



HAL
open science

Manifold learning for amyotrophic lateral sclerosis functional loss assessment

Vincent Grollemund, Gaétan Le Chat, Marie-Sonia Secchi-Buhour, François Delbot, Jean-François Pradat-Peyre, Peter Bede, Pierre-François Pradat

► **To cite this version:**

Vincent Grollemund, Gaétan Le Chat, Marie-Sonia Secchi-Buhour, François Delbot, Jean-François Pradat-Peyre, et al.. Manifold learning for amyotrophic lateral sclerosis functional loss assessment. *Journal of Neurology*, 2021, 268 (3), pp.825 - 850. 10.1007/s00415-020-10181-2 . hal-03401779

HAL Id: hal-03401779

<https://hal.science/hal-03401779v1>

Submitted on 25 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Manifold learning for amyotrophic lateral sclerosis functional loss assessment

Development and validation of a prognosis model

Vincent Grollemund^{1,2} · Gaétan Le Chat² · Marie-Sonia Secchi-Buhour² · François Delbot^{1,3} · Jean-François Pradat-Peyre^{1,3} · Peter Bede^{4,5,6} · Pierre-François Pradat^{4,5,7}

Received: 22 June 2020 / Revised: 5 August 2020 / Accepted: 6 August 2020 / Published online: 4 September 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Amyotrophic lateral sclerosis (ALS) is an inexorably progressive neurodegenerative condition with no effective disease-modifying therapy at present. Given the striking clinical heterogeneity of the condition, the development and validation of reliable prognostic models is a recognised research priority. We present a prognostic model for functional decline in ALS where outcome uncertainty is taken into account. Patient data were reduced and projected onto a 2D space using Uniform Manifold Approximation and Projection (UMAP), a novel non-linear dimension reduction technique. Information from 3756 patients was included. Development data were sourced from past clinical trials. Real-world population data were used as validation data. Predictors included age, gender, region of onset, symptom duration, weight at baseline, functional impairment, and estimated rate of functional loss. UMAP projection of patients showed an informative 2D data distribution. As limited data availability precluded complex model designs, the projection was divided into three zones defined by a functional impairment range probability. Zone membership allowed individual patient prediction. Patients belonging to the first zone had a probability of 83% ($\pm 3\%$) to have an ALSFRS score over 20 at 1-year follow-up. Patients within the second zone had a probability of 89% ($\pm 4\%$) to have an ALSFRS score between 10 and 30 at 1 year follow-up. Finally, patients within the third zone had a probability of 88% ($\pm 7\%$) to have an ALSFRS score lower than 20 at 1 year follow-up. This approach requires a limited set of features, is easily updated, improves with additional patient data, and accounts for results uncertainty. This method could therefore be used in a clinical setting for patient stratification and outcome projection.

Keywords ALS · Prognosis · UMAP · Manifold learning · Non-linear dimension reduction

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00415-020-10181-2>) contains supplementary material, which is available to authorized users.

✉ Vincent Grollemund
vincent.grollemund@lip6.fr

¹ LIP6, Sorbonne Université, Paris, France

² FRS Consulting, Paris, France

³ Modal'X, Nanterre Université, Nanterre, France

⁴ Laboratoire d'Imagerie Biomédicale, Sorbonne Université, Paris, France

Introduction

Amyotrophic lateral sclerosis (ALS) is a relentlessly progressive neurodegenerative condition involving both the upper and lower motor neurons, leading to progressive limb weakness and bulbar dysfunction. Mean survival from

⁵ Département de Neurologie, Pitié-Salpêtrière University Hospital, APHP, Paris, France

⁶ Computational Neuroimaging Group, Trinity College, Dublin, Ireland

⁷ Northern Ireland Center for Stratified Medicine, Biomedical Sciences Research Institute Ulster University, Londonderry, UK

symptom onset is typically 3–5 years [1], with death occurring secondary to respiratory failure. The disease is characterised by considerable clinical heterogeneity [2] and considerable differences in progression rates [3].

In clinical practice, accurate prognostic modelling is crucial for planning multidisciplinary interventions, recruitment into clinical trials and informing patients and caregivers. Disease heterogeneity is a major barrier to accurate patient stratification in clinical trials [4]. Previous cohort studies identified a number of negative prognostic indicators in ALS [5]. These include both demographic and clinical parameters and encompass age, bulbar involvement, short symptom onset to diagnosis interval, marked motor disability, cognitive impairment, fast progression rate, low vital capacity and c9orf72 status [6–9]. However, prognosis prediction at individual patient level is not reliable based on clinical and demographic variables alone [6]. There is a growing trend to develop prognostic tools based on a combination of prognostic indicators [5, 10], using supervised machine learning models such as random forests [11], neural networks with random forests [12], regression models [13], Royston–Parmar regression models [14] and boosting algorithms [15]. In recent years, prognosis research has focused more on patient survival [11, 12, 14, 15] than functional decline [15–17], although functional loss is central for personalised patient care.

Disease progression in ALS can be tracked by standardised clinical instruments [18], biofluid markers [19], or quantitative neuroimaging [20]. Functional decline is also predominantly appraised by the ALS Functional Rating Scale (ALSFRS) [21], a composite motor score evaluating key domains of mobility and daily living. The ALSFRS score is calculated by evaluating disability in the bulbar, respiratory, gross and fine motor functions. Accordingly, the ALSFRS scores only provide a broad overview of motor disability, and does not evaluate impairments in extrapyramidal, cerebellar, cognitive and behavioural domains. Currently, the revised version of ALSFRS (ALSFRS-R) [22] is the most widely used instrument to monitor motor disability in clinical practice and pharmacological trials. Clinical staging systems were designed to circumvent these limitations; stages are defined based on key clinical progression milestones. Three clinical staging systems have been proposed for ALS: King's [23], Milano-Torino System (MiToS) [24] and Fine Till 9 (FT9) [25]. King's staging system describes more specifically earlier stages of disease progression, while MiToS details later stages of disease progression with a higher resolution [26]. FT9 tries to balance both [25]. Stage membership has already been successfully used in patient stratification [27] and drug performance evaluation in clinical trial settings [28].

Unsupervised learning methods provide new modelling possibilities given their ability to evaluate data distributions

without strong underlying statistical hypothesis [29, 30]. These methods support data visualisation by projecting data onto a new low-dimensional space. Uniform Manifold Approximation and Projection (UMAP) [31] is a novel method based on non-linear dimension reduction which can be readily combined with probability assessments. The main objective of this study was to evaluate a UMAP-based 1-year functional loss prediction model in ALS, designed using three clinical trial datasets, and validated by a real-world (RW) dataset. ALS staging systems were analysed with regard to their ability to segregate patients from our datasets.

Methods

Patient population

Validation and test data for this research included a total of 3756 patients from four different datasets, three of which originated from clinical trials. The first dataset, which is referred to as Trophos, was a clinical trial for olesoxime, a drug developed by Trophos [32] which included 512 patients. After excluding samples with missing data, 357 patients remained. The second dataset, Exonhit, was a clinical trial for pentoxifylline, a drug produced by Exonhit Pharma [33] which included 400 patients. Following the exclusion of incomplete samples, data from 227 patients were retained. The third database was PRO-ACT, funded by the ALS Therapy Alliance and released in 2012 as part of the DREAM Phil Bowen ALS prediction Prize4Life competition. PRO-ACT consists of pooled data from 16 completed phase II–III clinical trials and one observational study [34]. The original PRO-ACT sample size was 10723, reduced to 2841 after discarding samples with missing data. The fourth dataset was population based and contained RW patient data. These data were obtained from the database of the Paris tertiary referral centre for ALS and collected between September 1999 and April 2008. The original sample size was 1377 which was reduced to 331 after the removal of incomplete samples. Baseline patient feature distribution for 1-year functional loss analysis is presented for each cohort in Table 1. Additional information on each dataset is provided as supplementary information.

Clinical predictors and outcome

The primary outcome was 1-year functional loss. Secondary outcomes include 1-year ALSFRS subscores as well as 1-year King's, MiToS and FT9 clinical stages. MiToS and FT9 clinical stages are calculated from the ALSFRS question set, while King's staging can be approximated using ALSFRS questions [35]. 1-year ALSFRS subscores and

Table 1 Predictor distribution per dataset

Features	PRO-ACT	Trophos	Exonhit	RW	Overall
<i>n</i>	2841	357	227	331	3756
Gender (male/female)	1811/1030	235/122	148/79	181/150	2375/1381
Onset (spinal/bulbar)	2263/578	291/66	176/51	256/75	2986/770
Age (years)	54.8 ± 11.5 (18:81)	55.7 ± 11.2 (26:79)	53.9 ± 12.1 (22:77.9)	60.3 ± 12.2 (27.7:88.4)	55.3 ± 11.7 (18:88.4)
Symptom duration (months)	21.2 ± 12.7 (0.5:102.5)	17 ± 8.1 (5:38)	26.1 ± 12.4 (4:58)	23.9 ± 22.8 (0:153.8)	21.2 ± 12.7 (0:153.8)
Baseline weight (kg)	76.5 ± 15.8 (30:148.6)	72.3 ± 12.8 (41:130)	72.7 ± 14.4 (45:130)	70.7 ± 12.7 (40:140)	76.5 ± 15.8 (30:148.6)
Baseline ALSFRS (score)	31.3 ± 5.1 (11:40)	32.8 ± 4 (16:40)	29.2 ± 6.1 (10:39)	30.9 ± 6.3 (8:40)	31.3 ± 5.1 (8:40)
Baseline ALSFRS decline rate (score/month)	−0.52 ± 0.42 (−5.65:0)	−0.49 ± 0.36 (−2.67:0)	−0.49 ± 0.33 (−2:−0.03)	−0.59 ± 0.48 (−2.93:0)	−0.52 ± 0.41 (−5.65:0)
Baseline upper limbs ALSFRS subscore (score)	5.9 ± 2 (0:8)	N.A.	5.4 ± 2.3 (0:8)	5.9 ± 2.3 (0:8)	5.9 ± 2.1 (0:8)
Baseline lower limbs ALSFRS subscore (score)	5.1 ± 2.2 (0:8)	N.A.	4.7 ± 2.3 (0:8)	4.9 ± 2.3 (0:8)	5 ± 2.2 (0:8)
Baseline trunk ALSFRS subscore (score)	5.9 ± 1.8 (0:8)	N.A.	5.4 ± 2.1 (0:8)	5.8 ± 2.1 (0:8)	5.9 ± 1.8 (0:8)
Baseline bulbar ALSFRS subscore (score)	10.5 ± 2 (0:12)	N.A.	10 ± 2.4 (0:12)	10.5 ± 2.1 (3:12)	10.5 ± 2.1 (0:12)
Baseline respiratory ALSFRS subscore (score)	3.8 ± 0.4 (1:4)	N.A.	3.7 ± 0.5 (3:4)	3.8 ± 0.4 (1:4)	3.8 ± 0.4 (1:4)

Numerical predictors are described using mean ± standard deviation (range). *N.A.* stands for non-applicable

1-year clinical stages were predicted for patients who monitored ALSFRS subscores, which excluded Trophos data. The choice of predictors was based on feature completeness after database cross referencing. Predictors for data reduction include gender, region of onset (spinal/bulbar), age, symptom duration, baseline ALSFRS score, baseline weight, and estimated functional decline rate [36]. The ALSFRS-R was not recorded for all of PRO-ACTs clinical trials, so functional loss was assessed with the ALSFRS. Functional loss subscores were excluded from data reduction to include Trophos data. Despite the potential role of respiratory measures as prognostic indicators, vital capacity could not be included in our model, as it was missing from real-life data and only inconsistently recorded in PRO-ACT. Analysis of functional loss subscores and of clinical stages was performed on patients when such information was available.

The functional decline rate was estimated using the following formula:

$$\text{decline rate} = \frac{\text{ALSFRS}_{\text{maximum}} - \text{ALSFRS}_{\text{baseline}}}{\text{symptom duration}}, \quad (1)$$

with $\text{ALSFRS}_{\text{baseline}}$, the ALSFRS score recorded at baseline, $\text{ALSFRS}_{\text{maximum}}$, the maximum score for the ALSFRS (which is 40) and *symptom duration*, time in months between symptom onset and baseline.

Table 2 provides an overview of patient outcome feature distribution. 1-year average ALSFRS was above 23 for all datasets. 1-year stage distribution was skewed for all three clinical staging systems: skewed towards stage 3 for King's, stages 0 and 1 for MiToS, and stages 2 and 3 for FT9.

Due to data collection differences between the cohorts, we did not perform missing data imputation and opted for complete case analysis. Further information on missing features in data is available in the [Supplementary material](#) section.

Data processing

Pre-processing was limited to predictor normalisation to the 0–1 range. Data transformation was carried out through non-linear dimension reduction, also called manifold learning. Dimension reduction was performed using UMAP [31]. UMAP works in two steps. First, a compressed embedding

Table 2 Outcome distribution per dataset

Features	PRO-ACT	Trophos	Exonhit	RW	Overall
<i>n</i> (functional loss)	2841	357	227	331	3756
1-year functional loss (score)	23 ± 8 (0:40)	25 ± 7 (6:38)	22 ± 8 (4:39)	24 ± 8 (4:40)	23 ± 8 (0:40)
1-Year upper limbs ALSFRS subscore (score)	4 ± 3 (0:8)	N.A.	3.8 ± 3 (0:8)	4.1 ± 3 (0:8)	4 ± 3 (0:8)
1-Year lower limbs ALSFRS subscore (score)	3.2 ± 2 (0:8)	N.A.	3.3 ± 2 (0:8)	3.2 ± 2 (0:8)	3.2 ± 2 (0:8)
1-Year trunk ALSFRS subscore (score)	3.9 ± 2 (0:8)	N.A.	3.7 ± 2 (0:8)	3.9 ± 2 (0:8)	3.9 ± 2 (0:8)
1-Year bulbar ALSFRS subscore (score)	8.8 ± 3 (0:12)	N.A.	8.3 ± 3 (0:12)	9 ± 3 (0:12)	8.8 ± 3 (0:12)
1-year respiratory ALSFRS subscore (score)	3.3 ± 1 (0:4)	N.A.	3.3 ± 1 (0:4)	3.5 ± 1 (1 : 4)	3.3 ± 1 (0:4)
<i>n</i> (1-year staging)	2682	N.A.	227	331	3240
1-Year King's stage distribution	360 749 1232 324 17	N.A.	24 58 117 28 0	57 95 151 26 2	441 902 1500 378 19
1-Year MiToS stage distribution	1133 1009 363 121 56	N.A.	98 76 43 9 1	147 139 41 2 2	1378 1224 447 132 59
1-Year FT9 stage distribution	70 439 806 759 608	N.A.	3 30 51 83 60	6 74 121 108 22	79 543 978 950 690

Numerical predictors are described using mean ± standard deviation (range). King's stage are 1, 2, 3, 4 and 4.5 while MiToS and FT9 stages go from 0 to 4. *N.A.* stands for non-applicable

of the input space (aka initial patient data) is generated through topological analysis of the data structure. Subsequently, a low-dimensional (in our case 2D) data embedding is created through a cross-entropy optimisation process. UMAP preserves data neighbourhood, distances and density. Development data were used to learn a 2D representation of patients. Validation data were projected using the learnt mapping. Information on the subject can be found in the [Supplementary material](#) section. Sample sizes of development and validation data for 1-year functional loss were, respectively, 3425 and 331. 1-year clinical stages sample sizes were 2909 and 331 for, respectively, development and validation data, after removing Trophos data and PRO-ACT samples with missing ALSFRS subscores.

1-year functional loss assessment zones were manually identified by dividing the UMAP projection space into three zones, based on observed patient distribution patterns, with regard to ALSFRS scores at year one. Zone division was motivated by observed spatial correlations with regard to our primary outcome. Within each zone, the distribution of 1-year ALSFRS score ranges was evaluated to identify the score range which included the majority of patients¹. The outcome probability for each zone was defined as belonging to a 1-year ALSFRS score range. Confidence bounds were derived at a 95% confidence level using the area sample size and the following formula [37]:

¹ For example, provided the distribution of 1-year ALSFRS score ranges was 5%,15%,50%, and 30% for, respectively, 1-year ALSFRS scores below 10, between 10 and 20, between 20 and 30 and above 30, 80% of the population had a 1-year ALSFRS score above 20 and the associated 1-year ALSFRS score range for that zone would be [20, 40].

$$\text{width} = 2z_{\alpha} \sqrt{\frac{P(1-P)}{N}}, \quad (2)$$

with $\alpha = 1 - \text{confidence}_{\text{level}}$, z_{α} , the value for two normal distribution, P , the outcome probability and N , the sample size.

Zone sample size directly influenced the functional loss assessment. The less populated a zone, the wider was the probability confidence interval and the less reliable was the analysis of zone membership. Validation data were projected onto the UMAP projection space to check if distribution patterns observed for development data still held. RW patients were then assigned to their corresponding functional loss assessment zone. Validation data zone assignment was assessed with regard to actual 1-year functional loss.

Results

Analysis of patient characteristics—input feature distribution

Development data were projected using UMAP in a 2D space, as shown in Fig. 1a. Initial plot of data did not show relevant patient stratification as all patients were clustered together. Plot analysis helped to identify strong correlations between projection coordinates and predictors. This was the case for age and baseline ALSFRS scores (Fig. 1c and f, respectively) and to a lesser extent for weight and estimated ALSFRS decline rate (Fig. 1e and g, respectively). Onset, gender, and symptom duration did not show a high degree of correlation as demonstrated in Fig. 1a, b and d as feature distribution appeared to be random with regard to the UMAP

projection. Further analysis of ALSFRS subscores showed that the correlation with baseline ALSFRS scores was partly explained by upper limb, lower limb and trunk subscore correlations with the projection, in respectively Fig. 1h, i, and j. Bulbar and respiratory subscores were predominantly close to their maximum value (respectively, 12 and 4) and did not seem to correlate with the projection as shown in Fig. 1k and l, respectively. Trophos data were not included in subscore analysis as these features were not available. Projection data seemed to be independent of cohort membership as patients from each source were evenly distributed in the projection space.

Analysis of patient outcomes—output feature distribution

Analysis of UMAP projection with regard to outcome variables showed spatial patterns as presented in Fig. 2. The 1-year ALSFRS score correlated well with the UMAP projection. Patients with a low 1-year ALSFRS score tended to be in the lower right part of the projection, while patients with a higher 1-year ALSFRS score seemed to cluster in the upper part of the projection. Unsurprisingly, the 1-year ALSFRS score, in Fig. 2a, correlated strongly with the baseline ALSFRS score (in Fig. 1f). A similar patient distribution was observed for the three clinical staging systems. Patients with lower clinical stages, hence with weaker functional impairment, tended to locate in the upper part of the projection, while advanced clinical stages seemed to be in the lower part of the projection as shown for King's, MiToS and FT9 staging systems in, respectively, Fig. 2b–d. 1-year ALSFRS subscore analysis showed that upper limb, lower limb and trunk subscores were correlated with total 1-year ALSFRS as shown, respectively, in Fig. 2e–g, similarly to baseline ALSFRS subscores shown in Fig. 1. 1-year bulbar and respiratory subscore analysis provided little information as 1-year subscores were overall close to their maximum values (respectively, 12 and 4) as shown in, respectively, Fig. 2h and i.

Analysis of projection space segmentation—zone division

As stated earlier, patients with marginal 1-year functional loss were mainly located in the upper pane of the projection space as seen in Fig. 2a. Further analysis of patient distribution was carried out by analysing patient subgroups based on ALSFRS score range and projecting these subgroups separately. The subgroups were: patients with a score lower than 10 in Fig. 3a, patients with a score between 10 and 20 in Fig. 3b, patients with a score between 20 and 30 in Fig. 3c and patients with a score higher than 30 in Fig. 3d. Plot analysis showed a progression pattern of uneven distribution

for patients. Healthier patients with marginal functional loss described by an ALSFRS score higher than 30 were located in the upper projection pane, while weaker patients with an ALSFRS score lower than 10 were in the lower pane of the projection. This motivated the segmentation of the projection space in three zones to isolate the three patient populations with marginal, intermediate or significant 1-year functional losses as shown in Fig. 3e. The patients in the upper projection pane were assigned to the marginal functional loss zone as 40% and 43% of patients in that zone had, respectively, an ALSFRS score higher than 30 and between 20 and 30. Overall for that zone, the functional loss assessment was that 83% ($\pm 3\%$) of the marginal loss zone patients had a functional loss above 20. The patients in the centre of the projection pane were assigned to the intermediate functional loss zone, as 47% and 42% of patients in that zone had, respectively, an ALSFRS between 20 and 30 and between 10 and 20. Overall for that zone, the functional loss assessment was that 89% ($\pm 4\%$) of the intermediate loss zone patients had a functional loss between 10 and 30. Patients in the lower projection pane were assigned to the significant functional loss zone, as 58% and 30% of patients in that zone had, respectively, an ALSFRS between 10 and 20 and below 10. Overall for that zone, the functional loss assessment was that 88% ($\pm 7\%$) of the significant loss zone patients had a functional loss between 10 and 20 and below 10. Predictors of patient population within each zone are presented in Table 3. Figure 3f–i show the distribution within each of the three zones for each ALSFRS score range subgroups. Table 4 presents zone membership distribution with regard to ALSFRS score range.

Patients in the marginal loss zone appeared to have a shorter symptom duration period, a higher initial ALSFRS score, and a slower ALSFRS decline rate with specifically higher upper limb, lower limb and trunk baseline ALSFRS subscores. Although differences were subtle, baseline bulbar and respiratory ALSFRS subscores were higher in the marginal functional loss group than in the significant functional loss group. Differences in onset distribution and gender distribution were not deemed relevant as group sizes were uneven: significant functional loss patients only made up 9% of the total patient population, marginal functional loss patients constituted 62% of the total patient population, and the remaining 29% going into the intermediate functional loss group. Analysis of ALSFRS score ranges with regard to zone membership, as presented in Table 5, showed that the marginal functional loss zone concentrated the majority of ALS patients with a 1-year ALSFRS score above 30, as 95% of these patients belonged to that zone. Similarly, but to a lesser extent, ALS patients with a low 1-year ALSFRS score tended to belong more to the significant functional loss zone than the two other zones: 46% with regard to 31% and 23% for, respectively, the intermediate and marginal functional

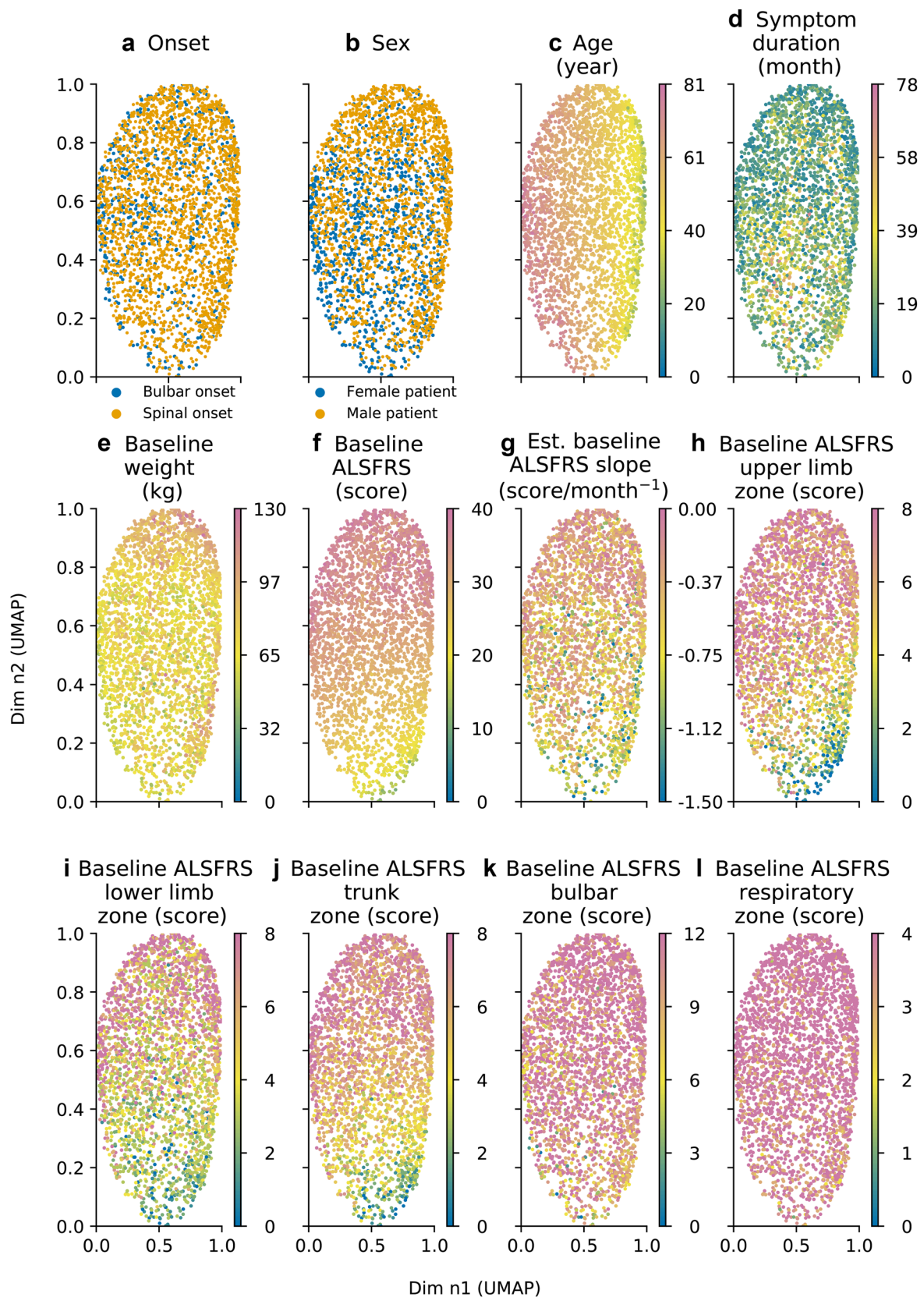


Fig. 1 Projection of the ALS population in a 2D space with UMAP. The projection is coloured with regard to predictor values: onset (a), sex (b), age (c), symptom duration in month (d), baseline weight in kg (e), baseline ALSFRS score (f), estimated ALSFRS loss rate (g), and baseline ALSFRS subscores, with respectively upper limb (h), lower limb (i), trunk (j), bulbar (k) and respiratory (l), distributions with regard to UMAP projection. Each point represents an individual patient. Age ranges between 18 and 81 years old (c), symptom duration ranges between 0.5 and 78 months (d), baseline weight ranges between 30 and 130 kg and (e), ALSFRS score ranges between 0 and 40 (f), estimated baseline ALSFRS slope ranges between 0.00 and -1.50 ALSFRS points per month (g), upper limb, lower limb and trunk ALSFRS subscores range between 0 and 8 (h–j), bulbar ALSFRS subscores ranges between 0 and 12 and respiratory ALSFRS subscore ranges between 0 and 4. Axes are dimensionless and come from UMAP dimension reduction

loss zones. Additional information on zone membership and staging systems are available in the [Supplementary material](#) section.

Analysis of the model with additional data—external validation

The prognosis model was assessed using external RW patient data provided by the Paris tertiary referral centre for ALS. Patient distribution within the projection space was examined with regard to specific outcome variables. The different trends for outcome variables identified in Fig. 2 remained valid with patient distribution being uneven for patients with marginal and significant functional loss within 1 year. Patients with a weaker loss tended to concentrate in the upper pane of the projection, as shown in Fig. 4a. Validation data showed that MiToS stage 0 patients tended to concentrate in the upper part of the projection, as presented in Fig. 4c. Little can be said about King's and FT9 distribution patterns, except that validation data seemed to reproduce loosely distribution patterns of development data as shown in, respectively, Fig. 4b and d. Additional plots on distribution differences and use cases are provided in the [Supplementary material](#) section.

Zone division—external data evaluation

Functional loss distribution of validation data with regard to zone membership is presented in Table 6. Trends identified for development data remained true, even slightly skewed: validation data patients with an ALSFRS score lower than 10 belonged at 54% to the significant functional loss zone as opposed to 46% for development data. The 331 patients were assigned to the three zones in the following way: 36, 93 and 202 assigned to, respectively, the significant, intermediate and marginal functional loss zones. Estimated functional loss within the marginal, intermediate, and significant loss zones were, respectively, 27 ± 6 , 20 ± 6 , and 13 ± 5 .

Adding 331 patients reduced the overall confidence bound for functional loss assessment relatively by 9% (from 3.18 to 2.88%), 9% (from 3.96 to 3.59%) and 10% (from 6.99 to 6.29%) for, respectively, marginal, intermediate and significant functional loss zones. Additional information on model performance comparison with two machine learning models can be found in the [Supplementary material](#) section.

Discussion

Our study demonstrates the utility of UMAP for functional loss analysis in ALS. We have successfully applied this non-linear dimension reduction method to ALS clinical trial data to predict 1-year functional decline. Our results indicate that limited patient information, collected early in the course of the disease, may be sufficient to obtain a relevant low-dimensional patient projection with regard to functional loss, an essential outcome variable for personalised patient care. These input features correlated with the different outcomes of interest, thus explaining the observed distribution patterns. These correlations were validated in an independent patient cohort. Distribution patterns in the projection were used to identify zones with distinct functional losses. We propose a simple 1-year functional loss estimation model. The benefit of our approach with regard to standard machine learning methods is threefold. First, our model is simple; it uses only simple probabilities and readily available clinical features. Second, we limit prognosis error by providing a rough prognosis estimate. Third, our model is easily updated and improves with additional data. No learning was required for our model to work as UMAP is a dimension reduction method.

Data constraints, when working with the available datasets, were challenging. PRO-ACT units are not systematically available, for instance vital capacity units are inconsistent reported either in litres or in per cent, and slow or forced vital capacities are used inconsistently. The datasets did not provide information on the genetic, cognitive or behavioural status, which both influence the prognosis [38–41]. Furthermore, 26% of PRO-ACT patients have missing weight units. As an example, a weight value of 99 could be interpreted both in kilograms or pounds. Inclusion criteria for all trials pooled within PRO-ACT are not fully documented as 6 out of the 23 pooled clinical trial names were not disclosed. As seen during baseline and 1-year ALSFRS subscores analysis, a majority of patients had little or no bulbar and respiratory involvement at baseline and after 1 year, which meant these features might not have contributed to the UMAP projection as they were uninformative.

Another data constraint was that several known prognosis biomarkers were not available within at least one of the four datasets, such as ALSFRS subscores, riluzole intake, vital

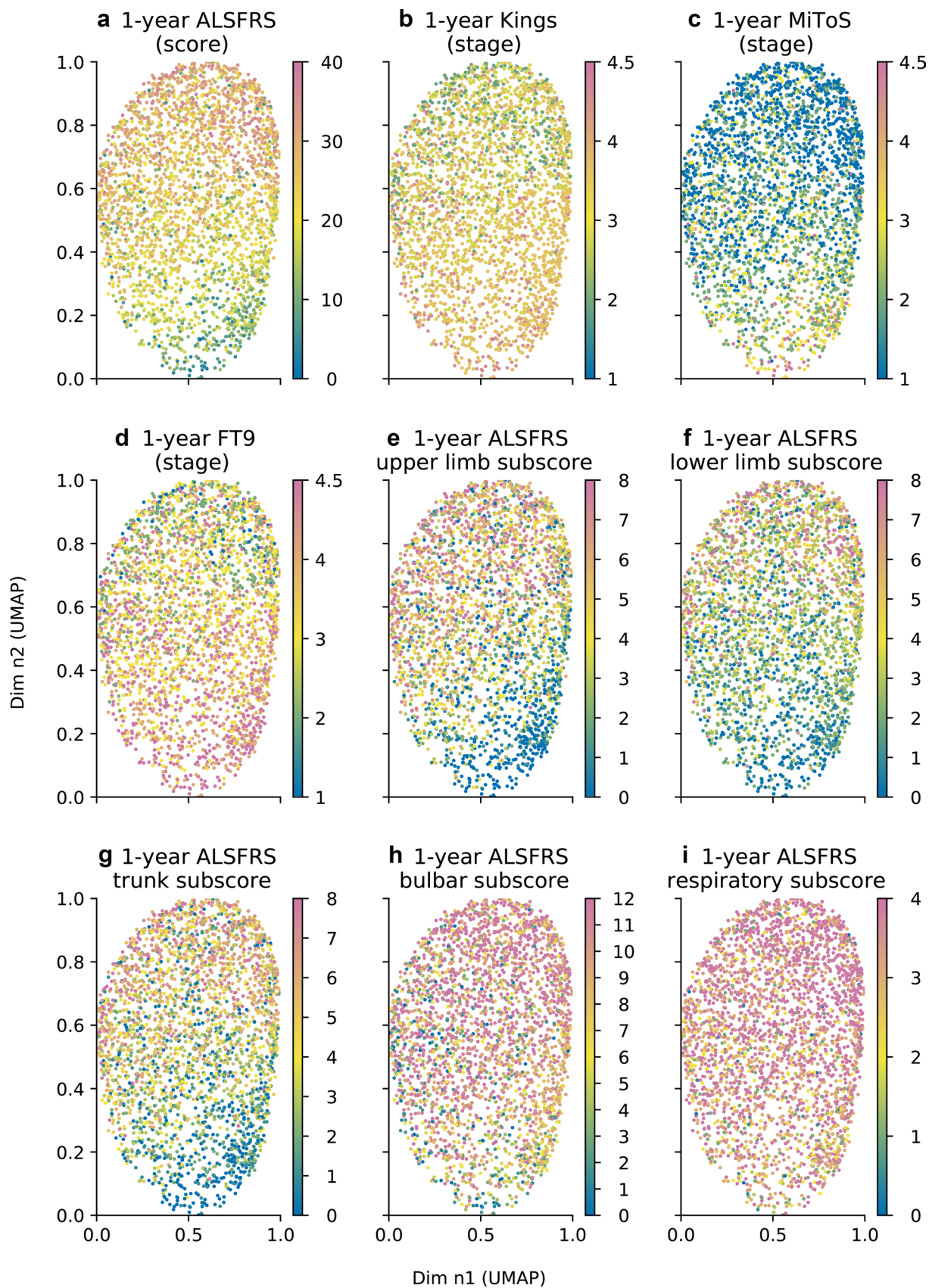


Fig. 2 Outcomes: 1-year ALSFRS (a); 1-year stages for respectively King's (b), MiToS (c) and FT9 (d) and 1-year ALSFRS subscores, for respectively upper limb (e), lower limb (f), trunk (g), bulbar (h) and respiratory (i) distribution with regard to UMAP projection. Each point represents an individual patient. For 1-year ALSFRS (a), functional loss ranges between 0 and 40. King's stages are 1,2,3,4,4.5 (b), while MiToS (c) and FT9 (d) stages are 0,1,2,3,4. Upper limb, lower limb and trunk ALSFRS subscores range between 0 and 8 (e–g), bulbar ALSFRS subscores (h) range between 0 and 12 and respiratory ALSFRS subscore (i) ranges between 0 and 4. Each data point colour is mapped to a specific ALSFRS score (a), clinical stage (b–d) or ALSFRS subscore (e–i). Axes are dimensionless and come from UMAP dimension reduction

capacity [42], time to generalisation [43] or weight loss, as it is thought to be more relevant than weight itself [44]. This limited the models ability to discriminate patients within the projection space. Additional clinical features, such as bulbar onset, or executive impairment [7], may be potential predictors to feed our model. Including biofluid markers (such as neurofilaments levels), genetic (C9orf72 status) [45], and imaging features [13, 46, 47] could have improved the accuracy of prognostic modelling [48]. Clinical staging analysis was limited, as one of the four datasets used needed to be excluded due to missing ALSFRS subscores. Given the significant proportion of samples with missing data, especially in the PRO-ACT datasets, meticulous data recording should be encouraged in future initiatives so that accrued datasets are not plagued by incomplete information.

Analysis of clinical staging systems with the available data showed that these systems did not capture patient evolution more effectively than ALSFRS. Patient selection in therapeutic studies is biased towards patients in the early stages of the disease. Most of these patients are assigned to a subset of stages without covering the full staging system spectrum. As a result, these systems are unable to describe patient differences with the relevant level of granularity. In the context of clinical trials, it suggests that ALSFRS may be the most sensitive index for progression monitoring as well as for patient stratification. In RW settings, where patient heterogeneity is significantly stronger and follow-up is longer, these new systems of staging remain pertinent to capture the clinical progression of the disease.

As a non-linear unsupervised learning model, UMAP can capture and characterise complex relationships between predictors. However, it is a black-box approach. Model interpretability cannot be obtained: the explicit relationship between UMAP input and output variables remains unavailable. Analysis of input feature distribution in the UMAP projection gives a broad overview of the importance of each variable with regard to the projection. Furthermore, similarly to other machine learning models, UMAP identifies underlying data correlations but cannot reveal causal relationships. Nevertheless, our model provides confidence intervals which most machine learning techniques such as

random forest, boosting or neural network methods do not ordinarily provide. This additional information can help clinicians to evaluate prognosis in finer detail.

Sample size was crucial, as it directly influenced the level of detail for the projection space division. Although our dataset size was considerable (3756), a larger data sample will allow defining more zones with distinct functional loss estimates. Dividing the projection space into three was deemed most appropriate given patient distribution and sample size. Given the available data, we had to deal with the trade-off between prognosis personalisation and narrow confidence bounds for functional loss. Testing on external RW data was necessary to assess model ability to scale up and model validity as it was designed using trial patients. As such, we have to acknowledge that the model was built on a specific cohort of ALS patients who met entry criteria to clinical trials, a cohort which may not be representative of all patients with ALS. Minor differences were observed when assessing zone membership. Additional RW data could correct this bias and limit the resulting over-optimistic prognosis it entails. Overall, we successfully implemented a simple 1-year functional loss estimation model using a novel non-linear unsupervised learning method. As in all rare diseases, the collection of large datasets remains challenging and is our strongest limitation. International collaborative studies are mandatory to collect large datasets to circumvent that issue and design precision models [49]. Additionally, this approach can be extended to other medical and non-medical contexts. Provided only limited data with strong data constraints are available, combining UMAP with a simple probabilistic approach can help uncover relevant correlations in data.

Supplementary material

Datasets

The first dataset, referred to as Trophos, was a 2014 double-blind, randomised, placebo-controlled phase II–III clinical trial for olesoxime, a drug produced by Trophos [32]. The trial lasted 18 months; recruitment started on May 4th 2009 and ended on March 31st 2010. The final follow-up was on September 15th 2011. Data were collected on 512 patients with clinical and biological data. Patients were recruited from 15 tertiary referral centres in five European countries. Inclusion criteria included probable or definite ALS on the El Escorial criteria, a slow vital capacity above or equal to 70% of the height-adjusted normative value, age between 18 and 80 years, symptom duration between 6 and 36 months and riluzole treatment (50 mg of riluzole, twice a day for at least a month). The trial was conducted according to the European guidelines for Good Clinical Practice.

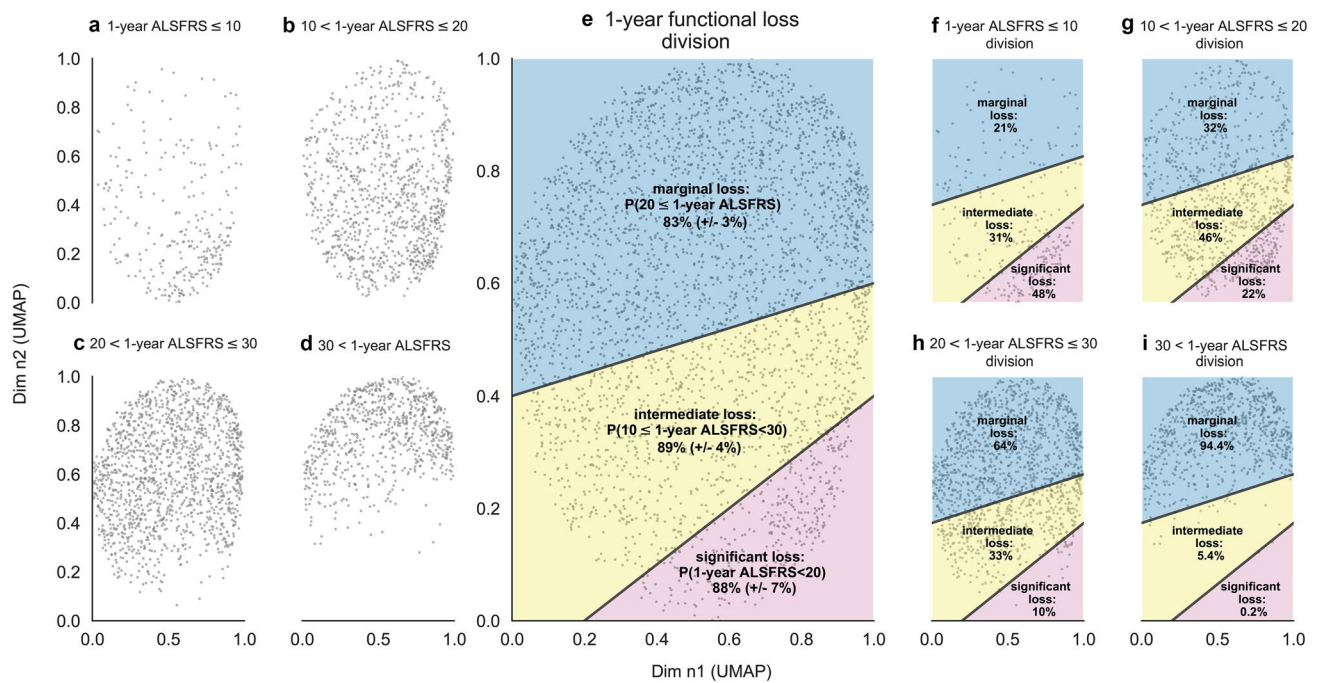


Fig. 3 One-year functional loss projection space segmentation: patients are divided into four groups based on their functional loss, which were patients with a 1-year functional loss lower than 10 (**a**), between 10 and 20 (**b**), between 20 and 30 (**c**) and above 30 (**d**). Each point represents an individual patient. The overall space is divided into three zones (**e**) based on functional loss distribution: the func-

tional loss estimated for each zone is calculated using patients from that zone. Distribution within each zone is detailed for patients with a 1-year functional loss lower than 10 (**f**), between 10 and 20 (**g**), between 20 and 30 (**h**) and above 30 (**i**). Axes are dimensionless and come from UMAP dimension reduction

Table 3 Predictor distribution per functional loss area for development data

Features	Marginal functional loss zone	Intermediate functional loss zone	Significant functional loss zone	Overall
n	2121	979	325	3425
Gender (male/female)	1440/681	502/477	186/139	2128/1191
Onset (spinal/bulbar)	1617/504	778/201	252/73	2647/672
Age (years)	55.6 ± 10.8 (18:81)	55.6 ± 12.5 (24:79)	47.8 ± 10.3 (26:78)	54.9 ± 11.5 (18:81)
Symptom duration (months)	18.4 ± 10.7 (2.9:72)	25.2 ± 14.3 (0.5:102.5)	24.5 ± 12 (5.1:61.8)	20.9 ± 12.4 (0.5:102.5)
Baseline weight (kg)	78.3 ± 15.3 (41.4:148.6)	70.3 ± 13.6 (30:137.8)	76.6 ± 17.6 (36.5:137)	75.9 ± 15.5 (30:148.6)
Baseline ALSFRS (score)	34.5 ± 2.5 (27:40)	28.7 ± 2.4 (19:35)	20.9 ± 3.1 (10:28.5)	31.4 ± 5.1 (10:40)
Baseline ALSFRS decline rate (score/month)	-0.38 ± 0.25 (-2.27:0)	-0.66 ± 0.47 (-5.65: -0.1)	-0.98 ± 0.53 (-4.15: -0.27)	-0.52 ± 0.41 (-5.65:0)
Baseline upper limbs ALSFRS subscore (score)	6.9 ± 1.3 (0:8)	5.1 ± 1.8 (0:8)	2.6 ± 1.8 (0:8)	5.9 ± 2 (0:8)
Baseline lower limbs ALSFRS subscore (score)	6 ± 2 (0:8)	4 ± 1.9 (0:8)	2.7 ± 1.4 (0:8)	5.1 ± 2.2 (0:8)
Baseline trunk ALSFRS subscore (score)	6.8 ± 1.1 (3:8)	5 ± 1.4 (0:8)	2.9 ± 1.4 (0:6)	5.9 ± 1.8 (0:8)
Baseline bulbar ALSFRS subscore (score)	10.9 ± 1.7 (3:12)	10.1 ± 2.3 (0:12)	9 ± 2.3 (0:12)	10.5 ± 2 (0:12)
Baseline respiratory ALSFRS subscore (score)	3.9 ± 0.4 (1.5:4)	3.7 ± 0.5 (1.5:4)	3.5 ± 0.6 (1:4)	3.8 ± 0.4 (1:4)
1-year ALSFRS (score)	26.6 ± 7.4 (0:40)	19.5 ± 6.4 (0:36)	12.7 ± 5.9 (0:37)	23.2 ± 8.4 (0:40)

Numerical predictors are described using mean ± standard deviation (range)

Table 4 Zone distribution with regard to functional loss for development data

Group	n	ALSFRS < 10	10 ≤ ALSFRS < 20	20 ≤ ALSFRS < 30	30 ≤ ALSFRS
Marg. FLZ	2121	2%	15%	43%	40%
Inte. FLZ	979	7%	42%	47%	4%
Sign. FLZ	325	30%	58%	11%	1%
Overall	3425	6%	27%	41%	26%

FLZ functional loss zone

Table 5 Patient functional loss distribution with regard to zones for development data

Group	n	Sign. FLZ	Inte. FLZ	Marg. FLZ
ALSFRS < 10	213	46%	31%	23%
10 ≤ ALSFRS < 20	906	21%	45%	34%
20 ≤ ALSFRS < 30	1406	2.5%	32.5%	65%
30 ≤ ALSFRS	900	0.2%	4.8%	95%
Overall	3425	9%	29%	62%

FLZ functional loss zone

The second dataset, named Exonhit, was a 2006 double-blind randomised, place-controlled phase II–III clinical trial for pentoxifylline, a drug developed by Exonhit Pharma [33]. The trial lasted 18 months, with recruitment starting on October 3rd 2002 and ending on February 12th 2003. Follow-up ended on August 6th 2004. Data were collected on the 400 patients and were limited to clinical features and muscle strength testing. Patients were recruited from 12 tertiary referral care centres from four countries. Patients were selected based on a probable or definite ALS using the El Escorial criteria, age between 18 and 80 years, a symptom duration between 6 and 47 months and riluzole treatment (established on 50 mg riluzole, twice a day). The trial was conducted according to European guidelines for Good Clinical Practice.

The third database is PRO-ACT, which is funded by the ALS Therapy Alliance and released in 2012 as part of the DREAM Phil Bowen ALS prediction Prize4Life competition. It consists of pooled data from 16 completed phase II–III clinical trials and one observational study [34]. Included clinical trials were conducted between 1990 and 2010 and lasted on average 12 months. Core patient data included clinical and biological laboratory results data. In December 2015, data from five clinical trials were added to PRO-ACT without additional information, totalling 22 different clinical trial sources and 10,723 patients. Patient age ranged from 18 to 88, 82% were on a riluzole treatment (with variable dosage), but no El Escorial categorisation or individual vital capacity values were made available.

The fourth dataset is population based and contains RW patient data. Data were obtained from the database of the Paris tertiary referral centre for ALS (Pitié Salpêtrière

Hospital Assistance Publique des Hôpitaux de Paris, France) between September 1999 and April 2008. The initial patient sample size was 1377 and included data on baseline ALSFRS with subscores, age, time interval since symptom onset, El Escorial criteria, and muscle strength values.

As patient treatment is not relevant for our current work, we refer the readers to the specific clinical trial reports for additional information.

Missing data

Missing feature analysis focused solely on baseline predictors and outcomes (1-year ALSFRS, 1-year King's, 1-year MiToS and 1-year FT9 stages). Table 7 presents missing data ratios per feature for all datasets. Features which were not available in all datasets, such as testing and biological laboratory results, were discarded. ALSFRS subscores were not recorded for Trophos patients. Outcome features may be missing due to loss to follow-up or death. Features at time $t+3$ were less available than at baseline. Data collection was not disclosed for PRO-ACT data which aggregates multiple clinical trials. This prevented the identification of missing data patterns. No specific details were available on the data collection process for PRO-ACT clinical trials which, amongst others, limited the use of vital capacity measurements. Working solely with baseline features was motivated both by missing data ratios and prognostic relevance, e.g. the ability to provide an accurate prognosis at the first visit.

Missing data imputation was omitted and our model was trained solely on complete case samples. Although generally recommended in medical settings, data imputation seemed hazardous in this specific data context, specifically working with PRO-ACT. Multiple imputation methods often assume that missingness patterns are missing at random, i.e. that they depend on other observed variables in the dataset. This information is difficult to verify and these data imputation methods are often performed on the biggest feature subset available so as to improve the odds of such a hypothesis being true. Given the differences in the data collection process and the limited feature subset shared between the different datasets, data imputation could not have been carried out on the global data structure. Data imputation at a dataset level would not have been productive and would have led to significant additional noise in data given small sample

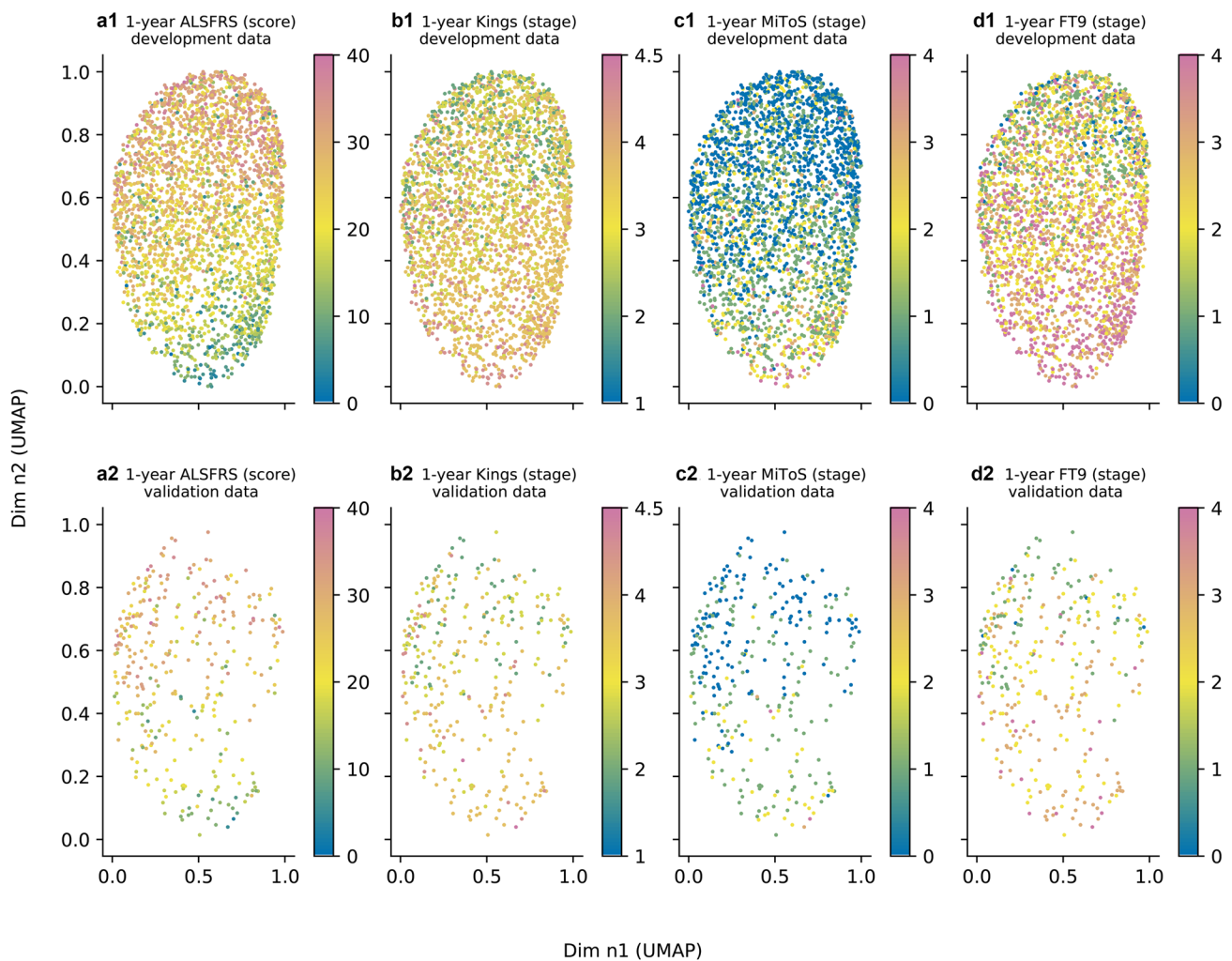


Fig. 4 Outcomes with regard to UMAP projection for development and validation data: 1-year functional loss for development (**a.1**) and validation (**a.2**) data, 1-year King's for development (**b.1**) and validation data (**b.2**), 1-year MiToS for development (**c.1**) and validation data (**c.2**) and 1-year FT9 for development (**d.1**) and validation data (**d.2**). Each point represents an individual patient. For 1-year ALS-

FRS (**a**), score range between 0 and 40. King's stages are 1,2,3,4 or 4.5 (**b**). MiToS and FT9 stages are 0,1,2,3 or 4 (**c,d**). The data point colour is mapped to a specific ALSFRS score (for **a**) or clinical stage (for **b-d**). Axes are dimensionless and come from UMAP dimension reduction. (**a.1-d.1**) represent development data plots; (**a.2-d.2**) represent validation data plots

Table 6 Patient functional loss distribution with regard to zones for validation data

Group	n	Sign. FLZ	Inte. FLZ	Marg. FLZ
ALSFRS < 10	13	54%	31%	15%
10 ≤ ALSFRS < 20	92	29.3%	42.4%	28.3%
20 ≤ ALSFRS < 30	130	1.5%	34.5%	64%
30 ≤ ALSFRS	90	0%	5%	95%
Overall	331	11%	28%	61%

FLZ functional loss zone

sizes and significant missing feature ratios for each dataset. Even advanced multiple imputation methods such as [51], which deal with missing data imputation at a study level

(for meta-analysis purposes), require knowing the collection process for each study in scope, which we cannot access for PRO-ACT as features could be missing due to loss to follow-up or due to clinical trial setup. Furthermore, as UMAP is a neighbourhood-based approach, data imputation can be seen as adding data where it is missing. This would have induced sample similarity in cases where little information was known on the subjects, creating visual artefacts of similar patients within the projection space and adding significant bias to the visual representation. Our spatial distribution approach would have had a more limited performance had we worked with imputed data that would have artificially created spatial proximity.

Table 7 Missing feature analysis per dataset

Feature	PRO-ACT	Exonhit	Trophos	Real world	Overall
Initial sample size (n)	10723	400	512	1377	13012
Gender	0%	0%	0%	0%	0%
Onset	12%	0%	0%	2%	10%
Age	28%	0%	0%	0%	23%
Symptom duration	36%	0%	0%	0%	30%
Baseline weight	39%	3%	1%	3%	33%
Baseline height	38%	0%	100%	3%	35%
Baseline ALSFRS	36%	2%	0%	0%	30%
Baseline ALSFRS upper limb subscore	39%	0%	0%	0%	36%
Baseline ALSFRS lower limb subscore	39%	0%	100%	0%	36%
Baseline ALSFRS bulbar subscore	39%	0%	100%	0%	36%
Baseline ALSFRS respiratory subscore	39%	0%	100%	0%	36%
Baseline ALSFRS trunk subscore	39%	0%	100%	0%	36%
Baseline pulse	32%	1%	100%	100%	41%
Baseline diastolic blood pressure	32%	1%	0%	100%	37%
Baseline systolic blood pressure	32%	1%	0%	100%	37%
Baseline vital capacity (L)	23%	1%	0%	100%	29%
Baseline vital capacity (%)	10%	1%	0%	100%	19%
1-year ALSFRS	66%	42%	30%	75%	65%
1-year King's stage	66%	41%	75%	100%	68%
1-year MiToS stage	66%	41%	75%	100%	68%
1-year FT9 stage	68%	42%	75%	100%	68%
Overall missing ratio	36%	8%	37%	47%	37%
Overall predictor missing ratio	29%	1%	28%	39%	30%
Overall outcome missing ratio	67%	41%	75%	83%	68%
Final sample size for 1-year survival (n)	2841	227	331	357	3756

UMAP

Supervised learning models usually require large amounts of data to avoid overfitting and lack of generalisation, which are not available in ALS research. Unsupervised learning methods have the advantage of capturing distribution patterns without data implications. Standard linear methods such as principal component analysis (PCA) [52] have been used in ALS for gene expression analysis [29]. Unfortunately, these conventional linear-based methods are not capable of describing non-linear relationships and have underperformed in this study context. Non-linear methods provide new modelling possibilities given their comprehensive ability to describe data correlations and have successfully been tried out for ALS phenotype identification on clinical trial data [30] with *t* Student Stochastic Neighbour Embedding (*t*-SNE), the current state-of-the-art manifold learning model [53]. UMAP outperforms *t*-SNE on the following aspects: it scales with regard to complexity (calculation-wise), allows dimension reduction for other purposes than visualisation (i.e. dimensions can be larger than 3), has a convex cost function (where *t*-SNE has a non-convex cost function that leads to initialisation based results) and preserves

neighbourhood, distances and density (and not only neighbourhood like *t*-SNE) [54]. All these assets make UMAP more suitable for clustering at a later stage than *t*-SNE.

UMAP is neighbourhood based and works in two steps. First, a compressed embedding of the input space is built through topological analysis of the data structure using simplexes². The compressed embedding is a simplicial simplex which can be seen as a neighbourhood graph of the input data that is built from the open cover³ of the simplexes. Second, a low-dimensional (in our case 2D) data embedding is found through a cross-entropy⁴ optimisation process. UMAP preserves data neighbourhood, distances and density. The

² In geometry, a simplex is defined as a set of points, where none is a barycentre of the remaining points. The convex hull of these points corresponds to the face of the simplex. In simpler terms, a *n*-simplex can be thought of as the generalisation of a triangle in the *n*th dimension.

³ An open cover is essentially just a family of sets whose union is the whole space [31].

⁴ In machine learning, cross entropy is frequently used as a cost function to compare two probability distributions (*p,q*): *p* is optimised to approximate *q* the fixed target distribution.

initial modelling step depends on whether the algorithm should focus on preserving the local or global input data structure. Data structure is estimated according to the size of the neighbourhood investigated. The second compression step is mainly defined using two parameters which are the output dimension size and the minimum distance permitted between two points in the output space, i.e. how compact the output projection can be.

Direct model specification is not possible as the UMAP projection function is a black-box approach based on nearest neighbours. Nonetheless, the UMAP projection function can be stored and used later to incorporate additional new data. As it is a black-box approach, projection features cannot be analysed to provide any interpretability. Output dimension analysis, as commonly performed for PCA, cannot be carried out. Analysis of input feature distribution in the UMAP projection space is an alternative as it gives a broad overview of variable importance with regard to the projection. In Fig. 1, age and baseline ALSFRS scorer appeared to be the variables which mattered most distance-wise in the output space. Both variables had a global incidence on the overall UMAP projection distribution. Other variables, which showed a weaker or limited impact on the overall UMAP projection distribution, may have had a more local impact in the output space distance-wise.

Performance was directly related to the observed sample size (and its ability to represent the overall ALS patient population), as the more the data collected, the finer was the split in the input space with a controlled confidence bound. UMAP was performed on a 2D plane, as the 1D projection led to uninterpretable results and the 3D projection led to results similar to those observed for 2D (with a potato-shaped form). Given the additional dimension, data density was lower and projection analysis and partitioning was more complex, so the 2D projection was preferred. Given the size of the dataset at our disposal, we could not explore finer splits without a significant increase of the confidence bound. As more data get collected, a finer division of the projection space can be performed, and a precision medicine approach can be implemented to provide clinicians with a more distinct patient division. Model updating is straightforward, as novel data points can be projected onto the 2D space using the learnt projection function.

UMAP projection is fitted using the clinical trial data. UMAP projection for additional data is not computed using the UMAP package, as proper mixed data type management (when dealing with both categorical and numerical features) has yet to be implemented in the official UMAP package. As a proxy, UMAP projection is learnt using a random forest model (with 10 trees). The random forest model serves as a proxy for now to estimate the 2D mapping of additional samples. Our random forest model reached a 99.3% coefficient

of determination (R^2) score. The random forest was trained using 80% of the clinical trial data and was tested using the remaining 20%.

Use case

Novel patient data, provided all baseline features are recorded, can be projected in the reduced UMAP space. The corresponding 2D coordinates determine zone membership to one of the three functional loss zones. Zone membership and the spatial positioning within the projection space provide a broad estimate of patient 1-year functional loss. The following examples are provided for more details:

- Patient A (ID 1735) is a 43-year-old man with a spinal onset, baseline weight 55 kg, baseline ALSFRS score 31, symptom duration estimated at 7 months, and hence estimated baseline ALSFRS decline rate is assessed at -1.16 ALSFRS points per month. This information is used to compute the spatial coordinates of patient A within the UMAP projection space. Patient A spatial coordinates in the UMAP projection space are (0.82, 0.52), which fall into the intermediate functional loss zone. Patient A has a 89% probability that his resulting 1-year functional loss will belong to the [10, 30] range.
- Patient B (ID 1859) is a 66-year-old woman with a spinal onset, baseline weight 50 kg, baseline ALSFRS score 37, and symptom duration estimated at 7 months, and hence estimated baseline ALSFRS decline rate is assessed at -0.42 ALSFRS points per month. This information is used to compute the spatial coordinates of patient B within the UMAP projection space. Patient B spatial coordinates in the UMAP projection space are (0.12, 0.70), which fall into the marginal functional loss zone. Patient B has an 83% probability that his resulting 1-year functional loss will belong to the [20, 40] range.
- Patient C (ID 2946) is a 54-year-old man with a spinal onset, baseline weight 73 kg, baseline ALSFRS score 22, and symptom duration estimated at 39 months, and hence estimated baseline ALSFRS decline rate is assessed at -0.45 ALSFRS points per month. This information is used to compute the spatial coordinates of patient C within the UMAP projection space. Patient C spatial coordinates in the UMAP projection space are (0.6, 0.15), which fall into the marginal functional loss zone. Patient C has an 88% probability that his resulting 1-year functional loss will belong to the [0, 20] range.

The use cases are displayed in Fig. 5. Subsequent analysis of patients' A, B and C status after 1 year are that patients A, B and C had an ALSFRS score of, respectively, 23, 35 and 17.

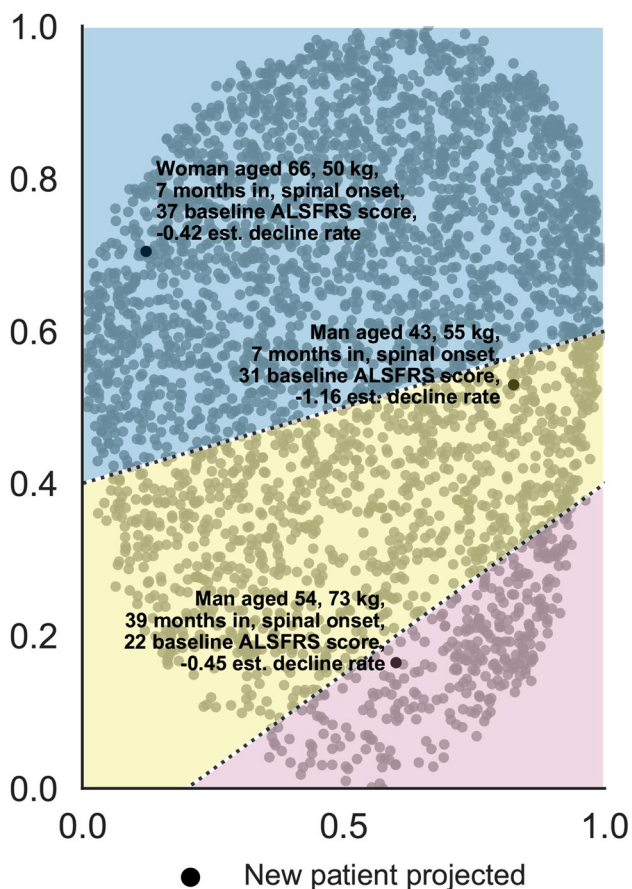


Fig. 5 One-year functional loss projection space segmentation: novel patient data is projected onto the reduced space and prognosis is estimated based on projection coordinates. Axes are dimensionless and come from UMAP dimension reduction

Distribution differences

Differences in distribution between development and validation datasets have been analysed with regard to each outcome in scope, for different outcome values. These were ALSFRS range for functional loss and stage value for clinical staging systems. Visual differences for these subsets are presented for functional loss and clinical stages in Figs. 6 and 7, respectively. Further analysis of distribution differences between development and validation datasets by ALSFRS range and zone membership is shown in Fig. 8. Little can be said on differences in distribution for King's clinical stages 4 and 4.5, MiToS stages 3 and 4 and FT9 stage 0 given the limited data available for such stages. Distribution according to the y-axis seemed similar for most subsets, while a left shift was frequently observed on the x-axis for validation data with respect to development data.

Model comparison

Our model was compared to logistic regression and random forest models. Models were trained on development data and tested on validation data. Our model returned a functional loss range probability and ALSFRS values overlapped between zones, as the significant and intermediate functional loss zones were associated with ALSFRS ranges between 0 and 20 and 10 and 30, respectively. As such, direct model comparison with regular machine learning models was not possible. We chose to calculate, for each zone, the number of patients within the ALSFRS range using the zone probability. The number of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN) could then be approximated for each zone and model performance within each zone could be assessed. The two machine learning models were trained to classify patients into four groups: 1-year ALSFRS between 0 and 10, between 10 and 20, between 20 and 30 and between 30 and 40. This output was then labelled to match with our model's output, that is, to say whether the right ALSFRS range was predicted.

The number of TP, FP, FN and TN were reported for each model. The following classification metrics were used: accuracy⁵, precision⁶, specificity⁷, recall⁸, balanced accuracy⁹ and F1-measure¹⁰. Patient distribution according to zone membership and ALSFRS range are presented in Table 8. Model performance is summarized in Table 9 for TP, FP, FN and TN and in Table 10 for classification metrics.

88% of the 36 patients associated with the significant functional loss zone were labelled as belonging to the [0,20] ALSFRS range (32). Remaining patients are labelled as belonging to the [20,40] ALSFRS range (4). 89% of the 93 patients belonging to the intermediate functional loss zone were labelled as belonging to the [10,30] ALSFRS range (83). The remaining patients are labelled as belonging to either the [0,10] or [30,40] ALSFRS ranges (10). 83% of the 202 patients assigned to the marginal functional loss zone were labelled as belonging to the [20,40] ALSFRS range (168). The remaining patients are labelled as belonging to the [0,20] ALSFRS range (34). Performance metrics were approximated based on these figures. Hence for the marginal functional zone, 196 patients were associated with the accurate ALSFRS range (TP+TN) with six patients wrongly

$$^5 \text{ Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$^6 \text{ Precision (positive predictive value)} = \frac{TP}{TP+FP}$$

$$^7 \text{ Specificity (true negative rate, selectivity)} = \frac{TN}{TN+FP}$$

$$^8 \text{ Recall (sensitivity, true positive rate)} = \frac{TP}{TP+FN}$$

$$^9 \text{ Balanced accuracy (average of specificity and recall)} = \frac{\text{Specificity} + \text{Recall}}{2}$$

$$^{10} \text{ F1-measure (harmonic mean of precision and recall)} = \frac{2 \cdot \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

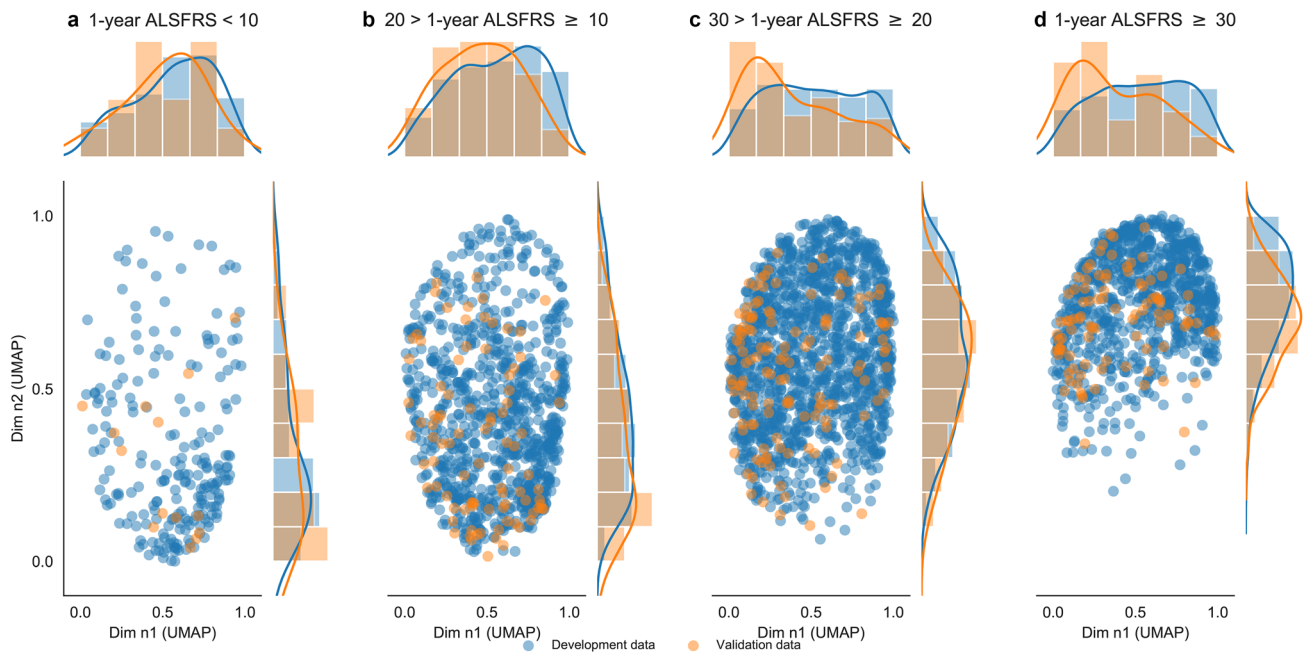


Fig. 6 Distribution patterns for development and validation data (**a**). Distribution patterns for specific subsets of data based on ALSFRS range: less than 10 (**a**), between 10 and 20 (**b**), between 20 and 30

(**c**) and above 30 (**d**). Axes are dimensionless and come from UMAP dimension reduction

labelled as not belonging to the target ALSFRS range (FN). Our proposed model outperformed the random forest and logistic regression models in all three zones and for most metrics specifically balanced accuracy and F1-measure. Due to the fact that functional loss ranges for each zone overlapped, model performance was analysed at a zone level and not on validation data as a whole. Given the low sample sizes for each zone and that performance metrics were calculated differently for our model and the other machine learning models, where individual predictions were available for all patients, performance results should be viewed with caution.

Clinical stage distribution patterns

Clinical stage distribution patterns for development data

Analysis of 1-year clinical stages distribution patterns was carried out in the same manner as functional loss. Fig. 9 presents patient distribution after zone division for each clinical staging system and for each stage similarly to what was

performed for ALSFRS in Fig. 3a–d. Figure 9a–e, respectively, represent patient distribution within the three zones for King’s stages 1–4.5, respectively. In the same manner, Fig. 9f–j, represent MiToS stages 0–4 and Fig. 9k–o represent FT9 stages 0–4.

King’s stage analysis showed that nearly all stage 1 patients (96.6%) in Fig. 10a and a majority of stage 2 patients (73%) in Fig. 10b belonged to the marginal functional loss zone, which demonstrated that King’s stages were fairly able to isolate patients with a weak functional impairment. Division for more advanced stages (2–4.5) was less informative, as patient distribution within the three zones was evenly divided between the marginal and intermediate functional loss zones with only a small portion of patients falling into the significant functional loss zone as summarised in Table 11 and shown in Fig. 10c–e. Analysis of zone membership with regard to patient stage is presented in Table 12. The significant functional loss zone was mostly made of stages 3 and 4 (90%) with a majority of stage 3 patients (67%). Intermediate functional loss zone analysis also showed that a majority of patients within that zone were

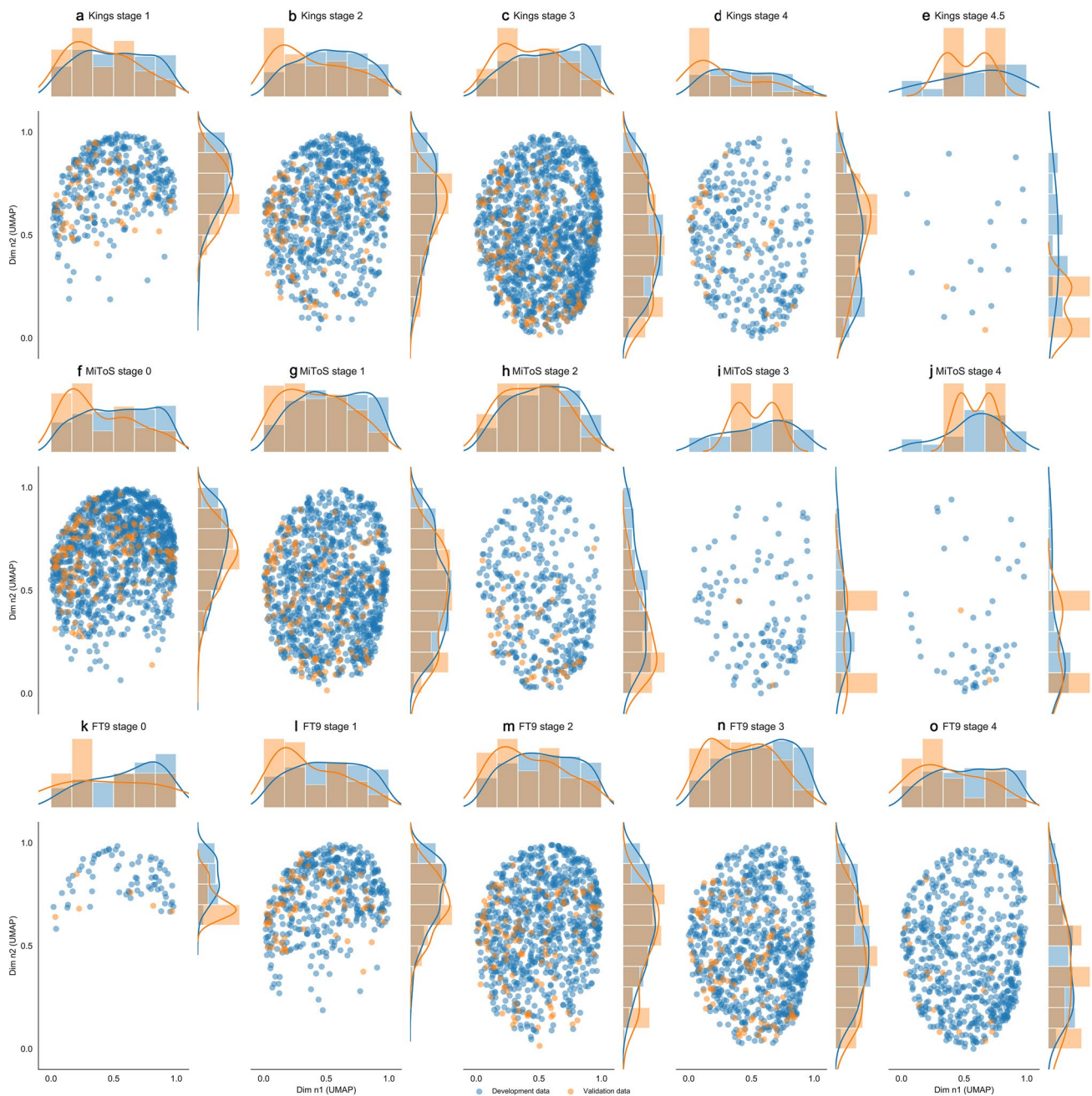


Fig. 7 Distribution patterns for development and validation data (a). Distribution patterns for specific subsets of data based on stage values, with 1-year King's for (a–e), 1-year MiToS for (f–j) and 1-year FT9 for (k–o). Axes are dimensionless and come from UMAP dimension reduction

stage 3 (61.5%), but overall patient distribution was more skewed towards lower stages.

The marginal functional loss zone was more evenly distributed as patient distribution seemed concentrated around stages 1–3 (90.5%). King's stage 3 patients outnumbered the other stages (46% of patient population was at stage 3 after one year), partially explaining why patients from that stage are found in significant proportions in all three zones. Stage 4.5 sample size was limited, hence little can be concluded

from the analysis of this stage. Overall, as King's stages increased, the proportion of patients in the marginal functional loss zone decreased, while the proportion of patients in the intermediate and significant functional loss zone increased.

MiToS stage analysis was less informative than King's, as 79.5% of the population was at stage 0 or 1 after 1 year, with only a minority of patients having undergone most of the disease course. Low MiToS stages concentrated in

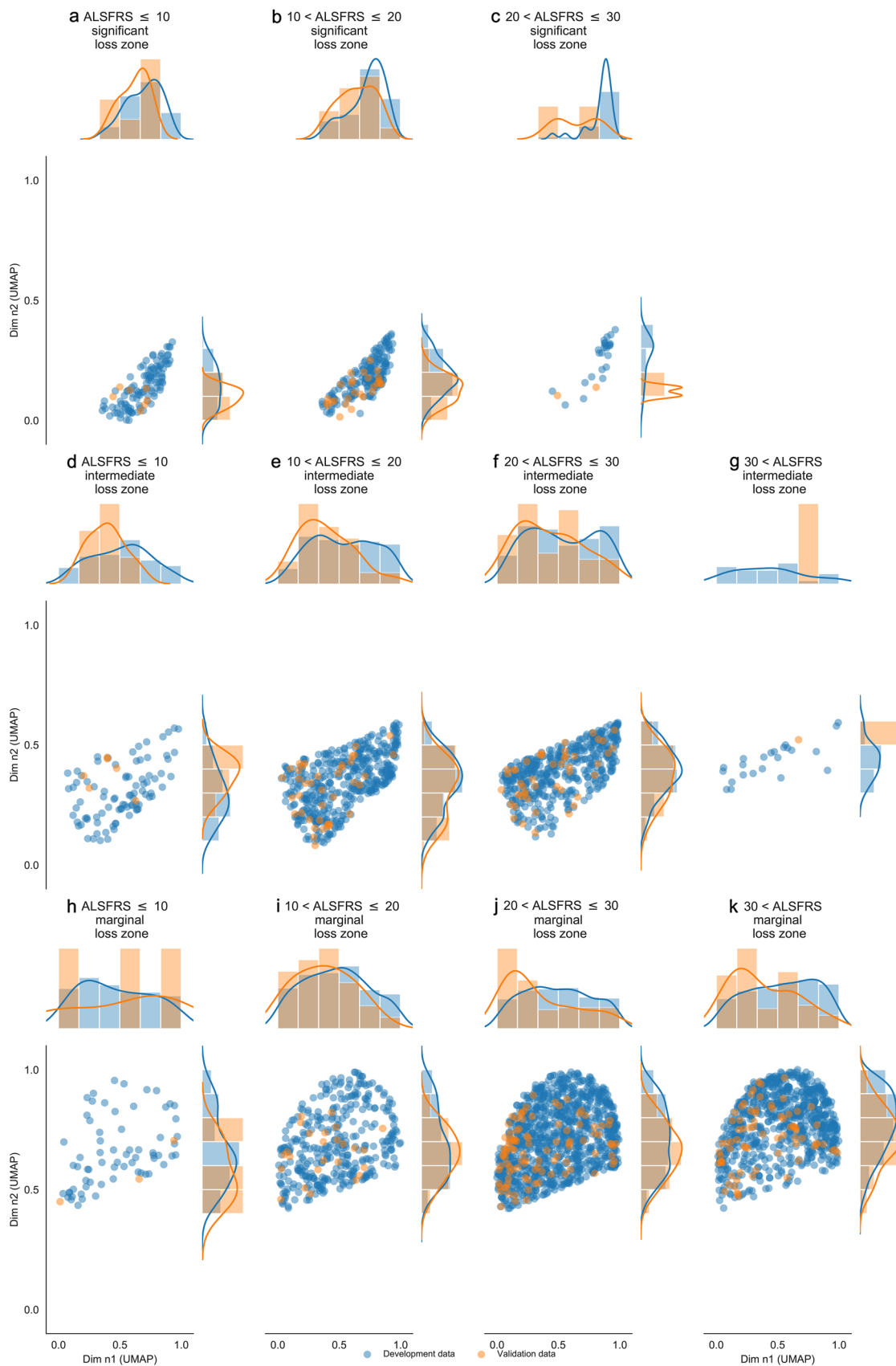


Fig. 8 Distribution patterns for development and validation data for 1-year functional loss. Distribution patterns for specific subsets of data based on functional loss and the zones identified for prognosis estimation. Patients that belong to the significant functional loss zone are separated in three subsets based on ALSFRS range in (a–c). Patients that belong to the intermediate functional loss zone are separated in four subsets based on ALSFRS range in (d–g). Patients that belong to the significant functional loss zone are separated in three subsets based on ALSFRS range in (h–k). Axes are dimensionless and come from UMAP dimension reduction

the marginal functional loss zone as for King’s, as 83% of MiToS stage 0 patients were associated with the marginal functional loss zone. These are presented in Fig. 10f and g. Table 13 presents stage distribution for each zone. As clinical stages increased, the proportion of patients within the marginal functional loss zone decreased. Initially, intermediate and significant functional loss zone membership increased together, but as stages progressed, only significant functional loss zone membership increased as shown in Fig. 10h–j. These results should be interpreted with caution as MiToS stage 2–4 patients were significantly underrepresented. Zone distribution with regard to MiToS stage is summarised in Table 14. The marginal functional loss zone was primarily made up of stage 0 patients. Unsurprisingly, as stage distribution was skewed towards stages 0 and 1, stage 1 patients were predominantly represented in the intermediate (51%) and significant (41.5%) functional loss zones. Little more can be said regarding MiToS added value with regard to patient stratification due to the overrepresentation of stage 0 and 1 patients.

FT9 stage distribution was the most balanced out of the three clinical stage systems, the results of which are presented in Table 15. Low stages 0 and 1 were concentrated in the marginal functional loss zone, as all patients from stage

0 and 94% of stage 1 patients were assigned to the marginal functional loss zone as shown in Fig. 10k and l. Results were less clear for more advanced stages as patients at stage 3 and 4 could be assigned to all three zones as shown in Fig. 10m–o. Zone distribution with regard to FT9 stage is summarised in Table 16. Overall, as FT9 stages increased, the proportion of patients within the intermediate and significant functional loss zones increased while the proportion of marginal functional loss zone assignments decreased.

Uneven patient distribution with regard to stages clearly stood out for King’s stage 4.5, MiToS stage 4 and FT9 stage 0. The ability for the marginal functional loss zone to concentrate King’s stage 1, MiToS stage 0 and FT9 stage 1 was quite apparent. Distribution per staging system and per stage showed, in this specific data context, that staging systems could only help discriminate patients with a weak functional impairment from the rest of the patient population.

Clinical stage distribution patterns for validation data

Stage analysis of zone assignments for validation data with regard to King’s, MiToS and FT9 staging systems are presented in Tables 17, 18 and 19, respectively. Trends

Table 9 Model comparison on validation data

Model	TP	FP	FN	TN
LR - Marg. FLZ	170	27	4	1
LR - Inte. FLZ	85	7	1	0
LR - Sign. FLZ	32	1	2	1
RF - Marg. FLZ	149	18	25	10
RF - Inte. FLZ	76	6	10	1
RF - Sign. FLZ	34	1	0	1
PM - Marg. FLZ	168	0	6	28
PM - Inte. FLZ	83	0	3	7
PM - Sign. FLZ	32	0	2	2

LR Logistic regression, RF random forest, PM proposed model which combines UMAP with spatial division, FLZ functional loss zone

Table 8 Zone distribution with regard to functional loss for validation data

Group	n	ALSFRS < 10	10 ≤ ALSFRS < 20	20 ≤ ALSFRS < 30	30 ≤ ALSFRS
Marg. FLZ	202	3	25	100	74
Inte. FLZ	93	6	40	46	1
Sign. FLZ	36	7	27	2	
Overall	331	16	92	148	75

FLZ functional loss zone

Table 10 Model comparison on validation data

Model	Accuracy	Precision	Specificity	Recall	Balanced accuracy	F1 measure
LR Marg. FLZ	85%	86%	4%	98%	51%	92%
LR Inte. FLZ	91%	92%	0%	99%	49%	96%
LR Sign. FLZ	92%	97%	50%	94%	72%	97%
RF Marg. FLZ	79%	89%	36%	86%	61%	87%
RF Inte. FLZ	83%	93%	14%	88%	51%	90%
RF Sign. FLZ	97%	97%	50%	100%	75%	99%
PM Marg. FLZ	97%	100%	100%	97%	98%	98%
PM Inte. FLZ	97%	100%	100%	97%	98%	98%
PM Sign. FLZ	94%	100%	100%	94%	97%	97%

LR Logistic regression, *RF* random forest, *PM* proposed model which combines UMAP with spatial division, *FLZ* functional loss zone

identified earlier with development data seemed to be confirmed, as patients at a lower stage tended to concentrate in the marginal functional zone as 95% of King's stage 1 and 74% of King's stage 2, 86% of MiToS stage 0 patients, 100% of FT9 stage 0 and 96% of FT9 stage 1 patients were assigned to that zone, which confirms what was observed for development data. Given the limited sample size in higher stages for all three stages (26 and 2 patients at, respectively, stages 4 and 4.5 for King's, 2 and 2 patients at, respectively, stages 3 and 4 for MiToS and 22 patients at stage 4 for FT9), little can be concluded from the obtained results. The ability to discriminate between marginal and non-marginal functional loss at a stage level can be observed, but patients with a significant functional impairment cannot be identified.

Computational environment

Programs were designed in Python using the following packages: numpy, pandas, matplotlib, seaborn, umap-learn and scikit-learn. Programs were run on:

- Dell laptop with Windows 10 Pro, Intel Core i7-9850H CPU, 64GB RAM,
- Dell remote server with Linux Fedora 28, 80 Intel Xeon Gold 6148 CPUs, 1.5TB RAM.

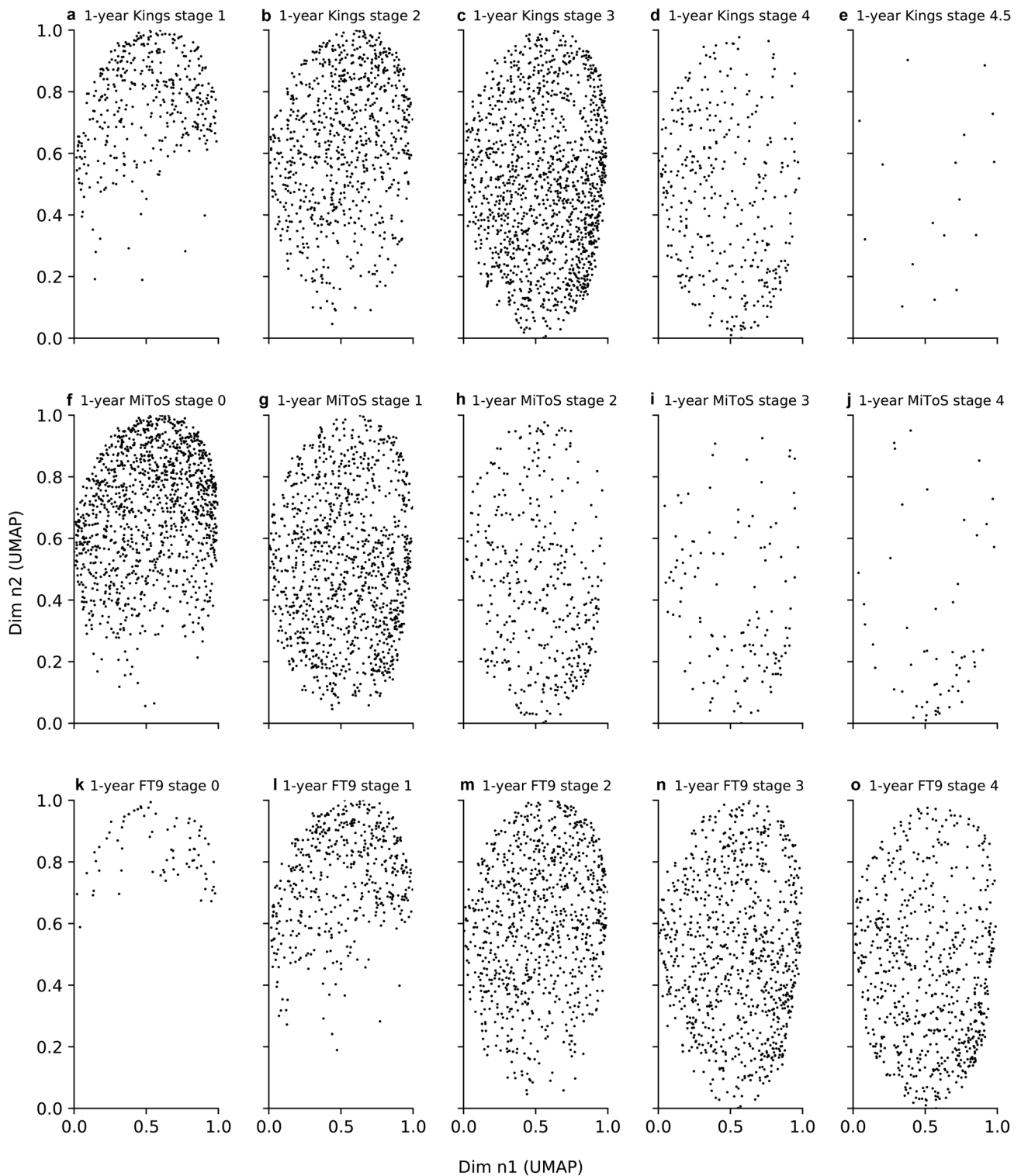


Fig. 9 One-year staging distribution: patients are divided into five groups based on their clinical stage: stages 1–4.5 for King’s (a–e), stages 0–4 for MiToS (f–j) and stages 0–4 for FT9 (k–o). Each point

represents an individual patient. Axes are dimensionless and come from UMAP dimension reduction

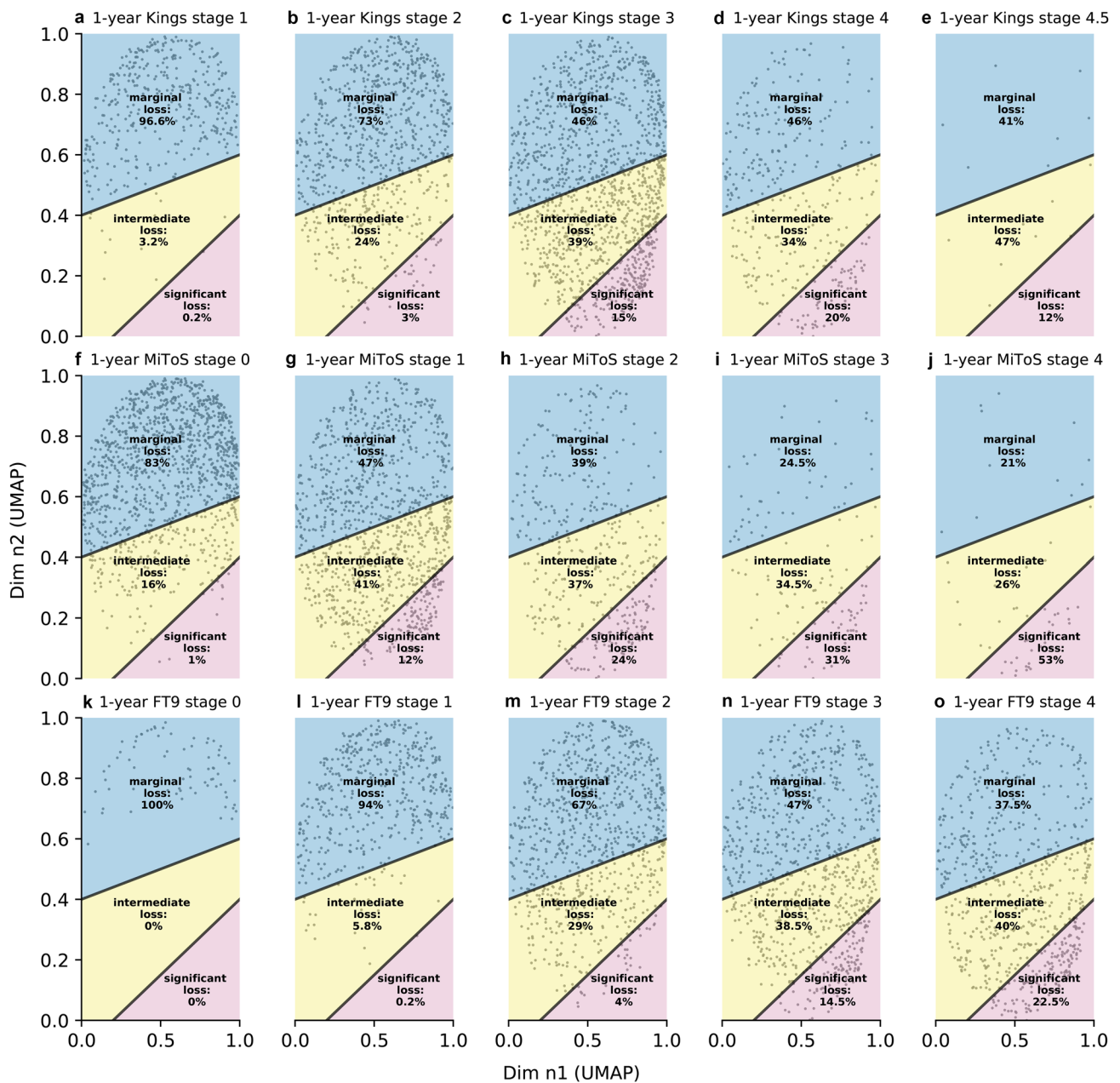


Fig. 10 One-year staging with functional loss space segmentation: patients are divided into five groups based on their clinical stage: stages 1–4.5 for King's (a–e), stages 0–4 for MiToS (f–j) and stages 0–4 for FT9 (k–o). Each point represents an individual patient. The overall space is divided into three zones based on functional loss dis-

tribution. Population distribution within the three zones is calculated from population from each clinical stage. For each staging system and each stage, the percentage of the patients at that clinical stage belonging to each of the three zones is displayed. Axes are dimensionless and come from UMAP dimension reduction

Table 11 Patient King's stage distribution with regard to zones for development data

	n	Sign. FLZ	Inte. FLZ	Marg. FLZ
KS1	384	0.2%	3.2%	96.6%
KS2	807	3%	24%	73%
KS3	1349	15%	39%	46%
KS4	352	20%	34%	46%
KS4.5	17	12%	47%	41%
Overall	2909	10.5%	30%	59.5%

KS King's Stage, FLZ functional loss zone

Table 12 Zone distribution with regard to King's stages for development data

	n	KS1	KS2	KS3	KS4	KS4.5
Sign. FLZ	306	0%	9%	67%	23%	1%
Inte. FLZ	870	1.5%	22%	61.5%	14%	1%
Marg. FLZ	1733	21.5%	34%	35%	9.5%	0%
Overall	2909	13%	28%	46%	12%	1%

KS King's Stage, FLZ functional loss zone

Table 13 Patient MiToS stage distribution with regard to zones for development data

	n	Sign. FLZ	Inte. FLZ	Marg. FLZ
MS0	1231	1%	16%	83%
MS1	1085	12%	41%	47%
MS2	406	24%	37%	39%
MS3	130	31%	34.5%	34.5%
MS4	57	53%	26%	21%
Overall	2909	10.5%	30%	59.5%

MS MiToS stage, FLZ functional loss zone

Table 14 Zone distribution with regard to MiToS stages for development data

	n	MS0	MS1	MS2	MS3	MS4
Sign. FLZ	306	3.5%	41.5%	32%	13%	10%
Inte. FLZ	870	24%	51%	18%	5%	2%
Marg. FLZ	1733	58%	29%	9%	3%	1%
Overall	2909	42.5%	37%	14%	4.5%	2%

MS MiToS stage, FLZ functional loss zone

Table 15 Patient FT9 stage distribution with regard to zones for development data

	n	Sign. FLZ	Inte. FLZ	Marg. FLZ
FS0	73	0%	0%	100%
FS1	469	0.2%	5.8%	94%
FS2	857	4%	29%	67%
FS3	842	14.5%	38.5%	47%
FS4	668	22.5%	40%	37.5%
Overall	2909	10.5%	30%	59.5%

FS FT9 Stage, FLZ functional loss zone

Table 16 Zone distribution with regard to FT9 stages for development data

	<i>n</i>	FS0	FS1	FS2	FS3	FS4
Sign. FLZ	306	0%	0%	11%	39.5%	49.5%
Inte. FLZ	870	0%	3%	29%	37%	31%
Marg. FLZ	1733	4%	26%	33%	23%	14%
Overall	2909	3%	16%	29%	29%	23%

FS FT9 Stage, FLZ functional loss zone

Table 17 Zone distribution with regard to King's stages for development data

	<i>n</i>	Sign. FLZ	Inte. FLZ	Marg. FLZ
KS1	57	0%	5%	95%
KS2	95	1%	25%	74%
KS3	151	20%	38%	42%
KS4	26	11%	31%	58%
KS4.5	2	50%	50%	0%
Overall	331	11%	28%	61%

KS King's stage, FLZ functional loss zone

Table 18 Zone distribution with regard to MiToS stages for validation data

	<i>n</i>	Sign. FLZ	Inte. FLZ	Marg. FLZ
MS0	147	1%	13%	86%
MS1	139	14%	37%	49%
MS2	41	34%	49%	17%
MS3	2	50%	50%	0%
MS4	2	50%	50%	0%
Overall	331	11%	28%	61%

MS MiToS stage, FLZ functional loss zone

Table 19 Zone distribution with regard to FT9 stages for validation data

	<i>n</i>	Sign. FLZ	Inte. FLZ	Marg. FLZ
FS0	6	0%	0%	100%
FS1	74	0%	4%	96%
FS2	121	7%	31%	62%
FS3	108	22%	39%	39%
FS4	22	18%	45.5%	36.5%
Overall	331	11%	28%	61%

FS FT9 stage, FLZ functional loss zone

Acknowledgements This article reported its findings as advised in the TRIPOD (Transparent Reporting of a multivariate prediction model for Individual Prognosis or Diagnosis) statement [50]. The following organisations and individuals within the PRO-ACT Consortium contributed to the design and implementation of the PRO-ACT Database

and/or provided data, but did not participate in the analysis of the data or the writing of this report: Neurological Clinical Research Institute (NCRI), Massachusetts General Hospital (MGH); Northeast ALS Consortium; Novartis; Prize4Life; Regeneron Pharmaceuticals, Inc., Sanofi; Teva Pharmaceuticals Industries, Ltd. Peter Bede is supported by the Spastic Paraplegia Foundation, Inc. (SPF), 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; he is the patron of the Irish Motor Neuron Disease Association (IMNDA).

Author Contributions VG contributed to the design of the study, analysed the data, and wrote the first draft of the manuscript. VG, GL, FD, J-FP-P, M-SS-B, PB and P-FP contributed to discussions regarding model testing and results, and to the revision of the manuscript. VG, GL, FD, J-FP-P, M-SS-B, PB and P-FP read and approved the final version.

Funding VG, GL, M-SS-B, J-FP-P, and FD: contributions were made within a SORBONNE UNIVERSITE/CNRS and FRS Consulting partnership which received funding from MESRI Grant CIFRE 2017/1051. Peter Bede is supported by the Health Research Board (HRB Ireland; HRB EIA-2017-019), Irish Institute of Clinical Neuroscience IICN and the Iris O'Brien Foundation.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical approval For this type of study formal consent is not required.

Availability of data and material Anonymised data are freely accessible from the public database of the Northeast ALS Consortium. Statistical analyses will be shared upon request of any qualified investigator.

Code availability Python packages used (UMAP, scikit-learn, pandas, numpy, matplotlib, seaborn, scipy) are freely accessible online.

References

1. Robberecht W, Philips T (2013) The changing scene of amyotrophic lateral sclerosis. *Nat Rev Neurosci* 14(4):248–264. <https://doi.org/10.1038/nrn3430>
2. Finegan E, Chipika RH, Shing SLH, Hardiman O, Bede P (2019) Primary lateral sclerosis: a distinct entity or part of the ALS spectrum? *Amyotrophic Lateral Scler Frontotemporal Degener* 20(3–4):133–145. <https://doi.org/10.1080/21678421.2018.1550518>

3. Swinnen B, Robberecht W (2014) The phenotypic variability of amyotrophic lateral sclerosis. *Nat Rev Neurol* 10(11):661–670. <https://doi.org/10.1038/nrneurol.2014.184>
4. Hiroshi M, Brooks Benjamin R, Vincenzo S (2014) Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol* 13(11):1127–1138. [https://doi.org/10.1016/s1474-4422\(14\)70129-2](https://doi.org/10.1016/s1474-4422(14)70129-2)
5. Elamin M, Bede P, Montuschi A, Pender N, Chio A, Hardiman O (2015) Predicting prognosis in amyotrophic lateral sclerosis: a simple algorithm. *J Neurol* 262(6):1447–1454. <https://doi.org/10.1007/s00415-015-7731-6>
6. Adriano C, Giancarlo L, Orla H, Robert S, Douglas M, Ettore B, Traynor Bryan G (2009) On behalf of the Eurals Consortium. Prognostic factors in ALS: a critical review. *Amyotrophic Lateral Scler* 10(5–6):310–323. <https://doi.org/10.3109/17482960802566824>
7. Elamin M, Phukan J, Bede P, Jordan N, Byrne S, Pender N, Hardiman O (2011) Executive dysfunction is a negative prognostic indicator in patients with ALS without dementia. *Neurology* 76(14):1263–1269. <https://doi.org/10.1212/wnl.0b013e318214359f>
8. Gordon PH, Salachas F, Lacomblez L, Le Forestier N, Pradat P-F, Bruneteau G, Elbaz A, Meininger V (2012) Predicting survival of patients with amyotrophic lateral sclerosis at presentation: a 15-year experience. *Neurodegenerat Dis* 12(2):81–90. <https://doi.org/10.1159/000341316>
9. Joachim W, Anton S, Whrle Johannes C, Frederick P, Nix Wilfred A, Matthias M, Grau Armin J (2014) Factors predicting one-year mortality in amyotrophic lateral sclerosis patients—data from a population-based registry. *BMC Neurol*. <https://doi.org/10.1186/s12883-014-0197-9>
10. Vincent G, Pierre-François P, Giorgia Q, François D, Gaétan LC, Jean-François P-P, Peter B (2019) Machine learning in amyotrophic lateral sclerosis: achievements, pitfalls, and future directions. *Front Neurosci*. <https://doi.org/10.3389/fnins.2019.00135>
11. Zhengnan H, Hongjiu Z, Jonathan B, Goutman Stephen A, Bhramar M, Dinov Ivo D, Yuanfang G et al (2017) Complete hazard ranking to analyze right-censored data: an als survival study. *PLoS Comput Biol* 13(12):e1005887
12. Beaulieu-Jones Brett K, Greene CS et al (2016) Semi-supervised learning of the electronic health record for phenotype stratification. *J Biomed Inf* 64:168–178
13. Christina S, Orla H, Peter B (2017) Survival prediction in amyotrophic lateral sclerosis based on MRI measures and clinical characteristics. *BMC Neurol*. <https://doi.org/10.1186/s12883-017-0854-x>
14. Henk-Jan W, Debray Thomas PA, Visser Anne E, van Eijk RPA, Rooney James PK, Andrea C, Sarah M, McDermott Christopher J, Thompson Alexander G, Susana P, Xenia K, Angela R, Beatrice S, Helma S, Middelkoop Bas M, Dekker Annelot M, van Vugt JJFA, van Rheenen W, Vajda A, Heverin M, Kazoka M, Hollinger H, Gromicho M, Krner S, Ringer Thomas M, Rdiger A, Gunkel A, Shaw Christopher E, Bredenoord Annelien L, van Es Michael A, Corcia P, Couratier P, Weber M, Grosskreutz J, Ludolph Albert C, Petri S, de Carvalho M, Van Damme P, Talbot K, Turner Martin R, Shaw Pamela J, Al-Chalabi A, Chiò A, Hardiman O, Moons Karel GM, Veldink Jan H, van den Berg Leonard H, (2018) Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol* 17(5):423–433. [https://doi.org/10.1016/s1474-4422\(18\)30089-9](https://doi.org/10.1016/s1474-4422(18)30089-9)
15. Mei-Lyn O, Fang TP, Holbrook Joanna D (2017) Predicting functional decline and survival in amyotrophic lateral sclerosis. *PLoS ONE* 12(4):e0174925
16. Taylor Albert A, Christina F, Meraida P, Liuxia W, Neta Z, Mike K, Glass Jonathan D, Ennist David L (2016) Predicting disease progression in amyotrophic lateral sclerosis. *Ann Clin Trans Neurol* 3(11):866–875. <https://doi.org/10.1002/acn3.348>
17. Hothorn T, Jung HH (2014) RandomForest4life: a random forest for predicting ALS disease progression. *Amyotrophic Lateral Scler Frontotemporal Degener* 15(5–6):444–452. <https://doi.org/10.3109/21678421.2014.893361>
18. Elamin M, Pinto-Grau M, Burke T, Bede P, Rooney J, O’Sullivan M, Lonergan K, Kirby E, Quinlan E, Breen N, Vajda A, Heverin M, Pender N, Hardiman O (2016) Identifying behavioural changes in ALS: validation of the beaumont behavioural inventory (BBI). *Amyotrophic Lateral Scler Frontotemporal Degener* 18(1–2):68–73. <https://doi.org/10.1080/21678421.2016.1248976>
19. Hannah CR, Eoin F, Stacey LHS, Orla H, Peter B, (mar, (2019) Tracking a fast-moving disease: longitudinal markers, monitoring, and clinical trial endpoints in ALS. *Front Neurol*. <https://doi.org/10.3389/fneur.2019.00229>
20. Bede P, Hardiman O (2017) Longitudinal structural changes in ALS: a three time-point imaging study of white and gray matter degeneration. *Amyotrophic Lateral Scler Frontotemporal Degener* 19(3–4):232–241. <https://doi.org/10.1080/21678421.2017.1407795>
21. Cedarbaum JM, Stambler N (1997) Performance of the amyotrophic lateral sclerosis functional rating scale (ALSFRS) in multicenter clinical trials. *J Neurol Sci* 152:s1–s9. [https://doi.org/10.1016/s0022-510x\(97\)00237-2](https://doi.org/10.1016/s0022-510x(97)00237-2)
22. Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, Nakanishi A (1999) The ALSFRS-r: a revised ALS functional rating scale that incorporates assessments of respiratory function. *J Neurol Sci* 169(1–2):13–21. [https://doi.org/10.1016/s0022-510x\(99\)00210-5](https://doi.org/10.1016/s0022-510x(99)00210-5)
23. Roche Jose C, Ricardo R-G, Scott Kirsten M, William S, Ellis Catherine E, Rachel B, Lokesh W, Turner Martin R, Nigel LP, Shaw Christopher E et al (2012) A proposed staging system for amyotrophic lateral sclerosis. *Brain* 135(3):847–852
24. Adriano C, Hammond Edward R, Gabriele M, Virginio B, Graziella F (2013) Development and evaluation of a clinical staging system for amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* 86(1):38–44
25. Thakore Nimish J, Lapin Brittany R, Kinzy Tyler G, Pioro Erik P (2018) Deconstructing progression of amyotrophic lateral sclerosis in stages: a markov modeling approach. *Amyotrophic Lateral Scler Frontotemporal Degener* 19(7–8):483–494
26. Ton F, Al KA, Stahl Daniel R, La Torre L, Claudia MC, Carolyn Y, Shaw Pamela J, Nigel LP, Ammar A-C (2017) Comparison of the king’s and MiToS staging systems for ALS. *Amyotrophic Lateral Sclerosis Frontotemporal Degener* 18(3–4):227–232. <https://doi.org/10.1080/21678421.2016.1265565>
27. Alessandro Z, Rosario V, Adriano C, Di CB, (apr, (2019) A dynamic bayesian network model for the simulation of amyotrophic lateral sclerosis progression. *BMC Bioinf*. <https://doi.org/10.1186/s12859-019-2692-x>
28. Ton F, Ahmad AK, Jacques-Henri M, Ashley J, Nigel LP, Gilbert B, Ammar A-C (2018) Stage at which riluzole treatment prolongs survival in patients with amyotrophic lateral sclerosis: a retrospective analysis of data from a dose-ranging study. *Lancet Neurol* 17(5):416–422
29. Taguchi Y, Iwade M, Umeyama H (2015) Heuristic principal component analysis-based unsupervised feature extraction and its application to gene expression analysis of amyotrophic lateral sclerosis data sets. In 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE. <https://doi.org/10.1109/cibcb.2015.7300274>
30. Ming T, Chao G, Goutman Stephen A, Alexandr K, Bhramar M, Yuanfang G, Dinov Ivo D (2019) Model-based and model-free techniques for amyotrophic lateral sclerosis diagnostic prediction and patient clustering. *Neuroinformatics* 17(3):407–421

31. McInnes L, Healy J, Melville J (2018) Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint [arXiv:1802.03426](https://arxiv.org/abs/1802.03426)
32. Lenglet T, Lacomblez L, Abitbol JL, Ludolph A, Mora JS, Wim R, Shaw PJ, Pruss RM, Cuvier V, Meininger V et al (2014) A phase ii–iii trial of olesoxime in subjects with amyotrophic lateral sclerosis. *Eur J Neurol* 21(3):529–536
33. Meininger V, Asselain B, Guillet P, Leigh PN, Ludolph A, Lacomblez L, Wim R et al (2006) Pentoxifylline in als: a double-blind, randomized, multicenter, placebo-controlled trial. *Neurology* 66(1):88–92
34. PRO-ACT (2015) Pro-act database. <https://nctu.partners.org/ProACT/Home/Index>. Accessed: 2020 Jan 01
35. Rubika B, Ashley J, Naheed J, Catherine K, Ellis Catherine M, Rachel B, Turner Martin R, Nigel LP, Shaw Christopher E, Ammar A-C (2014) Estimating clinical stage of amyotrophic lateral sclerosis from the als functional rating scale. *Amyotrophic Lateral Scler Frontotemporal Degener* 15(3–4):279–284
36. Querin G, El Mendili MM, Lenglet T, Delphine S, Marchand-Pauvert V, Benali H, Pradat P-F (2017) Spinal cord multi-parametric magnetic resonance imaging for survival prediction in amyotrophic lateral sclerosis. *Eur J Neurol* 24(8):1040–1046. <https://doi.org/10.1111/ene.13329>
37. Hulley Stephen B, Cummings Steven R, Browner Warren S, Grady Deborah G, Newman Thomas B (2006) Designing clinical research. Lippincott Williams & Wilkins, Philadelphia
38. Olney RK, Murphy J, Forshew D, Garwood E, Miller BL, Langmore S, Kohn MA, Lomen-Hoerth C (2005) The effects of executive and behavioral dysfunction on the course of ALS. *Neurology* 65(11):1774–1777. <https://doi.org/10.1212/01.wnl.0000188759.87240.8b>
39. Zhouwei X, Rafea A, Alruwaili S, Frequency of abnormality and effect on survival, (2017) Robert David Henderson, and Pamela Ann McCombe Screening for cognitive and behavioural impairment in amyotrophic lateral sclerosis. *J Neurol Sci* 376:16–23. <https://doi.org/10.1016/j.jns.2017.02.061>
40. Elamin M, Bede P, Byrne S, Jordan N, Gallagher L, Wynne B, O'Brien C, Phukan J, Lynch C, Pender N, Hardiman O (2013) Cognitive changes predict functional decline in ALS: a population-based longitudinal study. *Neurology* 80(17):1590–1597. <https://doi.org/10.1212/wnl.0b013e31828f18ac>
41. Goutman Stephen A, Chen Kevin S, Paez-Colasante Ximena, Feldman Eva L (2018) Emerging understanding of the genotype–phenotype relationship in amyotrophic lateral sclerosis. In: *Neurogenetics, Part II*, pages 603–623. Elsevier, <https://doi.org/10.1016/b978-0-444-64076-5.00039-9>
42. Alice P, De ME, Andrea L, Giulia S, Elena C, Fabrizio R, Valeria S, Christian L (2019) The prognostic value of spirometric tests in amyotrophic lateral sclerosis patients. *Clin Neurol Neurosurg* 184:105456. <https://doi.org/10.1016/j.clineuro.2019.105456>
43. Tortelli R, Copetti M, Panza F, Fontana A, Cortese R, Capozzo R, Introna A, D'Errico E, Zoccolella S, Arcuti S et al (2016) Time to generalization and prediction of survival in patients with amyotrophic lateral sclerosis: a retrospective observational study. *Eur J Neurol* 23(6):1117–1125
44. Moglia C, Calvo A, Grassano M, Canosa A, Manera U, D'Ovidio F, Bombaci A, Bersano E, Mazzini L, Mora G, Chiò A (2019) Early weight loss in amyotrophic lateral sclerosis: outcome relevance and clinical correlates in a population-based cohort. *J Neurol Neurosurg Psychiatry* 90(6):666–673. <https://doi.org/10.1136/jnnp-2018-319611>
45. Susan B, Marwa E, Peter B, Aleksey S, Cathal W, Bernie C, Mark H, Norah J, Kevin K, Catherine L, McLaughlin Russell L, Mahadeva IP, Caoimhe OB, Julie P, Brona W, Bokde Arun L, Bradley Daniel G, Niall P, Ammar A-C, Orla H (2012) Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a c9orf72 repeat expansion: a population-based cohort study. *Lancet Neurol* 11(3):232–240. [https://doi.org/10.1016/s1474-4422\(12\)70014-5](https://doi.org/10.1016/s1474-4422(12)70014-5)
46. Peter B, Iyer PM, Eoin F, Taha O, Orla H (2017) Virtual brain biopsies in amyotrophic lateral sclerosis: diagnostic classification based on in vivo pathological patterns. *NeuroImage* 15:653–658. <https://doi.org/10.1016/j.neuroimage.2017.06.010>
47. Querin G, El Mendili M-M, Bede P, Delphine S, Lenglet T, Marchand-Pauvert V, Pradat P-F (2018) Multimodal spinal cord MRI offers accurate diagnostic classification in ALS. *J Neurol Neurosurg Psychiatry* 89(11):1220–1221. <https://doi.org/10.1136/jnnp-2017-317214>
48. Agosta F, Spinelli EG, Riva N, Fontana A, Basaia S, Canu E, Castelnovo V, Falzone Y, Carrera P, Comi G, Filippi M (2019) Survival prediction models in motor neuron disease. *Eur J Neurol* 26(9):1143–1152. <https://doi.org/10.1111/ene.13957>
49. Bede P, Querin G, Pradat P-F (2018) The changing landscape of motor neuron disease imaging. *Curr Opin Neurol* 31(4):431–438. <https://doi.org/10.1097/wco.0000000000000569>
50. Collins Gary S, Reitsma Johannes B, Altman Douglas G, Moons Karel GM (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): the tripod statement. *BMC Med* 13(1):1
51. Quartagno M, Carpenter JR (2015) Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. *Stat Med* 35(17):2938–2954. <https://doi.org/10.1002/sim.6837>
52. Karl P (1901) Liii. on lines and planes of closest fit to systems of points in space. *Lond Edinburgh Dublin Philos Maga J Sci* 2(11):559–572
53. van der Maaten Laurens, Hinton Geoffrey, (2008) Visualizing data using t-sne. *J Mach Learn Res* 9:2579–2605
54. Schubert Erich, Gertz Michael (2017) Intrinsic t-stochastic neighbor embedding for visualization and outlier detection. In: *International Conference on Similarity Search and Applications*, Springer, Berlin pp 188–203

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com