Supplementary methods

This section provides additional technical details on the sample, the applied imaging and analysis steps and methods.

Structural imaging

Trained radiographers placed all participants in the same position to obtain high reproducibility across participants and time points. This minimized any bias related to potential gradient non-linearity over time. The potentially confounding effect of upgrading the MRI scanner was mitigated by virtue of acquiring data from patients and controls before and after upgrade, allowing us to account for the upgrade effect (common to both cohorts).

At all time points, a 3D T1w scan (MPRAGE) was acquired at 1 mm isotropic resolution with 176 partitions using the following parameters: field of view of 224 x 256 mm^2, matrix size of 224 x 256, repetition time (TR) of 2420 ms, echo time (TE) of 4.18 ms, inversion time (TI) of 960 ms, flip angle α=9°, and readout bandwidth of 150 Hz per pixel. Scan time was 9 minutes.

At all time points, three differently contrast weighted 3D multi-echo FLASH volumes were acquired with 1mm isotropic resolution and a field of view of 240 x 256 mm^2 (matrix size of 240 x 256) with 176 partitions in a total scan time of 23 minutes. This produced the quantitative MPM data. Parallel imaging with a speed up factor of 2 was used in the phase-encoding direction (anterior-posterior) using a generalized auto-calibration partially parallel acquisition algorithm (GRAPPA) and a partial Fourier acquisition with a 6/8 sampling factor was used in the partition direction (left-right) to minimize acquisition time. Predominantly T1-weighting was achieved with TR=25 ms and α=23° while PD-weighting was achieved with TR=25 ms and α=4°.

For both the PD-weighted (PDw) and T1w volumes eight echoes were acquired at 19.68 ms. To achieve magnetisation transfer weighting (MTw) (TR=37 ms, α=9°) an off-resonance RF pulse prior to non-selective excitation was used. For the MTw volumes, seven equidistantly spaced echoes were acquired with TE ranging from 2.46 ms to 17.22 ms. The readout bandwidth was 480 Hz per pixel. Scan time was 23 minutes.

Longitudinal Image Processing and Analysis

Neurodegeneration within the cervical cord over 2 years

We measured the cross-sectional cord area at the C2/C3 level as this level represents the most reliable assessment site for cord area measurements using semi- or fully-automated segmentation methods[14]. We used the JIM 6.0 software (Xynapse systems, Aldwincle, UK) to extract from the structural T1w volume 10 contiguous and reformatted axial slices of 3 mm. An active-surface model [15] (Figure 1A) was then applied to calculate the cross-sectional cord area automatically. The anterior-posterior width (APW (elliptical short axis)) and the left-right width (LRW (elliptical long axis)) were then extracted based on an ellipse fitted to the boundary of the cord area using in-house Matlab scripts (The Mathworks Inc., Natick, MA, USA). To assess myelin sensitive changes in the MT maps at the identical cervical cord level, we used in-house Matlab scripts based on nearest-neighbour region growing to define the cord volume and used the same ellipse fitting procedure as described above. This produced a region of interest, which was then superimposed on the MT maps to extract the mean MT from the cross-sectional area of the cervical cord.

Neurodegeneration within the brain over 2 years

Because longitudinal MR-based morphometry is prone to artifacts due to scanner

inhomogeneities, registration inconsistency, and subtle age-related deformations of the brains, it requires specific preprocessing pipelines in order to expose changes of interest and provide efficient estimates of effect sizes.

First, this pipeline begins with spatial registration for longitudinal MRI data[19]. Registration combines non-linear and rigid-body registration and further corrections for intensity bias artifacts. The procedure uses a generative model how the images are generated – and provides consistent estimates of within-subject brain deformations over the study period. The registration model also creates a midpoint T_1-weighted image for each subject and the corresponding deformation fields for every scan. The midpoint image encodes the average brain shape of all scans acquired from a participant.

Second, we applied SPM12's unified segmentation to each subject's midpoint T_1 image, which assumes every voxel is drawn from an unknown mixture of six distinct tissue classes: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF), bone, other tissue and air [20].

Third, nonlinear template generation and image registration[21] was applied to subject-specific midpoint GM and WM tissue maps – and the template was registered to MNI space using an affine transform. This iterative procedure optimises the template for the group in question. The resulting template (in MNI space) is then used for subsequent analysis. For longitudinal VBM, the normalised tissue segments from all subjects and time-points were modulated to account for local volume changes over time (explicitly modelled in step one). Since we were interested in tissue compartment-specific volume changes, ‘modulation’ is used to preserve the original volume following image deformation during normalization. This modulation entails multiplying the tissue segments with the local volume expansion or contraction (encoded in the determinant of the Jacobian tensor of the normalising deformation).

All quantitative MPM maps from all time-points were co-registered with the corresponding T_1-weighted images in native space and subsequently normalized to the MNI space using transformations obtained in previous steps.
Finally, morphometric images were smoothed using Gaussian kernels of 6mm full width at half maximum and quantitative parameter images were smoothed using a previously established tissue-weighted-smoothing procedure with a kernel of 3 mm, in order to preserve quantitative values within GM and WM tissue classes [22].
Subsequent modelling and analysis was performed for smoothed, normalised morphometric and quantitative parameter images within specific brain areas defined below.

Statistical analysis

Cord: Stata 13 (StataCorp LP, TX) was used to statistically assess recovery and change in spinal cord MRI indices. Rates of change of cord area, LRW, APW, mean cord MT (all subjects) and recovery (patients only) were estimated with linear mixed effects models – with the MRI parameter, clinical measure as response variable and time as predictor. In all models, group and group x time interaction terms were included to assess patient trajectories with respect to controls adjusting for potentially confounding effects of age and gender. Nonlinear clinical trajectories were modelled using a log time scale. To characterise structural trajectories, we modelled rate of change and time-dependant changes in the rate of change. Technically, these two effects were modelled using second order Taylor expansion of time. This is equivalent to a polynomial expansion in terms of linear and quadratic time effects. The first effect corresponds to a progressive change while a positive quadratic effect models a slowing down or deceleration of putative markers of neurodegeneration. Finally, we used regression models to identify associations between anatomical

changes by 6 months and 2-year clinical outcome measures, adjusting for potentially confounding effects of age and clinical change between six months and baseline. Note that for one patient the SCIM score was not available at 6 months and instead the 2 months score was used. Sub-acute MRI changes (at 6 months) were chosen because fastest decline was observed during this time period [5].

Brain: We used SPM12 to analyse group differences of structural trajectories (see http://www.fil.ion.ucl.ac.uk/spm/ for technical paper references and documentation). In particular, we applied a random-effects analysis using parametric t-tests of the rate of change parameter of the individual trajectories. This is follows the conservative two-stage summary statistics approach [25] commonly used in fMRI and longitudinal image analysis [26]. That means in a first stage (called 'fixed-effects-analysis'), we estimated individual quadratic trajectory models $y(t) = \beta_0 + \beta_1 t + \beta_2 t^2$ and obtained intercepts (β_0), rate of change (β_1), quadratic effects (β_2) and time since injury (t) for all subjects in the sample independently. The second stage ('random-effects-analysis') is then performed to make inferences on these parameters on the level of the population. In this second stage, we used two-sample parametric t-tests (for all voxels within each ROI), comparing the model parameters across groups, while adjusting for age and sex as covariates of no interest. More specifically, we applied one-tailed t-statistics to test for linear (e.g. $\beta_1 < 0$ indicating decline) and quadratic (e.g. $\beta_2 > 0$ indicating deceleration) structural changes over the study period. As implemented in SPM, the associated p-values from peak effects of the statistical parametric maps were corrected for multiple comparisons using random field theory (RFT) within each considered ROI. Cluster significance was tested after applying a conservative cluster-forming threshold of p=0.001. We additionally used SPM's multiple linear regression models to test for associations between brain changes, lesion level and clinical recovery in patients adjusting for potentially confounding

effects of age and clinical change between six months and baseline. The explanatory variables were lesion level and clinical outcome, while the response variables were changes in structural markers over the first six months.