Journal of the Neurological Sciences 432 (2022) 120079

Contents lists available at ScienceDirect

Journal of the Neurological Sciences

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

Phenotypic categorisation of individual subjects with motor neuron disease based on radiological disease burden patterns: A machine-learning approach

Peter Bede [a,b,*], Aizuri Murad [a], Jasmin Lope [a], Stacey Li Hi Shing [a], Eoin Finegan [a], Rangariroyashe H. Chipika [a], Orla Hardiman [a], Kai Ming Chang [a,c]

[a] Computational Neuroimaging Group, Biomedical Sciences Institute, Trinity College Dublin, Ireland
[b] Pitié-Salpètrière University Hospital, Sorbonne University, Paris, France
[c] Department of Electronics and Computer Science, University of Southampton, UK

ARTICLE INFO

Keywords:
Neuroimaging
Biomarkers
Machine-learning
Artificial neural networks
Amyotrophic lateral sclerosis
Primary lateral sclerosis
Motor neuron disease
Diffusion imaging
Clinical trials

ABSTRACT

Motor neuron disease is an umbrella term encompassing a multitude of clinically heterogeneous phenotypes. The early and accurate categorisation of patients is hugely important, as MND phenotypes are associated with markedly different prognoses, progression rates, care needs and benefit from divergent management strategies. The categorisation of patients shortly after symptom onset is challenging, and often lengthy clinical monitoring is needed to assign patients to the appropriate phenotypic subgroup. In this study, a multi-class machine-learning strategy was implemented to classify 300 patients based on their radiological profile into diagnostic labels along the UMN-LMN spectrum. A comprehensive panel of cortical thickness measures, subcortical grey matter variables, and white matter integrity metrics were evaluated in a multilayer perceptron (MLP) model. Additional exploratory analyses were also carried out using discriminant function analyses (DFA). Excellent classification accuracy was achieved for amyotrophic lateral sclerosis in the testing cohort (93.7%) using the MLP model, but poor diagnostic accuracy was detected for primary lateral sclerosis (43.8%) and poliomyelitis survivors (60%). Feature importance analyses highlighted the relevance of white matter diffusivity metrics and the evaluation of cerebellar indices, cingulate measures and thalamic radiation variables to discriminate MND phenotypes. Our data suggest that radiological data from single patients may be meaningfully interpreted if large training data sets are available and the provision of diagnostic probability outcomes may be clinically useful in patients with short symptom duration. The computational interpretation of multimodal radiology datasets herald viable diagnostic, prognostic and clinical trial applications.

Abbreviations: **AD**, axial diffusivity; **ALS**, amyotrophic lateral sclerosis; **ALSFRS-r**, revised amyotrophic lateral sclerosis functional rating scale; **ANN**, artificial neural network; **ASO**, antisense oligonucleotides; **ATR**, Anterior Thalamic Radiation; **AUC**, area under the curve; ***C9orf72***, chromosome 9 open reading frame 72; **CST**, corticospinal tract; **CT**, cortical thickness; **DC**, disease control; **DFA**, discriminant function analysis; **DTI**, diffusion tensor imaging; **EMM**, estimated marginal mean; **EPI**, echo-planar imaging; **FA**, fractional anisotropy; **FLAIR**, fluid-attenuated inversion recovery; **FOF**, fronto-occipital fasciculus; **FOV**, field of view; **FTD**, frontotemporal dementia; **FWE**, familywise error; **GAN**, generative adversarial networks; **GAN**, generative adversarial neural network; **GM**, grey matter; **HC**, healthy control; **ICP**, inferior cerebellar peduncle; **ILF**, inferior longitudinal fasciculus; **IR-SPGR**, inversion recovery prepared spoiled gradient recalled echo; **IR-TSE**, inversion recovery turbo spin echo sequence; **IVIG**, intravenous immunoglobulin; **KD**, Kennedy's disease; **LMN**, lower motor neuron; **LO**, lateral occipital; **Lt**, Left; **MCP**, middle cerebellar peduncle; **MD**, mean diffusivity; **ML**, machine-learning; **MLe**, Medial Lemniscus; **MND**, Motor neuron disease; **MNI152**, Montreal Neurological Institute 152 standard space; **MPM**, multilayer perceptron model; **NISALS**, Neuroimaging Society in ALS; **PBA**, pseudobulbar affect; **PCA**, principal component analysis; **PCC**, pathological crying and laughing; **PMC**, primary motor cortex; **PMML**, predictive model markup language file; **QC**, quality control; **RD**, Radial diffusivity; **ROC**, receiver operating characteristic curve; **ROI**, region of interest; **Rt**, right; **SBMA**, bulbar muscular atrophy; **SCI**, Spinal cord injury; **SCP**, superior cerebellar peduncle; **SD**, standard deviation; **SE-EPI**, spin-echo echo planar imaging; **SENSE**, sensitivity Encoding; **SLF**, superior longitudinal fasciculus; **SMA**, Spinal muscular atrophy; **SPIR**, spectral presaturation with inversion recovery; **SVD**, singular value decomposition; **SVM**, support vector machine; **T1w**, T1-weighted imaging; **TBSS**, tract-based spatial Statistics; **TE**, echo time; **TFCE**, threshold-free cluster enhancement; **TI**, inversion time; **TIV**, total intracranial volume; **TR**, repetition time; **UF**, uncinate fasciculus; **UMN**, upper motor neuron; **WM**, white matter; **XML**, extensible markup language file.

* Corresponding author at: Room 5.43, Computational Neuroimaging Group, Trinity Biomedical Sciences Institute, Trinity College Dublin, Pearse Street, Dublin 2, Ireland.

E-mail address: bedep@tcd.ie (P. Bede).

https://doi.org/10.1016/j.jns.2021.120079
Received 16 November 2021; Received in revised form 25 November 2021; Accepted 29 November 2021
Available online 2 December 2021

1. Introduction

Motor neuron diseases (MND) encompass a multitude of clinically heterogeneous conditions with strikingly different progression rates, survival and disability profiles [1–4]. Initial symptoms of MND phenotypes may be very similar and often only careful longitudinal follow-up helps to categorise individual patients into the appropriate diagnostic subgroups [5,6]. Patients initially presenting with lower limb spasticity and subtle pseudobulbar affect for example may represent 'UMN-predominant ALS' and these patients may progress to develop overt LMN sings consistent with ALS, or, they may retain a relatively pure UMN-predominant clinical profile long after symptom onset, consistent with PLS [7,8]. The relevance of accurately classifying single participants cannot be underestimated; these diagnostic labels are associated with markedly different survival outcomes, care needs, and accurate phenotypic classification is also indispensable for stratification for clinical trials [9–13]. The traditional strategy in the clinical setting is to follow patients carefully over time. Subsequent diagnostic criteria for PLS for example have all included a minimum symptom duration criteria before the diagnosis could be established [6,14,15]. UMN vs. LMN predominance is just one of the many dimensions of disease heterogeneity in MND. The degree of cognitive impairment, difference in progression rates, bulbar versus limb disability predominance, genetic status, the presence of apathy, deficits in social cognition, extra-pyramidal and cerebellar involvement are other important facets of disease heterogeneity [16–18]. The precision categorisation of individuals with MND has important ramifications for individualised multidisciplinary care, advising caregivers, resource allocation, rehabilitation and recruitment into relevant clinical trials [19]. From a therapeutic perspective, there is a marked shift from the notion of 'one-drug-for-all' to precision, genotype- and phenotype-specific therapeutic strategies. Recent clinical trials of antisense oligonucleotides (ASO) exemplify this important trend. The categorisation of individual MND patients currently relies on the careful observation of cardinal clinical features, progression rates and the consideration of supporting neurophysiological and neuropsychological findings [20–24]. In the clinical setting, neuroimaging has a limited role in confirming a suspected diagnosis or categorising patients with MND . In contrast, academic studies have consistently demonstrated the biomarker potential of computational imaging in ascertaining genotype-associated features, tracking progressive changes and meaningfully evaluating the entire neuroaxis from cortical changes, through infratentorial disease-burden to spinal cord alterations [25–28]. In contrast to group-level analyses [29], there has also been notable progress in interpreting single-patient datasets in MND, including epidemiology data, clinical variables and neuroimaging data [30–32]. Machine-learning (ML) algorithms have the potential to categorise individual patients into diagnostic, phenotypic or prognostic subgroups if large, uniformly acquired datasets are available for model training. Single-centre ML initiatives in MND often suffer from sample size limitations which can lead to model overfitting and data compiled from multiple sites are vulnerable to acquisition inconsistencies and missing data. A common shortcoming of early ML studies in MND is binary classification into ALS versus a healthy control group [33] instead of implementing multi-class models with clinically relevant output labels [34]. Another stereotyped drawback of classification studies is model testing on patients with an established diagnosis, often with long symptom duration, instead of model validation on 'early-stage patients' or patients with a suspected diagnosis. Finally, most ML studies in MND implement a single model instead of comparing the performance of several models on the same cohort of patients. Accordingly, the objective of this pilot study is the multi-class classification of patients with MND based on imaging data using a dual approach; a supervised model and an artificial neural network framework. An additional objective of the study is the classification of PLS and ALS patients with relatively short symptom duration to scrutinise the efficiency of a proposed classification strategy.

2. Methods

2.1. Participants

A total of 300 participants, 215 patients with amyotrophic lateral sclerosis ('ALS'), 42 patients with primary lateral sclerosis ('PLS') and 43 poliomyelitis survivors ('PMS') were included in a prospective, single-centre study. In accordance with the Ethics Approval of this research project (Beaumont Hospital, Dublin, Ireland), all participants gave informed consent. Exclusion criteria included comorbid neoplastic, paraneoplastic or neuroinflammatory diagnoses, prior cerebrovascular events, and known traumatic brain injury. Participating ALS patients were diagnosed according to the revised El Escorial criteria and PLS patients were diagnosed according to the new consensus criteria [35].

2.2. Magnetic resonance imaging

A uniform imaging protocol was used on a 3 Tesla Philips Achieva Magnetic resonance (MR) platform. T1-weighted (T1w) images were acquired with a 3D Inversion Recovery prepared Spoiled Gradient Recalled echo (IR-SPGR) sequence with the following parameters: spatial resolution of 1 mm^3, field-of-view (FOV) of 256 × 256 × 160 mm, flip angle = 8°, SENSE factor = 1.5, TR/TE = 8.5/3.9 ms, TI =1060 ms. Diffusion tensor images (DTI) were acquired using a spin-echo echo planar imaging (SE-EPI) pulse sequence with a 32-direction Stejskal-Tanner diffusion encoding scheme; 60 slices with no interslice gap, spatial resolution = 2.5 mm^3, FOV = 245 × 245 × 150 mm, TR/TE = 7639 / 59 ms, SENSE factor = 2.5, b-values = 0, 1100 s/mm^2, spectral presaturation with inversion recovery (SPIR) fat suppression and dynamic stabilisation. Fluid-attenuated inversion recovery (FLAIR) images were acquired with an Inversion Recovery Turbo Spin Echo (IR-TSE) sequence to rule out comorbid inflammatory or vascular pathologies. Imaging data from all participants were clinically reviewed for incidental findings prior to inclusion in quantitative analyses. Subsequent to radiological review, a comprehensive panel of imaging metrics, consisting of cortical grey matter indices, subcortical grey matter variables, and white matter integrity measures were systematically retrieved from an extensive list of anatomical regions in each participant. Following standardised pre-processing steps and spatial registration, 28 volume values, 68 cortical thickness measures and 120 white matter indices were uniformly retrieved from each subject; a total of 216 imaging measures.

2.3. Cortical thickness values

The standard pipeline of the FreeSurfer image analysis suite [36] was used for pre-processing, including non-parametric non-uniform intensity normalization, affine registration to the MNI305 atlas, intensity normalization, skull striping, automatic subcortical segmentation, linear volumetric registration, neck removal, tessellation of the grey matter-white matter boundary, surface smoothing, inflation to minimize metric distortion, and automated topology correction [37]. Subsequent to pre-processing, the anatomical labels of the Desikan-Killiany atlas [11] were utilised to estimate average cortical thickness in the following cortical regions in the left and right cerebral hemispheres separately; (1) banks superior temporal sulcus, (2) caudal anterior-cingulate cortex, (3) caudal middle frontal gyrus, (4) cuneus cortex, (5) entorhinal cortex, (6) frontal pole, (7) fusiform gyrus, (8) inferior parietal cortex, (9) inferior temporal gyrus, (10) insula, (11) isthmus–cingulate cortex, (12) lateral occipital cortex, (13) lateral orbitofrontal cortex, (14) lingual gyrus, (15) medial orbital frontal cortex, (16) middle temporal gyrus, (17) parahippocampal gyrus, (18) paracentral lobule, (19) pars opercularis, (20) pars orbitalis, (21) pars triangularis, (22) pericalcarine cortex, (23) postcentral gyrus (24) posterior-cingulate cortex, (25) precentral gyrus, (26) precuneus cortex, (27) rostral anterior cingulate cortex, (28) rostral middle frontal gyrus, (29) superior frontal gyrus, (30) superior parietal

cortex, (31) superior temporal gyrus, (32) supramarginal gyrus, (33) temporal pole, (34) transverse temporal cortex.

2.4. Volume metrics

A Bayesian segmentation algorithm was utilised to parcellate the brainstem into the medulla oblongata, pons and midbrain, which relies on a probabilistic brainstem atlas derived from 49 scans [38]. A total of 28 volume variables were estimated from each pre-processed T1-weighted dataset: (1) left cerebellar white matter volume, (2) left cerebellar cortex volume, (3) left thalamus volume, (4) left caudate volume, (5) left putamen volume, (6) left pallidum volume, (7) left hippocampus volume, (8) left amygdala volume, (9) left accumbens volume, (10) right cerebellar white matter volume, (11) right cerebellar cortex volume, (12) right thalamus volume, (13) right caudate volume, (14) right putamen volume, (15) right pallidum volume, (16) right hippocampus volume, (17) right amygdala volume, (18) right accumbens volume, (19) posterior corpus callosum volume, (20) middle corpus callosum volume, (21) central corpus callosum volume, (22) mid-anterior corpus callosum volume, (23) anterior corpus callosum volume, (24) medulla volume, (25) pons volume, (26) superior cerebellar peduncle volume, (27) midbrain volume, (28) total intracranial volume. Each volume value was converted as a percentage of the subject's total intracranial volume (TIV) to account for TIV variations.

2.5. White matter indices

The diffusivity profiles of 30 white matter regions were evaluated. Following quality control, eddy current corrections and skull removal were applied to DTI data before a tensor model was fitted to generate diffusivity maps of axial diffusivity (AD), fractional anisotropy (FA), mean diffusivity (MD), and radial diffusivity (RD). FMRIB's software library's tract-based statistics (TBSS) module was implemented for the non-linear registration of DTI images, skeletonisation and the creation of a mean FA mask. The study-specific white matter skeleton was masked in MNI space by the anatomical labels of the following white matter regions: left & right anterior thalamic radiation, left & right cerebellar white matter skeleton, left & right cingulum, left & right corticospinal tract, left & right external capsule, forceps major, forceps minor, fornix, left & right inferior cerebellar peduncle, left & right inferior fronto-occipital fasciculus, left & right inferior longitudinal fasciculus, left & right medial lemniscus, middle cerebellar peduncle, left & right posterior thalamic radiation, left & right superior cerebellar peduncle, left & right superior longitudinal fasciculus, left & right uncinate fasciculus. To generate spatial masks for the cerebellar peduncles, medial lemniscus, external capsule and posterior thalamic radiation, the labels of the ICBM-DTI-81 white-matter atlas [39,40] were used. To create masks for the forceps major, forceps minor, anterior thalamic radiation, uncinate, superior and inferior longitudinal fasciculi, cingulum, corticospinal tracts, and inferior fronto-occipital fasciculi, the labels of the JHU white-matter tractography atlas [41,42] were utilised. Label 2 of the MNI probabilistic atlas [43,44] was used to generate a white mask for the cerebellum and FMRIB's fornix template [45] was used to evaluate the diffusivity profile of the fornix.

2.6. Statistical interpretation

First, a multilayer perceptron model was implemented with one hidden layer containing 12 nodes (units). The hidden layer activation function was hyperbolic tangent and a batch-type training approach was used with a gradient descent optimisation algorithm. The diagnosis (ALS, PLS, PMS) was set as dependent variable, and retrieved imaging measures designated as covariates. Imaging values were rescaled by standardisation. The available data ($n = 300$) was split into a training sample ($n = 206$, 68.7%) and testing sample ($n = 94$, 31.3%). The predicted pseudo-probability of each diagnostic group was plotted in a bar chart to illustrate the accuracy of diagnostic classification and receiver operating characteristic curves (ROC) were also plotted. The synaptic weight estimates for each dependent variable from the perceptron model was exported to an extensible markup language (XML (PMML)) file which was subsequently applied to an independent sample of scans encompassing data from patients with short symptom duration. To rank the significance of individual imaging metrics in predicting group membership, an independent variable importance analysis was also performed. IBM's SPSS version 25 was utilised for statistical modelling. In a supplementary exploratory analysis, discriminant function analyses (DFA) were also run to evaluate classification accuracy in juxtaposition to ANN outcomes. Given the associations between the four diffusivity metrics, only FA was evaluated in each white matter regions and AD, MD and RD were not considered as input variables in the DFA. In addition to FA, volume and cortical metrics were included as initial input variables. All 300 participants were included in the original and cross-validated sample. First, standard descriptive statistics were run to calculate the mean and standard deviation of each input variable for in each group. Then, tests of equality of group means were performed to measure each independent variable's potential before the model is created. This is derived from a one-way ANOVA for the independent variable using the grouping variable as the factor. A significance value greater than 0.10 suggests that the variable is unlikely to contribute meaningfully to the model. Wilks' lambda is an alternative measure of a variable's potential, smaller values indicating superior discriminating potential. Subsequently, box's test of equality of covariance matrices was performed to test the assumption of homogeneity of variances and covariances. Only variables passing the tolerance criteria were entered in the final DFA model.

3. Results

The three groups, (1) amyotrophic lateral sclerosis ('ALS', $n = 215$, age: 62.953 ± 10.6, 141 male, 203 right handed), (2) primary lateral sclerosis ('PLS', $n = 42$, age: 61.45 ± 9.65, 26 male, 38 right handed) and (3) poliomyelitis survivors ('PMS', $n = 43$, age: 65.12 ± 6.48, 18 male, 41 right handed) were matched for age ($p = 0.23$) and handedness ($p = 0.56$), but not for sex ($p = 0.014$).

3.1. Artificial neural network outcomes

The available data set was split into a training sample ($n = 206$, 68.7%) and testing sample ($n = 94$, 31.3%). In the training sample, cross entropy error was 10.148, and 48.14 in the testing sample; incorrect predications were 0.5% in the training sample and 20.2% in the testing sample. Classification summary is presented in Table 1. The predicted pseudo-probability of diagnosis in each cohort (confirmed diagnosis) is presented in Fig. 1. Receiver operating characteristic (ROC) curves are presented in Fig. 2. Area under the curve values were 0.987 for ALS, 0.976 for PLS, and 0.99 for PMS. The normalised importance of the 20 most relevant imaging variables in predicting group membership is shown in Fig. 3.

3.2. Discriminant function analyses

Two canonical discriminant functions were used in the analysis and the canonical correlations for the dimensions one and two are 0.759 and 0.629, respectively. Classification outcomes are presented in Table 2. A graph of individuals was plot on the discriminant dimensions with the relevant group centroids (Fig. 4.). Case-wise statistics were inspected to review the clinical and demographic profiles of misclassified subjects individually.

3.3. Perceptron model testing on individual with short symptom duration

In an exploratory supplementary analysis, the artificial neural

Table 1
Classification outcomes of the multilayer perceptron model.

Sample	Established diagnosis	Classification			
		Predicted			
		ALS	PLS	PMS	Percent Correct
Training	ALS	152	0	0	100.0%
	PLS	0	26	0	100.0%
	PMS	1	0	27	96.4%
	Overall Percent	74.3%	12.6%	13.1%	99.5%
Testing	ALS	59	1	3	93.7%
	PLS	7	7	2	43.8%
	PMS	5	1	9	60.0%
	Overall Percent	75.5%	9.6%	14.9%	79.8%

network model was further tested on a sample of scans of "short-symptom-duration" patients, including 12 PLS patients with a less than 5-year symptom duration ($n = 12$, 8 males, mean age 55.08 ± 9.35) and 63 ALS patients with a less than 1-year symptom duration ($n = 63$, 44 males, mean age 62.82 ± 11.56). The synaptic weight estimates for each dependent variable from the original ANN model was applied to the independent sample of scans. Of the 12 PLS patients, the model has correctly categorised 9 (75%), 2 patients were labelled as "ALS" and one as PMS. Of the 63 ALS patients, 60 (95.2%) were correctly classified as ALS by the model, and 3 were misclassified and "PMS". Of the 43 poliomyelitis survivors, 36 were correctly categorised as PMS, 6 were labelled as "ALS", and one patient categorised as "PLS".

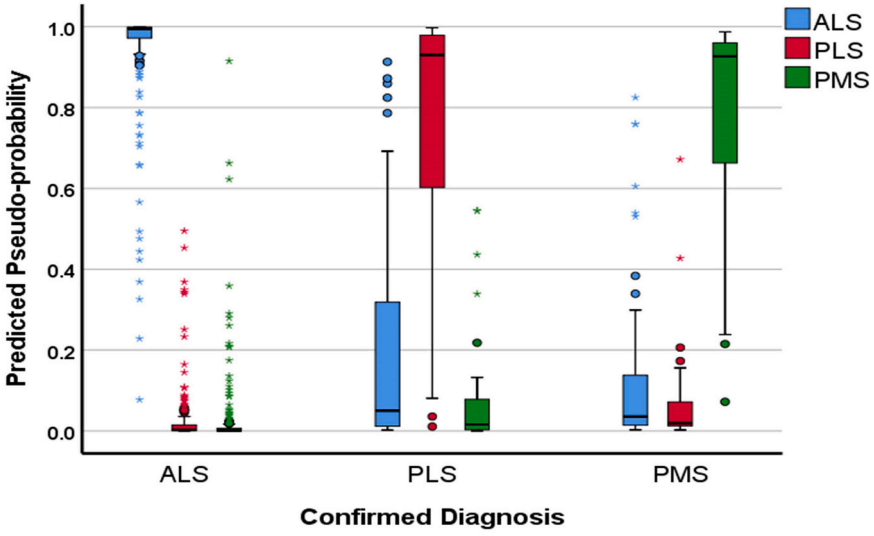

Fig. 1. The predicted pseudo-probability of diagnosis in each cohort (confirmed diagnosis).

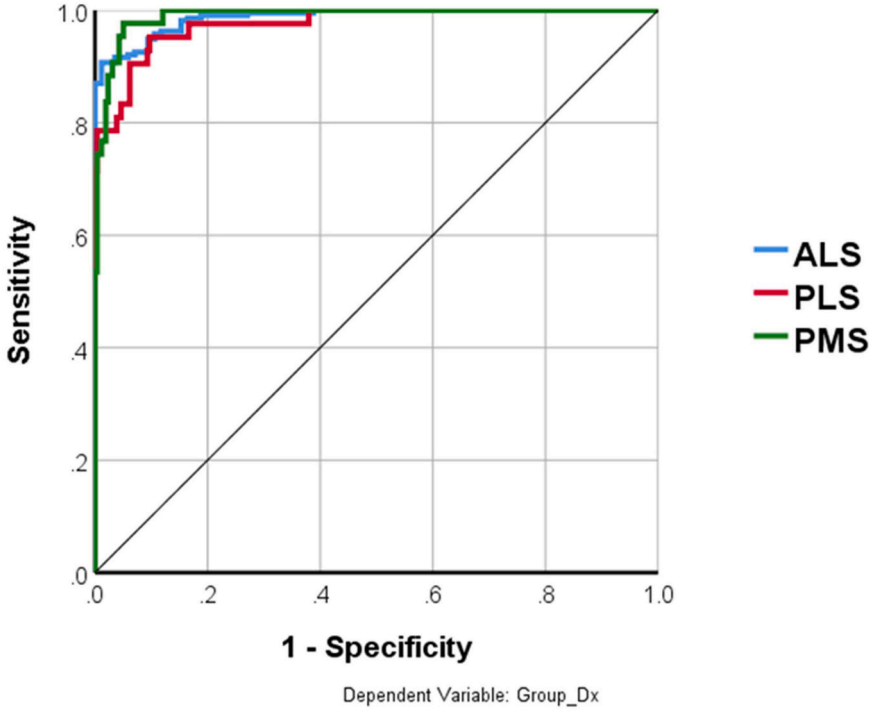

Fig. 2. Receiver operating characteristic (ROC) curves.

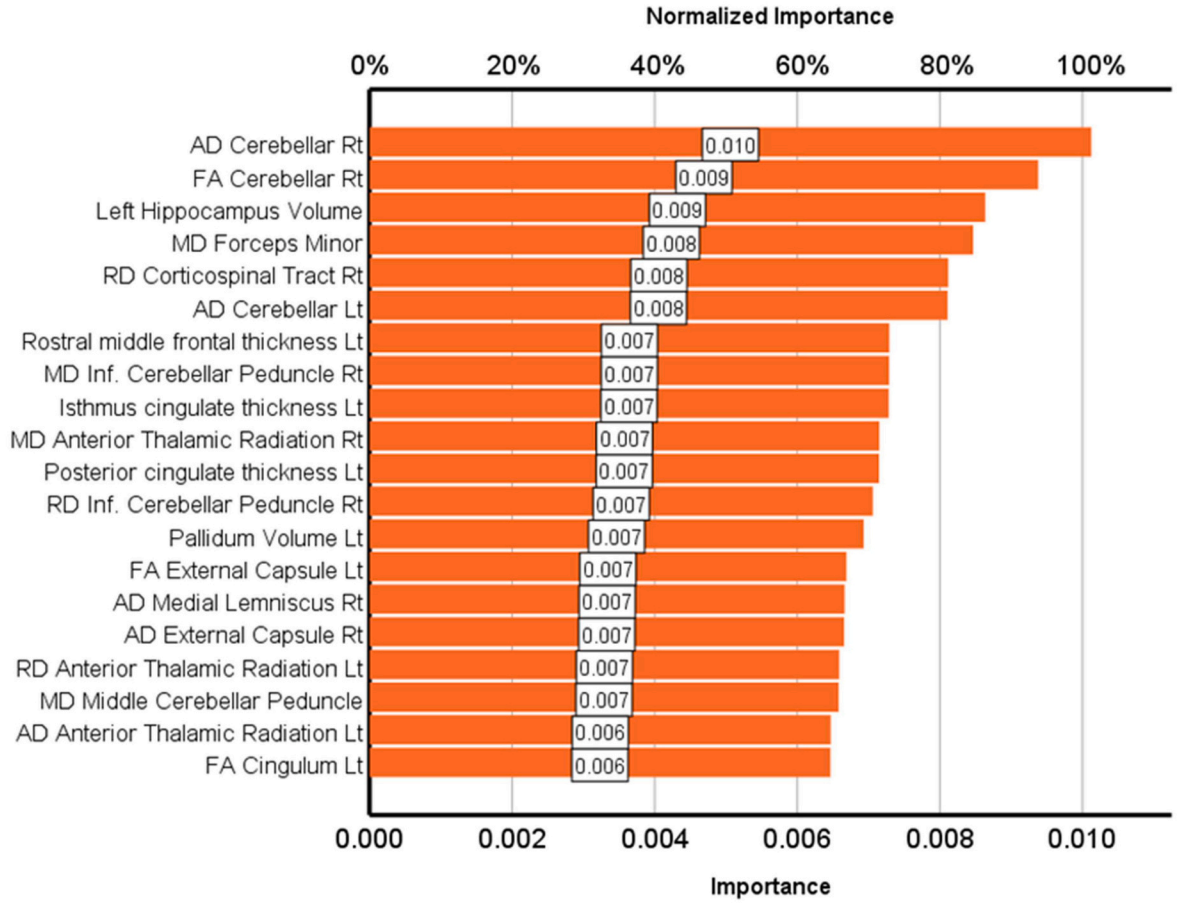

Fig. 3. The normalised importance of the 20 most relevant imaging variables in predicting group membership.

Table 2
Classification outcomes of the discriminant function analyses. 90.3% of original grouped cases correctly classified. b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case. c. 68.7% of cross-validated grouped cases correctly classified.

		Group	Predicted Group Membership			
			ALS	PLS	PMS	Percent Correct
Original	Count	ALS	207	3	5	96.3%
		PLS	6	36	0	85.7%
		PMS	14	1	28	65.1%
Cross-validated[b]	Count	ALS	174	13	28	80.9%
		PLS	22	18	2	42.9%
		PMS	28	1	14	32.6%

4. Discussion

Our findings demonstrate that multiple viable strategies exist to categorise patients with motor neuron disorders into phenotypic subgroups based on multimodal imaging data. These approaches interpret spatially coded imaging data without additional clinical information, complementary neurophysiology or wet biomarker data. Our results illustrate that discrimination is not only possible from healthy controls, but patient groups with relatively similar cerebral profiles may also be distinguished.

In our perceptron model, we achieved superior classification in accuracy in patients with ALS (93%) compared to patients with PLS (43%) or poliomyelitis survivors (60%). This highlights the challenges of discriminating ALS and PLS based on cerebral imaging data alone and supports to wider implementation of quantitative spinal protocols [46–49]. With the recent publication of seminal post mortem series in PLS [50], the lack of anterior horn involvement is increasingly evident, even in patients with long symptom duration. It is noteworthy that the model correctly identified 75% of early-stage PLS patients in an independent sample of MRI scans based on brain data alone. The model also performed well on a sample of ALS patients with short symptom duration (95%) demonstrating the potential clinical utility of applying ML frameworks to MR data of patients with suspected diagnoses. Recent studies of asymptomatic mutation carriers have demonstrated radiological changes long before symptom onset [51–53]. A future validation of the model would be to implement this strategy on MR data of pre-manifest mutation carriers to assess the sensitivity of the model to identify imaging traits consistent with a certain phenotype before symptom manifestation.

Feature importance ranking offers important insights into the discriminatory potential of specific anatomical regions. Most descriptive studies compare a specific MND phenotype to healthy controls to characterise core, phenotype-associated pathological signatures. Feature importance analysis between several MND phenotypes highlights discriminatory regions which may not be traditionally associated with the clinical profile of these conditions [54]. In our ANN model, cerebellar integrity metrics rank high, which is of interest as the characterisation of cerebellar degeneration in ALS is relatively recent [27,55,56], there are only sporadic reports of cerebellar involvement in PLS [57] and limited information on cerebellar involvement in poliomyelitis survivors [58–60]. Metrics of the cingulum also rank high and this region has been primarily implicated as a connectivity hub in fMRI studies of ALS [61]. External capsule integrity also helps to discriminate MND subtypes, which is interesting given that internal capsule

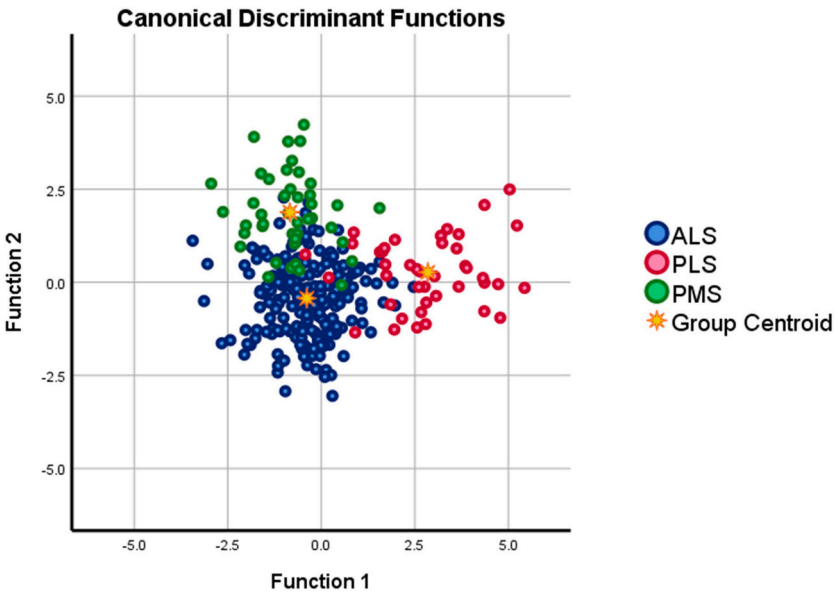

Fig. 4. Canonical discriminant functions, individual patients and group centroids.

degeneration is traditionally regarded as pathognomonic of ALS [62].

In this study we explored the classification efficiency of two statistical approaches, artificial neural networks and discriminant function analysis (DFA). DFA has a number of stringent data assumptions (normal distribution of predictor data, homoscedasticity of variables across groups, low degree of multicollinearity, relative lack of outliers etc.) which may restrict its application to large imaging datasets. Machine learning strategies are traditionally divided into 'supervised' and 'unsupervised' learning approaches. Each model has their respective advantages and disadvantages and the choice of a specific model should be primarily determined by data characteristics, sample size, number of candidate features, associations between candidate features, outliers, and missing values. A variety of innovative ML approaches have been applied to ALS datasets to date [19,31,32,63–71], but multiple models are seldom applied to the same datasets to compare their respective accuracy [72]. Alternative approaches have also been trialled in MND to interpret single MR scans, such as *z*-scoring imaging parameters with reference to normative radiology data [73–75]. There are also important lessons to consider from studies in other neurodegenerative conditions [76,77], and novel approaches such as dimensionality reduction [78], 'deep-learning' strategies [79,80], generative adversarial neural networks (GAN) etc. may also prove very efficient in MND.

The accuracy of classification models is often tested on patients with long symptom duration in a disease-stage by which they developed marked disease- or phenotype-associated patterns of pathology. From a clinical perspective, the categorisation of patients with a well-established diagnosis offers limited additional value and the real advantage of data-driven classification is to assign diagnostic labels to early-stage, short-symptom-duration patients or patients with a suspected diagnosis before meeting diagnostic criteria. Accordingly, a proposed ML model in MND should ideally be tested on patients soon after symptom onset. This is particularly important in PLS where diagnostic criteria are based on symptom duration and patients often fear transition to ALS which carries a worse prognosis. The specific diagnostic label provided by ANN models is just one of the outcomes to be considered, a clinically relevant output measure is the 'predicted probability', i.e. how certain the proposed outcome label is. This is very important to consider in real-life clinical situations i.e. a high probability score of group membership may inspire confidence with regards to the suggested diagnostic label.

This study is not without limitations; the pre-processing of the raw MR data relied on computationally intensive and time-consuming pipelines which in their current form are ill-suited to aid clinical decision making in real time. The value of including additional imaging indices, such as cortical volume metrics, thalamic nuclei volumes, amygdalar nuclei characteristics, hippocampal subfield measures etc. was not evaluated [81–83]. In this study we sought to test whether MR data alone can offer phenotypic categorisation, but the presented models could incorporate wet biomarker panels, functional data or clinical metrics which may improve patient stratification further [84,85]. The input variables in our models only encompassed imaging measures and individual demographic variables were not considered. Sexual dimorphism is widely recognised in neuroradiology, both in healthy populations and in MND [86,87]. To account for gender effects, separate training data sets could be generated for male and female participants, but this would require larger samples for robust model training. Finally, the inclusion of additional LMN-predominant cohorts, such as SBMA/KD, SMA, SCI would have helped to scrutinise our models further [47,88–91]. Notwithstanding these limitations, our findings demonstrate the feasibility of imaging-based patient categorisation in MND. Based on promising ML initiatives in neurodegeneration, the key priorities of future research include testing model accuracy on cohorts with short symptom duration, model validation on external data sets, multiclass model development instead of pursuing binary classification frameworks, transparent feature importance analyses, and testing proposed models on asymptomatic mutation carriers. While categorising individuals into diagnostic or phenotypic groups is of interest and may help to expedite the diagnosis, an obvious evolution of ML research in MND is the classification of patients into prognostic, survival, progression-rate, and disability-profile subcategories. Effective predictive modelling may also have a significant impact on clinical trial design [9,66,92].

5. Conclusions

Individual patients with MND can be categorised into phenotypic groups based on their imaging profile if large training datasets are available. Upper motor neuron predominant MND patients with short symptom duration may be accurately categorised as likely PLS. Computational neuroimaging offers precision patient stratification strategies in MND and emerging machine-learning studies herald viable diagnostic, prognostic and clinical trials applications.

Funding

This study was sponsored by the Spastic Paraplegia Foundation, Inc. (SPF). Professor Peter Bede and the Computational Neuroimaging Group are also supported by the Health Research Board (HRB EIA-2017-019), the EU Joint Programme – Neurodegenerative Disease Research (JPND), the Andrew Lydon scholarship, the Irish Institute of Clinical Neuroscience (IICN), and the Iris O'Brien Foundation.

Research ethics approval

This study was approved by the Ethics (Medical Research) Committee—Beaumont Hospital, Dublin, Ireland.

Author contribution

Conceptualisation of the study: PB, KMC.
Drafting the manuscript: PB, KMC.
Neuroimaging analyses: PB, AM, JL, KMC.
Revision of the manuscript for intellectual content: PB, AM, JL, SLHS, EF, RHC, OH, KMC.

Declaration of Competing Interest

The authors of this manuscript have no conflicts of interest to disclose.

References

[1] O. Hardiman, et al., Neurodegenerative Disorders: A Clinical Guide, 2016 ed., Springer Cham Heidelberg New York Dordrecht London© Springer International Publishing Switzerland 2016: Springer International Publishing, 2016, pp. 1–336.
[2] M.K. Floeter, T. Wu, Longitudinal evaluation of upper motor neuron burden scales in primary lateral sclerosis, Amyotroph Lateral Scler Frontotemporal Degener (2020) 1–7.
[3] F. Trojsi, et al., Functional overlap and divergence between ALS and bvFTD, Neurobiol. Aging 36 (1) (2015) 413–423.
[4] Y. Yunusova, et al., Clinical measures of bulbar dysfunction in ALS, Front. Neurol. 10 (2019) 106.
[5] R.H. Chipika, et al., Tracking a fast-moving disease: longitudinal markers, monitoring, and clinical trial endpoints in ALS, Front. Neurol. 10 (2019) 229.
[6] M.G. Clark, et al., Loss of functional connectivity is an early imaging marker in primary lateral sclerosis, Amyotroph Lateral Scler Frontotemporal Degener 19 (7–8) (2018) 562–569.
[7] E. Finegan, et al., The clinical and radiological profile of primary lateral sclerosis: a population-based study, J. Neurol. 266 (11) (2019) 2718–2733.
[8] E. Finegan, et al., Evolving diagnostic criteria in primary lateral sclerosis: the clinical and radiological basis of "probable PLS", J. Neurol. Sci. 417 (2020), 117052.
[9] H. Blasco, et al., A pharmaco-metabolomics approach in a clinical trial of ALS: identification of predictive markers of progression, PLoS One 13 (6) (2018), e0198116.
[10] C. Schuster, O. Hardiman, P. Bede, Survival prediction in amyotrophic lateral sclerosis based on MRI measures and clinical characteristics, BMC Neurol. 17 (1) (2017) 73.
[11] C. Schuster, O. Hardiman, P. Bede, Development of an automated MRI-based diagnostic protocol for amyotrophic lateral sclerosis using disease-specific pathognomonic features: a quantitative disease-state classification study, PLoS One 11 (12) (2016), e0167331.
[12] V. Grollemund, et al., Manifold learning for amyotrophic lateral sclerosis functional loss assessment : development and validation of a prognosis model, J. Neurol. 268 (3) (2020) 825–850, https://doi.org/10.1007/s00415-020-10181-2.
[13] H. Mitsumoto, B.R. Brooks, V. Silani, Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? Lancet Neurol. 13 (11) (2014) 1127–1138.
[14] H. Mitsumoto, et al., Phenotypic and molecular analyses of primary lateral sclerosis, Neurol. Genet. 1 (1) (2015), e3.
[15] H. Mitsumoto, et al., Primary lateral sclerosis (PLS) functional rating scale: PLS-specific clinimetric scale, Muscle Nerve 61 (2) (2020) 163–172.
[16] M. Feron, et al., Extrapyramidal deficits in ALS: a combined biomechanical and neuroimaging study, J. Neurol. 265 (9) (2018) 2125–2136.
[17] M. Abidi, et al., Neural correlates of motor imagery of gait in amyotrophic lateral sclerosis, J. Magn. Reson. Imaging 53 (1) (2021) 223–233.
[18] D. Lule, et al., Neuroimaging of multimodal sensory stimulation in amyotrophic lateral sclerosis, J. Neurol. Neurosurg. Psychiatry 81 (8) (2010) 899–906.
[19] V. Grollemund, et al., Machine learning in amyotrophic lateral sclerosis: achievements, pitfalls, and future directions, Front. Neurosci. 13 (2019) 135.
[20] D. Lule, et al., Emotional adjustment in amyotrophic lateral sclerosis (ALS), J. Neurol. 259 (2) (2012) 334–341.
[21] E. Finegan, et al., Extra-motor cerebral changes and manifestations in primary lateral sclerosis, Brain Imag. Behav. 15 (5) (2021) 2283–2296, https://doi.org/10.1007/s11682-020-00421-4.
[22] T. Burke, et al., Measurement of social cognition in amyotrophic lateral sclerosis: a population based study, PLoS One 11 (8) (2016), e0160850.
[23] T. Burke, et al., Discordant performance on the 'Reading the Mind in the Eyes' Test, based on disease onset in amyotrophic lateral sclerosis, Amyotroph Lateral Scler Frontotemporal Degener 17 (7–8) (2016) 467–472.
[24] S. Vucic, et al., Utility of threshold tracking transcranial magnetic stimulation in ALS, Clin. Neurophysiol. Pract. 3 (2018) 164–172.
[25] E.P. Pioro, M.R. Turner, P. Bede, Neuroimaging in primary lateral sclerosis, Amyotroph Lateral Scler Frontotemporal Degener 21 (sup1) (2020) 18–27.
[26] M. Abidi, et al., Adaptive functional reorganization in amyotrophic lateral sclerosis: coexisting degenerative and compensatory changes, Eur. J. Neurol. 27 (1) (2020) 121–128.
[27] S. Tu, et al., Cerebellar tract alterations in PLS and ALS, Amyotroph Lateral Scler Frontotemporal Degener 20 (3–4) (2019) 281–284.
[28] S. Tu, et al., Regional thalamic MRI as a marker of widespread cortical pathology and progressive frontotemporal involvement in amyotrophic lateral sclerosis, J. Neurol. Neurosurg. Psychiatry 89 (12) (2018) 1250–1258.
[29] C. Schuster, et al., The segmental diffusivity profile of amyotrophic lateral sclerosis associated white matter degeneration, Eur. J. Neurol. 23 (8) (2016) 1361–1371.
[30] V. Grollemund, et al., Development and validation of a 1-year survival prognosis estimation model for amyotrophic lateral sclerosis using manifold learning algorithm UMAP, Sci. Rep. 10 (1) (2020) 13378.
[31] H.J. Westeneng, et al., Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model, Lancet Neurol. 17 (5) (2018) 423–433, https://doi.org/10.1016/S1474-4422(18)30089-9.
[32] P.M. Ferraro, et al., Multimodal structural MRI in the diagnosis of motor neuron diseases, NeuroImage: Clinical 16 (2017) 240–247.
[33] P. Bede, et al., Virtual brain biopsies in amyotrophic lateral sclerosis: diagnostic classification based on in vivo pathological patterns, Neuroimage Clin 15 (2017) 653–658.
[34] P. Bede, A. Murad, O. Hardiman, Pathological neural networks and artificial neural networks in ALS: diagnostic classification based on pathognomonic neuroimaging features, J. Neurol. (2021), https://doi.org/10.1007/s00415-021-10801-5.
[35] M.R. Turner, et al., Primary lateral sclerosis: consensus diagnostic criteria, J. Neurol. Neurosurg. Psychiatry 91 (4) (2020) 373–377.
[36] B. Fischl, FreeSurfer, Neuroimage 62 (2) (2012) 774–781.
[37] B. Fischl, A.M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images, Proc. Natl. Acad. Sci. 97 (20) (2000) 11050–11055.
[38] J.E. Iglesias, et al., Bayesian segmentation of brainstem structures in MRI, Neuroimage 113 (2015) 184–195.
[39] S. Mori, et al., Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template, Neuroimage 40 (2) (2008) 570–582.
[40] S. Mori, et al., MRI Atlas of Human White Matter, Elsevier, The Netherlands, 2005.
[41] S. Wakana, et al., Reproducibility of quantitative tractography methods applied to cerebral white matter, Neuroimage 36 (3) (2007) 630–644.
[42] K. Hua, et al., Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification, Neuroimage 39 (1) (2008) 336–347.
[43] D.L. Collins, et al., Automatic 3-D model-based neuroanatomical segmentation, Hum. Brain Mapp. 3 (3) (1995) 190–208.
[44] J. Mazziotta, et al., A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM), Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 356 (1412) (2001) 1293–1322.
[45] C.A. Brown, et al., Development, validation and application of a new fornix template for studies of aging and preclinical Alzheimer's disease, Neuroimage Clin 13 (2017) 106–115.
[46] M.M. El Mendili, et al., Spinal cord imaging in amyotrophic lateral sclerosis: historical concepts-novel techniques, Front. Neurol. 10 (2019) 350.
[47] G. Querin, et al., The spinal and cerebral profile of adult spinal-muscular atrophy: a multimodal imaging study, Neuroimage Clin 21 (2019), 101618.
[48] G. Querin, et al., Presymptomatic spinal cord pathology in c9orf72 mutation carriers: a longitudinal neuroimaging study, Ann. Neurol. 86 (2) (2019) 158–167.
[49] G. Querin, et al., Multimodal spinal cord MRI offers accurate diagnostic classification in ALS, J. Neurol. Neurosurg. Psychiatry 89 (11) (2018) 1220–1221.
[50] I.R.A. Mackenzie, H. Brieberg, TDP-43 pathology in primary lateral sclerosis, Amyotroph Lateral Scler Frontotemporal Degener (2020) 1–7.
[51] D.E. Lulé, et al., Deficits in verbal fluency in presymptomatic C9orf72 mutation gene carriers-a developmental disorder, J. Neurol. Neurosurg. Psychiatry 91 (11) (2020) 1195–1200.
[52] R.A. Menke, et al., Increased functional connectivity common to symptomatic amyotrophic lateral sclerosis and those at genetic risk, J. Neurol. Neurosurg. Psychiatry 87 (6) (2016) 580–588.
[53] R.H. Chipika, et al., The presymptomatic phase of amyotrophic lateral sclerosis: are we merely scratching the surface? J. Neurol. 268 (12) (2020) 4607–4629, https://doi.org/10.1007/s00415-020-10289-5.
[54] E. Verstraete, et al., Mind the gap: the mismatch between clinical and imaging metrics in ALS, Amyotroph Lateral Scler Frontotemporal Degener 16 (7–8) (2015) 524–529.
[55] K. Bharti, et al., Involvement of the dentate nucleus in the pathophysiology of amyotrophic lateral sclerosis: a multi-center and multi-modal neuroimaging study, Neuroimage Clin 28 (2020), 102385.

[56] P. Bede, et al., Genotype-associated cerebellar profiles in ALS: focal cerebellar pathology and cerebro-cerebellar connectivity alterations, J. Neurol. Neurosurg. Psychiatry 92 (11) (2021) 1197–1205, https://doi.org/10.1136/jnnp-2021-326854.
[57] E. Canu, et al., Extramotor damage is associated with cognition in primary lateral sclerosis patients, PLoS ONE [Electronic Resource] 8 (12) (2013), e82017.
[58] Li Hi, S. Shing, et al., Post-polio syndrome: more than just a lower motor neuron disease, Front. Neurol. 10 (2019) 773.
[59] Li Hi, S. Shing, et al., Increased cerebral integrity metrics in poliomyelitis survivors: putative adaptation to longstanding lower motor neuron degeneration, J. Neurol. Sci. 424 (2021), 117361.
[60] Li Hi, S. Shing, et al., Extra-motor manifestations in post-polio syndrome (PPS): fatigue, cognitive symptoms and radiological features, Neurol. Sci. 42 (11) (2021) 4569–4581, https://doi.org/10.1007/s10072-021-05130-4.
[61] M. Proudfoot, P. Bede, M.R. Turner, Imaging cerebral activity in amyotrophic lateral sclerosis, Front. Neurol. 9 (2018) 1148.
[62] V. Rajagopalan, et al., Diffusion tensor imaging evaluation of corticospinal tract Hyperintensity in upper motor neuron-predominant ALS patients, J. Aging. Res. 2011 (2011), 481745.
[63] P. Bede, G. Querin, P.F. Pradat, The changing landscape of motor neuron disease imaging: the transition from descriptive studies to precision clinical tools, Curr. Opin. Neurol. 31 (4) (2018) 431–438.
[64] J. Gordon, B. Lerner, Insights into amyotrophic lateral sclerosis from a machine learning perspective, J. Clin. Med. 8 (10) (2019).
[65] K. Imamura, et al., Prediction model of amyotrophic lateral sclerosis by deep learning with patient induced pluripotent stem cells, Ann. Neurol. 89 (6) (2021) 1226–1233.
[66] H.K. van der Burgh, et al., Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis, Neuroimage Clin 13 (2017) 361–369.
[67] R.C. Welsh, L.M. Jelsone-Swain, B.R. Foerster, The utility of independent component analysis and machine learning in the identification of the amyotrophic lateral sclerosis diseased brain, Front. Hum. Neurosci. 7 (2013) 251.
[68] T. Fekete, et al., Multiple kernel learning captures a systems-level functional connectivity biomarker signature in amyotrophic lateral sclerosis, PLoS One 8 (12) (2013), e85190.
[69] G. E Elahi, et al., Texture classification of MR images of the brain in ALS using M-CoHOG: a multi-center study, Comput. Med. Imaging Graph. 79 (2020), 101659.
[70] A. Tena, et al., Detection of bulbar involvement in patients with amyotrophic lateral sclerosis by machine learning voice analysis: diagnostic decision support development study, JMIR Med. Inform. 9 (3) (2021), e21331.
[71] D. Beaulieu, et al., Development and validation of a machine-learning ALS survival model lacking vital capacity (VC-free) for use in clinical trials during the COVID-19 pandemic, Amyotroph Lateral Scler Frontotemporal Degener 22 (sup1) (2021) 22–32.
[72] A. Sengur, et al., Classification of amyotrophic lateral sclerosis disease based on convolutional neural network and reinforcement sample learning algorithm, Health Inf. Sci. Syst 5 (1) (2017) 9.
[73] M. Tahedl, et al., Cortical progression patterns in individual ALS patients across multiple timepoints: a mosaic-based approach for clinical use, J. Neurol. 268 (5) (2021) 1913–1926.
[74] M. Tahedl, et al., Evaluation and categorisation of individual patients based on white matter profiles: single-patient diffusion data interpretation in neurodegeneration, J. Neurol. Sci. 428 (2021), 117584.
[75] M. Tahedl, et al., Propagation patterns in motor neuron diseases: individual and phenotype-associated disease-burden trajectories across the UMN-LMN spectrum of MNDs, Neurobiol. Aging 109 (2021) 78–87.
[76] A.L. Young, et al., Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference, Nat. Commun. 9 (1) (2018) 4273.
[77] T. Tong, et al., Five-class differential diagnostics of neurodegenerative diseases using random undersampling boosting, Neuroimage Clin 15 (2017) 613–624.
[78] P.R. Raamana, et al., Three-class differential diagnosis among Alzheimer disease, frontotemporal dementia, and controls, Front. Neurol. 5 (71) (2014).
[79] J. Hu, et al., Deep learning-based classification and voxel-based visualization of frontotemporal dementia and Alzheimer's disease, Front. Neurosci. 14 (2020), 626154.
[80] D. Ma, et al., Differential diagnosis of frontotemporal dementia, Alzheimer's disease, and normal aging using a multi-scale multi-type feature generative adversarial deep neural network on structural magnetic resonance images, Front. Neurosci. 14 (2020) 853.
[81] R.H. Chipika, et al., "switchboard" malfunction in motor neuron diseases: selective pathology of thalamic nuclei in amyotrophic lateral sclerosis and primary lateral sclerosis, Neuroimage Clin 27 (2020), 102300.
[82] R.H. Chipika, et al., Amygdala pathology in amyotrophic lateral sclerosis and primary lateral sclerosis, J. Neurol. Sci. (2020), https://doi.org/10.1007/s10072-021-05130-4, 117039.
[83] F. Christidi, et al., Hippocampal pathology in amyotrophic lateral sclerosis: selective vulnerability of subfields and their associated projections, Neurobiol. Aging 84 (2019) 178–188.
[84] D. Devos, et al., A ferroptosis-based panel of prognostic biomarkers for amyotrophic lateral sclerosis, Sci. Rep. 9 (1) (2019) 2918.
[85] S. Dukic, et al., Patterned functional network disruption in amyotrophic lateral sclerosis, Hum. Brain Mapp. 40 (16) (2019) 4827–4842.
[86] P. Bede, et al., Sexual dimorphism in ALS: exploring gender-specific neuroimaging signatures, Amyotroph Lateral Scler Frontotemporal Degener 15 (3–4) (2014) 235–243.
[87] F. Trojsi, et al., Between-sex variability of resting state functional brain networks in amyotrophic lateral sclerosis (ALS), J. Neural Transm. (Vienna) 128 (12) (2021) 1881–1897.
[88] P.F. Pradat, et al., The French national protocol for Kennedy's disease (SBMA): consensus diagnostic and management recommendations, Orphanet. J. Rare Dis 15 (1) (2020) 90.
[89] G. Querin, et al., Biomarkers of spinal and bulbar muscle atrophy (SBMA): a comprehensive review, Front. Neurol. 9 (2018) 844.
[90] G. Querin, et al., The motor unit number index (MUNIX) profile of patients with adult spinal muscular atrophy, Clin. Neurophysiol. 129 (11) (2018) 2333–2340.
[91] M.V. Lebouteux, et al., Revisiting the spectrum of lower motor neuron diseases with snake eyes appearance on magnetic resonance imaging, Eur. J. Neurol. 21 (9) (2014) 1233–1241.
[92] N. Zhou, P. Manser, Does including machine learning predictions in ALS clinical trial analysis improve statistical power? Ann. Clin. Transl. Neurol. 7 (10) (2020) 1756–1765.