



Unraveling the Complexity of Amyotrophic Lateral Sclerosis Survival Prediction

Stephen R. Pfohl^{1,2}, Renaid B. Kim^{1,3}, Grant S. Coan^{1,4} and Cassie S. Mitchell^{1*}

¹ Department of Biomedical Engineering, Georgia Institute of Technology, Emory University School of Medicine, Atlanta, GA, United States, ² Department of Biomedical Informatics, Stanford University, Stanford, CA, United States, ³ Medical Scientist Training Program, University of Michigan Medical School, Ann Arbor, MI, United States, ⁴ School of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

Objective: The heterogeneity of amyotrophic lateral sclerosis (ALS) survival duration, which varies from <1 year to >10 years, challenges clinical decisions and trials. Utilizing data from 801 deceased ALS patients, we: (1) assess the underlying complex relationships among common clinical ALS metrics; (2) identify which clinical ALS metrics are the “best” survival predictors and how their predictive ability changes as a function of disease progression.

Methods: Analyses included examination of relationships within the raw data as well as the construction of interactive survival regression and classification models (generalized linear model and random forests model). Dimensionality reduction and feature clustering enabled decomposition of clinical variable contributions. Thirty-eight metrics were utilized, including Medical Research Council (MRC) muscle scores; respiratory function, including forced vital capacity (FVC) and FVC % predicted, oxygen saturation, negative inspiratory force (NIF); the Revised ALS Functional Rating Scale (ALSFRS-R) and its activities of daily living (ADL) and respiratory sub-scores; body weight; onset type, onset age, gender, and height. Prognostic random forest models confirm the dominance of patient age-related parameters decline in classifying survival at thresholds of 30, 60, 90, and 180 days and 1, 2, 3, 4, and 5 years.

Results: Collective prognostic insight derived from the overall investigation includes: multi-dimensionality of ALSFRS-R scores suggests cautious usage for survival forecasting; upper and lower extremities independently degenerate and are autonomous from respiratory decline, with the latter associating with nearer-to-death classifications; height and weight-based metrics are auxiliary predictors for farther-from-death classifications; sex and onset site (limb, bulbar) are not independent survival predictors due to age co-correlation.

Conclusion: The dimensionality and fluctuating predictors of ALS survival must be considered when developing predictive models for clinical trial development or in-clinic usage. Additional independent metrics and possible revisions to current metrics, like the ALSFRS-R, are needed to capture the underlying complexity needed for population and personalized forecasting of survival.

Keywords: predictive medicine, data complexity, survival analysis, neuromuscular disease, health informatics

OPEN ACCESS

Edited by:

Arjen van Ooyen,
VU University Amsterdam,
Netherlands

Reviewed by:

Julien Cohen-Adad,
École Polytechnique de Montréal,
Canada
Cristina Moglia,
Università degli Studi di Torino, Italy

***Correspondence:**

Cassie S. Mitchell
cassie.mitchell@bme.gatech.edu

Received: 01 January 2018

Accepted: 28 May 2018

Published: 14 June 2018

Citation:

Pfohl SR, Kim RB, Coan GS and Mitchell CS (2018) Unraveling the Complexity of Amyotrophic Lateral Sclerosis Survival Prediction. *Front. Neuroinform.* 12:36. doi: 10.3389/fninf.2018.00036

INTRODUCTION

The prolific 2014 ALS Association's Ice Bucket Challenge commenced the world-wide dumping of ice water on the heads of courageous supporters to bring awareness and research funding to a lesser-known yet fatal neurodegenerative disease, Amyotrophic Lateral Sclerosis (ALS) (Wicks, 2014). Despite recent identification of genotypes (Hrastelj and Robertson, 2016) and phenotypes (Hollinger et al., 2016), ALS remains elusive, largely due to its heterogeneous nature. The multifaceted etiology of ALS is reflected in its sporadic underpinnings, as less than 10% of patients have a familial form (Hrastelj and Robertson, 2016). Varied symptom onset types include extremity muscle deficits, the dominant "limb onset"; dysarthria or dysphagia symptoms, "bulbar onset"; and a small fraction have onset with respiratory or generalized weakness. Finally, post-onset ALS survival duration fluctuates from less than 1 year to over 10 years (Beghi et al., 2011).

Inherent heterogeneity of ALS makes it exceedingly difficult to predict and exacerbates anxiety among ALS patients and their families. Heterogeneity also increases the physicians' challenge to develop a timely intervention plan and is blamed for numerous failed clinical trials, as statistical variance drowns all but inconceivable treatment effect sizes (Beghi et al., 2011). The inability to accurately gauge varied outcomes makes the assessment of therapeutic efficacy problematic, slowing curative discovery (Rutkove, 2015). In short, the ability to account for heterogeneity and predict real-time ALS clinical outcomes is a necessity.

Recent crowdsourcing efforts (Atassi et al., 2014; Küffner et al., 2015; Zach et al., 2015) focused on developing computer models to predict clinical ALS progression. Namely, the DREAM-Phil Bowen ALS Prediction Prize4Life challenge (Küffner et al., 2015) used the PRO-ACT clinical trial data (Atassi et al., 2014) to forecast slope of decline of the Revised ALS Functional Rating Scale (ALSFRS-R) (Cedarbaum and Stambler, 1997; Cedarbaum et al., 1999; Castrillo-Viguera et al., 2010; Creemers et al., 2015), a clinical questionnaire that evaluates externally observable function. However, the ambiguity and heterogeneity associated with ALSFRS-R (Voustianiouk et al., 2008; Franchignoni et al., 2013, 2015; Mandrioli et al., 2015) as an outcome metric and the characteristics of the PRO-ACT cohort (Fournier and Glass, 2015) limit their translational usage. A very recently published multi-clinic study was one of the first to develop a personalized patient prediction model (Westeneng et al., 2018) for patients who did not receive ventilation support. While the latter study was yet another step forward, patient heterogeneity still greatly complicates personalized prediction, much less population-level models necessary for clinical trials.

Before truly representative and personalized or population-level prediction models can be a new standard for clinical trial development and in-clinic decision-making, we must first better understand the complex relationships among the clinical metrics used to characterize ALS progression and forecast survival. ALS is a diagnosis of exclusion, meaning there is not one single objective medical "test" (imaging, blood test, genotyping, etc.) that is applicable to positively identifying all ALS patients

nor is there one single "measurement" that can be used to objectively and precisely stage the greatly variable temporal ALS disease course. For this reason, a combination of assessments must be used to characterize clinical ALS. Understanding the relationships among existing clinical assessments and how they correlate to various disease stages is critical to optimize clinical ALS characterization, and ultimately, interventional and survival forecasting.

The primary goals of this study are to: (1) assess the underlying complex relationships among clinical ALS metrics; (2) identify the "best" ALS clinical predictors and how predictor ranking changes as a function of disease progression and/or disease stage. First, we utilize raw clinical ALS assessment metrics and patient characteristics to assess the complexity of the disease, including the underlying latent dimensionality. Secondly, we construct interactive models that directly utilize survival/lifespan as the outcome metric for ALS patient survival classification at various thresholds ranging from 30 days to 5 years. Our analytical findings provide key insight for developing better ALS progression metrics and resolving the heterogeneity of ALS that plagues clinical trials and blurs clinical decision-making.

MATERIALS AND METHODS

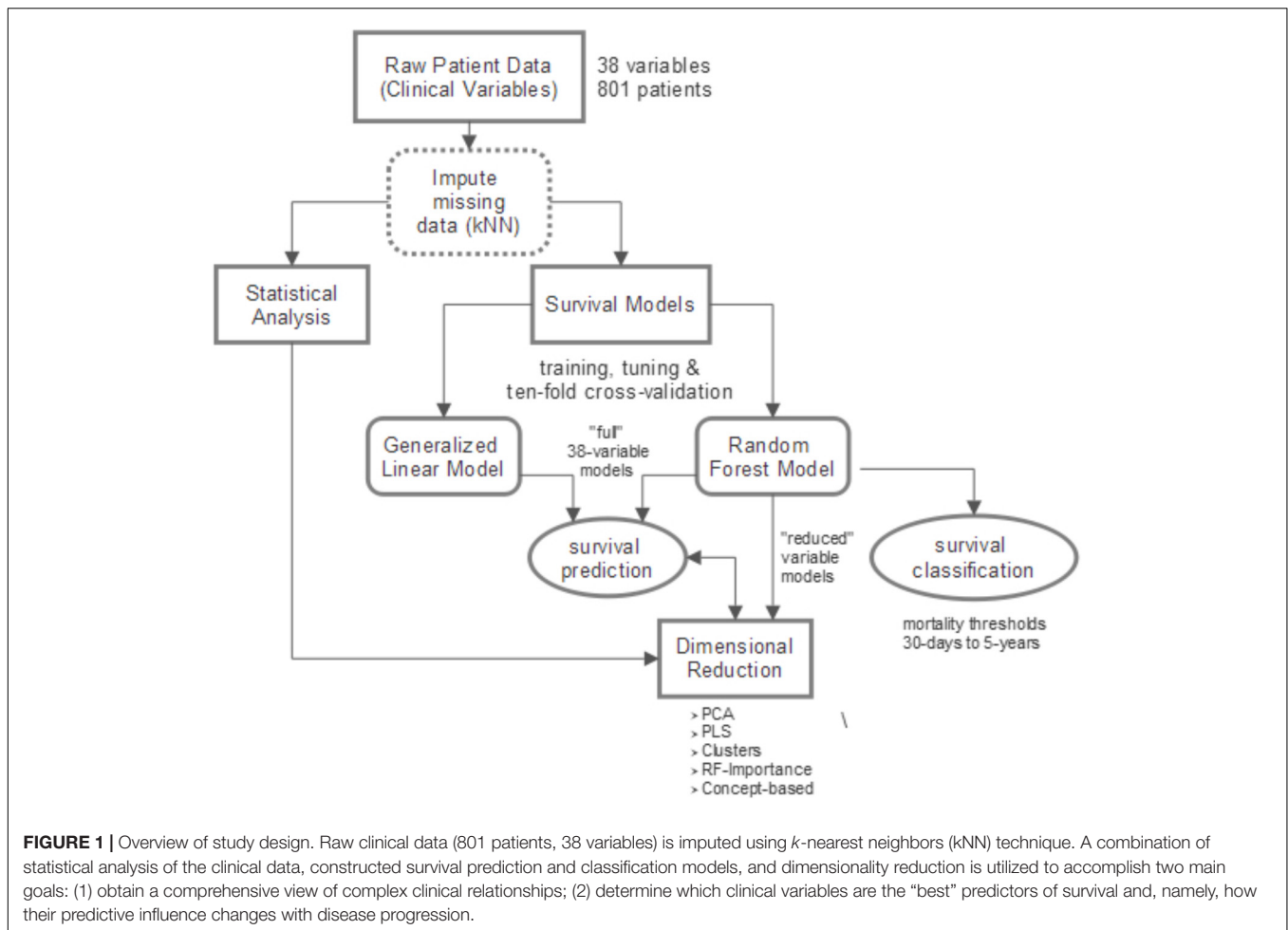
We performed a retrospective analysis of 8,028 clinical visit records collected from 1,585 patients at the Emory ALS Clinic (Emory University Hospital, Atlanta, GA, United States). Data collection, transcription, and quality control methods are as previously published (Mitchell et al., 2015a; Hollinger et al., 2016). All data was de-identified and anonymized. Institutional Review Board approval was obtained from Georgia Institute of Technology and Emory University.

We first performed an exploratory analysis of the raw clinical data. We then construct interactive models to predict and/or classify survival duration. Collectively, the analysis of raw data and the models are used to assess the complex relationships among clinical measures of ALS and their dynamic correlation with survival forecasting. **Figure 1** provides a diagrammatic overview of the study's methods with additional details presented below.

Data Inclusion Criteria and Missing Value Imputation

The 38 clinical variables utilized as predictors are listed in the left column of **Table 1**. Included patients had a known date of death and a complete set of Medical Research Council (MRC) muscular scores measured for at least one visit. The final data set (**Table 1**) for model development encompassed 3,918 visit records collected from 801 patients. Missing values in the original data set were imputed using the *k*-nearest neighbor (kNN) imputation technique. Specifically we utilized the mean value for the 5-nearest neighbors (e.g., kNN *k*-value of 5), with a Euclidean distance metric as described in Ishwaran et al. (2008).

Metrics included are: forced vital capacity (FVC); the observed percent of predicted FVC (% predict FVC); the negative inspiratory force (NIF); oxygen saturation (O₂ sat); left and



right sided MRC muscular function scores for eleven extremity muscles; height; weight; onset type (bulbar or limb); sex; age; the (ALSFRS-R) (Cedarbaum and Stambler, 1997; Cedarbaum et al., 1999), including the ALSFRS-R total score, the ALSFRS-R activities of daily living (ADL) sub-score, and the ALSFRS-R respiratory sub-score; estimated date of ALS symptom onset (as reported by patient); ALS diagnosis date; ALS “baseline” visit date (first visit at tertiary ALS clinic); and date of death.

Exploratory Clinical Data Statistical Analysis

The linear Pearson correlation was computed for each pair of visit-level variables for the 3,918 visit records and visualized as a correlation matrix using *corrplot* (Wei and Simko, 2016). To produce the dendrogram, correlations were re-calculated with time until death (the response variable) excluded. The resulting pairwise correlations were transformed to a distance metric such that distance was $(1 - \text{correlation})/2$ and clustered hierarchically with complete linkage.

The dimensionality of the imputed predictor dataset was reduced with principal component analysis (PCA) and results visualized with *FactoMineR* and *factoextra* (Lê et al., 2008). Additionally, a partial least square (PLS) model

(Dayal and MacGregor, 1997; Friedman et al., 2010) was fit with visit-level variables with time until death as the response variable. The projection of a variable onto a PLS component was taken as the corresponding element of PLS weight matrix [matrix **R** in (Dayal and MacGregor, 1997)].

Overview of Survival Prediction Models

Two primary predictive model types were constructed: a generalized linear model (GLM) where the regularization path is computed using the elastic net penalty (e.g., a linear combination between lasso and ridge regression) (Friedman et al., 2010), abbreviated in the figures as “Lasso,” and the full random forest model (Liaw and Wiener, 2002), depicted in the figures as “Full-RF. Both the GLM model and the Full-RF included all 38 variables illustrated in **Table 1**. The GLM and random forest model types were chosen based on their usage in prior ALS survival prediction literature (Fournier and Glass, 2015; Küffner et al., 2015), as well as their unique attributes that make them suitable to the goals of this study. The GLM enables *p*-values to be calculated to examine quantitative significance of clinical variables as predictors for survival. The main disadvantages of the GLM are that it does not account for complex interactions, and the GLM implicitly assumes the input variables fit a multivariate

TABLE 1 | Number of missing visit records by field.

Variable	Missing records – all (%)	Missing records – valid (%)
Bicep.Left	3238 (39.8)	1451 (37.0)
Bicep.Right	3237 (39.8)	1446 (36.9)
Delt.Left	3240 (39.9)	1456 (37.2)
Delt.Right	3236 (39.8)	1452 (37.1)
Edb.Left	4869 (59.9)	2374 (60.6)
Edb.Right	4851 (59.7)	2347 (59.9)
Gas.Left	3911 (48.1)	1905 (48.6)
Gas.Right	3896 (47.9)	1900 (48.5)
Ham.Left	3475 (42.7)	1639 (41.8)
Ham.Right	3480 (42.8)	1641 (41.9)
Intrin.Left	3580 (44.0)	1657 (42.3)
Intrin.Right	3573 (43.9)	1657 (42.3)
IP.Left	3414 (42.0)	1589 (40.6)
IP.Right	3419 (42.1)	1587 (40.5)
Quad.Left	3385 (41.6)	1569 (40.0)
Quad.Right	3383 (41.6)	1566 (40.0)
TA.Left	3464 (42.6)	1619 (41.3)
TA.Right	3449 (42.4)	1608 (41.0)
Tricep.Left	3299 (40.6)	1501 (38.3)
Tricep.Right	3290 (40.5)	1490 (38.0)
Wrist.Left	3458 (42.5)	1592 (40.6)
Wrist.Right	3445 (42.4)	1580 (40.3)
Time.Since.First.Visit	0 (0.0)	0 (0.0)
Time.Until.Death	3495 (43.0)	0 (0.0)
Duration.From.Diagnosis	467 (5.7)	59 (1.5)
Duration.From.Onset	105 (1.3)	39 (1.0)
Forced.Vital.Capacity	2718 (33.4)	1226 (31.3)
% Predict	2502 (30.8)	1108 (28.3)
NIF	3388 (41.7)	1459 (37.2)
O2.sat	5086 (62.6)	2345 (59.9)
Age	0 (0.0)	0 (0.0)
Patient.Weight	5205 (64.0)	2114 (54.0)
Patient.Info.Height	3931 (48.4)	1590 (40.6)
ALSFRS.r.total	5076 (62.4)	2611 (66.6)
ALSFRS.r.ADL	6615 (81.4)	2996 (76.5)
ALSFRS.r.Respiratory	6316 (77.7)	2818 (71.9)
Onset.Type	221 (2.7)	112 (2.9)
Sex	0 (0.0)	0 (0.0)

The Valid designation refers to those records belonging to patients in which the time of death was known and for which muscular measurements were recorded for at least one visit.

normal distribution. In contrast, the random forest models do account for complex interactions, do not require the assumption of a multivariate normal distribution, are non-parametric and robust under arbitrary re-scalings of the inputs (Svetnik et al., 2003). In summary, random forest models are simply better suited for examining temporal dynamics and for elucidating data structure complexity with model dimensionality reduction techniques (Strobl et al., 2008).

For the stated goals of this study, which focuses on examining complex clinical ALS metric relationships and their varying correlation with survival as a function of disease

progression, using independent records was preferable to using patient-specific histories (e.g., patient-specific variable slopes) to minimize over-fitting. Therefore, all survival models predict time until death using independent visit records. Specifically, time until death (in days) for each visit was considered as the response variable with the visit-level variables (Table 1) as predictors. Both predictive regression and classification random forest models of survival were constructed. Using standard random forest classification models, mortality was predicted as a binary outcome across mortality thresholds (Table 2) in which a positive result is assigned to visits with a time until death smaller than the threshold time. Random forest out of bag predictions for the final model were used to estimate sensitivity (recall), specificity, precision (positive predictive value) across the spectrum of internal decision cut-offs and mortality thresholds.

Reduced Models Using Dimensionality Reduction

In machine learning, dimensionality reduction is the process of reducing the number of model variables under consideration by obtaining a set of principal variables from the full set of variables that can be used to accurately predict outcome. Dimensional reduction is typically divided into feature selection and feature extraction. In short, the primary purpose of dimensional reduction in the present study is to better elucidate underlying characteristics of the data set and what variables most contribute to model performance (e.g., better survival prediction).

We used four standard dimensionality reduction settings to iteratively examine predictive performance in reduced random forest models with k components, “PCA-RF” utilized projections of data on to the principal components as predictors in the random forest (Tolosi and Lengauer, 2011); “PLS-RF” used the PLS scores (Dayal and MacGregor, 1997) as predictors in the random forest; “Importance” used rank-ordering of variable importance values from the Full-RF to reduce the number of predictors (Strobl et al., 2007); “Clusters” used the clustered features from supervised feature clustering (Tolosi and Lengauer, 2011) as predictors to random forest. No explicit specifications are made regarding the number of dimensions included in the reduced models. Instead, the relative performance of the models for a grid of potential reduced model component values, k , is performed and illustrated for all four dimensionality

TABLE 2 | Number of visits and patients across mortality thresholds.

Mortality threshold	Number of visits (%)	Number of patients (%)
30 days	91 (2.3)	85 (10.6)
60 days	194 (5.0)	172 (21.5)
90 days	324 (8.3)	283 (35.3)
180 days	721 (18.4)	491 (61.3)
1 year	1506 (38.4)	639 (79.8)
2 years	2565 (65.5)	732 (91.4)
3 years	3101 (79.1)	766 (95.6)
4 years	3396 (86.7)	782 (97.6)
5 years	3591 (91.7)	789 (98.5)

Data shown are the number of visits in the dataset with a time until death of less than associated mortality threshold and the patients associated with those visits.

reduction methods. This allows visualization of the necessary dimensionality to attain performance comparable to that of the full random forest model, “Full-RF” which uses all 38 clinical variables.

In addition to reduced models constructed using standard dimensionality reduction techniques as described above, a reduced model variable sub-set combination experiment was performed using concept-based variables grouped by clinical concept: “duration” (time from onset, diagnosis, and first visit); “age”; “respiratory” (O₂ saturation, NIF, FVC, and % predicted FVC); “muscle” (all individual muscle scores), “ALSFRS” (ALSFRS-R total, respiratory and ADL sub-scores); “weight and height”; “onset type and sex” (bulbar or limb; male or female). Concept-based variable grouping assists in understanding the importance of variable sub-sets that are clinically intuitive.

Survival Model Construction and Validation

All regression and classification models were built with the *caret* framework in the R programming language. Random forest models were built with the *randomForest* (Liaw and Wiener, 2002) package and generalized linear models were built with the *glmnet* package (Friedman et al., 2010). For each model, 10-fold cross validation was repeated ten times. The mean and standard deviation for model performance metrics are computed based on pooling the performance results across the ten repetitions of 10-fold cross validation, corresponding to 100 values.

For ten repetitions of 10-fold cross-validation, the original sample is first randomly partitioned into ten equal size subsamples. Of the ten subsamples, a single subsample is retained as the validation data for testing the model, and the remaining subsamples are used as training data. The cross-validation process is then repeated ten times (the folds), with each of the subsamples used exactly once for validation. The results from the folds are then combined to produce a mean and standard deviation for each performance metric. The advantage of this method is that it maximizes the analytical sample size as all observations are used for both training and validation, and each observation is used for validation exactly once. For this exploratory study to assess clinical ALS dimensionality and complexity, 10 repetitions of 10-fold cross-validation was deemed superior to a truly split training and validation data, the latter which is better for non-exploratory models solely focused on personalized prediction accuracy.

In regression cases, performance is defined as either the root mean squared error (RMSE) or the testing R^2 (squared correlation or percent variance explained) of the predicted response for testing samples with respect to their known value. In classification cases, the performance was defined as the area under the curve (AUC) of the receiver-operating-characteristic (ROC) for testing samples.

For each testing-training split of the cross-validation procedure, the missing data were imputed using kNN technique and scaled to zero-mean and unit-variance. This process was repeated over a grid search of the tuning parameters (α and λ for elastic net, m_{try} for random forest) to determine the parameters that result in optimal performance (minimal RMSE or maximal AUC). Other relevant random forest

parameters followed the defaults from the *randomForest* R package (Liaw and Wiener, 2002); that is, 500 trees were used for each forest and trees were grown to maximal depth subject to the constraints that terminal nodes had a minimum of size 5 for regression tasks and of size 1 for classification tasks. In each case, a final model was then generated with the optimal parameters for the full dataset after missing value imputation using kNN. If unspecified, data are presented as mean \pm standard deviation and the results visualized with *ggplot* (Wickham, 2009).

Assessment of Survival Model Variable Importance

The importance of variables in the random forest models was assessed as the percent increase in the out-of-bag mean squared error upon independent permutation of each variable for the final model following cross validation with the supervised feature clustering method used to correct for potential correlation bias (Tolosi and Lengauer, 2011). The significance of variables in the GLM model was assessed with the exact post-selection inference method for the lasso (Tibshirani et al., 2014), enabling p -value computation while controlling for implicit multiple comparison.

RESULTS

A total of 801 deceased patients, 3,918 clinic visits, and 38 assessed clinical variables (Table 1) were ultimately included in the results and analysis of this study. As depicted in Figure 1, this exploratory study consists of analysis of complex relationships within the raw data, development of survival prediction and classification models, and dimensionality reduction of the raw data and survival models. The resulting overlapping insights bring into focus a comprehensive view of the complex relationships among variables currently utilized to assess disease progression and predict ALS patient survival.

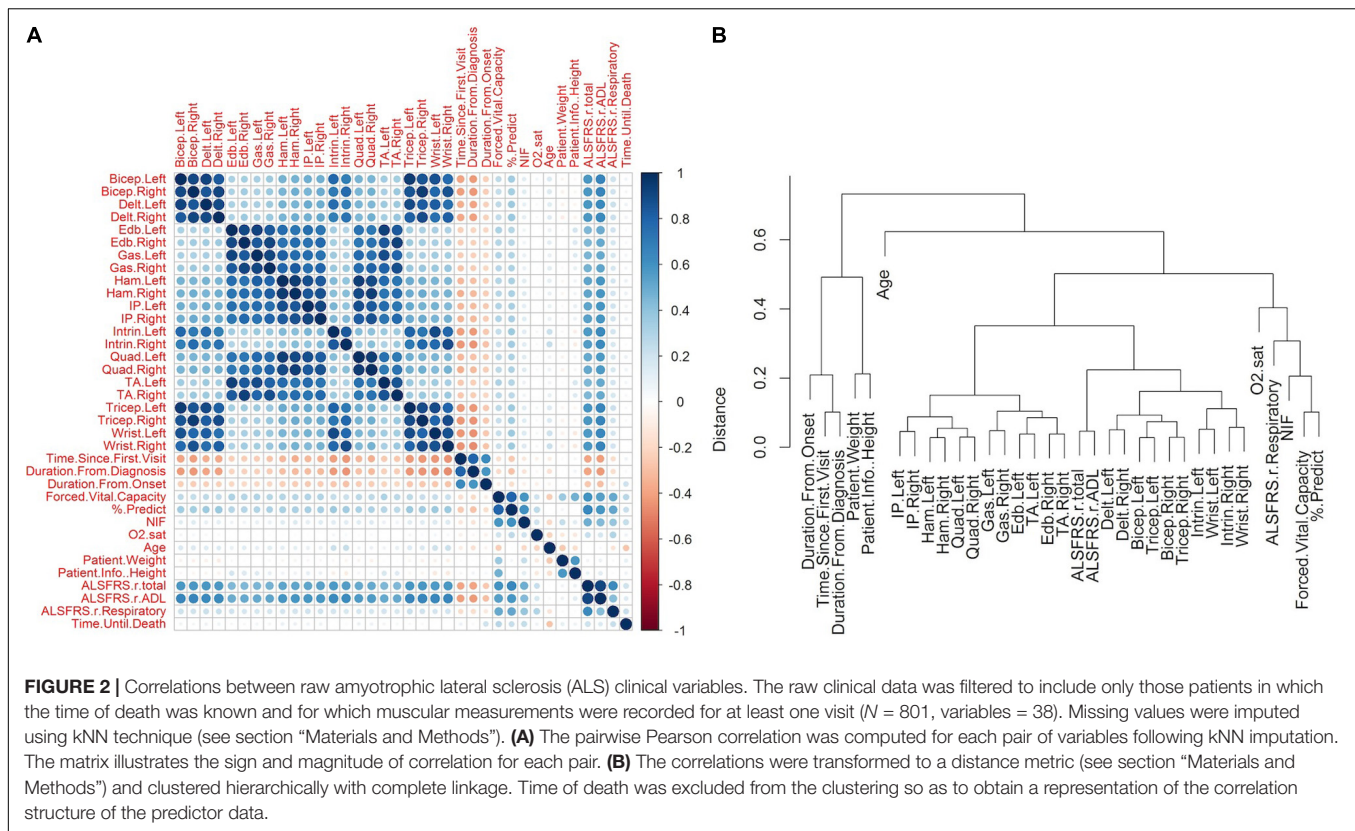
Exploratory Clinical Data Analysis

We begin with an exploratory analysis of the raw clinical data to examine relationships, correlations, complexity, and variance that can impart clinical and theoretical insight.

Correlation of Visit-Level Variables (Clinical Metrics)

The visualization of the correlation structure of the raw data reveals several high level relationships among the visit-level variables (Figures 2A,B). In particular, there is a high degree of correlated decline ($r \approx 0.7$ – 0.9) in the upper body muscles (bicep, tricep, deltoid, wrist, intrinsic), which occurs largely independently of an analogously correlated decline of the lower body muscles (hamstring, quadriceps, gastrocnemius, iliopsoas, extensor digitorum brevis, transverse abdominis), as is evident by the lesser correlation ($r \approx 0.3$ – 0.45) between the upper and lower scores.

The respiratory measures of FVC, % predicted of FVC, and NIF tend to correlate together ($r \approx 0.6$ – 0.8). The FVC and the % predicted of FVC show mild correlations with the muscle scores ($r \approx 0.25$ – 0.35), while the relationship between NIF and the



muscle scores is weaker ($r \approx 0.07$ – 0.15). The oxygen saturation shows a mild relationship with FVC, % predicted of FVC, and NIF ($r \approx 0.20$ – 0.25), but has very little relationship with muscular scores ($r \approx 0$ – 0.06).

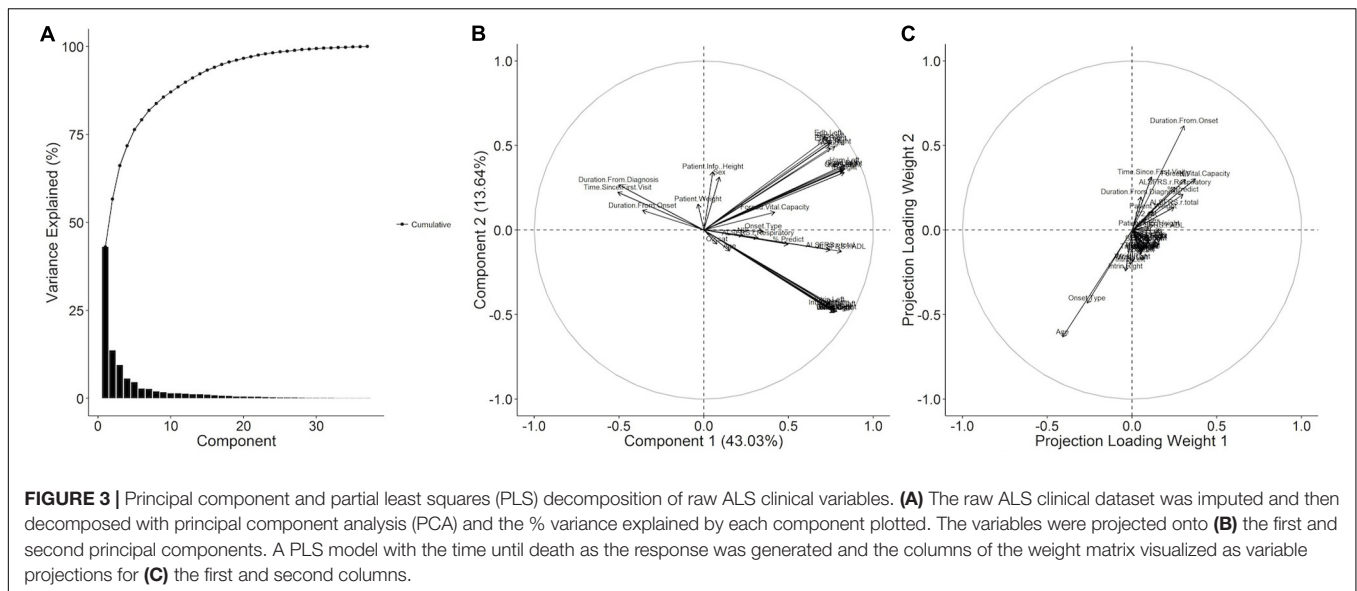
The ADL sub-score of ALSFRS-R moderately correlates with the muscular measures ($r \approx 0.5$ – 0.65), while the respiratory sub-scores show a weak relationship with the muscle scores ($r \approx 0.16$ – 0.22). Interestingly, the ADL and respiratory ALSFRS-R sub-scores correlate comparably with FVC, % predicted FVC, and NIF (ALSFRS-ADL: $r = 0.548$, 0.587 , 0.432 ; ALSFRS-Respiratory: $r = 0.504$, 0.519 , 0.386), with a slightly larger degree of correlation between the respiratory measures and ADL sub-score than with the respiratory measures and ALSFRS-R respiratory sub-score.

Negative correlations with the *disease start variables* (time since first visit, duration from diagnosis, duration from onset) implies that the variable in question declines after the start of disease. For each declining variable, the duration from onset shows a lesser magnitude of negative correlation than time since first visit (e.g., baseline). There is lesser correspondence between the visit-level variables and the time until death. The muscle scores, which showed a moderate degree of decline with respect to the disease start variables, have very little correlation with the time until death ($r \approx -0.03$ – 0.09). In contrast, the respiratory measures, which show a lesser degree of decline with respect to the disease start, have a somewhat stronger correlation with the time until death (FVC: $r = 0.28$; % predicted FVC: $r = 0.22$; NIF: $r = 0.24$). There are

mild correlations of the time until death with ALSFRS (total: $r = 0.21$; ADL: $r = 0.15$; respiratory: $r = 0.23$) and with age ($r = -0.27$). The lack of any large correspondence of any one variable with the time until death implies that a complex interaction of measures may be required for accurate survival prediction.

Exploratory Dimensionality Reduction of Raw Clinical Variables

Decomposition of the visit-level predictor variables (all visit-level variables with the exception of time until death) with PCA gives an alternate representation of the data in terms of a reduced number of orthogonal components. In this representation, the first and second principal components account for 43.03 and 13.64% of the variance with 91.06% of the variance accounted for by the first thirteen components (**Figure 3A**). The projections of the variables onto the first two components (**Figure 3B**) reveal that the majority of the variance in the predictor data is due to variance in the muscle scores and the ALSFRS-R total and its ADL sub-score. The third and fourth components (Supplementary Figure S1A) account for variance in the respiratory visit data and patient characteristics. An analogous partial-least decomposition of the variables was generated through a PLS model for time until death as the response variable. In this case, age, onset type, and the duration from onset have the most prominent contributions to the first four components (**Figure 3C** and Supplementary Figure S1B), implying that those variables may account for



much of the predictive signal in the context of the PLS model.

Survival Models to Explore Clinical Variables as Predictors

Explicitly, time until death from each visit was modeled as a continuous outcome with visit-level variables (Table 1) as predictors. Because the visits are treated independently, the resulting predictions are not attached to the known history of a given patient, making them suitable for exploratory population-level insight. The two primary predictive models, which included all 38 metrics shown in Table 1, included the GLM illustrated in Figure 4A as, “Lasso,” and the full random forest model, “Full-RF.” Additionally, dimensionality reduction techniques were used to construct reduced random forest models.

Survival Regression Model Performance

The full random forest regression model, “Full-RF” for time until death with all 38 visit-level variables (Table 1) attains a testing RMSE of 547.73 ± 45.94 days and a testing R^2 of 0.524 ± 0.065 over ten repetitions of 10-fold cross validation, greatly outperforming the GLM, “Lasso,” with a testing RMSE of 697.07 ± 47.10 days and a testing R^2 of 0.216 ± 0.037 (Figure 4A) for a mixing parameter $\alpha = 0.01$ and a regularization parameter $\lambda = 6.629$. The out-of-bag random forest variable importance for the full model in terms of the % increase in MSE (Figure 4B) indicates that age is the most important of the visit-level variables for the prediction of time until death. Following age, the duration from onset and diagnosis, patient height, % predicted of FVC, and NIF are the next most important variables. None of the muscle scores rank in the top ten in the variable importance list, although the ALSFRS total and ADL sub-score, which moderately correlate with muscle scores, are ranked at 10 and 11, respectively. The patient sex and onset type rank near the bottom of the variable importance list at rank 31 and 37, respectively.

Predictive Performance of Dimensionally Reduced Models

Examination of reduced models (models constructed using dimensionality reduction techniques, which utilize fewer variables to predict survival) provide additional insight into the full data structure and relationships among the underlying clinical metrics. The performance of the full random forest model, “Full-RF,” was analyzed under various dimensionality reduction settings to determine if a reduced representation can perform as well as the full model. The random forest models built off of principal component and PLS projections, “PCA-RF” and “PLS-RF,” fail to replicate the performance of the full random forest model (Figure 4A). The model built by stepwise inclusion of the variables by variable importance measures, “importance,” performs within one standard deviation of the full model with only four components (age, duration from onset, duration from diagnosis, patient height); applying supervised feature clustering, performance is consistent with the full model through 23 merging steps at which point there are fourteen variable clusters in the model as components, k (Figure 4A). Variable importance ranking for this reduced 14-component model (Figure 4C) compares favorably with the full model (Figure 4B).

Variable Combination Experiment Identifies Key Clinical Concepts

An experiment was performed to determine which concept-based groupings of clinical variables have the most impact in predicting survival: “duration”; “age”; “respiratory”; “muscle”; “ALSFRS”; “weight and height”; “onset type and sex.” See Section “Reduced Models Using Dimensionality Reduction” for concept cluster definitions. The evidence that the prediction of time until death is primarily driven by the age and disease duration variables is furthered by the results of the concept-based variable group combination experiment (Figure 5). In fact, a model consisting solely of duration and age variables outperforms all models in which only the duration or age is excluded. Beyond that inflection

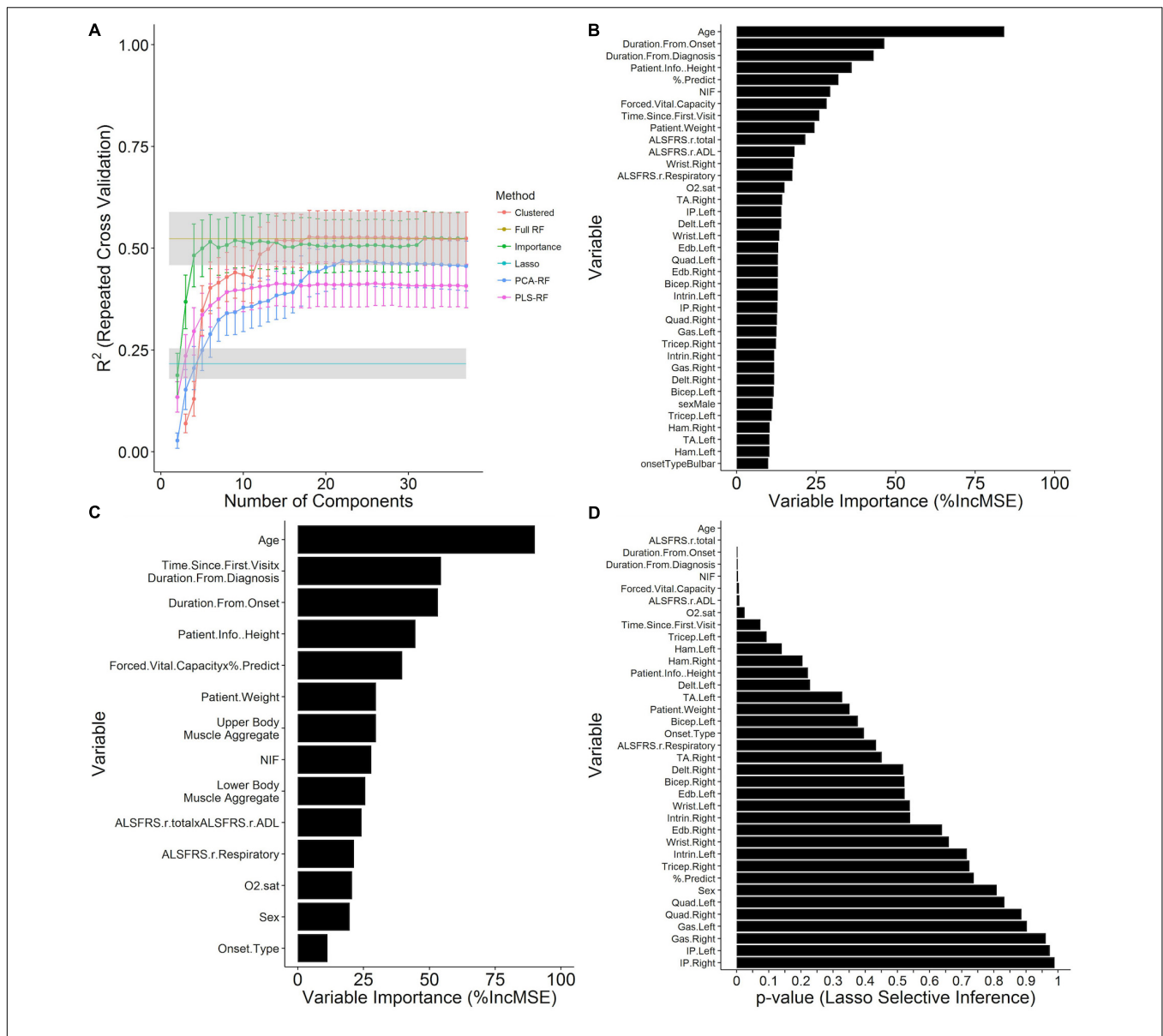


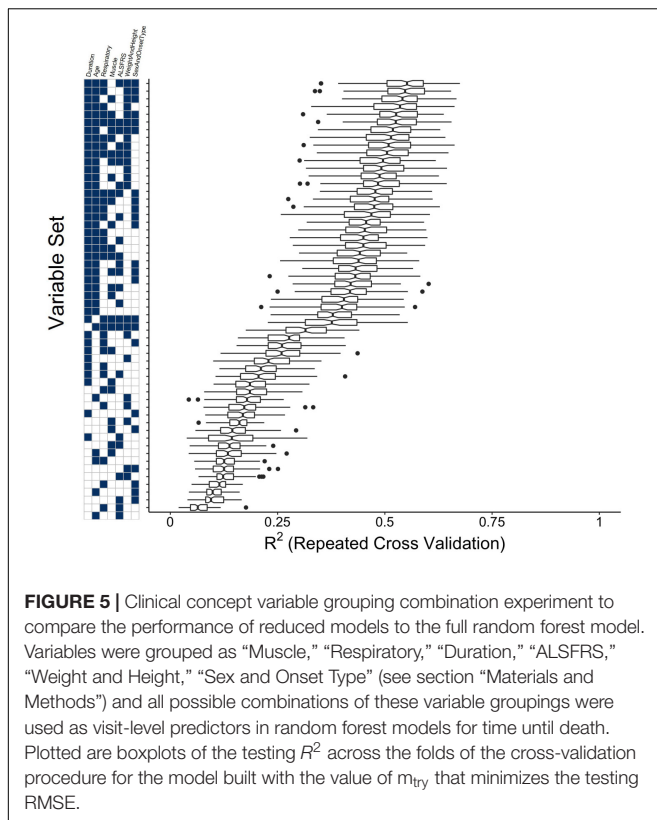
FIGURE 4 | Prediction of time until death using survival models constructed from independent visit records. Time until death from each visit was modeled as a continuous outcome with visit-level variables (Table 1) as predictors. Data shown are the mean \pm SD of the testing R^2 computed over ten repetitions of 10-fold cross validation with calculated missing values (imputed using kNN, see section “Materials and Methods”) embedded within the cross-validation procedure for (A) GLM and random forest models at the value of the tuning parameters that minimize the cross-validation RMSE and are plotted over the number of components in the model for several dimensionality reduction techniques (see section “Materials and Methods”). (B) The out-of-bag random forest variable importance results for the random forest with all predictors included in terms of the % increase in MSE of prediction upon permutation of the values of each variable. (C) Variable importance for a random forest model at the 23rd iteration of the variable clustering algorithm with 14 variable clusters. (D) Significance results for variable inclusion in the GLM, as computed by the selective inference method (see section “Materials and Methods”).

point in the trend, the relative gains in model performance upon inclusion of additional sets are marginal.

Model Comparisons Illustrate Importance of Complex Interactions

While the GLM model (“Lasso”) performance results do not compare favorably with the random forest (“Full-RF”) performance results, the p -values for variable inclusion (Figure 4D) obtained from the selective inference framework

(Tibshirani et al., 2014) are statistically interpretable. The results generally support the random forest variable importance, but a few specific departures, such as the insignificance of % predict FVC, likely owed to lasso-prone correlation bias (Tolosi and Lengauer, 2011), and the exaggerated significance of the ALSFRS-R total, collectively reveal the predictive importance of complex interactions and correlation structure accounted by the random forest. Thus, for the purpose of survival prediction, the random forest model is superior.



Survival Model Classification Task Performance

The random forest survival model regression task was reframed as multiple classification tasks where the target response is a binary indicator of whether the time of until death for the visit is less than the specified mortality threshold. Mortality thresholds of 30, 60, 90, and 180 days, and 1, 2, 3, 4, and 5 years were considered (Table 2). The AUC of the ROC estimate over the cross-validation procedure (Figure 6A) tends to increase as a function of the mortality threshold (30 days: 0.77 ± 0.0713 ; 1 year: 0.85 ± 0.0202 ; 5 years: 0.907 ± 0.028). The out-of-bag ROC estimate for the final model (Figure 6B) demonstrates the tradeoff between the true positive rate (sensitivity) and the false positive rate (1-specificity) for each mortality prediction threshold in general correspondence with the cross-validation ROC results. The analogous out-of-bag precision (positive predictive value) and recall (sensitivity) relationship (Supplementary Figure S2A) indicates severe tradeoffs for detection at 30, 60, and 90 days with roughly linear tradeoffs for thresholds at 1 and 2 years. Additionally, the optimal internal decision cutoff, the fraction of decision tree votes required for positive classification, as measured by the out-of-bag F_1 score (Supplementary Figure S2B), tends to increase as the mortality threshold increases due to the class imbalance at higher mortality thresholds.

The formulation of the model in terms of classification tasks across many time thresholds allows for the assessment of the variable importance for each of those thresholds to obtain a measure of the change in variable importance over the course

of disease. At 30 days (Figure 6C), the magnitudes of the variable importance measures are diminished with respect to those observed at 1 year (Figure 6D) or 5 years (Figure 6E) and are relatively uniform in value. At 30 days, the top ranking variables are primarily associated with the time since disease start, the ALSFRS-R total and ADL sub-score, and the FVC and % predicted of FVC. In contrast to other models, the age does not appear to be highly informative for classification at 30 days. At 1 year (Figure 6D), the age stands out as the most important variable, followed by the NIF, duration from diagnosis, % predicted of FVC, and the ALSFRS total. At 5 years, the influence of measured clinical variables seem to lose out in importance to time based variables, age, weight, and height. For this case, the FVC (rank 6) is most influential respiratory measure, followed by the % predicted of FVC (rank 8) and NIF (rank 9). In no case do any of the muscle scores seem to rank favorably in terms of variable importance and the onset type, and patient sex consistently rank as the least informative variables.

DISCUSSION

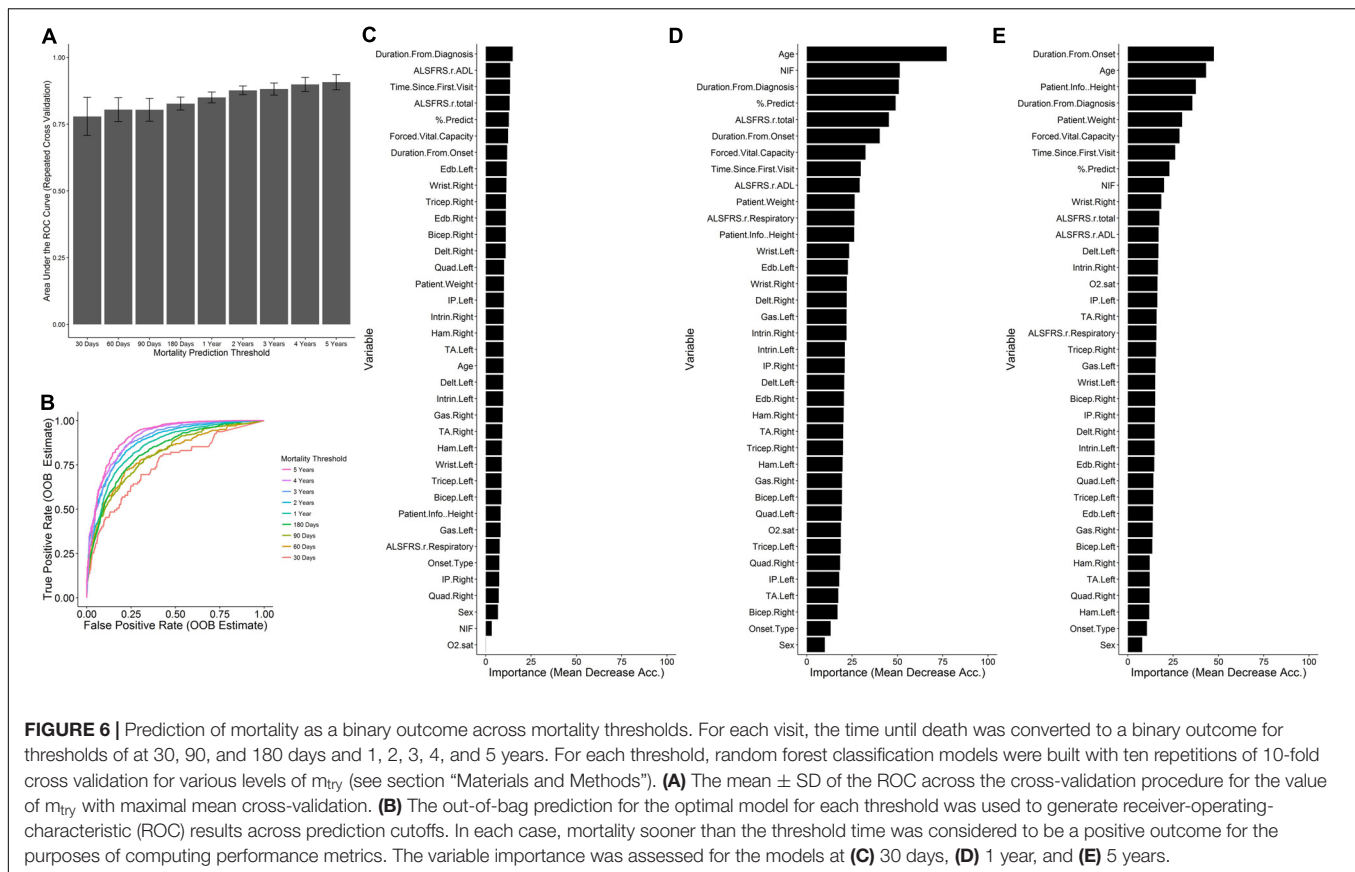
Our novel assessment of clinical ALS disease complexity using both raw clinical data analysis and interactive survival models identifies dynamic, changing correlations among ALS clinical metrics with disease progression. Knowing how the “best” predictors change with time or ALS disease progression is critical for developing future personalized survival models ready for real-time usage in the clinic or population models for clinical trial development. The relationships identified in the raw data and the predictive models illustrate key prognostic insights that can also lead to the development of better ALS assessment metrics.

Unlike most previous predictive computational models of ALS (Atassi et al., 2014; Küffner et al., 2015; Zach et al., 2015), we directly utilized survival and lifespan-based outcome metrics rather than specifying the ALSFRS-R slope as a measure of ALS progression *a priori*. Using survival as the predicted outcome removes any associated interpretative ambiguity. By utilizing survival as the outcome metric, we were able to evaluate the predictive value of the ardently contested ALSFRS-R alongside other in-clinic metrics, such as MRC muscle scores, clinical respiratory measures, and patient characteristics.

Below we summarize our major findings and their clinical significance. We close by discussing how future advances should focus on the identification of independent metrics of ALS progression that acknowledge its underlying dimensional complexity.

Time Variables Dominate ALS Survival Prediction

The time variables (patient age, time since onset, time since diagnosis, time since baseline visit, disease duration) are the dominant predictors, particularly for extended survival classification (beyond 1 year). Time variables consistently ranked highly in every model scenario, and



the clustered variable experiment (Figure 5) showed that they are the single most-important variable group needed to accurately predict survival; that is, time variables provide information that cannot be ascertained from other metrics. As shown in previous work (Fournier and Glass, 2015), time since baseline visit was the most consistent of the time predictors.

Dominant Clinical Predictors Change With Progression

The best predictors change over time and disease progression (Figure 6). For example, functional muscle degeneration (ADL sub-score of the ALSFRS-R, MRC muscle scores) correlates with early disease progression whereas quantifiable respiratory losses (namely % predict FVC but also NIF, O_2 sat, etc.) correlate better with survival. Height and weight do not factor in to survival classification early on, but are more influential for predicting survival beyond 2 years. Patient age is not important for survival classification at 30 days, but is important at 1 year and beyond. In summary, considering temporal changes in variable importance is key for developing accurate predictions at each disease stage or survival classification. Finally, changes of variable importance with time or disease progression could explain discourse in the field over the importance of specific predictors.

ALSFRS-R Is Not Necessarily the “Gold Standard”

Revised ALS Functional Rating Scale is typically quite significant in linear models that do not consider complex interactions (Figure 4D), which is how its impact has been previously assessed by those citing its veracity (Cedarbaum et al., 1999; Kimura et al., 2006; Kollewe et al., 2008). However, others (Voustianiouk et al., 2008; Franchignoni et al., 2013, 2015; Mandrioli et al., 2015) have acknowledged the concerns of ALSFRS-R validity as a primary outcome metric or as stand-in for survival.

The ALSFRS-R is not highly ranked in any of the random forest models with the exception of the random forest for survival classification at 30 days (Figure 6C). Our results reveal that the ALSFRS-total is more discriminatory for patients near death, and its clinical significance is very much tied to the slope of decline in individual patients versus a population-level effect. As such, ALSFRS-R should not be used as the primary or solo outcome metric for population-level models.

Notably, there are complex incongruences in the ALSFRS-total, and especially the ALSFRS-R respiratory sub-score, which is consistently lower ranking or insignificant (Figures 4, 6) despite clear, consistent significance in the % predict FVC. Our findings corroborate the inability of the ALSFRS-R total score (Franchignoni et al., 2013, 2015; Mandrioli et al., 2015), and even the ALSFRS-R respiratory sub-score

(Pinto and de Carvalho, 2015) to represent true progression due to multidimensionality (Franchignoni et al., 2013, 2015) and heterogeneity (Voustianiouk et al., 2008; Mandrioli et al., 2015). Re-evaluation of how sub-scores or ALSFRS-R questions are combined could increase the significance of this universal ALS clinical metric.

Respiratory, Upper and Lower Muscle Function Losses Are Independent

The raw clinical correlations and the results of all models clearly reveal that upper motor and lower motor losses occur independently, a result that had been theorized (Ravits et al., 2007) but not yet fully embraced by the field. Moreover, the decline in respiratory function occurs predominantly independently from the upper and lower functional muscle degeneration. As noted above, survival prediction is tied more strongly to respiratory metrics, especially % predict FVC, whereas early disease progression is tied more closely to functional muscle degeneration. Respiratory decline has received nearly unanimous support by the field as key survival predictor (Magnus et al., 2002; Fournier and Glass, 2015; Küffner et al., 2015; Mandrioli et al., 2015).

It was of particular interest to assess the predictive value of the MRC muscle scores, as they were not recorded in the PRO-ACT (Atassi et al., 2014) or prior analyses (Kent-Braun et al., 1998; Magnus et al., 2002). Our cumulative findings suggest that no single muscle or group or muscles is independently predictive of survival. However, the independent segregation of the upper and lower muscle scores, as well as the correlation of muscle degeneration to onset, means that individual or aggregate muscle scores could still be valuable for the development of early-stage ALS progression assessments.

Interestingly, a recent meta-analysis of SOD1-G93A ALS mouse data showed that while some treatments, like those targeting oxidative stress, could decrease functional decline by an average of 59.6%, only a 10% average difference in survival was noted; this also suggests that muscle function, while important to quality of life and a key target of ALS disease, is likely not the primary, direct contributor to survival (Bond et al., 2018).

Onset Type and Gender Are Not Independent Predictors of Survival

Qualitative observations suggest bulbar patients and/or males have shorter disease duration (Magnus et al., 2002; Testa et al., 2004; Mandrioli et al., 2006; Küffner et al., 2015). However, the raw data and all model scenarios reveal that onset type and gender are not independent predictors of survival. Like the DREAM models (Küffner et al., 2015), onset type and gender always rank near the bottom in variable importance. Onset type closely aligns with the information provided by age (shown in Figure 3); given bulbar onset ALS patients tend to be older, this finding is intuitive (Coan and Mitchell, 2015). Similarly, females tend to have a longer lifespan than males even in the general [non-ALS] population (Mitchell et al., 2015b); thus, studies citing gender differences (McCombe and Henderson, 2010) are likely tied to

either age and/or gender-biased treatment etiology (Pfohl et al., 2015), rather than independent gender effects.

Height and Weight Have Varying Significance

Height and weight have varying effects, which could explain contested literature (Paganoni et al., 2011; Reich-Slotky et al., 2013; Küffner et al., 2015). Height and weight are, unsurprisingly, intertwined and appear to be especially helpful for extended survival classification, such as at 5 years (Figure 6E). In the presented population-level survival prediction models, height appears to be important and thus, could either be a stand-alone population-level predictor (e.g., perhaps due to its co-correlation with motoneuron length) or, in contrast, height could simply be a stand-in proxy for weight. Weight, itself, has little predictive value in the illustrated analysis. However, in the present analysis, weight was based upon independent visit records without historical context from a specific patient (e.g., slope of weight decline). Other studies have stated that the slope of weight decline with ALS disease progression is likely more important (Küffner et al., 2015; Hollinger et al., 2016) in terms of predictive power. The impact of weight is complicated by the fact that obesity is tied to earlier ALS onset age (Hollinger et al., 2016).

Complexity Is the Key to Unlock ALS

A viewpoint change is needed to acknowledge, based on increasing evidence, that ALS is a multi-faceted, inherently complex and system-level disease that has no one single root cause (Coan and Mitchell, 2015; Irvin et al., 2015; Kim et al., 2015; Hollinger et al., 2016). A complex system cannot necessarily be represented by the sum of its individual parts. We illustrated that the ALSFRS-R score, itself, is a low-level example of a metric flawed by the multidimensional asymmetric sum of its parts (Franchignoni et al., 2013, 2015). The identification of a complete set of independent predictors that fully explain ALS “system” variance is paramount. Other variables shown to have quantifiable commonalities among ALS sub-populations, such as antecedent disease (Hollinger et al., 2016), genetic make-up (Hrastelj and Robertson, 2016), blood chemistry (Küffner et al., 2015), physical activity (Gallo et al., 2016), environmental factors (McCombe and Henderson, 2010), and the prognostic value of imaging as identified by recent studies (Baldaranov et al., 2017; Grolez et al., 2018; Menke et al., 2018) could account for remaining variance.

We are still far from developing a mechanistic representation of the complex system of multi-factorial pathophysiological processes (Kim et al., 2015) involved in the progression of ALS, although dynamics and regulatory instability (Irvin et al., 2015; Hollinger et al., 2016) have been hypothesized to be key to treatment prediction (Mitchell and Lee, 2012). A complete disease model is inherently hierarchical (Selinger et al., 2003) and typical observational clinical measures exist as emergent properties at the highest level of that hierarchy with no clear explanatory relationship to the complex underlying pathology. For the practical purposes

of personalized clinical prediction of ALS, full specification of all aspects of that complete hierarchical system may not be essential, as recently shown with a personalized prediction model (Westeneng et al., 2018). However, any model built solely off observational clinical measures may be fundamentally deficient in fully explaining the heterogeneity in ALS progression, regardless of modeling strategy. Future work must link lower-level mechanisms to higher-level emergent predictors. Such linking will require identification of common system-level relationships and instabilities (Irvin et al., 2015; Hollinger et al., 2016) that are present in the disease regardless of ALS genotype, phenotype, or initiating mechanism.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations and approval of the Institutional Review Board of Georgia Institute of Technology and Emory University. A waiver for consent was granted for this retrospective study, which utilized de-identified and anonymized data from deceased patients.

REFERENCES

- Atassi, N., Berry, J., Shui, A., Zach, N., Sherman, A., Sinani, E., et al. (2014). The PRO-ACT database: design, initial analyses, and predictive features. *Neurology* 83, 1719–1725. doi: 10.1212/WNL.0000000000000951
- Baldaranov, D., Khomenko, A., Kobar, I., Bogdahn, U., Gorges, M., Kassubek, J., et al. (2017). Longitudinal diffusion tensor imaging-based assessment of tract alterations: an application to amyotrophic lateral sclerosis. *Front. Hum. Neurosci.* 11:567. doi: 10.3389/fnhum.2017.00567
- Beghi, E., Chio, A., Couratier, P., Esteban, J., Hardiman, O., Logroscino, G., et al. (2011). The epidemiology and treatment of ALS: focus on the heterogeneity of the disease and critical appraisal of therapeutic trials. *Amyotroph. Lateral Scler.* 12, 1–10. doi: 10.3109/17482968.2010.502940
- Bond, L., Bernhardt, K., Madria, P., Sorrentino, K., Scelsi, H., and Mitchell, C. S. (2018). A metadata analysis of oxidative stress etiology in preclinical amyotrophic lateral sclerosis: benefits of antioxidant therapy. *Front. Neurosci.* 12:10. doi: 10.3389/fnins.2018.00010
- Castrillo-Viguera, C., Grasso, D. L., Simpson, E., Shefner, J., and Cudkowicz, M. E. (2010). Clinical significance in the change of decline in ALSFRS-R. *Amyotroph. Lateral Scler.* 11, 178–180. doi: 10.3109/17482960903093710
- Cedarbaum, J. M., and Stambler, N. (1997). Performance of the amyotrophic lateral sclerosis functional rating scale (ALSFERS) in multicenter clinical trials. *J. Neurol. Sci.* 152(Suppl. 1), S1–S9. doi: 10.1016/S0022-510X(97)00237-2
- Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., et al. (1999). The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J. Neurol. Sci.* 169, 13–21. doi: 10.1016/S0022-510X(99)00210-5
- Coan, G., and Mitchell, C. S. (2015). An assessment of possible neuropathology and clinical relationships in 46 sporadic amyotrophic lateral sclerosis patient autopsies. *Neurodegener. Dis.* 15, 301–312. doi: 10.1159/000433581
- Creemers, H., Grupstra, H., Nollet, F., Van Den Berg, L. H., and Beelen, A. (2015). Prognostic factors for the course of functional status of patients with ALS: a systematic review. *J. Neurol.* 262, 1407–1423. doi: 10.1007/s00415-014-7564-8
- Dayal, B. S., and MacGregor, J. F. (1997). Improved PLS algorithms. *J. Chemom.* 11, 73–85. doi: 10.1002/(SICI)1099-128X(199701)11:1<73::AID-CEM435>3.0.CO;2-#
- Fournier, C., and Glass, J. D. (2015). Modeling the course of amyotrophic lateral sclerosis. *Nat. Biotechnol.* 33, 45–47. doi: 10.1038/nbt.3118

AUTHOR CONTRIBUTIONS

SP contributed to study design, model construction, statistical analysis, results interpretation, drafted the initial manuscript, and review of content. RK assisted in data collection, analysis, results interpretation, and review of content. GC managed data quality control, assisted in results interpretation, and reviewed the content. CM contributed to project oversight, study conception, study design, data collection protocol, results interpretation, wrote the final manuscript, and review of content.

FUNDING

This study was funded by National Institutes of Health grants NS081426, NS069616, and NS098228 to CM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fninf.2018.00036/full#supplementary-material>

- Franchignoni, F., Mandrioli, J., Giordano, A., and Ferro, S. (2015). A further Rasch study confirms that ALSFRS-R does not conform to fundamental measurement requirements. *Amyotroph. Lateral Scler. Frontotemporal Degener.* doi: 10.3109/21678421.2015.1026829 [Epub ahead of print].
- Franchignoni, F., Mora, G., Giordano, A., Volanti, P., and Chio, A. (2013). Evidence of multidimensionality in the ALSFRS-R Scale: a critical appraisal on its measurement properties using Rasch analysis. *J. Neurol. Neurosurg. Psychiatry* 84, 1340–1345. doi: 10.1136/jnnp-2012-304701
- Friedman, J. H., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22. doi: 10.18637/jss.v033.i01
- Gallo, V., Vanacore, N., Bueno-De-Mesquita, H. B., Vermeulen, R., Brayne, C., Pearce, N., et al. (2016). Physical activity and risk of amyotrophic lateral sclerosis in a prospective cohort study. *Eur. J. Epidemiol.* 31, 255–266. doi: 10.1007/s10654-016-0119-9
- Grolez, G., Kyheng, M., Lopes, R., Moreau, C., Timmerman, K., Auger, F., et al. (2018). MRI of the cervical spinal cord predicts respiratory dysfunction in ALS. *Sci. Rep.* 8:1828. doi: 10.1038/s41598-018-19938-2
- Hollinger, S. K., Okosun, I. S., and Mitchell, C. S. (2016). Antecedent Disease and Amyotrophic Lateral Sclerosis: What Is Protecting Whom? *Front Neurol* 7:47. doi: 10.3389/fneur.2016.00047
- Hrastelj, J., and Robertson, N. P. (2016). Ice bucket challenge bears fruit for amyotrophic lateral sclerosis. *J. Neurol.* 263, 2355–2357. doi: 10.1007/s00415-016-8297-7
- Irvin, C. W., Kim, R. B., and Mitchell, C. S. (2015). Seeking homeostasis: temporal trends in respiration, oxidation, and calcium in SOD1 G93A amyotrophic lateral sclerosis mice. *Front. Cell. Neurosci.* 9:248. doi: 10.3389/fncel.2015.00248
- Ishwaran, H., Kogalur, U. B., Blackstone, E. H., and Lauer, M. S. (2008). Random survival forests. *Ann. Appl. Stat.* 2, 841–860. doi: 10.1214/08-AOAS169
- Kent-Braun, J. A., Walker, C. H., Weiner, M. W., and Miller, R. G. (1998). Functional significance of upper and lower motor neuron impairment in amyotrophic lateral sclerosis. *Muscle Nerve* 21, 762–768. doi: 10.1002/(SICI)1097-4598(199806)21:6<762::AID-MUS8>3.0.CO;2-5
- Kim, R. B., Irvin, C. W., Tilva, K. R., and Mitchell, C. S. (2015). State of the field: An informatics-based systematic review of the SOD1-G93A amyotrophic lateral sclerosis transgenic mouse model. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 17, 1–14. doi: 10.3109/21678421.2015.1047455

- Kimura, F., Fujimura, C., Ishida, S., Nakajima, H., Furutama, D., Uehara, H., et al. (2006). Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. *Neurology* 66, 265–267. doi: 10.1212/01.wnl.0000194316.91908.8a
- Kollewe, K., Mauss, U., Krampfl, K., Petri, S., Dengler, R., and Mohammadi, B. (2008). ALSFRS-R score and its ratio: a useful predictor for ALS-progression. *J. Neurol. Sci.* 275, 69–73. doi: 10.1016/j.jns.2008.07.016
- Küffner, R., Zach, N., Norel, R., Hawe, J., Schoenfeld, D., Wang, L., et al. (2015). Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. *Nat. Biotechnol.* 33, 51–57. doi: 10.1038/nbt.3051
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R News* 2, 18–22.
- Magnus, T., Beck, M., Giess, R., Puls, I., Naumann, M., and Toyka, K. V. (2002). Disease progression in amyotrophic lateral sclerosis: Predictors of survival. *Muscle Nerve* 25, 709–714. doi: 10.1002/mus.10090
- Mandrioli, J., Biguzzi, S., Guidi, C., Sette, E., Terlizzi, E., Ravasio, A., et al. (2015). Heterogeneity in ALSFRS-R decline and survival: a population-based study in Italy. *Neurol. Sci.* 36, 2243–2252. doi: 10.1007/s10072-015-2343-6
- Mandrioli, J., Faglioni, P., Nichelli, P., and Sola, P. (2006). Amyotrophic lateral sclerosis: prognostic indicators of survival. *Amyotroph. Lateral Scler.* 7, 211–220. doi: 10.1080/17482960600947648
- McCombe, P. A., and Henderson, R. D. (2010). Effects of gender in amyotrophic lateral sclerosis. *Genet. Med.* 7, 557–570. doi: 10.1016/j.genm.2010.11.010
- Menke, R. A. L., Proudfoot, M., Talbot, K., and Turner, M. R. (2018). The two-year progression of structural and functional cerebral MRI in amyotrophic lateral sclerosis. *Neuroimage Clin.* 17, 953–961. doi: 10.1016/j.nicl.2017.12.025
- Mitchell, C. S., Cates, A., Kim, R. B., and Hollinger, S. K. (2015a). Undergraduate biocuration: developing tomorrow's researchers while mining today's data. *J. Undergrad. Neurosci. Educ.* 14, A56–A65.
- Mitchell, C. S., Hollinger, S. K., Goswami, S. D., Polak, M. A., Lee, R. H., and Glass, J. D. (2015b). Antecedent disease is less prevalent in amyotrophic lateral sclerosis. *Neurodegener. Dis.* 15, 109–113. doi: 10.1159/000369812
- Mitchell, C. S., and Lee, R. H. (2012). *Dynamic Meta-Analysis as a Therapeutic Prediction Tool for Amyotrophic Lateral Sclerosis Amyotrophic Lateral Sclerosis*. Available at: <http://www.intechopen.com/books/amyotrophic-lateral-sclerosis/dynamic-meta-analysis-as-a-therapeutic-prediction-tool-for-amyotrophic-lateral-sclerosis>
- Paganoni, S., Deng, J., Jaffa, M., Cudkowicz, M. E., and Wills, A. M. (2011). Body mass index, not dyslipidemia, is an independent predictor of survival in amyotrophic lateral sclerosis. *Muscle Nerve* 44, 20–24. doi: 10.1002/mus.22114
- Pfohl, S. R., Halicek, M. T., and Mitchell, C. S. (2015). Characterization of the contribution of genetic background and gender to disease progression in the SOD1 G93A mouse model of amyotrophic lateral sclerosis: a meta-analysis. *J. Neuromuscul. Dis.* 2, 137–150. doi: 10.3233/JND-140068
- Pinto, S., and de Carvalho, M. (2015). The R of ALSFRS-R: does it really mirror functional respiratory involvement in amyotrophic lateral sclerosis? *Amyotroph. Lateral Scler. Frontotemporal Degener.* 16, 120–123. doi: 10.3109/21678421.2014.952641
- Ravits, J., Paul, P., and Jorg, C. (2007). Focality of upper and lower motor neuron degeneration at the clinical onset of ALS. *Neurology* 68, 1571–1575. doi: 10.1212/01.wnl.0000260965.20021.47
- Reich-Slotky, R., Andrews, J., Cheng, B., Buchsbaum, R., Levy, D., Kaufmann, P., et al. (2013). Body mass index (BMI) as predictor of ALSFRS-R score decline in ALS patients. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 14, 212–216. doi: 10.3109/21678421.2013.770028
- Rutkove, S. B. (2015). Clinical measures of disease progression in amyotrophic lateral sclerosis. *Neurotherapeutics* 12, 384–393. doi: 10.1007/s13311-014-0331-9
- Selinger, D. W., Wright, M. A., and Church, G. M. (2003). On the complete determination of biological systems. *Trends Biotechnol.* 21, 251–254. doi: 10.1016/S0167-7799(03)00113-6
- Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 9:307. doi: 10.1186/1471-2105-9-307
- Strobl, C., Boulesteix, A.-L., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8:25. doi: 10.1186/1471-2105-8-25
- Svetnik, V., Liaw, A., Tong, C., Culbertson, J. C., Sheridan, R. P., and Feuston, B. P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* 43, 1947–1958. doi: 10.1021/ci034160g
- Testa, D., Lovati, R., Ferrarini, M., Salmoiraghi, F., and Filippini, G. (2004). Survival of 793 patients with amyotrophic lateral sclerosis diagnosed over a 28-year period. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* 5, 208–212.
- Tibshirani, R. J., Taylor, J., Lockhart, R., and Tibshirani, R. (2014). *Exact Post-Selection Inference for Sequential Regression Procedures. ArXiv e-prints [Online]*. Available at: <http://adsabs.harvard.edu/abs/2014arXiv1401.3889T> [accessed January 1, 2014].
- Tolosi, L., and Lengauer, T. (2011). Classification with correlated features: unreliability of feature ranking and solutions. *Bioinformatics* 27, 1986–1994. doi: 10.1093/bioinformatics/btr300
- Voustaniouk, A., Seidel, G., Panchal, J., Sivak, M., Czaplinski, A., Yen, A., et al. (2008). ALSFRS and appel ALS scores: discordance with disease progression. *Muscle Nerve* 37, 668–672. doi: 10.1002/mus.20977
- Wei, T., and Simko, V. (2016). *corrplot: Visualization of a Correlation Matrix". 0.77*. Available at: <https://cran.r-project.org/package=corrplot>
- Westengen, H. J., Debray, T. P. A., Visser, A. E., van Eijk, R. P. A., Rooney, J. P. K., and Calvo, A. (2018). Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol.* 17, 423–433. doi: 10.1016/S1474-4422(18)30089-9
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer. doi: 10.1007/978-0-387-98141-3
- Wicks, P. (2014). The ALS ice bucket challenge - can a splash of water reinvigorate a field? *Amyotroph. Lateral Scler. Frontotemporal Degener.* 15, 479–480. doi: 10.3109/21678421.2014.984725
- Zach, N., Ennist, D. L., Taylor, A. A., Alon, H., Sherman, A., Kueffner, R., et al. (2015). Being PRO-ACTive: What can a clinical trial database reveal about ALS? *Neurotherapeutics* 12, 417–423. doi: 10.1007/s13311-015-0336-z

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pfohl, Kim, Coan and Mitchell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.