Virtual brain biopsies in amyotrophic lateral sclerosis: Diagnostic classification based on in vivo pathological patterns

Peter Bede*, Parameswaran M. Iyer, Eoin Finegan, Taha Omer, Orla Hardiman

Quantitative Neuroimaging Group, Academic Unit of Neurology, Biomedical Sciences Institute, Trinity College Dublin, Ireland

ABSTRACT

Background: Diagnostic uncertainty in ALS has serious management implications and delays recruitment into clinical trials. Emerging evidence of presymptomatic disease-burden provides the rationale to develop diagnostic applications based on the evaluation of in-vivo pathological patterns early in the disease.

Objectives: To outline and test a diagnostic classification approach based on an array of complementary imaging metrics in key disease-associated anatomical structures.

Methods: Data from 75 ALS patients and 75 healthy controls were randomly allocated in a ‘training’ and ‘validation’ cohort. Spatial masks were created for anatomical foci which best discriminate patients from controls in the ‘training sample’. In a virtual ‘brain biopsy’, data was then retrieved from these key disease-associated brain regions. White matter diffusivity indices, grey matter T1-signal intensity values and basal ganglia volumes were evaluated as predictor variables in a canonical discriminant function.

Results: Following predictor variable selection, a classification specificity of 85.5% and sensitivity of 89.1% was achieved in the training sample and 90% specificity and 90% sensitivity in the validation sample.

Discussion: This study evaluates disease-associated imaging measures in a dummy diagnostic application. Although larger samples will be required for robust validation, the study confirms the potential of multimodal quantitative imaging in future clinical applications.

1. Introduction

Patients with ALS often describe tripping, intermittent slurred speech, muscle cramping, leg stiffness years before their formal diagnosis. Delays in diagnosis compromises timely enrolment in clinical trials, and limit the neuroprotective potential of new therapeutic agents. It is now also widely recognized that a long presymptomatic phase precedes symptom manifestation (Eisen et al., 2014). The few existing presymptomatic studies ALS confirm that the pathological changes can be captured well before symptom onset (Carew et al., 2011). The diagnosis of ALS remains primarily clinical, despite the unprecedented advances in quantitative neuroimaging and the uniquely specific imaging signature of ALS. After years of descriptive MRI studies in ALS, a number of studies have now emerged outlining diagnostic and prognostic applications based on pattern recognition (Bede and Hardiman, 2014). While methods of machine learning and data mining have been extensively used in marketing, information technology, banking, traffic control and metrology, they have only been recently applied to neurodegeneration (Orru et al., 2012). Many of the pioneering radiology studies come from cancer screening applications developed for the automated detection of malignancies. The automated analysis of imaging data of neurodegenerative conditions poses unique methodological challenges. Age-related variability, overlap of several neurodegenerative conditions, co-existing vascular white matter alterations are just some of the confounders. The most commonly used statistical approaches in neurodegeneration include logistic regression and support vector machines. Discriminant function analysis allows the determination of probability of group membership based on predictor variables. While the assumptions of discriminant analysis may be less flexible than those of logistic regression, it offers relatively high classification accuracy.

We hypothesized that based on the distinct neuroimaging signature of ALS a robust diagnostic framework can be outlined using discriminant function analysis. Our objective was to describe a diagnostic classification approach which incorporates multiple, ALS-defining, cortical, basal ganglia and white matter measures. The overarching
hypothesis of this study is that ‘data biopsies’ from blinded data sets can be used to accurately categorize individual participants as patients or controls based on normative and disease-specific training data.

2. Methods

All participants of this prospective biomarker study provided informed consent in accordance with the Medical Ethics Approval of the project (Beaumont Hospital, Dublin, Ireland). ALS patients and controls were randomly allocated in a ‘training group’ and a ‘testing group’ for validation. In total, 55 ALS patients and 55 controls were allocated to the ‘training group’ and 20 patients and 20 controls to the ‘testing group’. Healthy controls have been specifically recruited for this ALS biomarker study. The demographic and clinical profile of study participants are presented in Table 1. The participants of the training and validation cohorts were matched for age, gender and education.

2.1. Methods overview

The schematic overview of the development of the diagnostic protocol is presented in Fig. 1. First, comparative analyses were carried out between the patients and controls of the training group to establish core ALS-associated patterns of degeneration. Whole-brain tract-based spatial statistics (TBSS), voxel based morphometry and basal ganglia volumetrics were carried out correcting for age and gender, the methods of which have been described previously described (Bede et al., 2013a; Bede et al., 2015; Bede et al., 2013b; Bede et al., 2013c). Additionally, the volumes of brain stem, left and right thalamus, caudate, pallidum, hippocampus, amygdala, putamen and accumbens nuclei were compared between the controls and ALS patients of the training group correcting for age. The main ‘disease-defining’ anatomical structures identified by these comparisons included the corticospinal tracts (CST), corpus callosum (CC), primary motor cortex (PMC), thalamus, caudate nuclei, accumbens nuclei, amygdala and the hippocampus. Accordingly, region-of-interest (ROIs) anatomical maps were created for these structures, including voxels which showed statistically significant differences between patients and controls in the corticospinal tracts, corpus callosum, and precentral gyrus.

Imaging data of both the training and testing groups have been spatially co-registered to the MNI reference system and averaged MR metrics were retrieved from each participant in the above ROIs. In summary, our standardized ‘virtual brain biopsy’ consisted of retrieving quantitative MRI metrics from key, ALS-associated brain regions in each participant which were then evaluated in a discriminant function model. The outcomes of the initial comparisons and the atlas-based ROI masks created based on these contrasts are presented in Fig. 2.

2.2. MRI pulse sequence descriptions

MR data were acquired on a 3 Tesla Philips Achieva system with a gradient strength of 80 mT/m and slew rate of 200 T/m/s using an 8-channel receive-only head coil. T1-weighted images were obtained using a three-dimensional inversion recovery prepared spoiled gradient recalled echo (IR-SPGR) sequence with FOV = 256 × 256 × 160 mm, spatial resolution = 1 mm³, TR/TE = 8.5/3.9 ms, TI = 1060 ms, flip

---

Table 1

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Training cohort</th>
<th>Validation cohort</th>
<th>p - value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALS</td>
<td>Healthy controls</td>
<td>ALS</td>
</tr>
<tr>
<td>n</td>
<td>55</td>
<td>55</td>
<td>20</td>
</tr>
<tr>
<td>Age (Mean/Std.Dev.)</td>
<td>9.4579/9.1528</td>
<td>63.850/10.2346</td>
<td>0.202</td>
</tr>
<tr>
<td>Gender (Male/Female)</td>
<td>34/21</td>
<td>28/27</td>
<td>14/6</td>
</tr>
<tr>
<td>ALSFRS-r (Mean/Std.Dev.)</td>
<td>38.29/5.398</td>
<td>37.80/5.854</td>
<td>0.734</td>
</tr>
<tr>
<td>Disease duration (Mean/Std.Dev.)</td>
<td>23.65/7.399</td>
<td>24.65/6.409</td>
<td>0.596</td>
</tr>
</tbody>
</table>

Fig. 1. A schematic outline of model development, testing and validation.
angle = 8°, SENSE factor = 1.5. DTI images were acquired using a spin-echo planar imaging (SE-EPI) sequence with a 32-direction Stejskal-Tanner diffusion encoding scheme: FOV = 245 × 245 × 150 mm, 60 slices with no interslice gap, TR/TE = 7639/59 ms, SENSE factor = 2.5, b-values = 0, 1100 s/mm², with SPIR fat suppression and dynamic stabilization in an acquisition time of 5 min 41 s.

2.3. MRI processing pipelines

Raw diffusion tensor imaging data sets underwent eddy current corrections, motion correction and brain-tissue extraction using FMRIB's software library (FSL) (Smith et al., 2004). A diffusion tensor model was then fitted, generating maps of axial diffusivity (AD), radial diffusivity (RD) and fractional anisotropy (FA). Tract-based spatial statistics (TBSS) and permutation-based nonparametric inference was used for group comparisons in a study-specific white matter template applying the threshold-free cluster enhancement (TFCE) method (Smith et al., 2006; Smith and Nichols, 2009). The comparisons of the ALS patients and healthy controls of the ‘training group’, which were corrected for age and gender and statistical significance was set at p < 0.05 FWE.

Grey matter analyses were carried out using voxel based morphometry (VBM) tool box if FMRIB's software library (FSL) (Douaud et al., 2007). Following brain extraction, and tissue-type segmentation, grey matter partial volume images were aligned to the Montreal Neurological Institute 152 standard space; first a standard 12 degrees of freedom registration, then a 12 degrees of freedom registration to a MNI152 sub-cortical mask. A model-based approach is used by FSL-FIRST for the segmentation of subcortical structures. The models incorporated in FIRST are constructed from manually segmented images from 336 subjects provided by the Center for Morphometric Analysis (CMA), MGH, Boston. Registration and segmentation were individually reviewed and visually verified for all subjects. Following registration and segmentation, subcortical mesh and volumetric outputs were generated by FSL-FIRST using automatic boundary corrections.

2.4. Basal ganglia volumes

Volumes of subcortical structures were estimated using the subcortical segmentation and registration tool FIRST, part of the FMRIB's Software Library (FSL). T1-weighted structural data first underwent brain extraction and the resulting images were individually verified. FSL-FIRST uses a two-stage affine registration algorithm of the input T1 data sets to MNI152 (Montreal Neurological Institute 152) standard space; first a standard 12 degrees of freedom registration, then a 12 degrees of freedom registration to a MNI152 sub-cortical mask. A model-based approach is used by FSL-FIRST for the segmentation of subcortical structures. The models incorporated in FIRST are constructed from manually segmented images from 336 subjects provided by the Center for Morphometric Analysis (CMA), MGH, Boston. Registration and segmentation were individually reviewed and visually verified for all subjects. Following registration and segmentation, subcortical mesh and volumetric outputs were generated by FSL-FIRST using automatic boundary corrections.

2.5. Region of interest maps

The grey matter region of interest (ROI) for the left and right primary motor cortex was created based on the Harvard-Oxford
probabilistic atlas (Desikan et al., 2006) and defined based on voxels of the primary motor cortex which showed statistically significant differences at $p < 0.05$ between controls and patients in the ‘training’ cohort. White matter ROIs were defined based on the ICBM-DTI-81 white matter labels atlas (Mori et al., 2008; Oishi et al., 2008). White matter ROIs were defined based on statistically significant differences in AD, RD and FA between the controls and ALS patients of the ‘training’ cohort in left corticospinal tract, right corticospinal tract and corpus callosum. While fractional anisotropy (FA) was evaluated, radial diffusivity (RD) and axial diffusivity (AD) were regarded as superior predictor variables. This is because, AD ($\lambda_1$) and RD ($(\lambda_2 + \lambda_3) / 2$) are mathematically independent white matter metrics and thought to reflect on axonal and myelin integrity respectively (Budde et al., 2009; Song et al., 2005).

The initial list of candidate predictor variables included RD, AD and FA in the left and right corticospinal tract and corpus callosum ROIs, T1-signal intensity in the left and right motor cortex ROIs, the laterally averaged volumes the thalamus and caudate, left and right hippocampus, nucleus accumbens, amygdala, brain stem and total intracranial volume (ICV). Candidate predictor variables were excluded from the model based on tests of equality of group means if Wilk’s lambda was not statistically significant. Predictor variables were also excluded if they were highly correlated with each other in the pooled within-groups matrices. Based on the evaluation of candidate predictors, the final variables in the model included the AD and RD of the left and right CST and CC, left and right PMC ROIs, and volumes of the left and right hippocampus and nucleus accumbens, the left amygdala, and thalamus and caudate volumes (Table 2). The constant, canonical discriminant function coefficients and group centroids were calculated based on these variables. No leave-one-out cross-validation was performed in the training group, as separate-groups covariance matrices were used. Classification accuracy was tested in the ‘training’ group and in the independent validation sample separately to appraise the discriminatory power and generalizability of the model.

3. Results

The canonical discriminant function reached an Eigen value of 0.871, a canonical correlation value of 0.682, Wilks lambda of 0.534, Chi-square 62.983 and a significance of $< 0.001$. The model showed an overall classification sensitivity of 89.1% and specificity of 85.5% in the ‘training’ group. In the ‘testing sample’ both sensitivity and specificity reached 90%. Classification function outcomes, histograms and scatter plots of discriminant score distributions are shown in Fig. 3.

<table>
<thead>
<tr>
<th>Predictor variables after candidate predictor assessment and selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average RD in the “right RD corticospinal tract ROI”</td>
</tr>
<tr>
<td>Average RD in the “left RD corticospinal tract ROI”</td>
</tr>
<tr>
<td>Average RD in the “RD corpus callosum ROI”</td>
</tr>
<tr>
<td>Average AD in the “right AD corticospinal tract ROI”</td>
</tr>
<tr>
<td>Average AD in the “left AD corticospinal tract ROI”</td>
</tr>
<tr>
<td>Average AD in the “AD corpus callosum ROI”</td>
</tr>
<tr>
<td>T1-signal intensity in the “right motor cortex ROI”</td>
</tr>
<tr>
<td>T1-signal intensity in the “left motor cortex ROI”</td>
</tr>
<tr>
<td>Averaged volume of the left and right thalamus</td>
</tr>
<tr>
<td>Averaged volume of the left and right caudate</td>
</tr>
<tr>
<td>Volume of the left hippocampus</td>
</tr>
<tr>
<td>Volume of the right hippocampus</td>
</tr>
<tr>
<td>Volume of the left nucleus accumbens</td>
</tr>
<tr>
<td>Volume of the right nucleus accumbens</td>
</tr>
<tr>
<td>Volume of the left amygdala</td>
</tr>
</tbody>
</table>

4. Discussion

Neuroimaging in ALS has been very successful in characterizing group-level, disease-specific changes (Agosta et al., 2010; Turner et al., 2012; Prell and Grosskreutz, 2013; Bede et al., 2016), and also in describing phenotype (Bede et al., 2013a; Machts et al., 2015; Floeter et al., 2014) and genotype-specific (Cistarò et al., 2014; Floeter et al., 2016) neuroimaging signatures. More recently, ALS MRI studies captured presymptomatic imaging alterations (Carew et al., 2011; Benatar and Wu, 2012; Menke et al., 2016). Despite the significant advances in group-level, cross-sectional (Muller et al., 2016) and longitudinal analyses (Menke et al., 2014; Agosta et al., 2009a) in neurodegenerative conditions, the interpretation of single data sets remains methodologically challenging and a striking paucity of such studies persists in ALS (Welsh et al., 2013; Foerster et al., 2013; Ben Bashat et al., 2011; Schuster et al., 2016). From a biomarker perspective, the meaningful interpretation of data from single individuals is paramount for the development of viable clinical applications (Agosta et al., 2010; Turner et al., 2013). In ALS, the gap between the number of ‘group-level’ descriptive studies and ‘individual-level’ classification studies continues to widen. Admittedly, ‘individual-level’ analyses require large training and validation data sets, robust multicenter validation, but even small pilot studies are surprisingly scarce in ALS.

The main limitation of the study lies in its sample size; however it indicates that discriminant function analyses provide relatively accurate diagnostic classification. The classification protocol outlined in this study should to be replicated in larger data sets and tested using data from multiple MRI platforms. Large imaging repositories such as that of the neuroimaging society of ALS (NISALS) are invaluable resources to test and validate such frameworks on multi-site data (Turner et al., 2011). While our patients had relatively high ALSFRS-r, indicating early stage disease, their disease duration was relatively long. The accurate classification of patients with significant disability or long disease duration says relatively little about the sensitivity of a proposed diagnostic algorithm. The validation of classification algorithms should ideally be carried out using data from patients scanned immediately after their diagnoses or presymptomatic mutation carriers. In this study, we have outlined a binary classification approach integrating multiple imaging indices from multiple anatomical sites. However, in order to simulate real-life diagnostic dilemmas, multinomial (multiclass) follow-up studies are required to evaluate algorithms distinguishing between ALS patients, healthy individual and mimic conditions. Such studies should ideally include conditions which are pathologically, clinically or radiologically similar to ALS, for example hereditary spastic paraplegia (HSP), primary lateral sclerosis (PLS), post-polio or Kennedy’s patients. (Querin et al., 2017) Another limitation of the study is that there is no quantitative spinal cord MR indices were included. Spinal measures are likely to improve diagnostic classification further as they have been previously shown to be sensitive longitudinal and prognostic markers in ALS (Agosta et al., 2009b; El Mendili et al., 2015; El Mendili et al., 2014; Bede et al., 2012).

A range of statistical methods are available to classify blinded data in neurodegeneration, and each method is associated with unique advantages and limitations. Despite the relatively restrictive assumptions of discriminant function analyses, the advantage of this approach is that it readily integrates multiple measures, and provides diagnostic probabilities in addition to discriminant scores. Until robust, multi-center classification studies are published in ALS, it remains to be seen which approach is best suited to reliably capture early pathological changes in vivo. The concept of ROI-based, spatial reference system guided ‘data biopsies’ is also applicable to longitudinal analyses and potentially for the development of future monitoring markers.

Ethics approval and consent to participate

All participants provided written informed consent in accordance to
the Medical Ethics approval of the research project (Ethics (Medical Research) Committee - Beaumont Hospital, Dublin, Ireland).

**Funding**

The study was supported by the Health Research Board (HRB-Ireland), the Irish Institute of Clinical Neuroscience IICN − Novartis Ireland Research Grant, the Iris O'Brien Foundation, the Perrigo Clinician-Scientist Research Fellowship, the Research Motor Neuron (RMN-Ireland) Foundation and the EU-Joint Programme for Neurodegeneration (JPND) SOPHIA project. The sponsors of the study had no role in study design, writing of the report or decision to submit this work for publication.

**Competing interests**

Peter Bede, Parameswaran M. Iyer, Eoin Finegan, Taha Omer reports no disclosures.

Prof. Hardiman has received speaking honoraria from Janssen Cilag, Biogen Idec, Sanofi Aventis and Merck-Serono, she has been a member of advisory panels for Biogen Idec, Allergen, Cytokinetiks, Ono Pharmaceuticals and Sanofi Aventis.

**Authors’ contributions**

All authors contributed to data collection, study design, analyses and drafting the manuscript.

**Acknowledgements**

We are grateful for the generosity and kindness of all of our patients and their caregivers who have kindly participated in this study.

**References**


Bede, P., Iyer, P.M., Schuster, C., et al., 2016. The selective anatomical vulnerability of


