Investigating the Impact of Environmental Data on ALS Prognosis with Survival Analysis

Notebook for the iDPP Lab on Intelligent Disease Progression Prediction at CLEF 2023

Ruben Branco*, Diogo F. Soares, Andreia S. Martins, Joana Barros Valente, Eduardo N. Castanho, Sara C. Madeira and Helena Aidos

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

Abstract

Amyotrophic lateral sclerosis (ALS) is characterized by rapid motor neuron degeneration and subsequent loss of motor function, typically leading to death by respiratory failure. As evidence of environmental pollutants playing a role on ALS incidence surfaces, iDPP CLEF 2023 challenge sought to evaluate the predictive power of these pollutants on prognosis. As such, we have trained four survival prediction models to rank patients based on the risk of reaching end-stage events: (a) initiation of non-invasive ventilation (NIV), (b) initiation of percutaneous endoscopic gastrostomy (PEG) and (c) death. Baseline models were trained with clinical and demographic data, and compared to models considering pollutant exposure using (1) a 6 month window and (2) all available records. The temporal aspect of environmental and clinical data was captured through feature statistics. Using Harrell's Concordance Index (C-Index) and Area Under the Receiver Operating Characteristic Curve (AUROC) as performance metrics, it was concluded that most of the developed models had some predictive power. However, the inclusion of environmental variables led to performance degradation when compared to the baseline model. Further work on capturing the temporality of environmental exposure is therefore required to understand the role of pollutants on ALS prognosis.

Keywords

amyotrophic lateral sclerosis, environmental data, disease prognosis, survival analysis

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder that affects the neuron cells responsible for controlling muscle movement, causing a progressive loss of motor functions, such as walking, speaking, and eating [1, 2]. Eventually, breathing difficulties arise, which are the leading cause of death in ALS patients [3].

Since there is no known cure for ALS, treatments focus on retarding the natural progression of the disease and improving the patient's quality of life [4]. This is a challenge since ALS is a complex and heterogeneous disease, in aspects such as affected areas (both at onset and

CLEF 2023: Conference and Labs of the Evaluation Forum, September 18–21, 2023, Thessaloniki, Greece

*Corresponding author.

rmbranco@ciencias.ulisboa.pt (R. Branco); dfsoares@ciencias.ulisboa.pt (D. F. Soares); amartins@lasige.di.fc.ul.pt (A. S. Martins); jfvalente@lasige.di.fc.ul.pt (J. B. Valente); ejcastanho@ciencias.ulisboa.pt (E. N. Castanho); sacmadeira@ciencias.ulisboa.pt (S. C. Madeira); haidos@ciencias.ulisboa.pt (H. Aidos)

throughout the disease course), with significant variability when evaluating disease progression and prognosis among individuals [5].

Machine learning techniques have emerged as powerful tools in understanding disease mechanisms and guiding clinical decision-making [6]. These methods were applied to identify biomarkers in both diagnosis and prognosis [6], prediction of survival and end-stage events [7], as well as the discovery of patient groups with similar disease progression courses [8] or treatment response [9].

The iDPP CLEF 2023 challenge[1] uses highly curated data from real MS and ALS patients, followed at clinical institutions in Lisbon, Portugal, Turin and Pavia, Italy. The focus of this challenge is the prognostic prediction in both MS and ALS.

The role of environmental pollutants in ALS has gained attention due to their increasing association and the observation of higher ALS rates, representing a potential disease risk factor [10, 11, 12]. Thus, the objective of task 3 of the challenge is to evaluate the impact of pollutants on the risk of impairment in ALS [13]. This task is based on task 1 of iDPP CLEF 2022 challenge[2], with the addition of the environmental data [14, 2].

In our contribution to iDPP, we evaluated the impact of including environmental pollutants in predicting end-stage events: initiation of non-invasive ventilation (NIV), percutaneous endoscopic gastrostomy (PEG), and death. We considered two possible time windows (6 months and all data) to analyze the temporal environmental data and compute statistical features for each pollutant. Four predictive models (Cox Proportional-Hazards, Random Survival Forest, Survival SVM, and Gradient Boosting) were evaluated for each task.

The paper is organized as follows: Section 2 introduces related works; Section 3 describes our approach and our experimental setup; Section 4 discusses our main findings; finally, Section 5 draws some conclusions and outlooks for future work.

2. Related Work

Amyotrophic lateral sclerosis does not manifest itself in a straightforward manner, as several clinical presentations regarding onset and progression can be observed. Additionally, ALS patients are typically faced with a short life expectancy [15]. Thus, prognosis prediction and identification of predictive attributes are fundamental areas of research in ALS.

Survival analysis techniques are some of the most widely used for these tasks. They are typically performed retrospectively, and by studying the impact of clinical, biological and lifestyle factors on survival times. A systematic survey carried by Chió et al. [15] identified several studies reporting that survival times were affected by factors such as age, site of onset and forced vital capacity. Other aspects have since received additional focus, like BMI [16] and cholesterol levels [17], as well as cognitive and behavioural impairment [18].

Studies on the effect of therapeutics like non-invasive ventilation (NIV) and percutaneous endoscopic gastrostomy (PEG) on survival were sparse at the time Chió et al. [15] conducted their review but typically indicated a positive effect. More recently, Ackrivo et al. [19] applied a multivariable Cox proportional hazards model to evaluate the role of NIV on survival. It

[1] http://brainteaser.dei.unipd.it/challenges/idpp2023/
[2] http://brainteaser.dei.unipd.it/challenges/idpp2022/

was determined that NIV usage was associated with a 26% reduction in death rate, and that patients with longer daily usage (> 4h) had improved survival. Spittel et al. [20] have conducted a survival study in a German cohort, in which Kaplan-Meier estimates of survival were higher in patients ventilated either non-invasively or through tracheostomy (TIV), with latter providing greater effect. Through similar methods, Spataro et al. [21] identified higher survival rates in dysphagic ALS patients submitted to PEG, when compared to those not receiving the treatment. Given the beneficial effect of these therapies, time-to-event prediction may also be applied to them, providing insight not only on prognosis but also on the best time for administration.

Research on building and evaluating predictive survival and time-to-event models, however, is still underdeveloped in comparison to survival analysis of single cohorts. Ackrivo et al. [7] developed a multivariable logistic regression model for the prediction of respiratory insufficiency or death using features collected at the first appointment. It was externally validated with promising results. Pfohl et al. [22] have constructed interactive survival regression and classification models, with variable thresholds, discovering that predictive factors of ALS survival are time-fluctuating. Thus, the introduction of temporal data to survival prediction is valuable, but poses challenges, such as capturing in a representative manner all the possible disease courses ALS may take. This was addressed by Soares et al. [23, 24] and Martins et al. [25], which built models to predict the need for NIV by encoding temporal dependencies through triclustering and sequential pattern mining, respectively. Recently triclustering methods also proved to be effective in understanding prognostics of other relevant critical endpoints [26].

Discussion on the influence of environmental factors in ALS has been gaining attention, with several studies being dedicated to the relation between pollution levels and ALS prevalence. Saucier et al. [11] conducted a series of systematic reviews, indicating a higher risk of ALS in association with several factors. Air pollutants, in particular, have shown evidence of being associated to higher risk of developing ALS [27], and to increased emergency department visits by ALS patients [28]. The impact of environmental variables on ALS prognosis is, however, still an open topic. Goutman et al. have studied the influence on survival of persistent organic pollutants [29] and self-reported exposure based on occupation [30], but environmental-based ALS survival prediction models are, to our knowledge, still in early development.

Thus, we propose to develop time-to-event models for the prediction of death, NIV and PEG, taking into account longitudinal records for both clinical and environmental exposure data.

3. Methodology

3.1. Data Preprocessing

The data made available with this challenge contains information on ALS patients and comprises three components: static (data collected at the patient's first visit to the center), clinical assessments (from patient follow-up), and environmental records (concentration of some pollutants in the air). Data was preprocessed to obtain a dataset with the shape needed to be the input of our method. Thus, we selected a set of the most relevant features from the literature [31] to be used in our models. We computed new features from the temporal and environmental ones and used them together with the static as model input. The selected and computed features are depicted in Table 1. We considered three datasets, one for each subtask according to the event

Table 1
Features

Component	Feature	Type	Computed?
Static	Sex	Binary	
	Ethnicity	Categorical	
	Age at Onset	Integer	
	Occupation	Categorical	
	Disease Onset	Categorical	
	UMN vs LMN	Categorical	
Visits (Temporal)	ALSFRS-R (total score):		
	Slope	Float	✓
	Median value of assessments		✓
	Std value of assessments		✓
	Bulbar score:		
	Slope	Float	✓
	Median value of assessments		✓
	Std value of assessments		✓
	Motor score:		
	Slope	Float	✓
	Median value of assessments		✓
	Std value of assessments		✓
	Respiratory score:		
	Slope	Float	✓
	Median value of assessments		✓
	Std value of assessments		✓

prediction [2].

Regarding the computed features, we focus on the temporal dynamics of the ALSFRS-R scale, looking for the slope, median and standard deviation of the patient's records for the total score and each subscore: bulbar, motor, and respiratory. When incorporating environmental data, we employ a similar strategy. For each pollutant, we filter the records of each pollutant to be within an appropriate time window. We consider two-time windows: 6 months, as described in the iDPP challenge [32, 33], and all the available records (no filtering). We then calculate the following features, for each pollutant: maximum, minimum, median, and standard deviation. These features, which describe the statistics of the time series of each pollutant, are then given to the model, in addition to the clinical data.

3.2. Modeling

In the previous section, Section 3.1, we described the pre-processing done to the data by selecting the most important static features and computing features to handle the temporal section of the data.

Further pre-processing is necessary at the event level to proceed with the modeling. There are three sub-tasks within Task 3: A, B, and C. They regard ranking subjects in terms of the risk of early occurrence of different endpoints:

A. Non-Invasive Ventilation (NIV) or Death, whichever occurs first;
B. Percutaneous Endoscopic Gastrostomy (PEG) or Death, whichever occurs first;
C. Death

Task 3A and 3B have two possible events requiring multi-event survival analysis. However, because for any patient, the event considered is whichever occurs first, the occurrence is independent. This allows us to formulate the problem as a single-event survival analysis, where the patient will have one of the events happen or none. We know the risk that either of them will happen, but we do not effectively know which event will happen.

We used `sksurv` [34] Python library to learn a model for each sub-task, experimenting with four models: Cox Proportional-Hazards, Random Survival Forest, Survival SVM, and Gradient Boosting.

We considered a standard machine learning pipeline: train/test split (90% train, 10% test), grid search hyperparameter optimization with Repeated Stratified K-Fold Cross Validation using the train split, and assessing the quality of the models using the test split, which is the same for each model.

Due to time constraints, we opted for a grid search hyperparameter search, with a relatively narrow search space. The search space is presented in Table 2.

Table 2
Hyperparameter space for each model. $\mathrm{np.geomspace}$ is a method that generates num numbers between the lower and upper bound, geometrically distributed.

Model	Hyperparameter	Search Values
CoxPHSurvivalAnalysis	alpha	$\mathrm{np.geomspace}(0.1, 3, \mathrm{num} = 50)$
RandomSurvivalForest	n_estimators	$[50, 100, 150, 200, \ldots, 1000]$
FastSurvivalSVM(max_iter=100000)	alpha	$\mathrm{np.geomspace}(0.001, 5, \mathrm{num} = 30)$
GradientBoostingSurvivalAnalysis	n_estimators	$[50, 100, 250, 500, 750, 1000]$
	learning_rate	$\mathrm{np.geomspace}(0.001, 5, \mathrm{num} = 10)$
	max_depth	$[1, 3, 5, 10]$

We optimize the models using Harrell's Concordance Index metric [35] (c-index), with the evaluation on the test set being the deciding factor on the models to submit to the challenge. The optimization was done for each sub-task and for three different setups in regards to environmental data: no environmental data, environmental data with a time window of six months, and all the environmental data available.

We conducted experiments using the datasets made available by the organization for Task 3 [32, 33]. All the experiments were run on a Desktop Computer with an AMD Ryzen 9 7950X 16-Core with 64GB of RAM and Ubuntu 22.04.2. The code was run using Python 3.10.11.

4. Results

In this section, we provide the results obtained for subtask A (Section 4.1), B (Section 4.2) and C (Section 4.3), along with a discussion. The particular focus of the discussion, and of this task overall, is to understand whether the inclusion of environmental data in the model inputs provides a richer context to improve the prognostic capabilities of the models.

Two evaluation metrics are presented to evaluate the models: Harrell's Concordance Index (C-Index) [35] and Area Under the Receiver Operating Characteristic Curve (AUROC). Both assess the predictive power but in different ways. C-Index measures the predictive power of a model in terms of its ability to rank and predict outcomes of groups of individuals, measured from 0 (worst) to 1 (best).

AUROC, on the other hand, measures a model's predictive power in terms of a binary classification problem by plotting the true positive rate against the false positive rate, considering different classification thresholds. The area under this curve yields a value between 0 (worst) and 1 (best). In the context of the challenge, the AUROC regards predicting event times at each time interval (12, 18, 24, 30, and 36 months).

Table 3
Hyperparameter values for each model used in subtask A (Value A column), subtask B (Value B column) and subtask C (Value C column), found through hyperparameter optimization. *: indicates a floating point with additional decimal places.

Model	Hyperparameter	Environmental Data	Value A	Value B	Value C
CoxPHSurvivalAnalysis	alpha	None	1.606*	2.798*	1.498*
		6 months	3.0	3.0	3.0
		All	3.0	3.0	3.0
RandomSurvivalForest	n_estimators	None	350	1000	1000
		6 months	450	200	800
		All	50	1000	800
FastSurvivalSVM(max_iter=3500)	alpha	None	0.001*	0.081*	0.001*
		6 months	0.001	0.001	0.001
		All	0.001	0.001	0.001
GradientBoostingSurvivalAnalysis	n_estimators	None	500	750	750
		6 months	500	500	1000
		All	500	500	750
	learning_rate	None	0.017*	0.017*	0.006*
		6 months	0.002*	0.006*	0.002*
		All	0.017*	0.006*	0.017*
	max_depth	None	1	1	3
		6 months	3	1	1
		All	1	1	1

4.1. Subtask A

As described in Section 3, we experimented with four models by performing hyperparameter optimization and testing them on a test set split derived from the provided training set. The

Table 4
Subtask A challenge results, with the following metrics: C-Index, Area Under the Receiving Operator Characteristic (AUROC), measured at different time intervals (12, 18, 24, 30 and 36 months). The report includes the estimated value for the metrics and the 95% confidence intervals underneath them. The models are Random Survival Forest (RSF) and Gradient Boosting Survival Analysis (GBSA). *: Best value for the metric (row) for the model RSF. **: Best value for the metric (row) for the model GBSA.

	None		**6 months**		**All**	
Metric	RSF	GBSA	RSF	GBSA	RSF	GBSA
C-Index	0.682* [0.651; 0.712]	0.691** [0.659; 0.723]	0.549 [0.468; 0.630]	0.613 [0.542; 0.683]	0.531 [0.458; 0.603]	0.572 [0.492; 0.653]
AUROC (12m)	0.748* [0.691; 0.804]	0.774** [0.718; 0.830]	0.603 [0.449; 0.756]	0.655 [0.521; 0.790]	0.553 [0.406; 0.700]	0.609 [0.451; 0.767]
AUROC (18m)	0.762* [0.711; 0.812]	0.780** [0.731; 0.829]	0.536 [0.397; 0.675]	0.631 [0.501; 0.760]	0.521 [0.385; 0.656]	0.598 [0.465; 0.731]
AUROC (24m)	0.778* [0.729; 0.827]	0.777** [0.728; 0.827]	0.545 [0.412; 0.677]	0.639 [0.507; 0.772]	0.489 [0.355; 0.624]	0.586 [0.455; 0.717]
AUROC (30m)	0.768* [0.716; 0.820]	0.757** [0.704; 0.811]	0.579 [0.446; 0.712]	0.686 [0.555; 0.817]	0.571 [0.434; 0.708]	0.609 [0.481; 0.736]
AUROC (36m)	0.764* [0.709; 0.818]	0.743** [0.686; 0.801]	0.562 [0.423; 0.701]	0.670 [0.534; 0.805]	0.450 [0.303; 0.596]	0.593 [0.460; 0.725]

hyperparameter found during the search for each model are presented in Table 3 (Value A column).

For the submissions, we considered the three best-performing models, out of the four, due to submission limits. In Table 4, we present the results for two of them: Fast Survival SVM (FSSVM) and Gradient Boosting Survival Analysis (GBSA). In terms of their overall performance, only in two instances have the estimated values for a metric fall below random chance (0.5), in the AUROC (24m) and AUROC (36m) for RSF using all the environmental data available. This indicates a reasonable predictive power by the models. However, in several instances, the 95% confidence interval contains 0.5. All of those instances are models trained with environmental data. Indeed, when looking at the results, we see an evident degradation in performance when providing the model with environmental and clinical data. For instance, c-index value is 0.682 and 0.691 for RSF and GBSA, respectively, when no environmental data is provided to the models. With the addition of the environemental data the c-index values, in both models, decreased as the time window of environmental data considered increase, namely 0.531 and 0.572 for RSF and GBSA, respectively.

4.2. Subtask B

Table 3 (Value B column) contains the hyperparameters through optimization for each of the considered models. Out of those models, we present the results for the Fast Survival SVM (FSSVM) and Gradient Boosting Survival Analysis (GBSA) in Table 5.

While subtask A pertains to the prediction of Non-Invasive Ventilation (NIV) or Death, subtask B pertains to Percutaneous Endoscopic Gastronomy (PEG) or Death. Despite this difference, the same patterns emerge in the results. Despite all metrics being scored above the random

Table 5
Subtask B challenge results, with the following metrics: C-Index, Area Under the Receiving Operator Characteristic (AUROC), measured at different time intervals (12, 18, 24, 30 and 36 months). The report includes the estimated value for the metrics and the 95% confidence intervals underneath them. The models are Fast Survival SVM (FSSVM) and Gradient Boosting Survival Analysis (GBSA). *: Best value for the metric (row) for the model FSSVM. **: Best value for the metric (row) for the model GBSA.

	None		**6 months**		**All**	
Metric	FSSVM	GBSA	FSSVM	GBSA	FSSVM	GBSA
C-Index	0.669* [0.639; 0.699]	0.679** [0.651; 0.708]	0.601 [0.530; 0.671]	0.641 [0.578; 0.704]	0.606 [0.535; 0.677]	0.647 [0.588; 0.705]
AUROC (12m)	0.736* [0.674; 0.797]	0.748** [0.690; 0.807]	0.652 [0.489; 0.816]	0.694 [0.557; 0.831]	0.654 [0.493; 0.816]	0.712 [0.590; 0.834]
AUROC (18m)	0.766* [0.719; 0.813]	0.768** [0.722; 0.813]	0.673 [0.557; 0.790]	0.698 [0.585; 0.810]	0.675 [0.562; 0.789]	0.710 [0.602; 0.818]
AUROC (24m)	0.741* [0.693; 0.790]	0.765** [0.719; 0.810]	0.679 [0.572; 0.785]	0.739 [0.641; 0.837]	0.695 [0.59; 0.800]	0.751 [0.656; 0.846]
AUROC (30m)	0.722* [0.672; 0.773]	0.744** [0.694; 0.794]	0.632 [0.514; 0.749]	0.687 [0.576; 0.798]	0.639 [0.52; 0.759]	0.673 [0.562; 0.785]
AUROC (36m)	0.719* [0.666; 0.771]	0.749** [0.697; 0.801]	0.577 [0.451; 0.703]	0.641 [0.518; 0.764]	0.585 [0.456; 0.715]	0.653 [0.529; 0.776]

baseline and generally achieving better results, hinting at PEG being easier to predict than NIV, the models' performance when trained with environmental data worsened. Although, in this subtask, the increase in the time window of environmental data considered, does not reflect a decrease in the c-index value.

4.3. Subtask C

The hyperparameters for each model used for the experiments in subtask C, are found in Table 3 (Value C column). Following the same structures as the previous tasks, Table 6 contains the results for this subtask, using Fast Survival SVM (FSSVM) and Gradient Boosting Survival Analysis (GBSA) models.

The results show no difference from the previous subtasks in the sense that the models show reasonable predictive power, and in general, the addition of environmental data did not introduce any advantage; however, there is an exception to the rule in the metric AUROC (12m), where GBSA fared better when leveraging clinical data and environmental data with a time window of 6 months when compared with no environmental data. Additionally, supplying all the available environmental data obtained better results than the none baseline. However, for other metrics, this pattern does not repeat itself. This could indicate that environmental data could be a productive predictor for predictions over small timespans.

4.4. Discussion

In the previous sections, we covered the results obtained for each subtask of Task 3 of the iDPP challenge. The overall takeaway is that we do not find evidence that, using the datasets provided,

Table 6
Subtask C challenge results, with the following metrics: C-Index, Area Under the Receiving Operator Characteristic (AUROC), measured at different time intervals (12, 18, 24, 30 and 36 months). The report includes the estimated value for the metrics and the 95% confidence intervals underneath them. The models are Fast Survival SVM (FSSVM) and Gradient Boosting Survival Analysis (GBSA). *: Best value for the metric (row) for the model FSSVM. **: Best value for the metric (row) for the model GBSA.

	None		**6 months**		**All**	
Metric	FSSVM	GBSA	FSSVM	GBSA	FSSVM	GBSA
C-Index	0.651* [0.622; 0.680]	0.664** [0.635; 0.692]	0.583 [0.515; 0.651]	0.641 [0.577; 0.706]	0.605 [0.538; 0.671]	0.634 [0.566; 0.703]
AUROC (12m)	0.738* [0.677; 0.798]	0.747 [0.685; 0.809]	0.709 [0.572; 0.845]	0.773** [0.656; 0.889]	0.721 [0.589; 0.853]	0.753 [0.626; 0.881]
AUROC (18m)	0.717* [0.668; 0.767]	0.737** [0.689; 0.785]	0.659 [0.545; 0.772]	0.714 [0.605; 0.823]	0.667 [0.555; 0.779]	0.706 [0.597; 0.816]
AUROC (24m)	0.702* [0.655; 0.749]	0.733** [0.688; 0.779]	0.654 [0.554; 0.754]	0.704 [0.608; 0.799]	0.684 [0.586; 0.782]	0.690 [0.592; 0.787]
AUROC (30m)	0.692* [0.644; 0.740]	0.712** [0.666; 0.759]	0.636 [0.529; 0.742]	0.672 [0.567; 0.776]	0.664 [0.56; 0.769]	0.643 [0.536; 0.751]
AUROC (36m)	0.699* [0.650; 0.748]	0.727** [0.679; 0.774]	0.553 [0.435; 0.671]	0.655 [0.541; 0.769]	0.601 [0.485; 0.718]	0.661 [0.541; 0.781]

environmental data provide additional insights to improve their predictive performance. We see the contrary, adding entropy to the models such that their performance degrades.

We posit this could be down to two factors:

- The statistics computed for the environmental data might not be informative enough to evidence patterns associated with event occurrences. Processing the data differently, or even resorting to deep learning to process the raw time series, might yield a different result;
- As discussed in Section 2, the degree of association between environmental data and the prognosis of ALS is still up for debate. It could be that for most of the cohort, there is a weak link between the two and, as such, yields no benefit in using such data.

5. Conclusions and Future Work

As no known cure for ALS exists, individuals anticipate a life expectancy limited to a few years when diagnosed with this condition. Therefore, the primary objective of treatments is to impede the progression of the disease to ensure a high standard of living for the patient. The objective of iDPP CLEF 2023 challenge is to evaluate the role of environmental pollutants in the risk of impairment in ALS.

In our work, we evaluated the performance of four predictive models (Cox Proportional-Hazards, Random Survival Forest, Survival SVM, and Gradient Boosting) in the prediction of end-stage events: initiation of non-invasive ventilation (NIV), percutaneous endoscopic gastrostomy (PEG), and death.

While our results suggest that our models have a reasonable predictive power (since in most instances the measures fall above random chance), the inclusion of environmental data does not improve the predictive performance of the models. In fact, in several instances, the models trained with environmental data degraded the performance compared to models trained only with clinical data. This degradation has two possible reasons: either the statistical analysis of the environmental data may not provide sufficient information to reveal patterns related to the occurrences of events, or the link between ALS and the pollutant data is weak in the cohort data. Therefore, alternative approaches should be used for processing the raw time series such as deep learning [36, 37] or biclustering/triclustering [38, 24, 39, 26] techniques for temporal analysis and classification.

Acknowledgments

The authors are funded by Fundação para a Ciência e a Tecnologia (FCT) through LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020), AIpALS project (PTDC/CCI-CIF/4613/2020), and PhD Research Scholarships to RB (2022.10727.BD), DFS (2020.05100.BD) and ENC (2021.07810.BD); and by BRAINTEASER project, that has received funding from the European Union's Horizon 2020 research and innovation program, under grant agreement No. 101017598.

References

[1] S. Zarei, K. Carr, L. Reiley, K. Diaz, O. Guerra, P. F. Altamirano, W. Pagani, D. Lodin, G. Orozco, A. Chinea, A comprehensive review of amyotrophic lateral sclerosis, Surg. Neurol. Int. 6 (2015) 171.

[2] H. Aidos, R. Bergamaschi, P. Cavalla, A. Chiò, A. Dagliati, B. D. Camillo, M. A. de Carvalho, N. Ferro, P. Fariselli, J. M. G. Dominguez, S. C. Madeira, E. Tavazzi, iDPP@CLEF 2023: The intelligent disease progression prediction challenge, in: Lecture Notes in Computer Science, Springer Nature Switzerland, 2023, pp. 491–498. doi:10.1007/978-3-031-28241-6_57.

[3] O. Hardiman, A. Al-Chalabi, A. Chio, E. M. Corr, G. Logroscino, W. Robberecht, P. J. Shaw, Z. Simmons, L. H. van den Berg, Amyotrophic lateral sclerosis, Nature Reviews Disease Primers 3 (2017) 17071. doi:10.1038/nrdp.2017.71.

[4] J. Dorst, A. Ludolph, A. Huebers, Disease-modifying and symptomatic treatment of amyotrophic lateral sclerosis, Therapeutic Advances in Neurological Disorders (2018).

[5] M. A. van Es, O. Hardiman, A. Chio, A. Al-Chalabi, R. J. Pasterkamp, J. H. Veldink, L. H. van den Berg, Amyotrophic lateral sclerosis, Lancet 390 (2017) 2084–2098.

[6] V. Grollemund, P.-F. Pradat, G. Querin, F. Delbot, G. Le Chat, J.-F. Pradat-Peyre, P. Bede, Machine Learning in Amyotrophic Lateral Sclerosis: Achievements, Pitfalls, and Future Directions, Frontiers in Neuroscience 13 (2019) 135. doi:10.3389/fnins.2019.00135.

[7] J. Ackrivo, J. Hansen-Flaschen, E. P. Wileyto, L. Schwab, Richard J.and Elman, S. M. Kawut, Development of a prognostic model of respiratory insufficiency or death in amyotrophic lateral sclerosis, European Respiratory Journal 53 (4) (2019).

[8] R. AKueffner, N. Zach, M. Bronfeld, et al., Stratification of amyotrophic lateral sclerosis patients: a crowdsourcing approach, Sci Rep 9 690 (2019).
[9] H. Seibold, A. Zeileis, T. Hothorn, Model-based recursive partitioning for subgroup analyses, The International Journal of Biostatistics 12 (1) (2016) 45–63.
[10] F.-C. Su, S. A. Goutman, S. Chernyak, B. Mukherjee, B. C. Callaghan, S. Batterman, E. L. Feldman, Association of environmental toxins with amyotrophic lateral sclerosis, JAMA Neurol. 73 (2016) 803–811.
[11] D. Saucier, P. P. W. Registe, M. Bélanger, C. O'Connell, Urbanization, air pollution, and water pollution: Identification of potential environmental risk factors associated with amyotrophic lateral sclerosis using systematic reviews, Front. Neurol. 14 (2023) 1108383.
[12] B. Oskarsson, D. K. Horton, H. Mitsumoto, Potential environmental factors in amyotrophic lateral sclerosis, Neurologic Clinics 33 (2015) 877–888. doi:https://doi.org/10.1016/j.ncl.2015.07.009, motor Neuron Disease.
[13] M. J. Mohammadi, K. Zarea, N. Hatamzadeh, A. Salahshouri, A. Sharhani, Toxic air pollutants and their effect on multiple sclerosis: A review study, Frontiers in Public Health 10 (2022). doi:10.3389/fpubh.2022.898043.
[14] A. Guazzo, I. Trescato, E. Longato, E. Hazizaj, D. Dosso, G. Faggioli, G. Di Nunzio, G. Silvello, M. Vettoretti, E. Tavazzi, et al., Overview of idpp@ clef 2022: the intelligent disease progression prediction challenge, in: CLEF, 2022, pp. 1613–0073.
[15] A. Chiò, G. Logroscino, O. Hardiman, R. Swingler, D. Mitchell, E. Beghi, B. G. Traynor, O. B. of the Eurals Consortium, Prognostic factors in als: A critical review, Amyotrophic Lateral Sclerosis 10 (2009) 310–323. doi:https://doi.org/10.3109/17482960802566824.
[16] P. Ning, B. Yang, S. Li, X. Mu, Q. Shen, F. Hu, Y. Tang, X. Yang, Y. Xu, Systematic review of the prognostic role of body mass index in amyotrophic lateral sclerosis, Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 20 (2019) 356–367. doi:https://doi.org/10.1080/21678421.2019.1587631.
[17] C. Ingre, L. Chen, Y. Zhan, J. Termorshuizen, L. Yin, F. Fang, Lipids, apolipoproteins, and prognosis of amyotrophic lateral sclerosis, Neurology 94 (2020) e1835–e1844. doi:https://doi.org/10.1212/WNL.0000000000009322.
[18] C. Nguyen, J. Caga, C. J. Mahoney, M. C. Kiernan, W. Huynh, Behavioural changes predict poorer survival in amyotrophic lateral sclerosis, Brain and Cognition 150 (2021). doi:https://doi.org/10.1016/j.bandc.2021.105710.
[19] J. Ackrivo, J. Hsu, J. Hansen-Flaschen, L. Elman, S. Kawut, Noninvasive ventilation use is associated with better survival in amyotrophic lateral sclerosis, Ann Am Thorac Soc 18 (2021) 486–494. doi:https://doi.org/10.1513/annalsats.202002-169oc.
[20] S. Spittel, A. Maier, D. Kettemann, B. Walter, B. Koch, K. Krause, J. Norden, C. Münch, T. Meyer, Non-invasive and tracheostomy invasive ventilation in amyotrophic lateral sclerosis: Utilization and survival rates in a cohort study over 12 years in germany, Eur J Neurol 28 (2021) 1160–1171. doi:https://doi.org/10.1111/ene.14647.
[21] R. Spataro, L. Ficano, F. Piccoli, V. La Bella, Percutaneous endoscopic gastrostomy in amyotrophic lateral sclerosis: Effect on survival, Journal of the Neurological Sciences 304 (2011) 44–48. doi:https://doi.org/10.1016/j.jns.2011.02.016.
[22] S. R. Pfohl, R. B. Kim, G. S. Coan, C. S. Mitchell, Unraveling the complexity of amyotrophic lateral sclerosis survival prediction, Frontiers in Neuroinformatics 12 (2018). doi:https:

//doi.org/10.3389/fninf.2018.00036.

[23] D. Soares, R. Henriques, M. Gromicho, S. Pinto, M. de Carvalho, S. Madeira, Towards triclustering-based classification of three-way clinical data: A case study on predicting non-invasive ventilation in als, in: G. Panuccio, M. Rocha, F. Fdez-Riverola, M. Mohamad, R. Casado-Vara (Eds.), Practical Applications of Computational Biology & Bioinformatics, 14th International Conference (PACBB 2020). PACBB 2020. Advances in Intelligent Systems and Computing, vol 1240, Springer, Cham., 2021. doi:https://doi.org/10.1007/978-3-030-54568-0_12.

[24] D. F. Soares, R. Henriques, M. Gromicho, M. de Carvalho, S. C. Madeira, Learning prognostic models using a mixture of biclustering and triclustering: Predicting the need for non-invasive ventilation in amyotrophic lateral sclerosis, Journal of Biomedical Informatics 134 (2022) 104172.

[25] A. S. Martins, M. Gromicho, S. Pinto, M. de Carvalho, S. C. Madeira, Learning prognostic models using disease progression patterns: Predicting the need for non-invasive ventilation in amyotrophic lateral sclerosis, IEEE/ACM Transactions on Computational Biology and Bioinformatics 19 (2021) 2572–2583. doi:https://doi.org/10.1109/TCBB.2021.3078362.

[26] D. F. Soares, R. Henriques, M. Gromicho, M. de Carvalho, S. C. Madeira, Triclustering-based classification of longitudinal data for prognostic prediction: targeting relevant clinical endpoints in amyotrophic lateral sclerosis, Scientific Reports 13 (2023) 6182.

[27] M. Seelen, R. Toro Campos, J. Veldink, A. Visser, G. Hoek, B. Brunekreef, A. van der Kooi, M. de Visser, J. Raaphorst, L. van den Berg, R. Vermeulen, Long-term air pollution exposure and amyotrophic lateral sclerosis in netherlands: A population-based case-control study, Environ Health Perspect 27 (2017). doi:https://doi.org/10.1289/EHP1115.

[28] W. Myung, H. Lee, H. Kim, Short-term air pollution exposure and emergency department visits for amyotrophic lateral sclerosis: A time-stratified case-crossover analysis, Environment International 123 (2019) 467–475. doi:https://doi.org/10.1016/j.envint.2018.12.042.

[29] S. Goutman, B. J, A. Patterson, et al., High plasma concentrations of organic pollutants negatively impact survival in amyotrophic lateral sclerosis, Journal of Neurology, Neurosurgery & Psychiatry 90 (2019) 907–912. doi:http://dx.doi.org/10.1136/jnnp-2018-319785.

[30] S. A. Goutman, J. Boss, C. Godwin, B. Mukherjee, E. L. Feldman, S. A. Batterman, Occupational history associates with als survival and onset segment, Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 24 (2023) 219–229. doi:https://doi.org/10.1080/21678421.2022.2127324.

[31] S. Pires, M. Gromicho, S. Pinto, M. Carvalho, S. C. Madeira, Predicting non-invasive ventilation in als patients using stratified disease progression groups, in: 2018 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE, 2018, pp. 748–757.

[32] G. Faggioli, A. Guazzo, S. Marchesin, L. Menotti, I. Trescato, H. Aidos, R. Bergamaschi, G. Birolo, P. Cavalla, A. Chiò, A. Dagliati, M. de Carvalho, G. M. Di Nunzio, P. Fariselli, J. M. García Dominguez, M. Gromicho, E. Longato, S. C. Madeira, U. Manera, G. Silvello, E. Tavazzi, E. Tavazzi, M. Vettoretti, B. Di Camillo, N. Ferro, Intelligent Disease Progression Prediction: Overview of iDPP@CLEF 2023, in: A. Arampatzis, E. Kanoulas, T. Tsikrika,

S. Vrochidis, A. Giachanou, A. Li, D. abd Aliannejadi, M. Vlachos, G. Faggioli, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the Fourteenth International Conference of the CLEF Association (CLEF 2023), Lecture Notes in Computer Science (LNCS), Springer, Heidelberg, Germany, 2023.

[33] G. Faggioli, A. Guazzo, S. Marchesin, L. Menotti, I. Trescato, H. Aidos, R. Bergamaschi, G. Birolo, P. Cavalla, A. Chiò, A. Dagliati, M. de Carvalho, G. M. Di Nunzio, P. Fariselli, J. M. García Dominguez, M. Gromicho, E. Longato, S. C. Madeira, U. Manera, G. Silvello, E. Tavazzi, E. Tavazzi, M. Vettoretti, B. Di Camillo, N. Ferro, Overview of iDPP@CLEF 2023: The Intelligent Disease Progression Prediction Challenge, in: M. Aliannejadi, G. Faggioli, N. Ferro, M. Vlachos (Eds.), CLEF 2023 Working Notes, CEUR Workshop Proceedings (CEUR-WS.org), ISSN 1613-0073., 2023.

[34] S. Pölsterl, scikit-survival: A library for time-to-event analysis built on top of scikit-learn, Journal of Machine Learning Research 21 (2020) 1–6.

[35] F. E. Harrell, R. M. Califf, D. B. Pryor, K. L. Lee, R. A. Rosati, Evaluating the yield of medical tests, Jama 247 (1982) 2543–2546.

[36] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, P.-A. Muller, Deep learning for time series classification: a review, Data Min. Knowl. Discov. 33 (2019) 917–963.

[37] M. Längkvist, L. Karlsson, A. Loutfi, A review of unsupervised feature learning and deep learning for time-series modeling, Pattern Recognit. Lett. 42 (2014) 11–24.

[38] R. Henriques, S. C. Madeira, FleBiC: Learning classifiers from high-dimensional biomedical data using discriminative biclusters with non-constant patterns, Pattern Recognition 115 (2021) 107900. doi:10.1016/j.patcog.2021.107900.

[39] E. N. Castanho, H. Aidos, S. C. Madeira, Biclustering fMRI time series: a comparative study, BMC Bioinformatics 23 (2022). doi:10.1186/s12859-022-04733-8.