion when standing: implications for afferent control of lateral sway. *J Physiol* 469: 479-499.

Edden, R. A., Bonekamp, D., Smith, M. A., Dubey, P. &Barker, P. B. (2007). Proton MR spectroscopic imaging of the medulla and cervical spinal cord. *J Magn Reson Imaging* 26(4): 1101-1105.

Farrell, J. A., Smith, S. A., Gordon-Lipkin, E. M., Reich, D. S., Calabresi, P. A. &van Zijl, P. C. (2008). High b-value q-space diffusion-weighted MRI of the human cervical spinal cord in vivo: feasibility and application to multiple sclerosis. *Magn Reson Med* 59(5): 1079-1089.

Farrell, J. A., Zhang, J., Jones, M. V., Deboy, C. A., Hoffman, P. N., Landman, B. A., Smith, S. A., Reich, D. S., Calabresi, P. A. &van Zijl, P. C. (2010). q-space and conventional diffusion imaging of axon and myelin damage in the rat spinal cord after axotomy. *Magn Reson Med* 63(5): 1323-1335.

Fischer, J. S., Rudick, R. A., Cutter, G. R. &Reingold, S. C. (1999). The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult Scler* 5(4): 244-250.

Fox, R. J., Thompson, A., Baker, D., Baneke, P., Brown, D., Browne, P., Chandraratna, D., Ciccarelli, O., Coetzee, T., Comi, G., Feinstein, A., Kapoor, R., Lee, K., Salvetti, M., Sharrock, K., Toosy, A., Zaratin, P. &Zuidwijk, K. (2012). Setting a research agenda

for progressive multiple sclerosis: the International Collaborative on Progressive MS. *Mult Scler* 18(11): 1534-1540.

Gasparovic, C., Song, T., Devier, D., Bockholt, H. J., Caprihan, A., Mullins, P. G., Posse, S., Jung, R. E. &Morrison, L. A. (2006). Use of tissue water as a concentration reference for proton spectroscopic imaging. *Magn Reson Med* 55(6): 1219-1226.

Goodkin, D. E., Hertsgaard, D. &Seminary, J. (1988). Upper extremity function in multiple sclerosis: improving assessment sensitivity with box-and-block and nine-hole peg tests. *Arch Phys Med Rehabil* 69(10): 850-854.

Hancu, I. (2009). Optimized glutamate detection at 3T. *J Magn Reson Imaging* 30(5): 1155-1162.

Henning, A., Schar, M., Kollias, S. S., Boesiger, P. &Dydak, U. (2008). Quantitative magnetic resonance spectroscopy in the entire human cervical spinal cord and beyond at 3T. *Magn Reson Med* 59(6): 1250-1258.

Hobart, J. C., Riazi, A., Lamping, D. L., Fitzpatrick, R. &Thompson, A. J. (2003). Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). *Neurology* 60(1): 31-36.

Horsfield, M. A., Sala, S., Neema, M., Absinta, M., Bakshi, A., Sormani, M. P., Rocca, M. A., Bakshi, R. &Filippi, M. (2010). Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. *Neuroimage* 50(2): 446-455.

Hurd, R., Sailasuta, N., Srinivasan, R., Vigneron, D. B., Pelletier, D. &Nelson, S. J. (2004). Measurement of brain glutamate using TE-averaged PRESS at 3T. *Magn Reson Med* 51(3): 435-440.

Kaiser, L. G., Schuff, N., Cashdollar, N. &Weiner, M. W. (2005). Age-related glutamate and glutamine concentration changes in normal human brain: 1H MR spectroscopy study at 4 T. *Neurobiol Aging* 26(5): 665-672.

Kearney, H., Yiannakas, M. C., Abdel-Aziz, K., Wheeler-Kingshott, C. A., Altmann, D. R., Ciccarelli, O. &Miller, D. H. (2013). Improved MRI quantification of spinal cord atrophy in multiple sclerosis. *J Magn Reson Imaging*.

Kearney, H., Yiannakas, M. C., Samson, R. S., Wheeler-Kingshott, C. A., Ciccarelli, O. &Miller, D. H. (2014). Investigation of magnetization transfer ratio-derived pial and subpial abnormalities in the multiple sclerosis spinal cord. *Brain* 137(Pt 9): 2456-2468.

Khaleeli, Z., Ciccarelli, O., Manfredonia, F., Barkhof, F., Brochet, B., Cercignani, M., Dousset, V., Filippi, M., Montalban, X., Polman, C., Rovaris, M., Rovira, A., Sastre-Garriga, J., Vellinga, M., Miller, D. &Thompson, A. (2008). Predicting progression in primary progressive multiple sclerosis: a 10-year multicenter study. *Ann Neurol* 63(6): 790-793.

Khaleeli, Z., Sastre-Garriga, J., Ciccarelli, O., Miller, D. H. &Thompson, A. J. (2007). Magnetisation transfer ratio in the normal appearing white matter predicts progression of disability over 1 year in early primary progressive multiple sclerosis. *J Neurol Neurosurg Psychiatry* 78(10): 1076-1082.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33(11): 1444-1452.

Losseff, N. A., Webb, S. L., O'Riordan, J. I., Page, R., Wang, L., Barker, G. J., Tofts, P. S., McDonald, W. I., Miller, D. H. &Thompson, A. J. (1996). Spinal cord atrophy and disability in multiple sclerosis. A new reproducible and sensitive MRI method with potential to monitor disease progression. *Brain* 119 (Pt 3): 701-708.

Marliani, A. F., Clementi, V., Albini-Riccioli, L., Agati, R. &Leonardi, M. (2007). Quantitative proton magnetic resonance spectroscopy of the human cervical spinal cord at 3 tesla. *Magnetic Resonance in Medicine* 57(1): 160-163.

Marliani, A. F., Clementi, V., Albini Riccioli, L., Agati, R., Carpenzano, M., Salvi, F. &Leonardi, M. (2010). Quantitative cervical spinal cord 3T proton MR spectroscopy in multiple sclerosis. *AJNR Am J Neuroradiol* 31(1): 180-184.

Moffett, J. R., Ross, B., Arun, P., Madhavarao, C. N. &Namboodiri, A. M. (2007). N-Acetylaspartate in the CNS: from neurodiagnostics to neurobiology. *Prog Neurobiol* 81(2): 89-131.

Muhlert, N., Atzori, M., De Vita, E., Thomas, D. L., Samson, R. S., Wheeler-Kingshott, C. A., Geurts, J. J., Miller, D. H., Thompson, A. J. &Ciccarelli, O. (2014). Memory in multiple sclerosis is linked to glutamate concentration in grey matter regions. *J Neurol Neurosurg Psychiatry*.

Oh, J., Saidha, S., Chen, M., Smith, S. A., Prince, J., Jones, C., Diener-West, M., van Zijl, P. C., Reich, D. S. &Calabresi, P. A. (2013). Spinal cord quantitative MRI discriminates between disability levels in multiple sclerosis. *Neurology* 80(6): 540-547.

Ourselin, S., Roche, A., Prima, S. &Ayache, N. (2000).Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, Vol. 1935, 557-566 (Eds S. Delp, A. DiGoia and B. Jaramaz). Springer Berlin Heidelberg.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ* 316(7139): 1236-1238.

Polman, C. H., Reingold, S. C., Edan, G., Filippi, M., Hartung, H. P., Kappos, L., Lublin, F. D., Metz, L. M., McFarland, H. F., O'Connor, P. W., Sandberg-Wollheim, M., Thompson, A. J., Weinshenker, B. G. &Wolinsky, J. S. (2005). Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol* 58(6): 840-846.

Provencher, S. W. (1993). Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magn Reson Med* 30(6): 672-679.

Provencher, S. W. (2014).LCModel & LCMgui User's Manual. Vol. 2014http://s-provencher.com/pub/LCModel/manual/manual.pdf.

Ramio-Torrenta, L., Sastre-Garriga, J., Ingle, G. T., Davies, G. R., Ameen, V., Miller, D. H. &Thompson, A. J. (2006). Abnormalities in normal appearing tissues in early primary progressive multiple sclerosis and their relation to disability: a tissue specific magnetisation transfer study. *J Neurol Neurosurg Psychiatry* 77(1): 40-45.

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology* 1(1): 43-46.

Runmarker, B. &Andersen, O. (1993). Prognostic factors in a multiple sclerosis incidence cohort with twenty-five years of follow-up. *Brain* 116 (Pt 1): 117-134.

Sastre-Garriga, J., Ingle, G. T., Chard, D. T., Ramio-Torrenta, L., McLean, M. A., Miller, D. H. &Thompson, A. J. (2005). Metabolite changes in normal-appearing gray and white matter are linked with disability in early primary progressive multiple sclerosis. *Arch Neurol* 62(4): 569-573.

Schneider, T., Ciccarelli, O., Kachramanoglou, C., Thomas, D. L. &Wheeler-Kingshott, C. A. M. (2011).Reliability of tract-specific q-space imaging metrics in healthy spinal cord. In *ISMRM*Montreal.

Smith, S. A., Levante, T. O., Meier, B. H. &Ernst, R. R. (1994). Computer-Simulations in Magnetic-Resonance - an Object-Oriented Programming Approach. *Journal of Magnetic Resonance Series A* 106(1): 75-105.

Solanky, B. S., Abdel-Aziz, K., Yiannakas, M. C., Berry, A. M., Ciccarelli, O. &Wheeler-Kingshott, C. A. (2013). In vivo magnetic resonance spectroscopy detection of combined glutamate-glutamine in healthy upper cervical cord at 3 T. *NMR Biomed* 26(3): 357-366.

Stein, J., Narendran, K., McBean, J., Krebs, K. &Hughes, R. (2007). Electromyography-controlled exoskeletal upper-limb-powered orthosis for exercise training after stroke. *Am J Phys Med Rehabil* 86(4): 255-261.

Svens, B. &Lee, H. (2005). Intra- and inter-instrument reliability of Grip-Strength Measurements: GripTrack™ and Jamar® hand dynamometers. *The British Journal of Hand Therapy* 10(2): 47-55.

Tench, C. R., Morgan, P. S. &Constantinescu, C. S. (2005). Measurement of cervical spinal cord cross-sectional area by MRI using edge detection and partial volume correction. *J Magn Reson Imaging* 21(3): 197-203.

Thompson, A. J., Montalban, X., Barkhof, F., Brochet, B., Filippi, M., Miller, D. H., Polman, C. H., Stevenson, V. L. &McDonald, W. I. (2000). Diagnostic criteria for primary progressive multiple sclerosis: a position paper. *Ann Neurol* 47(6): 831-835.

Wansapura, J. P., Holland, S. K., Dunn, R. S. &Ball, W. S. (1999). NMR relaxation times in the human brain at 3.0 tesla. *Journal of Magnetic Resonance Imaging* 9(4): 531-538.

Wheeler-Kingshott, C. A., Stroman, P. W., Schwab, J. M., Bacon, M., Bosma, R., Brooks, J., Cadotte, D. W., Carlstedt, T., Ciccarelli, O., Cohen-Adad, J., Curt, A., Evangelou, N., Fehlings, M. G., Filippi, M., Kelley, B. J., Kollias, S., Mackay, A., Porro, C. A., Smith, S., Strittmatter, S. M., Summers, P., Thompson, A. J. &Tracey, I. (2014). The current state-of-the-art of spinal cord imaging: Applications. *Neuroimage* 84: 1082-1093.

Wilm, B. J., Svensson, J., Henning, A., Pruessmann, K. P., Boesiger, P. &Kollias, S. S. (2007). Reduced field-of-view MRI using outer volume suppression for spinal cord diffusion imaging. *Magnetic Resonance in Medicine* 57(3): 625-630.

Wu, Y. C., Field, A. S., Duncan, I. D., Samsonov, A. A., Kondo, Y., Tudorascu, D. &Alexander, A. L. (2011). High b-value and diffusion tensor imaging in a canine model of dysmyelination and brain maturation. *Neuroimage* 58(3): 829-837.

Yiannakas, M. C., Kearney, H., Samson, R. S., Chard, D. T., Ciccarelli, O., Miller, D. H. &Wheeler-Kingshott, C. A. (2012). Feasibility of grey matter and white matter segmentation of the upper cervical cord in vivo: A pilot study with application to magnetisation transfer measurements. *Neuroimage* 63(3): 1054-1059.

Zackowski, K. M., Smith, S. A., Reich, D. S., Gordon-Lipkin, E., Chodkowski, B. A., Sambandan, D. R., Shteyman, M., Bastian, A. J., van Zijl, P. C. &Calabresi, P. A. (2009). Sensorimotor dysfunction in multiple sclerosis and column-specific magnetization transfer-imaging abnormalities in the spinal cord. *Brain* 132(Pt 5): 1200-1209.

Figure 1: Planning of spectroscopy voxel and DWI volume. Above: sagittal (A) and coronal (B) T2w images of the cervical cord with spectroscopy voxel centred on C2/3 intervertebral disc. Below: sagittal (C) and coronal (D) T1w image of the cervical cord showing DWI volume coverage centred on the C2/3 disc.

Figure 2: Differences in height and width of the dPDF from the posterior and lateral columns between a healthy control and patient are shown on the far left. Grouped P0xy maps, FWHMxy maps and post-processed spectra from 3 controls (central) and 3 patients (far right) demonstrate lower probability of zero net displacement (P0xy) and increased diffusion distribution (FWHMxy) in the patients. The spectra show reduced tNAA and Glx levels in the patients compared to the controls.

Supplementary figure 1: Illustration of gradient direction scheme used for x and y QSI encoding. The QSI gradient directions are chosen to maximise the diffusion encoding gradient strength in the perpendicular plane to the spinal cord (red arrows).

Supplementary figure 2: Q-space imaging processing pathway. From top left to bottom right: The raw data points per voxel are re-gridded and then extrapolated using a bi-exponential fit. The inverse Fourier transformation is performed to give the probability density function, from which summary statistics are derived.

Supplementary figure 3: Axial b0 image of the cervical spinal cord showing the location of regions of interest (ROIs) placed in the anterior (A), right lateral (R), left lateral (L) and posterior (P) columns. After ROI's were drawn on the b0 images, they were overlaid onto the QSI and ADC maps.

Supplementary figure 4: Scatter graphs showing correlation between spinal tNAA concentration and whole cord P0xy (left), FWHMxy (centre) and ADCxy (right)

1

	Healthy Controls (n = 24)	PPMS Patients (n = 21)
Mean age (SD)	42.1 (11.5) years	48 (7.9) years
Gender	19F: 5M	12F: 9M
Mean CSA (SD)	81.8 (8.1) mm^2	77.5 (9.6) mm^2
Mean GMVF (SD)	0.48 (0.01)	0.47 (0.01)
Mean WMVF (SD)	0.34 (0.01)	0.33 (0.01)
Mean brain parenchymal fraction (SD)	0.82 (0.02)	0.80 (0.02)
Mean T2 lesion volume (SD)		11.6 (9.4) ml
Mean disease duration (SD)		3.9 (1.5) years
Median EDSS (range)		5.0 (3.0 - 6.5)
Mean TWT (SD)		8.1 (5.9) seconds
Mean MSWS-12 (SD)		44.4 (11.4)
Mean summated MAS (SD)		7.2 (9.3)
Mean HPT (SD)		30.0 (13.3) seconds
Mean grip strength (SD)		50.2 (26.4) lbs force
Mean vibration perception threshold (SD)		10.7 (10.6)
Mean sway, 32cm, EO (SD)		0.87 (0.37) deg/s
Mean sway, 32cm, EC (SD)		1.07 (0.46) deg/s
Mean sway, 4cm, EO (SD)		0.98 (0.38) deg/s
Mean sway, 4cm, EC (SD)		1.28 (0.58) deg/s

Abbreviations: 9 hole peg test (HPT); 25ft timed walk test (TWT); Cord surface area (CSA); Expanded disability status scale (EDSS); Eyes closed (EC); Eyes open (EO); Grey matter volume fraction (GMVF); Modified Ashworth score (MAS); MS walking scale (MSWS); Standard deviation (SD); White matter volume fraction (WMVF).

Table 1: Demographic, clinical and radiological characteristics of patients and volunteers

Metabolite	Healthy Controls (n = 24)	Patients without C1-3 lesion (n = 9)	Patients with C1-3 lesion (n= 12)	All Patients (n = 21)
tNAA (mmol/L)	5.31 (1.47)	4.23 (0.86) P=0.206	3.89 (1.31) P=0.102	**4.01 (1.16) P=0.020**
tCho (mmol/L)	1.31 (0.41)	1.12 (0.22) P=0.241	1.33 (0.38) P=0.852	1.26 (0.34) P=0.610
tCr (mmol/L)	3.76 (1.13)	3.04 (0.35) P=0.099	4.22 (1.73) P=0.908	3.79 (1.48) P=0.963
Ins (mmol/L)	4.49 (1.23)	4.25 (1.17) P=0.287	**6.26 (1.84) P=0.006**	5.55 (1.88) P=0.081
Glx (mmol/L)	5.93 (1.66)	5.01 (1.90) P=0.170	**4.50 (0.71) P=0.047**	**4.65 (1.11) P=0.043**
ADCxy ($\mu m^2/ms$)	0.390 (0.09)	0.421 (0.05) P=0.151	**0.481 (0.10) P=0.002**	0.454 (0.08) P=0.006
ADCz ($\mu m^2/ms$)	1.783 (0.10)	0.183 (0.01) P=0.123	0.183 (0.02) P=0.119	1.834 (0.14) P=0.123
FWHMxy ($\mu m \times 10^2$)	0.236 (0.02)	**0.251 (0.01) P=0.020**	**0.276 (0.04) P<0.001**	**0.265 (0.03) P=0.001**
FWHMz ($\mu m \times 10^2$)	0.550 (0.03)	0.553 (0.03) P=0.427	**0.560 (0.03) 0.019**	0.557 (0.03) P=0.120
P0xy (a.u)	0.202 (0.02)	**0.188 (0.01) P=0.025**	**0.174 (0.03) 0.001**	**0.180 (0.02) P=0.001**
P0z (a.u)	0.113 (0.004)	0.112 (0.003) P=0.278	0.113 (0.004) P=0.481	0.113 (0.004) P=0.470

Abbreviations: total N-acetylaspartate (tNAA); Choline containing compounds (tCho); myo-Inositol (Ins); Glutamate-Glutamine (Glx); Creatine + phosphocreatine (tCr). P values obtained using a linear regression analysis, correcting for age, gender and CSA.

Table 2: Summary of mean (SD) metabolite concentrations and QSI indices from the cervical cord of patients and controls and P-values for adjusted group comparisons after correcting for age, gender and mean cord cross-sectional area.

Region of interest	Diffusion measure		PPMS Patients (n=21)	Healthy Controls (n=24)	P-value
Anterior Column	ADCxy	(μm^2/ms)	0.497 (0.15)	0.388 (0.12)	**0.028**
	ADCz	(μm^2/ms)	1.921 (0.20)	1.867 (0.17)	0.331
	FWHMxy	($\mu m \times 10^2$)	0.266 (0.04)	0.229 (0.02)	**0.001**
	FWHMz	($\mu m \times 10^2$)	0.562 (0.03)	0.600 (0.04)	0.159
	P0xy	(a.u)	0.180 (0.03)	0.210 (0.02)	**0.002**
	P0z	(a.u)	0.111 (0.004)	0.111 (0.01)	0.411
Posterior Column	ADCxy	(μm^2/ms)	0.458 (0.16)	0.368 (0.10)	**0.017**
	ADCz	(μm^2/ms)	2.181 (0.26)	2.092 (0.14)	**0.050**
	FWHMxy	($\mu m \times 10^2$)	0.261 (0.06)	0.229 (0.03)	**0.029**
	FWHMz	($\mu m \times 10^2$)	0.610 (0.04)	0.602 (0.04)	0.122
	P0xy	(a.u)	0.185 (0.03)	0.208 (0.03)	**0.018**
	P0z	(a.u)	0.102 (0.004)	0.103 (0.004)	**0.045**
Mean Lateral Columns	ADCxy	(μm^2/ms)	0.416 (0.11)	0.319 (0.10)	**0.001**
	ADCz	(μm^2/ms)	1.989 (0.26)	1.979 (0.12)	0.581
	FWHMxy	($\mu m \times 10^2$)	0.254 (0.04)	0.214 (0.02)	**< 0.001**
	FWHMz	($\mu m \times 10^2$)	0.579 (0.02)	0.579 (0.03)	0.318
	P0xy	(a.u)	0.189 (0.03)	0.224 (0.03)	**< 0.001**
	P0z	(a.u)	0.108 (0.007)	0.106 (0.004)	0.757

Table 3: Summary of mean (SD) Q-space imaging (QSI) indices and apparent diffusion coefficients (ADC) from the major white matter columns of patients and controls. P-values given for adjusted group comparisons after correcting for age, gender and CSA.

Clinical Score	Spinal Cord Measure	Regression Coefficient	95 % Confidence Interval	P-value
EDSS	tNAA	-0.41	-1.06, 0.34	0.01 < P < 0.05 *
Summated MAS	tNAA	-3.78	-16.49, 2.16	0.01 < P < 0.05 *
	P0xy	-283.72	-444.26, -123.19	0.002
	FWHMxy	191.30	86.36, 269.24	0.001
	ADCxy	63.64	18.89, 103.39	0.008
Mean grip	P0xy	435.96	-61.67, 933.59	0.081
Vibration perception	tNAA	-4.37	-8.08, -0.66	0.021
threshold	P0xy	-344.27	-512.61, -175.93	0.001
	FWHMxy	226.49	115.73, 337.26	0.001
	ADCxy	88.06	49.02, 127.10	< 0.001

Abbreviations: 9 hole peg test (HPT); Choline containing compounds (tCho); Expanded disability status scale (EDSS); Modified Ashworth score (MAS); MS walking scale (MSWS); total N-acetylaspartate (tNAA).

Table 4: Associations between whole cord measures (predictors) and clinical scores (response variables). Unstandardised regression coefficients for imaging measures are reported with 95% confidence intervals and p-values. The regression models were adjusted for age, gender and mean cord area. * Bootstrap P-values.

	Sway Coefficient (95% CI; P-value)	Pitch Coefficient (95% CI; P-value)	Roll Coefficient (95% CI; P-value)	Romberg Quotient Coefficient (95% CI; P-value)
tNAA	**P < 0.001** 4EO: -0.057 (-0.188, 0.074; P = 0.393) 4EC: -0.192 (-0.416, 0.033; P = 0.094) 32EO: -0.072 (-0.118,-0.022; P = 0.004) 32EC: -0.086 (-0.221, 0.048; P = 0.208)	**P < 0.0001** 4EO: -0.036 (-0.144, 0.073; P = 0.517) 4EC: -0.137 (-0.277, 0.003; P = 0.056) 32EO: -0.047 (-0.099, 0.006; P = 0.082) 32EC: -0.063 (-0.192, 0.066; P = 0.338)	**P = 0.005** 4EO: -0.039 (-0.111, 0.033; P = 0.286) 4EC: -0.103 (-0.256, 0.049; P = 0.184) 32EO: -0.044 (-0.069,-0.018; P = 0.001) 32EC: -0.046 (-0.083,-0.009; P = 0.014)	**P = 0.003** 4cm: -0.112 (-0.191,-0.033; P = 0.006) 32cm: -0.191 (-0.123, 0.085; P = 0.718)
Glx	**P = 0.017** 4EO: -0.171 (-0.301, -0.042; P = 0.010) 4EC: -0.320 (-0.564, -0.076; P = 0.010) 32EO: -0.057 (-0.127, 0.014; P = 0.114) 32EC: -0.163 (-0.310,-0.015; P = 0.031)	**P < 0.001** 4EO: -0.110 (-0.227, 0.007; P = 0.065) 4EC: -0.204 (-0.360, -0.049; P = 0.010) 32EO: -0.033 (-0.103, 0.038; P = 0.366) 32EC: -0.134 (-0.279, 0.010; P = 0.069)	**P = 0.012** 4EO: -0.108 (-0.171, -0.046; P = 0.001) 4EC: -0.207 (-0.367, -0.046; P = 0.012) 32EO: -0.039 (-0.076,-0.002; P = 0.039) 32EC: -0.062 (-0.104, 0.020; P = 0.004)	**P < 0.0001** 4cm: -0.134 (-0.236,-0.032; P = 0.010) 32cm: -0.140 (0.232,-0.048; P = 0.003)
Ins	**P < 0.0001** 4EO: 0.062 (-0.061, 0.185; P = 0.324) 4EC: 0.100 (-0.102, 0.303; P = 0.332) 32EO: 0.019 (-0.065, 0.103; P = 0.660) 32EC: 0.090 (-0.010, 0.191; P = 0.078)	**P = 0.014** 4EO: 0.068 (-0.324, 0.170; P = 0.184) 4EC: 0.062 (-0.078, 0.201; P = 0.388) 32EO: 0.026 (-0.045, 0.097; P = 0.477) 32EC: 0.093 (0.004, 0.183; P = 0.040)	**P = 0.440**	**P = 0.046** 4cm: 0.031 (-0.052, 0.113; P = 0.467) 32cm: 0.086 (-0.019, 0.153; P = 0.012)
Cho	**P = 0.545**	**P = 0.144**	**P = 0.979**	**P = 0.113**
Cr	**P = 0.306**	**P = 0.237**	**P = 0.821**	**P = 0.363**

ADCxy	4EC: 2.48 (0.44, 4.51; P = 0.017)	32EO: 0.93 (0.34, 1.52; P = 0.002) 32EC: 1.24 (0.28, 2.21; P = 0.011)	4cm: 1.41 (0.24, 2.59; P = 0.023)
FWHMxy	4EC: 5.57 (0.02, 11.12; P = 0.049)	32EO: 2.25 (0.56, 3.94; P = 0.009)	
P0xy		32EO: -3.80 (-6.14, -1.46; P = 0.001) 32EC: -5.01 (-8.84,-1.18; P = 0.010)	

Abbreviations: regression coefficient (Coef.); 95% confidence interval (CI); total N-acetylaspartate (tNAA); Choline containing compounds (tCho); myo-Inositol (Ins); Glutamate-Glutamine (Glx); Creatine + phosphocreatine (Cr); Stance width of 32cm, eyes open (32EO); Stance width of 32cm, eyes closed (32EC); Stance width of 4cm, eyes open (4EO); Stance width of 4cm, eyes closed (4EC)

Table 5: Associations between whole cord imaging measures and truncal stability. A multivariate analysis was used to assess associations between metabolite predictors and the multiple stability scores as response variables. P-values <0.05 for the joint test of the metabolite predictor are shown in **bold**. Metabolite regression coefficients, 95% confidence intervals and p-values are shown for the individual stability variables, and these are only shown where the joint test was significant. The regression models adjusted for age, gender and mean cord area.

Clinical Score	Region of interest	Diffusion Measure	Regression Coefficient	95 % Confidence Interval	P-value
Summated MAS	Lateral Column	P0xy	-205.16	-322.34, -87.97	0.002
		FWHMxy	171.80	91.34, 252.27	< 0.001
		ADCxy	50.66	16.50, 84.83	0.006
	Anterior Column	P0xy	-185.96	-316.30, -55.62	0.008
		FWHMxy	157.38	64.03, 250.72	0.002
	Posterior Column	P0xy	-136.56	-267.09, -6.03	0.04
Vibration	Lateral Column	P0xy	-268.54	-421.03, -116.05	0.002
		FWHMxy	223.44	125.91, 320.90	< 0.001
		ADCxy	67.20	28.07, 106.33	0.002
	Anterior Column	P0xy	-239.01	-384.52, -93.50	0.003
		FWHMxy	200.74	106.07,295.41	0.001
		ADCxy	40.73	7.82, 73.64	0.02
	Posterior Column	P0xy	-219.76	-353.71, -85.80	0.003
		FWHMxy	101.77	32.47, 171.08	0.007
		ADCxy	46.29	22.24, 70.34	0.001
32EO Sway	Anterior Column	ADCxy	0.81	0.12, 1.50	0.021
32EO Roll	Posterior Column	P0xy	-2.16	-4.07, -0.26	0.026
32EO Pitch	Anterior Column	ADCxy	0.69	0.15, 1.23	0.013
32EC Roll	Posterior Column	P0xy	-3.85	-6.58, -1.12	0.006
32EC Pitch	Anterior Column	ADCxy	1.04	0.12, 1.94	0.026
4EO Pitch	Anterior Column	ADCxy	1.02	0.13, 1.91	0.025

4EC Pitch	Anterior Column	ADCxy	1.54	0.47, 2.63	0.005
4cm Romberg	Lateral Column	ADCxy	1.18	0.20, 2.16	0.023

Table 6: Showing associations between column-specific diffusion indices (predictors) and clinical scores (response variable). Unstandardised regression coefficients for imaging measures are reported with 95% confidence intervals and p-values. The regression models were adjusted for age, gender and mean cord area.