Describing and characterising variability in ALS disease Progression

MUZAMMIL ARIF DIN ABDUL JABBAR[1,2], LING GUO[3], YANG GUO[3], ZACHARY SIMMONS[4], ERIK P. PIORO[5], SAVITHA RAMASAMY[3] & CRYSTAL JING JING YEO[2,5,6,7,8,9]

[1]University of Cambridge, Cambridge, United Kingdom of Great Britain and Northern Ireland, [2]Agency for Science, Technology and Research (A_STAR), Singapore, Singapore, [3]Institute for Infocomm Research (I2R), A_STAR, Singapore, Singapore, [4]Department of Neurology, Pennsylvania State University College of Medicine, University Park, USA, [5]Department of Neurology, Feinberg School of Medicine, Northwestern University, Chicago, USA, [6]Lee Kong Chian School of Medicine, Imperial College London and NTU, Singapore, Singapore, [7]School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK, [8]National Neuroscience Institute, Singapore, Singapore, and [9]Duke-NUS Medical School, Singapore, Singapore

Abstract

Background, Objectives: Decrease in the revised ALS Functional Rating Scale (ALSFRS-R) score is currently the most widely used measure of disease progression. However, it does not sufficiently encompass the heterogeneity of ALS. We describe a measure of variability in ALSFRS-R scores and demonstrate its utility in disease characterization. Methods: We used 5030 ALS clinical trial patients from the Pooled Resource Open-Access ALS Clinical Trials database to calculate variability in disease progression employing a novel measure and correlated variability with disease span. We characterized the more and less variable populations and designed a machine learning model that used clinical, laboratory and demographic data to predict class of variability. The model was validated with a holdout clinical trial dataset of 84 ALS patients (NCT00818389).

Results: Greater variability in disease progression was indicative of longer disease span on the patient-level. The machine learning model was able to predict class of variability with accuracy of 60.1–72.7% across different time periods and yielded a set of predictors based on clinical, laboratory and demographic data. A reduced set of 16 predictors and the holdout dataset yielded similar accuracy.

Discussion: This measure of variability is a significant determinant of disease span for fast progressing patients. The predictors identified may shed light on pathophysiology of variability, with greater variability in fast-progressing patients possibly indicative of greater compensatory reinnervation and longer disease span. Increasing variability alongside decreasing rate of disease progression could be a future aim of trials for faster-progressing patients.

Keywords: ALS, motor neurone disease, machine learning

1. Introduction

Plateaus and reversals have been described in amyotrophic lateral sclerosis (ALS) patients and are not well characterized (1). Plateaus are periods where the Revised ALS Functional Rating Scale (ALSFRS-R) score does not change, whereas reversals are periods where the slope of the ALSFRS-R score is greater than zero (1). Reversals have been hypothesized to be due to ALS mimic syndrome, an endogenous compensatory mechanism, a harmful environmental influence that was removed or undocumented treatment (1). Compensation may due to a polymorphism resulting in motor neuron regeneration at a greater rate than degeneration, such as in ephrin type-A receptor 4 (Epha4), a recognized disease modifier (2). While the US Food and Drug Administration (FDA) guidance for clinical trials recommends therapeutic effect be demonstrated by a slower decrease, improvement or stabilization of the ALSFRS-R or similar scales (3), considering only the numerical decrease per month would mean the loss of information on such plateaus and reversals. Therefore, we devise a new quantitative measure of variability, which describes how large the deviations of ALSFRS-R score are from uniform decrease. We hypothesize that this would be a complementary tool for disease characterization and help us understand influencers of disease variability.

We aim to:

a. Describe a quantitative measure of variability applicable to individual patients within and outside trials
b. Correlate this measure with disease span, the duration from onset of disease symptoms to death
c. Design and validate machine learning (ML) models to predict variability across different time periods, as a complementary tool for prognostication alongside models predicting
d. ALSFRS-R slope (4,5).
e. Characterize the strength of each parameter as a predictor in the ML models to gain insights on their importance

2. Methods

2.1. Data source

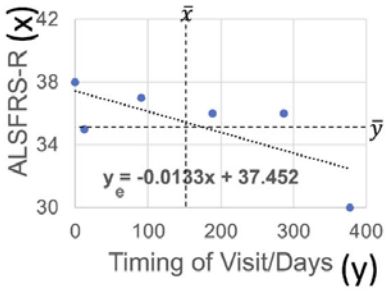
The Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) database comprising more than 10,700 patient records from 23 patient trials (6) was used. ALSFRS/ALSFRS-R scores and time of death data were used to quantify variability and disease span respectively, while demographic, clinical, adverse event and lab parameters were used by the ML model to predict variability. The date of symptom onset was the starting point for calculating time-dependent parameters including disease span. Patient characteristics are given in Table 1. ALSFRS scores (2753 patients) were multiplied by 1.2 to scale with ALSFRS-R scores (2277 patients) (7,8).

Table 1. Description of patient group characteristics.

Total number of patients included	5030
Patients with records in ALSFRS score	2753
Patients with records in ALSFRS-R score	2277
Age/years	55.4 ± 11.5
Percentage male (%)	62.0
Percentage limb onset (%)	72.0
Percentage bulbar onset (%)	22.6
Baseline ALSFRS-R score	36.0 ± 6.7
Baseline FVC percentage(%)	78.0 ± 18.2

Patients with records in ALSFRS score had their scores multiplied with 1.2 to scale with those with ALSFRS-R score.

(A) Trajectory of Patient 533



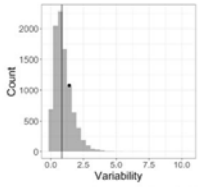
(B) Calculations for Each Visit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Visit No.	Date of Visit/Days	ALSFRS-R Score	Difference between (1) and mean	Difference between (2) and mean	Square of (3)	Multiplication of (3) and (4)	Calculation of expected values based on *m* and *c*	Difference between (2) and (7)	Square of (8), followed by mean and square root
	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(x-\bar{x})(y-\bar{y})$	$y_e = mx + c$	$y-y_e$	$(y-y_e)^2$
1	0	38	-159.50	2.67	25440.25	-425.33	37.45	0.55	0.30
2	12	35	-147.50	-0.33	21756.25	49.17	37.29	-2.29	5.26
3	91	37	-68.50	1.67	4692.25	-114.17	36.24	0.76	0.57
4	189	36	29.50	0.67	870.25	19.67	34.94	1.06	1.12
5	287	36	127.50	0.67	16256.25	85.00	33.64	2.36	5.57
6	378	30	218.50	-5.33	47742.25	-1165.33	32.43	-2.43	5.91
	Mean, $\bar{x}$ = 159.5	Mean, $\bar{y}$ = 35.3			Sum = $\sum(x-\bar{x})^2$ = 116757.50	Sum = $\sum(x-\bar{x})(y-\bar{y})$ - 1551.00			
					$m = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$ = -0.0133	$c = \bar{y} - m\bar{x} =$ **37.45**			

(C) Determining Variability Class for Each Time Period

6-month Period (Days 0 to 183)

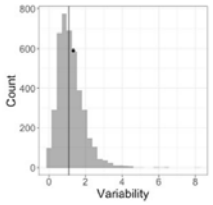
Consider visits 1 to 3:
Mean of Column (8) = 2.04
Variability = $\sqrt{2.04}$ = **1.42**



1.42 is greater than the median of all 6-month periods for all patients (0.831), thus patient 533 is classified as **more variable** for this 6-month period

9-month Period (Days 0 to 273)

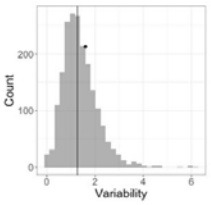
Consider Visits 1 to 4:
Mean of Column (8) = 1.81
Variability = $\sqrt{2.564}$ = **1.34**



1.34 is greater than the median of all 6-month periods for all patients (1.09), thus patient 533 is classified as **more variable** for this 9-month period

12-month Period (Days 0 to 365)

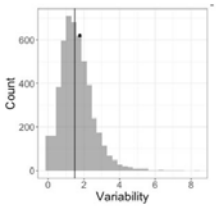
Consider visits 1 to 5:
Mean of Column (8) = 2.654
Variability = $\sqrt{2.564}$ = **1.60**



1.60 is greater than the median of all 12-month periods for all patients (1.27), thus patient 533 is classified as **more variable** for this 12-month period

Entire trial period

Consider all 6 visits:
Mean of Column (8) = 3.12
$\overline{12}$ = **1.77**



1.77 is greater than the median for all patients (1.49), thus patient 533 is classified as **more variable**

Figure 1. Calculation of variability. (A) Sample Trajectory of Patient ID 533, with line of best fit drawn through points. (B) Calculations for each visit, where the points are fitted to a line of best fit and the difference between the line of best fit and expected values for decrease per unit time is calculated. (1) and (2) are the raw data collected by clinicians. From these, a line of best fit is calculated in (7), with the intermediate steps shown in (3) to (6). The difference between the observed (2) and expected (7) is then calculated in (8), and the square is obtained in (9). (C) Based on various time periods, the root-mean-square is calculated by considering all the visits within the specified time period, and then comparing it to the entire pool of variability scores obtained from the PROACT database for that time period.

2.2. Quantification of variability

We quantified variability as the root mean square error (RMSE) of the difference between the ALSFRS-R scores over time and that of a linear model fit of the points. A non-variable patient with a constant decrease over time would have an RMSE of 0 and a more variable patient a larger RMSE. The population median was used to demarcate ‘more variable’ from ‘less variable’ (Figure 1). Notably, this method can only be calculated after at least three ALSFRS-R scores have been collected during a certain period, summarizing the data into a single measure, similar to the D50 method (9), where D50 is the timepoint at which ALSFRS-R is 50% of the maximum.

2.3. Characterization of plateaus and reversals

To understand disease progression patterns over various periods, we calculated the change in ALSFRS-R score during the first and second halves of 4 different time periods, 6, 12, 18 and 24 months, for each patient. Subsequently, we categorized the number of 6-month, 9-month and 12-month windows in which there was no change in ALSFRS-R score as a ‘plateau’ and where there was an increase involving at least 4 data points as a ‘reversal’, in line with a past study (1). There were insufficient (fewer than 4) data points in most patients for the 3-month window.

2.4. Disease span

The number of days from disease onset (time of first muscle weakness) to death, or ‘disease span’, was calculated and correlated to variability. To reduce the confounding effect of ALSFRS-R slope, patients were stratified into slow, intermediate or fast (10,11), using thresholds of 0.5 and 1.5 points per month.

2.5. Machine learning models

To predict future variability, ML models were designed. Patient records with less than half the required data were removed. Missing data was imputed two ways: linear interpolation if other data from the same patient was available from a separate visit, or k-nearest neighbor otherwise. All demographic and clinical parameters, relevant adverse event counts and all lab parameters which were available for at least 50% of patients were included (full list and prevalence given in Supplementary Table 1). Training was performed on 80% of randomly selected patients, while the remaining 20% were used for testing purposes.

We designed several models:

A. Categorical predictive models with observation periods: given the variability in number of visits per time period, the mean, decrease per unit time and standard deviation (SD) for

all parameters and the variability of ALSFRS-R, were calculated for all visits within an observation window of 6, 9 or 12 months to determine the input parameters. The output parameter was the variability of ALSFRS-R for the subsequent 6, 9 or 12 months. The input parameters are thus predictive biomarkers for future variability. As a control, we also built a model using features from the entire trial period but excluding variability, to predict predict variability across the entire trial. This tells us how well other features other than past variability predict ALSFRS-R variability.

B. Categorical predictive models based on single time points: Similar to (A), but based on a single time-point, instead of an observation window
C. Regression predictive models with observation windows: Similar to (A), but predicting the absolute value of variability instead of the binary category.
D. Categorical monitoring model: An additional model where the input parameters and outcome variables were both from the entire trial period. As the input and output variables are from the same time period, the input parameters are monitoring, rather than predictive, biomarkers. SD of ALSFRS-R was not an input parameter as it is highly correlated with RMSE of ALSFRS-R score, but the mean and decrease in ALSFRS-R score were included.
E. Reduced parameter model: To simulate conditions available to clinicians for patients outside trials, we trained an additional model on a reduced set of parameters of 16 parameters, comprising only demographic, clinical and adverse event data (Supplementary Table 1). Due to the shorter time period of about three months for some trials including NCT00818389(12) which we used for holdout validation, this model was trained based on a single time point for prediction of variability over the next three months.
F. Holdout validation model: Similar to (E), but on an holdout dataset of 84 patients from NCT00818389, the lithium ALS trial (12). This dataset was obtained and processed separately from the PRO-ACT database. The training and holdout have minimal patient overlap.

Extreme gradient boost (XGBoost) was used for both the classification and regression models. XGBoost is a gradient boosted decision tree, converting weak decision trees to strong classifiers (13) with an objective function comprising training loss to improve its predictive capability and regularization to avoid overfitting. Thus, XGBoost reduces bias and variance and has shown improved classification accuracy over other models (14). The models were evaluated on their accuracy, referring to the percentage of test samples that were correctly classified by the model, and sensitivity and specificity referring to the percentage of the more and less variable test samples that were correctly identified respectively. In addition, to evaluate the importance of individual input features, XGBoost generates a numerical gain value for each feature, referring to the improvement in accuracy brought by the feature to the branches it is on. The sum of gain values for all features of a single model is 1.

The code to process the data and train the machine learning models is available at github. com/arifjabbar1/ALSVariability.

2.6. Data availability statement

PRO-ACT database available at https://ncri1.partners.org/proact. NCT00818389 dataset is provided through DTUA2021A009305 with Massachusetts General Hospital, Harvard Medical School.

3. Results

3.1. ALS patients have variable progression

Given it has previously been suggested that the rate of decrease in ALSFRS-R is not constant over time (1,15), we investigated how the rates of disease progression changed over different time periods. Both plateaus (Figure 2(A)) and reversals (Figure 2(B)) were observed in the PROACT database and patterns of progressions differ (Figure 2(C)), with. some patients showing consistent progressions or improvements during an entire 12-month window, whereas others progress for a limited period followed by improvement in the next period, or vice versa. Patterns of progression may differ despite the same linear rate of decrease (Figure 2(D)).

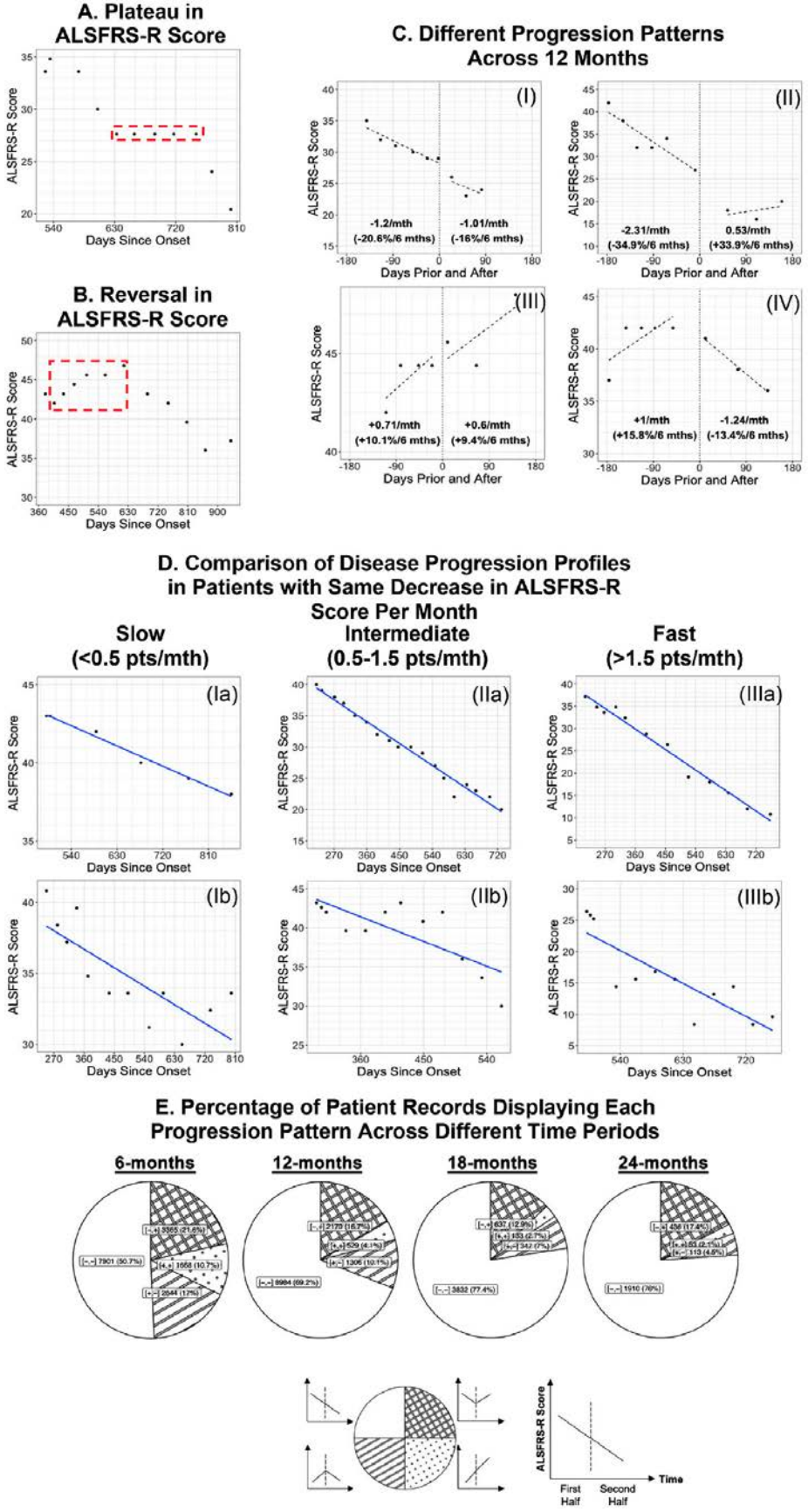

A. Plateau in ALSFRS-R Score
ALSFRS-R Score
Days Since Onset
B. Reversal in ALSFRS-R Score
ALSFRS-R Score
Days Since Onset
C. Different Progression Patterns Across 12 Months
(I)
-1.2/mth (-20.6%/6 mths)
-1.01/mth (-16%/6 mths)
(II)
-2.31/mth (-34.9%/6 mths)
0.53/mth (+33.9%/6 mths)
(III)
+0.71/mth (+10.1%/6 mths)
+0.6/mth (+9.4%/6 mths)
(IV)
+1/mth (+15.8%/6 mths)
-1.24/mth (-13.4%/6 mths)
Days Prior and After
D. Comparison of Disease Progression Profiles in Patients with Same Decrease in ALSFRS-R Score Per Month
Slow (<0.5 pts/mth)
Intermediate (0.5-1.5 pts/mth)
Fast (>1.5 pts/mth)
(Ia)
(Ib)
(IIa)
(IIb)
(IIIa)
(IIIb)
Days Since Onset
E. Percentage of Patient Records Displaying Each Progression Pattern Across Different Time Periods
6-months
12-months
18-months
24-months
ALSFRS-R Score
Time
First Half
Second Half

Figure 2. Each of these plots depicts the progression of ALSFRS-R score, on the vertical axis, against time in days, on the horizontal axis. Reversals refer to periods where the ALSFRS-R score is increasing, progressions where it is decreasing and plateaus where it is constant. Lines where plotted depict lines of best fit through the points. (A) An instance of a plateau in the ALSFRS-R scores, in this case lasting from day 633 to 717 (PRO-ACT database subject ID: 12890). (B) An instance of a reversal in the ALSFRS-R scores, of 4.8 points, lasting from day 407 to 617 (PRO-ACT database subject ID: 903574). (C) Four different individuals (PRO-ACT database subjects with IDs: I 162601, II 127889, III 229601, IV 254204) with varying disease progression patterns have been depicted. subject I shows progression throughout the entire 12 months, Subject II displays progression in the first 6 months followed by improvement in the next 6 months, subject III shows improvement throughout the entire period and subject IV shows improvement in the first 6 months followed by progression in the next. (D) A comparison of corresponding patients with the same numerical rate of decrease of ALSFRS-R score across their enrollment in the trial, for slow, intermediate and fast progressors. Plots Ia and Ib both depict patients with a slow progression, of –0.43 points per month. Plots IIa and IIb depict patients with intermediate progression of –1.16 points per month. Plots IIIa and IIIb depict patients with a fast progression of –1.53 and –1.52 points per month respectively. In all these three cases, patients Ia, IIa and IIIa have progressions that are highly consistent over time, leading to a plot where individual data points lie very closely to the line of best fit, whereas patients Ib, IIb and IIIb have more variable progressions, and the points lie further away from the lines of best fit (PRO-ACT database subject IDs: Ia 146089, Ib 155551, IIa 103471, IIb 116812, IIIa 144196, IIIb 107564). (E) Figure describes the changes in ALSFRS-R score within periods of 6, 12, 18 and 24 months, with the disease progression in the past 3, 6, 9 and 12 months respectively on the horizontal axis and disease progression in the future 3, 6, 9 and 12 months on the vertical axis.

3.2. ALS patients show different rates of disease progression over different time periods

The percentage of patient records showing a decrease in ALSFRS-R score throughout the entire monitoring period increases with length of the period, from 50.7% for the 6-month monitoring period to 76.0% for the 24-month period (Figure 2(E)). Nonetheless, plateaus and reversals persist in longer time periods as shown by the 24% of patient records with an increase or no change in score across at least 12 months. Given the significant proportion of apparent plateaus or reversals, we looked specifically at whether increased variability was correlated with plateaus and reversals.

3.3. More variable ALS patients show more reversals

For all three periods, the more variable group was more likely to show reversals (Table 2), suggesting that variability may correlate with better disease condition. To determine if increased variability is correlated with increased survival, we looked at whether variability is correlated with increased observed disease duration.

Table 2. Distribution of plateaus and/or reversals among more and less variable patients, with p values shown.

	Less variable	More variable	Z statistic	p value
6-month period				
Plateau only	84	37	4.3268	$p < 0.001$
Reversal only	562	736	−5.6397	$p < 0.001$
Both	101	22	7.2149	$p < 0.001$
None	1687	1639	1.4788	$p = 0.13888$
9-month period				
Plateau only	43	24	2.3374	$p = 0.01928$
Reversal only	376	441	−2.4929	$p = 0.01278$
Both	82	9	7.725	$p < 0.001$
None	1933	1960	−0.9669	$p = 0.33204$
12-month period				
Plateau only	45	22	2.8294	$p = 0.00466$
Reversal only	310	329	−0.8064	$p = 0.41794$
Both	80	9	7.5957	$p < 0.001$
None	1999	2074	−2.908	$p = 0.00362$

The p value is obtained by a z score test.

3.4. More variable progression correlated with greater observed disease duration in fast progressing patients

There was no significant difference in disease span, the time from onset of symptoms to death, between less and more variable patients when considered as an entire group as the differences were not as pronounced in the slow and medium progressing groups. However, among fast progressors, the more variable patients had a significantly greater mean disease span by 74 days (Figure 3(A)) suggesting that they survived longer in the disease state. To understand the differences between the more and less variable patients and correlate with known prognostic markers for ALS, we looked at various demographic and clinical markers including site of onset, FVC, age, gender and riluzole use.

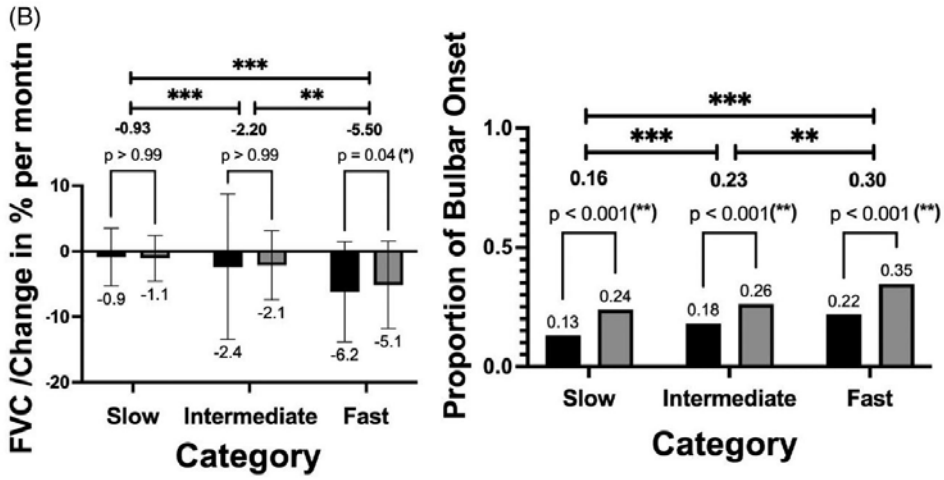
3.5. More variable progression correlated with bulbar onset and slower FVC decrease

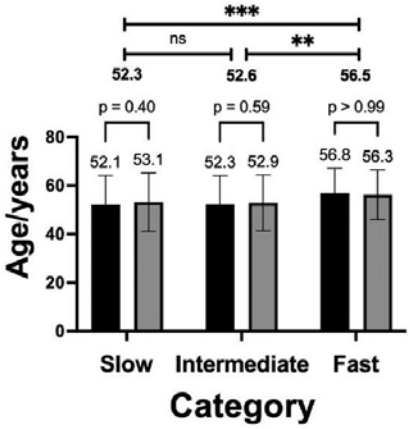
The more variable and less variable patient groups showed similar characteristics except for site of onset, where the more variable group was more likely to be of bulbar onset for all three progression classes, and within the fast group the more variable patients had slower decrease in FVC percentage per month (Figure 3(B)). Given these differences, we looked at whether our machine learning models could accurately classify patients as more or less variable.

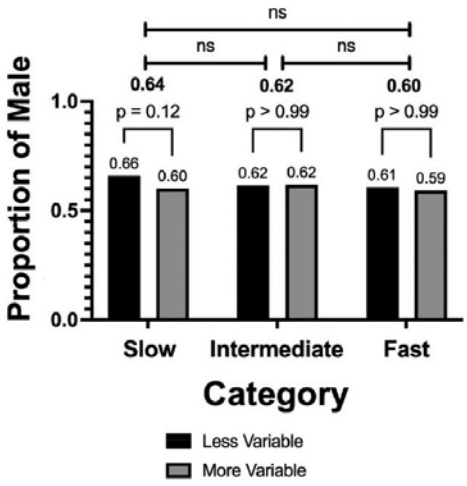
(A)

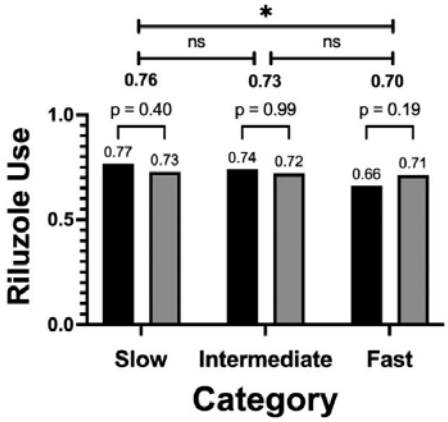
	Disease Span/days (n=number of patients)		
Category	**Slow**	**Intermediate**	**Fast**
Less Variable	1154 (n=81)	1068 (n=246)	782 (n=268)
More Variable	1152 (n=30)	1098 (n=281)	856 (n=478)
z-statistic	0.014679	-0.86672	-2.7674
***p*-value**	0.774	0.3861	**0.005651 (**)**

(B)









Number of Patients

	Less Variable	More Variable	Total
Slow	1028	391	1419
Intermediate	1007	1251	2258
Fast	399	792	1191

Figure 3. (A) Mean disease spans, referring to time from disease onset to death, for slow, intermediate and fast progressing patients. p values are comparison of distributions of days to death from disease onset between more and less variable groups. (B) Demographic data of more and less variable patients for slow, intermediate and fast categories. Means of the more and less variable patients within each category are indicated above bars, and error bars where shown are standard deviations of continuous variables. Bolded numbers above plots are the overall means for all patients within each category of progression. p values are obtained from 2-tailed z test of distributions for each category with Bonferroni corrections. _refers to adjusted p values below 0.05, __ below 0.001 and ___ below 0.0001. ns refers to adjusted p values above 0.05.

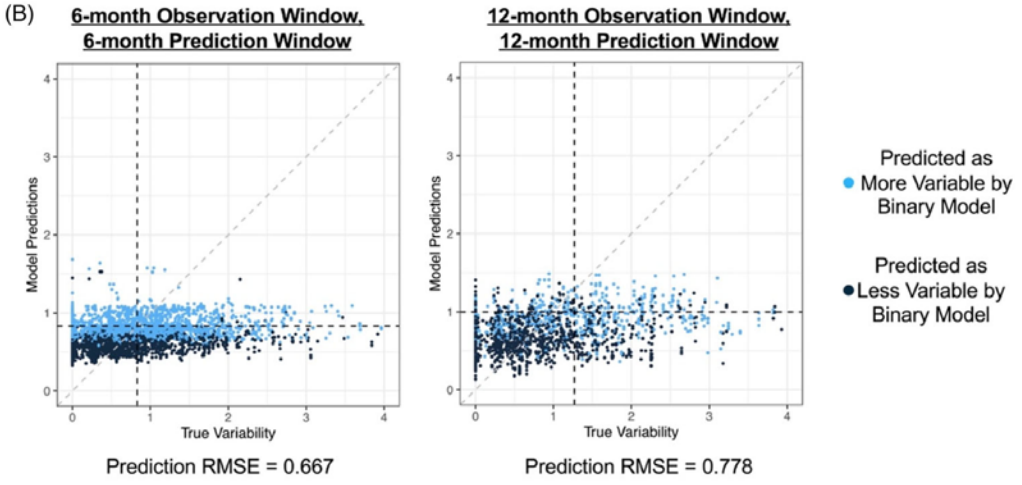
3.6. Machine learning models show good accuracy on internal and holdout datasets, and in full and reduced sets of parameters

To study if observation windows could predict future variability, we designed such models which showed good overall accuracy of 60.1–72.7%, slightly greater for longer timespans (Figure 4(A)). As single timepoint models require less prior information, they have larger applicability thus we designed them as well however accuracy is lower at 53.5–57.8%. Similarly, models predicting the absolute value of variability could be more informative however the root mean squared error (RMSE) of their predictions is 0.667 and 0.778 for 6-month and 12-month models (Figure 4(B)) which are 47.0% and 48.6% of the median variabilities respectively, demonstrating their relatively poor predictive power especially at high and low values of variability. To determine the applicability of the model for settings where less information may be available such as outside trials, a reduced parameter model was designed, showing similar accuracy as the full model (Figure 4(C)). In addition, to determine its applicability on new datasets, we tested the model on data from NCT00818389, which showed similar accuracy as well (Figure 4(C)). Given the good accuracy of our models, we studied key predictors generated by the four best-performing models with observation periods to gain insights into the biological basis for variability.

(A)

Observational Window	Single Time-Point	Prior 6 months	Single Time-Point	Prior 9 months	Single Time-Point	Prior 12 months	Entire Trial Period (387 ± 196 days)
Prediction Window	Next 6 months		Next 9 months		Next 12 months		
Accuracy	57.8%	61.3%	55.2%	60.1%	53.5%	63.1%	72.7%
Sensitivity	54.9%	61.3%	51.5%	57.6%	47.3%	72.0%	71.2%
Specificity	61.8%	61.4%	61.5%	61.7%	65.4%	58.6%	74.3%

(B)



(C)

Dataset	PRO-ACT Test Data		Lithium Trial
Number of Parameters	48	16	
Observational Window	Single Time-Point		
Prediction Window	Next 3 months		
Accuracy	60.6%	59.0%	58.4%
Sensitivity	59.5%	58.2%	63.8%
Specificity	61.5%	61.0%	56.4%

Figure 4. (A) Accuracy of various machine learning models. Accuracy data refers to the percentage of test samples that were correctly classified by the model. Of these, sensitivity and specificity refer to the percentage of the more and less variable test samples that were correctly identified respectively. The performance values under 'Entire Trial Period' refer to the how well features excluding variability across the entire trial, predict variability across the entire trial. B) Accuracy of the 3-month prediction model, based on a single time point for the full and reduced number of parameters. The PRO-ACT database is the internal validation set and lithium trial the holdout validation set. (C) Predictions of absolute value model. Points falling near the diagonal line are more accurate predictions while the horizontal and vertical lines are the thresholds for more and less variable for 6 months (0.831) and 12 months (1.27).

3.7. Past ALSFRS-R scores, number of adverse events and site of onset are possible predictors of variability

Among the top ten predictors, past ALSFRS-R scores, number of adverse events and site of onset were the top clinical parameters, and phosphorus, albumin and sodium were the top biochemical markers (Table 3). Density plots were generated for each of the top predictive features in the four models (Figure 5).

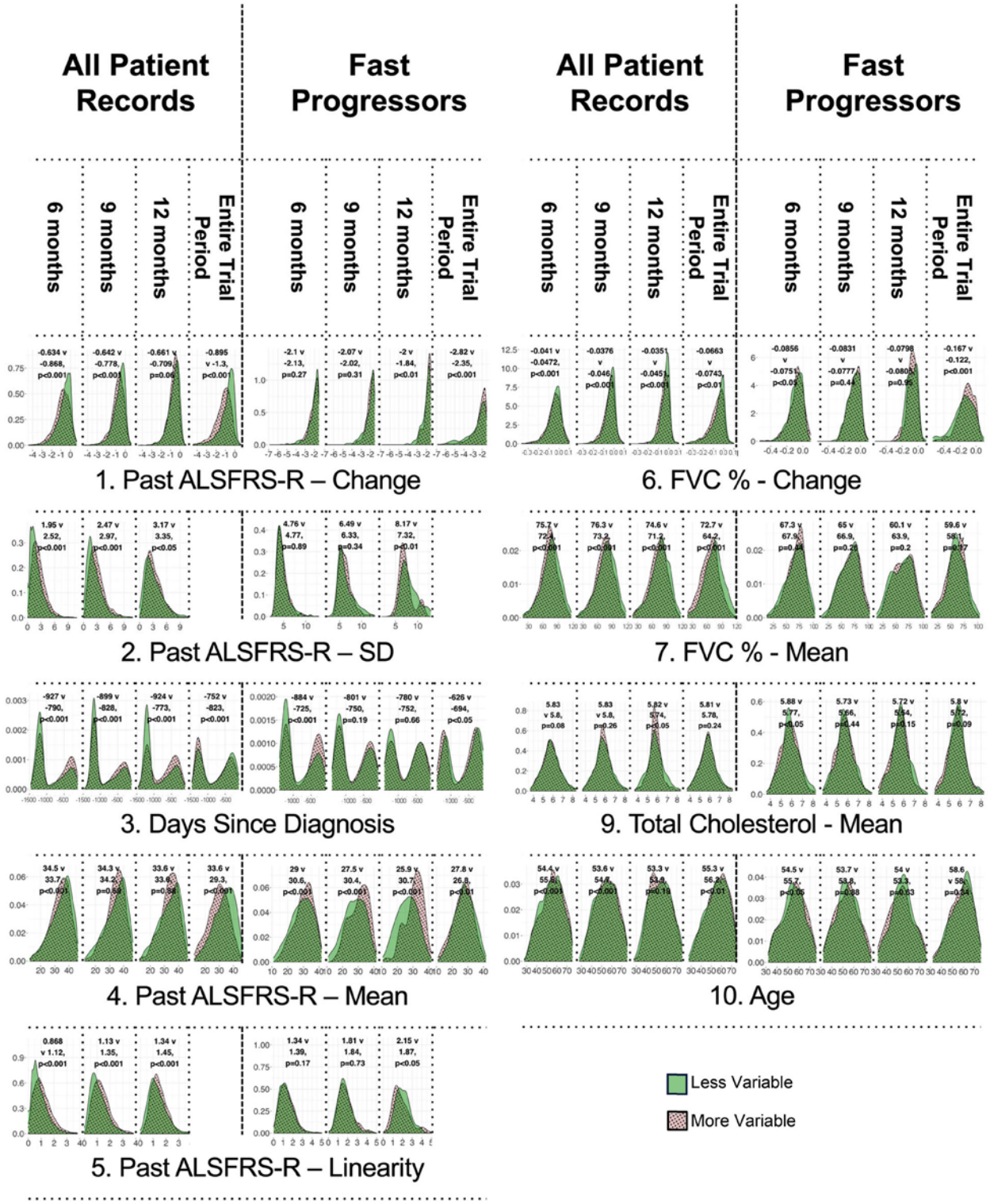


Figure 5. Density plots depicting the properties of the more and less variable patients for each model, for the top 10 features with highest gains. Feature 2, number of adverse events and 5, site of onset, could not be plotted as they are discrete rather than continuous variables. Plots were constructed depicting all patients as well as only fast patients in order to reduce the effect of differences in disease progression, which may be confounding in some cases. Above each density plot, the mean of the feature in less and more variable patients respectively is shown, as well as the p value of the comparison of the feature in these two populations.

Table 3. Importance of different biomarkers in predicting variability.

Overall rank	Feature	Predictive biomarkers: Rank in 6-month model	Predictive biomarkers: Rank in 9-month model	Predictive biomarkers: Rank in 12-month model	Predictive biomarkers: Total gain (across 3 models)	Monitoring biomarkers: Entire trial period rank	Monitoring biomarkers: Gain	Total gain across 4 models
1	Past ALSFRS-R change				0.000	1	0.456	0.000
2	Past ALSFRS-R SD	1	2	2	0.384	–	–	0.384
3	Days since diagnosis	2	7	1	0.372			0.372
4	Past ALSFRS-R mean	6	4	3	0.177	2	0.108	0.177
5	Past ALSFRS-R linearity	3	1	7	0.229	–	–	0.229
6	FVC % change	5		5	0.115			0.115
7	FVC % mean	9	5		0.047	3	0.060	0.047
8	Site (Limb)	4			0.089			0.089
9	Total cholesterol mean			4	0.082			0.082
10	Age			6	0.067			0.067

4. Discussion

4.1. Utility of our measure

Variability is not routinely studied or measured due to the lack of a suitable metric, with previous papers describing reversals as “arbitrary”(1). The contribution of our method to the field is allowing for calculation of variability in all patients with at least three ALSFRS-R scores. Although our study was based on the PROACT database, the measure can be studied and potentially applied to patients outside clinical trials.

Current methods of characterizing reversals and plateaus (1,16,17) search for groups of points while ignoring the remaining points, and may therefore include artefactual apparent reversals or plateaus caused by symptomatic treatment instead of true reversal in disease pathophysiology (1) or errors in measurement of ALSFRS-R score which has a test–retest variability of about 4.3 points by some estimates (18). Our measure is holistic and considers all points in disease progression, making it less sensitive to one-off fluctuations.

The difference in disease span in fast progressors between more and less variable patients of 2.4

months is slightly greater than the treatment effect of riluzole of 2.3 months (19) at 100 mg/day. Thus, we hypothesize that therapeutics which increase variability may also increase disease span in fast progressors, indicating potential value of tracking variability during clinical trials.

However, a similar trend was not observed for the intermediate and slow progressors. Disease pathophysiology may differ in these groups, or the effects of variability less obvious due to the disease progression being less drastic. Moreover, slow and intermediate progressors were less likely to die during trials and thus we were unable to calculate disease span for many of them. Of those who died during the trial and included in the analysis, they were likely enrolled at a later disease stage, where the capacity for compensatory reinnervation, discussed below, is slower (20) thus the effect on disease span less obvious.

4.2. Possible mechanisms for reinnervation

We hypothesize that compensatory reinnervation may explain the increased disease span in fast progressors. This phenomenon has been observed in some ALS patients through Motor Unit Number Estimation (MUNE) and compound motor action potential (CMAP) measurements.

These patients had a large decrease in MUNE but a slower decline in CMAP (21), suggesting that motor neuron death was compensated by muscle fibers reinnervation. Muscle biopsies have also revealed a significantly higher presence of reinnervated neuromuscular junctions in longer surviving patients (22).

While ALSFRS-R has been regarded as less sensitive than other measures of disease progression, including MScan (21), multipoint incremental MUNE (23), and CMAP (24), we chose to study it as it remains the recommended standard for clinical trials (3). Our method can calculate variability in these measures as well by substituting ALSFRS-R for the relevant measure (Figure 2), or other functional measures such as the Rasch-Built Overall Amyotrophic Lateral Sclerosis Disability Scale (ROADS) (25) should a new standard emerge.

4.3. Utility of machine learning model

As calculating our variability measure requires at least three scores in a certain observation period, our accurate ML models will allow for future variability to be predicted and for fast patients, correlated to disease span. The longer term 12-month model was more accurate than the shorter term 6- and 9- month models, as it has input from a longer period, and made predictions over a longer period where small fluctuations in ALSFRS-R score were less likely to cause errors in classification. Similarly, single time point models showed reduced accuracy as past progression patterns could not be studied. Shorter prediction windows had greater accuracy, suggesting it is more difficult to predict further in the future, where the cumulative likelihood of unexpected adverse events is greater and volatility in progression (26) makes prediction further into the future more difficult. Our models are highly generalizable to other trials such as NCT00818389 and where fewer parameters are available, which could be the case for patients outside trials.

As we are interested in understanding the trends of variability in disease progression, we believe a classification model is more appropriate than a regression model, further supported by the lower accuracies of the regression model. For fast progressing patients, our models can complement existing machine learning models which predict rate of ALSFRS-R score decline (4,27,28) or overall survival (29). This could be useful for measuring effects of therapeutics attempting to increase reinnervation, such as miRNA-206 which promotes neuromuscular synapse regeneration (30) and neuregulin-1(Nrg1) which enhances muscle reinnervation (31).

4.4. Top clinical features identified by model

Greater variability, mean and SD of past ALSFRS-R score was predictive of greater future variability, suggesting that patients displaying compensatory reinnervation in the past are more likely to do so in the future. Moreover, more variable patients were more likely to have bulbar onset, and less likely to have limb onset, although this may be artefactual due to symptomatic management of sialorrhea and pseudobulbar affect by anticholinergic drugs (32) and dextromethorphan/quinidine (33) respectively in bulbar patients. Finally, more variable patients had fewer days from diagnosis to current time indicating earlier disease stages, as reinnervation capacity appears to decline in late disease (20). Days since diagnosis may be underestimated in more variable patients as their increased reinnervation potential may mask initial motor neuron death and delay symptoms such as weakness, which may require motor neuron pool depletion by 50% or more (34).

4.5. Top biochemical biomarkers identified

More variable patients had significantly lower total cholesterol levels, which have been previously associated with poorer nutrition and prognosis (35). Given that more variable fast-progressing patients have a longer disease span, this is counterintuitive and may suggest novel pathways or that lipid metabolism may be altered in patients with increased reinnervation.

4.6. Limitations

Limitations of this study include: firstly, there is heterogeneity between the various clinical trials in the PRO-ACT database and it may not represent the true ALS patient population, hence the population median demarcating more from less variable patients may not be widely generalizable. Future studies could involve finding appropriate medians for patients outside trials. Secondly, genomic and cognitive data, potential predictors of variability, was not available. Finally, a linear model was used even though typical ALS progression is curvilinear (15). Using a curvilinear model requires more than one fitting parameter, thus variability would not have been summarized as a single value, but future work could explore the correlation between different fitting parameters and disease span.

5. Conclusion

From a clinical and research perspective, the mathematical method for quantification of variability may be used for disease characterization alongside ALSFRS-R slope. ML models can predict future variability and stratify patients for therapeutic intervention in addition to identifying important biomarkers. The mathematical basis of this method can be extended to individual components of the ALSFRS-R score, or other clinical scores and biomarkers.

Acknowledgements

Our thanks to Dr James Berry, Massachusetts General Hospital, and the NEALS consortium for kindly providing the Lithium trial dataset. Our thanks also to all contributors to the ProACT database. Finally, we thank all patients who have contributed their information to the patient databases and enabled us to work towards improving patient care and understanding ALS through our research using these databases.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

There is no funding to be declared.

References

1. Bedlack RS, Vaughan T, Wicks P, Heywood J, Sinani E, Selsov R, et al. How common are ALS plateaus and reversals? Neurology 2016;186:808–12.

2. Van Hoecke A, Schoonaert L, Lemmens R, Timmers M, Staats KA, Laird AS, et al. EPHA4 is a disease modifier of amyotrophic lateral sclerosis in animal models and in humans. Nat Med. 2012;18:1418–22.

3. Center for Drug Evaluation and Research (CDER) & Center for Biologics Evaluation and Research (CBER), Food and Drug Administration (FDA). Amyotrophic lateral sclerosis: developing drugs for treatment guidance for industry.

4. K€uffner R, Zach N, Norel R, Hawe J, Schoenfeld D, Wang L, et al. Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. Nat Biotechnol. 2015;33:51–7.

5. Ong ML, Tan PF, Holbrook JD. Predicting functional decline and survival in amyotrophic lateral sclerosis. PLoS One. 2017;12:e0174925.

6. Atassi N, Berry J, Shui A, Zach N, Sherman A, Sinani E, et al. The PRO-ACT database: design, initial analyses, and predictive features. Neurology 2014;83:1719–25.

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

8. The Amyotrophic Lateral Sclerosis Functional Rating Scale: assessment of activities of daily living in patients with amyotrophic lateral sclerosis. Arch Neurol. 1996;53:141–7.

9. Stubendorff B, Grehl T, Neuwirth C, R€odiger A, Gunkel A, Radscheidt M, et al. P 48 The time of the ALSFRS-R to decrease to 50% (D50) in a sigmoidal decay model sufficiently describes the complete disease course of amyotrophic lateral sclerosis. Clin Neurophysiol. 2017;128:e353–4–e354.

10. Steinacker P, Huss A, Mayer B, Grehl T, Grosskreutz J, Borck G, et al. Diagnostic and prognostic significance of neurofilament light chain NF-L, but not progranulin and S100B, in the course of amyotrophic lateral sclerosis: data from the German MND-net. Amyotroph Lateral Scler Frontotemporal Degener. 2017;18:112–9

11. Huber RG, Pandey S, Chhangani D, Rincon-Limas DE, Staff NP, Yeo CJJ. Identification of potential pathways and biomarkers linked to progression in ALS. Ann Clin Transl Neurol. 2023;10:150–65. 10.1002/acn3.51697

12. Aggarwal SP, Zinman L, Simpson E, McKinley J, Jackson KE, Pinto H, et al. Safety and efficacy of lithium in combination with riluzole for treatment of amyotrophic lateral sclerosis: a randomised, double-blind, placebocontrolled trial. Lancet Neurol. 2010;9:481–8.

13. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. San Francisco California USA: ACM; 2016:785–94 [cited 2022 Nov 4]. Available at: 10.1145/2939672.2939785

14. Ramaneswaran S, Srinivasan K, Vincent PMDR, Chang CY. Hybrid inception v3 XGBoost Model for acute lymphoblastic leukemia classification. Comput Math Methods Med. 2021;2021:e2577375–10.

15. Gordon PH, Cheng B, Salachas F, Pradat PF, Bruneteau G, Corcia P, et al. Progression in ALS is not linear but is curvilinear. J Neurol. 2010;257:1713–7.

16. Vasta R, D'Ovidio F, Canosa A, Manera U, Torrieri MC, Grassano M, et al. Plateaus in amyotrophic lateral sclerosis progression: results from a population-based cohort. Eur J Neurol. 2020;27:1397–404.

17. Hu N, Shen D, Yang X, Cui L, Liu M. Plateaus and reversals evaluated by different methods in patients with limb-onset amyotrophic lateral sclerosis. J Clin Neurosci. 2022;97:93–8.

18. Bakker LA, Schr€oder CD, Tan HHG, Vugts SMAG, van Eijk RPA, van Es MA, et al. Development and assessment of the inter-rater and intra-rater reproducibility of a self administration version of the ALSFRS-R. J Neurol Neurosurg Psychiatry. 2020;91:75–81.

19. Baldinger R, Katzberg HD, Weber M. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). Cochrane Database Syst Rev. 2012;2012:CD004157.

20. Krarup C. Lower motor neuron involvement examined by quantitative electromyography in amyotrophic lateral sclerosis. Clin Neurophysiol. 2011;122:414–22.

21. Jacobsen AB, Bostock H, Tankisi H. Following disease progression in motor neuron disorders with 3 motor unit number estimation methods. Muscle Nerve. 2019;59:82–7.

22. Bruneteau G, Simonet T, Bauch_e S, Mandjee N, Malfatti E, Girard E, et al. Muscle histone deacetylase 4 upregulation in amyotrophic lateral sclerosis: potential role in reinnervation ability and disease progression. Brain. 2013;136:2359–68.

23. Jagtap SA, Kuruvilla A, Govind P, Nair MD, Sarada C, Varma RP. Multipoint incremental motor unit number estimation versus amyotrophic lateral sclerosis functional rating scale and the medical research council sum score as an outcome measure in amyotrophic lateral sclerosis. Ann

Indian Acad Neurol. 2014;17:336–9.

24. Liu XX, Zhang J, Zheng JY, Zhang S, Xu YS, Kang DX, et al. Stratifying disease stages with different progression rates determined by electrophysiological tests in patients with amyotrophic lateral sclerosis. Muscle Nerve. 2009; 39:304–9.

25. Johnson SA, Burke KM, Scheier ZA, Keegan MA, Clark AP, Chan J, et al. Longitudinal comparison of the self-entry Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-RSE) and Rasch-Built Overall Amyotrophic Lateral Sclerosis Disability Scale (ROADS) as outcome measures in people with amyotrophic lateral sclerosis. Muscle Nerve. 2022;66:495–502.

26. Simon NG, Turner MR, Vucic S, Al-Chalabi A, Shefner J, Lomen-Hoerth C, et al. Quantifying disease progression in amyotrophic lateral sclerosis. Ann Neurol. 2014;76:643–57.

27. Gomeni R, Fava M, Pooled Resource Open-Access ALS Clinical Trials Consortium. Amyotrophic lateral sclerosis disease progression model. Amyotroph Lateral Scler Frontotemporal Degener. 2014;15:119–29.

28. Hothorn T, Jung HH. RandomForest4Life: a Random Forest for predicting ALS disease progression. Amyotroph Lateral Scler Frontotemporal Degener. 2014;15:444–52.

29. Westeneng HJ, Debray TPA, Visser AE, van Eijk RPA, Rooney JPK, Calvo A, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. Lancet Neurol. 2018;17:423–33.

30. Williams AH, Valdez G, Moresi V, Qi X, McAnally J, Elliott JL, et al. MicroRNA-206 delays ALS progression and promotes regeneration of neuromuscular synapses in mice. Science. 2009;326:1549–54.

31. Mancuso R, Mart_ınez-Muriana A, Leiva T, Gregorio D, Ariza L, Morell M, et al. Neuregulin-1 promotes functional improvement by enhancing collateral sprouting in SOD1(G93A) ALS mice and after partial muscle denervation. Neurobiol Dis. 2016;95:168–78.

32. Garuti G, Rao F, Ribuffo V, Sansone VA. Sialorrhea in patients with ALS: current treatment options. DNND. 2019;9:19–26.

33. Brooks BR, Thisted RA, Appel SH, Bradley WG, Olney RK, Berg JE, et al. Treatment of pseudobulbar affect in ALS with dextromethorphan/quinidine: A randomized trial. Neurology. 2004;63:1364–70.

34. Gunes T, Sirin NG, Sahin S, Kose E, Isak B. Use of CMAP, MScan fit-MUNE, and MUNIX in understanding neurodegeneration pattern of ALS and detection of early motor neuron loss in daily practice. Neurosci Lett. 2021;741:135488.

35. Dorst J, K€uhnlein P, Hendrich C, Kassubek J, Sperfeld AD, Ludolph AC. Patients with elevated triglyceride and cholesterol serum levels have a prolonged survival in amyotrophic lateral sclerosis. J Neurol. 2011;258:613–7.