NeuroImage: Clinical 11 (2016) 408–414

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Structural and diffusion imaging versus clinical assessment to monitor amyotrophic lateral sclerosis

Arturo Cardenas-Blanco[a], Judith Machts[a], Julio Acosta-Cabronero[a], Joern Kaufmann[b], Susanne Abdulla[a,b,c], Katja Kollewe[c], Susanne Petri[c], Stefanie Schreiber[a,b], Hans-Jochen Heinze[a,b,d], Reinhard Dengler[c], Stefan Vielhaber[a,b], Peter J. Nestor[a,*]

[a]German Center for Neurodegenerative Diseases (DZNE), Leipziger Strasse 44, 39120 Magdeburg, Germany
[b]Department of Neurology, Otto-von-Guericke University, Leipziger Strasse 44, 39120 Magdeburg, Germany
[c]Department of Neurology, Hannover Medical School, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany
[d]Leibniz Institute for Neurobiology, Brenneckestrasse 6, 39118 Magdeburg, Germany

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form 24 February 2016
Accepted 14 March 2016
Available online 16 March 2016

Keywords:
Longitudinal
Diffusion
MRI
ALSFRS-R
ALS
Biomarker

ABSTRACT

Amyotrophic lateral sclerosis is a progressive neurodegenerative disease that affects upper and lower motor neurons. Observational and intervention studies can be tracked using clinical measures such as the revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) but for a complete understanding of disease progression, objective in vivo biomarkers of both central and peripheral motor pathway pathology are highly desirable. The aim of this study was to determine the utility of structural and diffusion imaging as central nervous system biomarkers compared to the standard clinical measure, ALSFRS-R, to track longitudinal evolution using three time-point measurements. $N = 34$ patients with ALS were scanned and clinically assessed three times at a mean of three month time intervals. The MRI biomarkers were structural T1-weighted volumes for cortical thickness measurement as well as deep grey matter volumetry, voxel-based morphometry and diffusion tensor imaging (DTI). Cortical thickness focused specifically on the precentral gyrus while quantitative DTI biomarkers focused on the corticospinal tracts. The evolution of imaging biomarkers and ALSFRS-R scores over time were analysed using a mixed effects model that accounted for the scanning interval as a fixed effect variable, and, the initial measurements and time from onset as random variables. The mixed effects model showed a significant decrease in the ALSFRS-R score, ($p < 0.0001$, and an annual rate of change (AROC) of -7.3 points). Similarly, fractional anisotropy of the corticospinal tract showed a significant decrease ($p = 0.009$, AROC $= -0.0066$) that, in turn, was driven by a significant increase in radial diffusivity combined with a trend to decrease in axial diffusivity. No significant change in cortical thickness of the precentral gyrus was found ($p > 0.5$). In addition, deep grey matter volumetry and voxel-based morphometry also identified no significant changes. Furthermore, the availability of three time points was able to indicate that there was a linear progression in both clinical and fractional anisotropy measures adding to the validity of these results. The results indicate that DTI is clearly a superior imaging marker compared to atrophy for tracking the evolution of the disease and can act as a central nervous biomarker in longitudinal studies. It remains, however, less sensitive than the ALSFRS-R score for monitoring decline over time.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of unknown aetiology that causes degeneration of upper and lower motor neurons. Establishing biomarkers to help in the development of therapeutic agents, has become a key goal of clinical research. In recent years, 14 longitudinal MRI studies focused on tracking disease progression in ALS. Eight studies used diffusion tensor imaging (DTI) (Agosta et al., 2009a; Agosta et al., 2010; Blain et al., 2007; Keil et al., 2012; Menke et al., 2012; Muller et al., 2012; van der Graaff et al., 2011; Zhang et al., 2011). Four studies focused on structural measures using

* Corresponding author.
E-mail addresses: Arturo.Cardenas@dzne.de (A. Cardenas-Blanco), Judith.Machts@dzne.de (J. Machts), Julio.Acosta@dzne.de (J. Acosta-Cabronero), jkauf@neuro2.med.uni-magdeburg.de (J. Kaufmann), susanne.abdulla@med.ovgu.de (S. Abdulla), Kollewe.Katja@mh-hannover.de (K. Kollewe), Petri.Susanne@mh-hannover.de (S. Petri), stefanie.schreiber@med.ovgu.de (S. Schreiber), Hans-Jochen.Heinze@med.ovgu.de (H.-J. Heinze), Dengler.Reinhard@mh-hannover.de (R. Dengler), Stefan.Vielhaber@dzne.de (S. Vielhaber), Peter.Nestor@dzne.de (P.J. Nestor).

http://dx.doi.org/10.1016/j.nicl.2016.03.011

various approaches: cortical thickness (Schuster et al., 2014; Verstraete et al., 2012); tensor based morphometry (TBM) (Agosta et al., 2009b); and volumetry of deep grey matter structures (Westeneng et al., 2015). Only two studies compared structural and diffusion imaging in the same cohort (Kwan et al., 2012; Menke et al., 2014). Results are quite inconsistent across studies at present. A particular problem, potentially, is that all published studies used only two time-points making it difficult to judge if decline is linear or otherwise, and moreover, if some conflicting results may have been spurious—with greater than two time-points, in contrast, one can assess whether a coherent pattern of change is emerging over time as opposed to apparently 'significant' changes that may represent random error. Furthermore, longitudinal studies to date have typically contained small numbers—only five (Menke et al., 2014; Schuster et al., 2014; Verstraete et al., 2012; Verstraete et al., 2014; Westeneng et al., 2015) of the above studies listed included 20 or more patients. Finally, with most studies investigating only a single biomarker in isolation, the question arises of whether an apparently significant change in such instances adds true value—i.e. a change may be statistically significant, but, if the effect is order(s) of magnitude below other outcome measures, of little value. For further discussion of the challenges facing imaging studies in ALS, see Verstraete et al. (2015).

With all of these considerations in mind, the present study investigated $n = 34$ patients with ALS who were scanned on three occasions at an average of approximately three month intervals. The performance of DTI metrics and structural imaging—analysed using both the cortical thickness approach, VBM and automated deep grey matter volumetry—was assessed and also contrasted to the standard clinical outcome measure, the revised ALS Functional Rating Scale (ALSFRS-R) (Cedarbaum et al., 1999).

2. Materials and methods

2.1. Participants

Patients were recruited from specialist ALS clinics as part of a prospective study that has recruited $N = 125$ cases to date. Clinical diagnosis of ALS was made according to the revised El Escorial criteria (Brooks et al., 2000) with all patients fulfilling criteria for clinically definite or probable ALS. Patients suffering from flail limb or upper motor neuron only and/or showing symptoms of any of the frontotemporal lobar degeneration syndromes were excluded; this was established on clinical grounds, including caregiver interview. The Montreal cognitive assessment (MoCA) (Nasreddine et al., 2005) score was used to assess cognitive performance. Patients needed to have had three MRI examinations using the research scan protocol. Of those meeting these criteria, $N = 38$, four were excluded because they lacked at least one image modality at one time point, leaving $N = 34$ patients with complete data sets at each time-point, ALS-TP1, ALS-TP2, ALS-TP3. $N = 31$ patients had a classic (Charcot's) phenotype and $N = 3$ had a pyramidal phenotype (Chiò et al., 2011). ALSFRS-R (Cedarbaum et al., 1999) severity score was assessed by the same experienced neurologist on all occasions (SV) and to derive ALS Milano-Torino Staging (ALS-MITOS) scores (Chiò et al., 2015). Demographics are summarised in Table 1.

For some imaging comparisons, 29 healthy control subjects were recruited and screened to exclude neurological illness and cognitive impairment (MoCA ≥ 26). All subjects gave written informed consent; the ethics committee of Otto-von-Guericke University approved the study.

2.2. Image acquisition

All MRI scans were performed on the same Siemens Verio 3T system (Siemens Medical Systems, Erlangen, Germany) equipped with a gradient coil capable of 45 mT/m and 200 T/m/s slew rate. A standard 32-channel phased array imaging coil was used in receive mode. The field of view was aligned in all cases to the anterior commissure–posterior commissure line.

The DTI acquisition had a resolution of $2 \times 2 \times 2$ mm^3 and consisted of diffusion weighted data along 30 non-collinear diffusion directions with b = 1000 s/mm^2, and one scan without diffusion weighting (b = 0 s/mm^2). Full details of the acquisition scheme have been previously published (Cardenas-Blanco et al., 2014). T1-weighted, high-resolution structural MRI images were obtained using a three dimensional magnetization prepared rapid acquisition gradient-echo (MPRAGE) sequence with the following parameters: echo time/repetition time = 4.82/2500 ms, inversion time = 1100 ms, flip angle = 7°, receiver bandwidth = 140 Hz/pixel, distance factor 50% and a matrix size of $256 \times 256 \times 192$, yielding an isotropic resolution of 1 mm^3. A T2-weighted FLASH sequence acquired during the same session was used to exclude vascular pathology (no vascular lesions were identified in the dataset).

2.3. Image analysis

2.3.1. Diffusion tensor imaging

Diffusion tensor images were processed using The Oxford Centre for Functional MRI of the Brain (FMRIB) software library (Smith et al., 2004). Each diffusion weighted volume was affined-aligned to its corresponding b0 image using FMRIB's linear image co-registration tool (FLIRT v5.4.2) (Jenkinson and Smith, 2001) to correct for motion artefacts and eddy-current distortions. A binary brain mask of each b0 image was created, using the brain-extraction tool (BET v2.1) (Smith, 2002) with fractional threshold f = 0.1 and vertical gradient g = 0. FMRIB's diffusion toolbox (FDT v2.0) was used to fit the tensor and compute the eigenvalues L1 (axial diffusivity), L2 and L3 at each brain voxel and generate mean diffusivity (MD), fractional anisotropy (FA) and radial diffusivity (RD). Whole-brain analyses were performed using tract-based spatial statistics (TBSS). Spatial normalisation was achieved by warping all FA images to the $1 \times 1 \times 1$ mm^3 FMRIB58_FA standard template (FMRIB, University of Oxford, UK) in MNI152 space (Montreal Neurological Institute, McGill University, Canada) using FMRIB's nonlinear registration tool (FNIRT v1.0).

A cross-sectional analysis comparing ALS-TP3 and controls was completed to map the distribution of DTI changes. All warped ALS-TP3 and control FA maps were averaged to create a mean FA template, from which the mean FA skeleton was derived, using FA > 0.2. Finally, all spatially normalised FA, axial diffusivity (L1), RD and MD data were projected onto the skeleton and non-parametric statistics applied, where 10,000 permutations were run using randomize v2.1 with threshold free cluster

Table 1
Study participant demographics.

	Controls ($N = 29$)	ALS-TP1 ($N = 34$)	ALS-TP2 ($N = 34$)	ALS-TP3 ($N = 34$)
M/F	23/6	22/12	22/12	22/12
Age (years)	61.8 (10)	57.3 (9.9)	57.6 (9.9)	58.0 (9.9)
Symptom duration (mo)	–	23.6 (21.0)	27.0 (20.8)	31.3 (21.3)
ALSFRS-R score (/48)	–	40.2 (4.4)	37.9 (5.3)	35.1 (6.4)
ALS-MITOS stage 1[a]	–	2	3	7
MOCA (/30)	27.0 (0.8)	25.5 (2.1)	25.9 (3.0)	26.9 (2.7)

[a] Number of patients with an ALS-MITOS score of 1, all other patients scored zero; no patients progressed beyond a score of 1 over the course of the study.

enhancement (TFCE) enabled. The TFCE output was corrected for multiple comparisons by controlling the family wise error rate (FWE). The threshold level used was $p_{FWE} < 0.05$.

For the longitudinal analysis, a region of interest (ROI) comprising the left and the right corticospinal tracts was manually delineated in standard space, using the FMRIB58_FA template following a previously published method (Cardenas-Blanco et al., 2014). This mask was then intersected with the TBSS mean skeleton generated from all subjects (FA > 0.2) and mean values for FA, L1, RD and MD were extracted for each time point.

2.3.2. Structural imaging

2.3.2.1. Cortical thickness measurement. Cortical thickness measurements were obtained using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/; version 5.3.0) (Fischl and Dale, 2000). A whole brain cross-sectional analysis was done to contrast ALS-TP3 and controls using a Gaussian smoothing kernel of 20 mm, regressing out the effects of age and applying a clusterwise correction for multiple comparisons (Hagler et al., 2006). The longitudinal Freesurfer pipeline was used to analyse intra-subject cortical thinning (Reuter et al., 2010) over the three time points in the ALS cohort. To reduce bias from a specific time point, a template volume using all time points from each subject was generated and run through FreeSurfer and was later used as an initial estimate for surface reconstruction and cortical segmentation for each time-point in each subject (Reuter et al., 2012). The longitudinal analysis of cortical thickness was focused on the mean of the right and left precentral gyrus, which was extracted as an ROI using the Desikan-Killiany parcellation atlas (Desikan et al., 2006).

2.3.2.2. Voxel based morphometry. Voxel based morphometry measurements were carried out using two different software packages, FSL and SPM—both methods were undertaken to address the reported method dependent nature of VBM analyses (Diaz-de-Grenu et al., 2014, Rajagopalan et al., 2014). In both cases, the analysis contrasted ALS-TP3 and ALS-TP1.

2.3.2.2.1. FSL-VBM. This analysis precisely followed that of a previous study that reported widespread VBM changes in a longitudinal ALS analysis (Menke et al., 2014). Each subject's TP1 and TP3 MPRAGE images were bias corrected, brain extracted and linear transformation matrices to a halfway space were generated. TP1 and TP3 images in native space were then linearly registered to their corresponding halfway space and averaged. The bias-field corrected TP1 and TP3 images were then non-linearly registered to their corresponding average image in halfway space (Technical report TR07JA1, http://www.fmrib.ox.ac.uk/analysis/techrep). In parallel, partial volume grey matter images were obtained after applying FAST4 segmentation algorithm to the bias-field corrected TP1 and TP3 images in native space (Zhang et al., 2001). Grey matter partial volume images were then first linearly registered to MNI space and then non-linearly registered to MNI space and averaged, to obtain a study specific longitudinal template in standard space. Grey matter TP1 and TP3 images in native space were non-linearly registered into halfway space and their average non-linearly registered to the study specific template. A composition of the last two non-linear warps was used to register the native TP1 and TP3 grey matter images into the longitudinal study specific grey matter template. The Jacobian determinant of the composition of the two warp fields was used to modulate the partial volume grey matter images. Finally, the modulated segmented partial volume images were then smoothed with an isotropic Gaussian kernel of 3 mm standard deviation. Within-subject smoothed and modulated image differences were estimated by subtracting the TP1 from TP3. A one-sample *T*-test, across subjects using GLM was applied with 5000 permutations using randomize v2.1 with threshold free cluster enhancement (Smith and Nichols, 2009) (TFCE) enabled. The TFCE output was corrected for multiple comparisons by controlling the FWE, using a threshold $p < 0.05$.

2.3.2.2.2. SPM-VBM. Structural TP1 and TP3 images were segmented with the *new segment* method and then imported into DARTEL (SPM12) (http://www.fil.ion.ucl.ac.uk/spm/), where all grey and white matter segments were simultaneously registered to a study specific template. A new template defined as the average of the inverse transformation of all images into the template was repeated iteratively 18 times before modulation. Modulated grey matter segments were smoothed using a 7 mm full width at half maximum Gaussian kernel. Note the size of the smoothing kernel was adjusted to make it equal to the one used in FSL-VBM, given that the full width at half maximum of the kernel equals 2.35 × standard deviation. As for FSL-VBM, within subject smoothed and modulated image differences were estimating using the TP1 and TP3. A one-sample *T*-test was completed using an uncorrected threshold of $p < 0.05$.

2.3.2.3. Volumetric measurements of deep grey matter structures. Volumetric measurements of deep grey matter structures were measured because a previous study (Menke et al., 2014) had reported extensive longitudinal changes in the basal ganglia and diencephalon with FSL-VBM. The volumetric measures, therefore, served as an independent (of VBM) method to verify such findings. Volumes were extracted for caudate nucleus, thalamus, putamen, pallidum, amygdala and hippocampus. Volumes were obtained using the segmentation and registration tool FIRST (Patenaude et al., 2011). Following segmentation, measured volumes were corrected for total intracranial volume (TIV) using the covariance method (Jack et al., 1989). TIV was estimated as the sum of the first three tissue classes, grey matter, white matter and CSF, thresholded at 0.5 (Pengas et al., 2009).

2.4. Statistical analyses

Quantitative ROI data (DTI metrics (FA, MD, RD, L1), cortical thickness and deep grey matter volumetry) was tested for normality using the Lilliefors test. The results of the test for $\alpha = 0.05$ indicated that RD was not normally distributed in ALS, therefore, for consistency, the Kruskal-Wallis H test was used thereafter. In those cases where a significant difference among groups was found, a nonparametric Mann-Whitney *U* test was used to compare all derived metrics from independent sample groups (Mann and Whitney, 1947). Bonferroni correction for multiple comparisons was applied.

Assessment of longitudinal change in the quantitative ROI data and disease severity score employed a mixed-effects model, designed to take account of the scanning interval as a fixed effect variable and the initial measurements, as well as the time from onset as random variables; the significance level for the longitudinal analyses was set to $p < 0.05$.

Where significant longitudinal changes in ALS were identified, a power calculation was completed to estimate the minimum number of subjects needed in a hypothetical study to measure a 25% effect size following the procedure described by Diggle et al. (2002), using the rate of change derived from the longitudinal analyses as input.

3. Results

3.1. Clinical score

A Kruskal-Wallis H test was statistically significant ($p = 0.001$) for the three time-points of the ALSFRS-R score (Fig. 1). Post-hoc Mann-Whitney tests found significant differences between ALS-TP1 and ALS-TP3 ($p = 0.0004$). The mixed effect model showed a significant ($p < 0.0001$) annual rate of change of −7.3 points (Table 2).

3.2. Diffusion tensor imaging

3.2.1. Cross-sectional analysis, controls vs ALS-TP3

Contrasting the quantitative DTI metrics from ALS-TP3 and controls at $p < 0.05$ corrected for multiple comparisons as a whole brain analysis found no significant differences. Relaxing the statistical threshold to

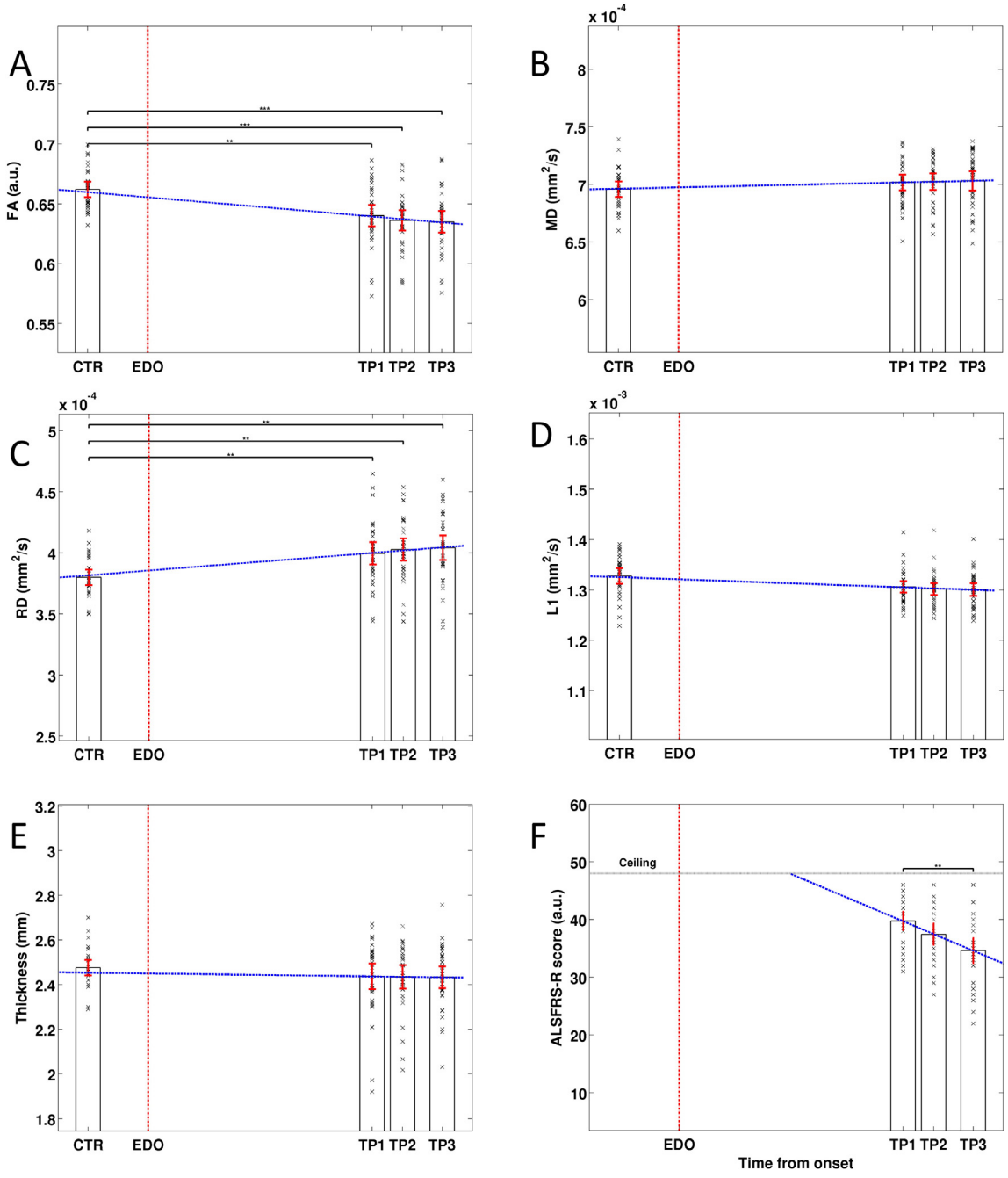

Fig. 1. Evolution of DTI, cortical thickness and ALSFRS-R over time. A–D: DTI metrics from the corticospinal ROI; E: cortical thickness of the precentral gyrus; F: ALSFRS-R score. Bars denote means, whiskers the 95% confidence interval and crosses are individual data points. *p <= 0.002, **$p < 0.001$, ***$p < 0.0001$. The x-axis is drawn to scale such that the three ALS time-points (TP1–3) represent the mean (in time) of each measurement spaced relative to each other and to the mean estimated disease onset (EDO). Blue lines indicate slope of change from TP1–3.

$p < 0.02$ (uncorrected) identified reduced FA values in ALS in the corticospinal tracts (Fig. 2).

3.2.2. Longitudinal corticospinal tract ROI analysis

The Kruskal-Wallis H test showed statistically significant differences between controls, ALS-TP1, ALS-TP2 and ALS-TP3 in FA ($p = 0.002$) and RD ($p < 0.0001$). No significant differences were found in MD ($p = 0.4$) while a trend to significance in L1 ($p = 0.006$) did not survive Bonferroni correction. Post-hoc Mann-Whitney tests found significant reductions in FA and increases in RD values between controls and all three ALS time points (all $p < 0.0001$). No significant differences were detected in any DTI metric within ALS time points (Fig. 1).

The mixed effects model (Table 2) showed a significant decrease in FA in the corticospinal tract ($p = 0.009$) with a predicted annual rate of change of −0.0066. Similarly, a significant increase in RD in the corticospinal tract was detected ($p = 0.04$). No significant effects were detected for MD while there was non-significant trend for L1 ($p = 0.07$).

3.3. Structural imaging

3.3.1. Cortical thickness measurements

No significant differences in cortical thickness were found at a whole brain level between ALS-TP3 and controls after correcting for multiple comparisons. The Kruskal-Wallis H test of the longitudinal ALS data

Table 2
Summary of linear mixed effects models predicting the ratio of change. AROC: annual rate of change; SE: standard errors (SE); DF: degrees of freedom.

Measure		AROC	SE	t-stat	DF	*P*-value
ALSFRS-R score (/48)		−7.3	0.73	−7.87	100	<0.000[a]
Diffusion:	FA (a. u.)	−0.0066	0.002	−2.65	100	0.009[a]
cortico-spinal tract	MD (mm^2/s)	1.7×10^{-6}	2.8×10^{-6}	0.59	100	0.55
	L1 (mm^2/s)	-7.5×10^{-6}	4.1×10^{-6}	−1.82	100	0.07
	RD (mm^2/s)	6.1×10^{-6}	2.0×10^{-6}	2.09	100	0.038[a]
Precentral gyrus thickness (mm)		−0.02	0.03	−0.62	100	0.54
Volumetry (mm^3)	Amygdala	−25.6	69.4	−0.41	100	0.68
	Caudate	182.5	109.5	1.69	100	0.09
	Hippocampus	11.0	98.6	0.12	100	0.90
	Pallidum	11.0	51.1	0.28	100	0.78
	Putamen	120.5	102.2	1.15	100	0.25
	Thalamus	142.7	84.0	−1.67	100	0.09

[a] Significant changes.

showed no significant changes in cortical thickness measurements in the precentral gyrus between controls and ALS subjects at different time points at the Bonferroni-corrected threshold of $p = 0.002$ (Fig. 1). The mixed effects model showed no significant changes in cortical thickness in the precentral gyrus over time (Table 2).

3.3.2. Voxel based morphometry

Neither the SPM nor the FSL analysis yielded any significant differences in grey matter density between ALS-TP1 and ALS-TP3.

3.3.3. Deep grey matter volumetry

No significant changes in volume between groups were detected at $p = 0.002$ Bonferroni corrected (Supplementary Fig. 1). No significant differences in volume of deep grey matter structures were detected across the three ALS time points using the mixed effects model (Table 2).

3.4. Power analysis

Using the significant annual rate of change for ALSFRS-R and FA obtained by means of the mixed effects model, the sample size needed to detect a hypothetical treatment effect of 25% change in slope in ALSFRS-R considering 80% power and an alpha level of 0.05 in a two arm clinical trial would be 94 subjects per arm; for FA, the number of subjects needed per arm would be 567.

4. Discussion

Using three time-point measurements over a mean total interval of six months to track progression in ALS, DTI metrics of the corticospinal tract emerged as sensitive biomarkers to detect change. In contrast, structural imaging, be that cortical thickness of the entire isocortex or motor cortex in particular, deep grey matter volumetry or whole brain VBM, was unable to detect change. Although DTI, especially FA, revealed a significant, linear decline over the three measurements, the ratio between the annual rate of change and the standard errors, estimated using the mixed effects model, as well as the results from the power analysis indicated that it was considerably less sensitive than the standard clinical outcome measure, the ALSFRS-R.

Longitudinal DTI results have been inconsistent in past studies though this almost certainly can be explained by technical development. The earliest studies, were insensitive to disease progression (Blain et al., 2007; Agosta et al., 2009a; Agosta et al., 2010) while later studies have generally reported DTI to be sensitive to change (van der Graaff et al., 2011; Zhang et al., 2011; Keil et al., 2012; Menke et al., 2012; Muller et al., 2012). The negative results from earlier studies came from datasets using lower field strength (1.5 T) with poorer resolution and signal to noise ratio; arguably, analysis methods for DTI have also become more reliable in recent years. The present finding that FA was the most sensitive DTI metric in tracking change is consistent with previous cross-sectional ALS data (Cardenas-Blanco et al., 2014): FA reduction was predominantly driven by increased RD but there was also a slight decrease in L1 overtime, their opposite behaviours are additive in accentuating reductions in FA.

Turning to cortical thickness of the motor cortex, the present results agree with two previous studies of $N = 51$ (Schuster et al., 2014) and $N = 20$ (Verstraete et al., 2012) classic ALS subjects in showing no significant progression of cortical thinning over time. One of these reports speculated that the absence of significant progressive thinning in the precentral gyrus could be indicative of a floor effect at the first time-point resulting in a lack of sensitivity to further atrophy detection over time (Schuster et al., 2014). The absence of significant difference in cortical thickness between controls and ALS-TP3 in the present study suggests the floor effect is an unsustainable argument. Furthermore, absence of motor cortical thinning in ALS is anticipated from pathology. Although degeneration of Betz cells in the motor cortex is a very well documented finding, these cells make up only a tiny fraction of the motor cortex. Past quantitative post-mortem analysis in patients dying of ALS (therefore, more advanced than those typically participating in imaging studies) showed no evidence of reduction in motor cortex volume or thickness compared to age-matched controls (Toft et al., 2005).

The lack of significant differences using VBM in the present study, although consistent with a tensor-based morphometry study (Agosta et al., 2009b), is strikingly at odds with a recent study that reported

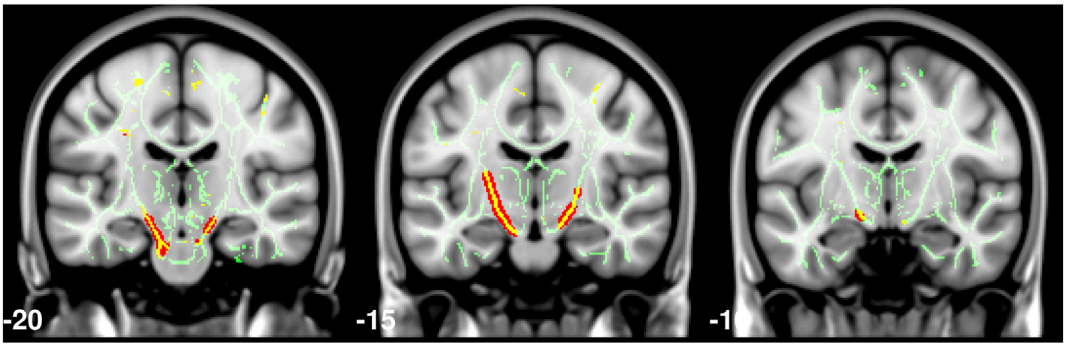

Fig. 2. Differences in FA between controls and ALS-TP3 at $p < 0.02$ (yellow/orange) uncorrected for multiple comparisons. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

extensive longitudinal changes in grey matter density using the FSL-VBM method (Menke et al., 2014). Differences in scanning interval might have contributed to this discrepancy between studies, in that the study in question had a mean time span of 16 months between the two scans compared to a mean of six months from first to third scans in the present study. There are, however, several reasons to suggest that these previously reported results using VBM may possibly be spurious. Firstly, although only statistically significant differences (but no effect sizes) were reported, it appeared that VBM was far more sensitive to detect longitudinal change than DTI, yet in the present study, that had the advantage of three time-points, DTI changes were not only more sensitive, but followed a highly linear trajectory over time. Secondly, using exactly the same method, FSL-VBM, and analysis steps as those reported by Menke et al. (2014) but also repeating the analysis using SPM-VBM, no significant effects were noted from ALS-TP1 to ALS-TP3. Thirdly, Menke et al. (2014) emphasized significant longitudinal change in deep grey matter of caudate and thalamus, yet in the present study using the independent method of FIRST volumetry to examine these structures, no significant longitudinal change was found. Furthermore, an absence of significant volumetric changes in these structures in a comparable time interval—5.5 months—was recently reported in another study (Westeneng et al., 2015). Similarly, a previous cross-sectional study was unable to demonstrate volumetric changes in these structures in non-demented ALS patients compared to controls (Machts et al., 2015). Time will tell with further studies, but perhaps the most important point is that studies on longitudinal biomarkers that only report statistical differences between two time points are very hard to interpret—a systematic artefact between measurements can also yield highly "significant" statistical effects. Plotting effect sizes to ensure biological plausibility, and, moreover, including more than two time points to ensure a plausible trajectory are, therefore, highly desirable.

To date, only a couple of longitudinal studies have examined more than one metric in the same study (Kwan et al., 2012, Menke et al., 2014) yet comparing across measures is an important step in putting findings in an overall context. The present study completed a head to head comparison of several MR-based measurements but also, importantly, compared findings to the standard clinical outcome measure, the ALSFRS-R. The results clearly indicated that although FA of the corticospinal tract is a robust biomarker, it was considerably less sensitive to change than the ALSFRS-R. This was exemplified by the power calculations that showed that to detect a comparable treatment effect, one would require a more than five-fold increase in numbers using FA when compared to ALSFRS-R. This is unsurprising considering the ALSFRS-R is a clinical severity summary score that can be influenced by both upper and lower motor neuron degeneration as well as all levels of the motor system whereas the FA biomarker is only measuring the central corticospinal tract. One caveat is that the follow-up ALSFRS-R scores in this study were not collected blind to preceding measurements; this could have introduced some bias due to examiner anticipation of decline. That said, the rate of ALSFRS-R decline in the present study was of the same order as that previously described (1.02 ± 2.3 points/month) from a pooled analysis of over 8000 clinical trial participants (Atassi et al., 2014). In contrast, the ALS-MITOS staging was not very sensitive in this relatively mild cohort—only 5 subjects moved one level over the course of the three time-points; in other words, 85% of patients' staging remained static. The power calculation for FA decline in the present study also suggested that more than double the number of subjects would be required to detect significant effect compared to a previous estimate for FA in the corticospinal tract (Zhang et al., 2011). This discrepancy is likely explained by the methodology in the earlier study in which the authors based their power calculation post-hoc on the most statistically significant sub-region of the corticospinal tract—clearly choosing the most significant sub-region will improve a power calculation, but the post-hoc approach required to do this would not be desirable for a trial.

5. Conclusion

Although the clinical scale was superior to FA change in the corticospinal tract, the robustness of the FA effect to track change argues that it still is a useful biomarker for mechanistic understanding. It would be desirable to confirm the site of action along the motor system for an intervention and this cannot be answered by a clinical measure that gives a summary score of the motor system. Ideally, biomarkers along the entire motor pathway are needed for this goal, and to this end, FA of the central corticospinal pathway can serve as one. Furthermore, imaging biomarkers rely on hardware sensitivity, which obviously goes hand in hand with technological developments. Future technological improvements, especially those capable of increasing resolution and signal to noise ratio are likely to translate to an increase in sensitivity.

Acknowledgments

We gratefully thank all the participants in this research study together with our radiographers for their dedicated efforts in scanning standardisation and quality assurance. The work was supported by the German Center for Neurodegenerative Diseases (DZNE).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nicl.2016.03.011.

References

Agosta, F., Rocca, M.A., Valsasina, P., Sala, S., Caputo, D., Perini, M., et al., 2009a. A longitudinal diffusion tensor MRI study of the cervical cord and brain in amyotrophic lateral sclerosis patients. J. Neurol. Neurosurg. Psychiatry 80, 53–55.

Agosta, F., Gorno-Tempini, M.L., Pagani, E., Sala, S., Caputo, D., Perini, M., et al., 2009b. Longitudinal assessment of grey matter contraction in amyotrophic lateral sclerosis: a tensor based morphometry study. Amyotroph. Lateral Scler. 10, 168–174.

Agosta, F., Pagani, E., Petrolini, M., Caputo, D., Perini, M., Prelle, A., et al., 2010. Assessment of white matter tract damage in patients with amyotrophic lateral sclerosis: a diffusion tensor MR imaging tractography study. AJNR Am. J. Neuroradiol. 31, 1457–1461.

Atassi, N., Berry, J., Shui, A., Zach, N., Sherman, A., Sinani, E., et al., 2014. The PRO-ACT database: design, initial analyses, and predictive features. Neurology 83, 1719–1725.

Blain, C.R., Williams, V.C., Johnston, C., Stanton, B.R., Ganesalingam, J., Jarosz, J.M., et al., 2007. A longitudinal study of diffusion tensor MRI in ALS. Amyotroph. Lateral Scler. 8, 348–355.

Brooks, B.R., Miller, R.G., Swash, M., Munsat, T.L., World Federation of Neurology Research Group on Motor Neuron Diseases, 2000. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. Other Motor Neuron. Disord. 1, 293–299.

Cardenas-Blanco, A., Machts, J., Acosta-Cabronero, J., Kaufmann, J., Abdulla, S., Kollewe, K., et al., 2014. Central white matter degeneration in bulbar- and limb-onset amyotrophic lateral sclerosis. J. Neurol. 261, 1961–1967.

Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., et al., 1999. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). J. Neurol. Sci. 169, 13–21.

Chiò A, Calvo A, Moglia C, Mazzini L, Mora G, PARALS study group. Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study. J. Neurol. Neurosurg. Psychiatry 2011; 82: 740–46.

Chiò, A., Hammond, E.R., Mora, G., Bonito, V., Filippini, G., 2015. Development and evaluation of a clinical staging system for amyotrophic lateral sclerosis. J. Neurol. Neurosurg. Psychiatry 86, 38–44.

Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., et al., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31, 968–980.

Diaz-de-Grenu, L.Z., Acosta-Cabronero, J., Chong, Y.F., Pereira, J.M., Sajjadi, S.A., Williams, G.B., et al., 2014. A brief history of voxel-based grey matter analysis in Alzheimer's disease. J. Alzheimers Dis. 38, 647–659.

Diggle, P.J., Heagerty, P., Liang, K.-Y., Zeger, S.L., 2002. Analysis of Longitudinal Data (Second Edition). Oxford University Press, Oxford.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97, 11050–11055.

Hagler Jr., D.J., Saygin, A.P., Sereno, M.I., 2006. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. NeuroImage 33, 1093–1103.

Jack Jr., C.R., Twomey, C.K., Zinsmeister, A.R., Sharbrough, F.W., Petersen, R.C., Cascino, G.D., 1989. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiology 172, 549–554.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.

Keil, C., Prell, T., Peschel, T., Hartung, V., Dengler, R., Grosskreutz, J., 2012. Longitudinal diffusion tensor imaging in amyotrophic lateral sclerosis. BMC Neurosci. 13, 141.

Kwan, J.Y., Meoded, A., Danielian, L.E., Wu, T., Floeter, M.K., 2012. Structural imaging differences and longitudinal changes in primary lateral sclerosis and amyotrophic lateral sclerosis. NeuroImage Clin. 2, 151–160.

Machts, J., Loewe, K., Kaufmann, J., Jakubiczka, S., Abdulla, S., Petri, S., et al., 2015. Basal ganglia pathology in ALS is associated with neuropsychological deficits. Neurology 85, 1301–1309.
Mann, H.B., Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. 18, 50–60.
Menke, R.A., Abraham, I., Thiel, C.S., Filippini, N., Knight, S., Talbot, K., et al., 2012. Fractional anisotropy in the posterior limb of the internal capsule and prognosis in amyotrophic lateral sclerosis. Arch. Neurol. 69, 1493–1499.
Menke, R.A., Korner, S., Filippini, N., Douaud, G., Knight, S., Talbot, K., et al., 2014. Widespread grey matter pathology dominates the longitudinal cerebral MRI and clinical landscape of amyotrophic lateral sclerosis. Brain 137, 2546–2555.
Muller, H.P., Unrath, A., Huppertz, H.J., Ludolph, A.C., Kassubek, J., 2012. Neuroanatomical patterns of cerebral white matter involvement in different motor neuron diseases as studied by diffusion tensor imaging analysis. Amyotroph. Lateral Scler. 13, 254–264.
Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al., 2005. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. 53, 695–699.
Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56, 907–922.
Pengas, G., Pereira, J.M., Williams, G.B., Nestor, P.J., 2009. Comparative reliability of total intracranial volume estimation methods and the influence of atrophy in a longitudinal semantic dementia cohort. J. Neuroimaging 19, 37–46.
Rajagopalan, V., Yue, G.H., Pioro, E.P., 2014. Do preprocessing algorithms and statistical models influence voxel-based morphometry (VBM) results in amyotrophic lateral sclerosis patients? A systematic comparison of popular VBM analytical methods. J. Magn. Reson. Imaging 40, 662–667.
Reuter, M., Rosas, H.D., Fischl, B., 2010. Highly accurate inverse consistent registration: a robust approach. NeuroImage 53, 1181–1196.
Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 61, 1402–1418.
Schuster, C., Kasper, E., Machts, J., Bittner, D., Kaufmann, J., Benecke, R., et al., 2014. Longitudinal course of cortical thickness decline in amyotrophic lateral sclerosis. J. Neurol. 261, 1871–1880.
Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.
Smith, S.M., Nichols, T.E., 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44, 83–98.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 (Suppl. 1), S208–S219.
Toft, M.H., Gredal, O., Pakkenberg, B., 2005. The size and distribution of neurons in the motor cortex in amyotrohic lateral sclerosis. J. Anat. 207, 399–407.
van der Graaff, M.M., Sage, C.A., Caan, M.W., Akkerman, E.M., Lavini, C., Majoie, C.B., et al., 2011. Upper and extra-motoneuron involvement in early motoneuron disease: a diffusion tensor imaging study. Brain 134, 1211–1228.
Verstraete, E., Veldink, J.H., Hendrikse, J., Schelhaas, H.J., van den Heuvel, M.P., van den Berg, L.H., 2012. Structural MRI reveals cortical thinning in amyotrophic lateral sclerosis. J. Neurol. Neurosurg. Psychiatry 83, 383–388.
Verstraete, E., Veldink, J.H., van den Berg, L.H., van den Heuvel, M.P., 2014. Structural brain network imaging shows expanding disconnection of the motor system in amyotrophic lateral sclerosis. Hum. Brain Mapp. 35, 1351–1361.
Verstraete, E., Turner, M.R., Grosskreutz, J., Filippi, M., 2015. Benatar M; attendees of the 4th NiSALS meeting. Mind the gap: the mismatch between clinical and imaging metrics in ALS. Amyotroph. Lateral Scler. Frontotemporal Degener. 16, 524–529.
Westeneng, H.J., Verstraete, E., Walhout, R., Schmidt, R., Hendrikse, J., Veldink, J.H., et al., 2015. Subcortical structures in amyotrophic lateral sclerosis. Neurobiol. Aging 36, 1075–1082.
Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57.
Zhang, Y., Schuff, N., Woolley, S.C., Chiang, G.C., Boreta, L., Laxamana, J., et al., 2011. Progression of white matter degeneration in amyotrophic lateral sclerosis: a diffusion tensor imaging study. Amyotroph. Lateral Scler. 12, 421–429.