Title: Validation of Serum Neurofilaments as Prognostic & Potential Pharmacodynamic Biomarkers for ALS

Authors: Michael Benatar MD PhD [1], Lanyu Zhang MS [1], Lily Wang PhD [1], Volkan Granit MD [1], Jeffrey Statland MD [2], Richard Barohn MD [2], Andrea Swenson MD [3], John Ravits MD [4], Carlayne Jackson MD [5], Ted M Burns MD [6], Jaya Trivedi MD [7], Erik P Pioro MD PhD [8], James Caress MD [9], Jonathan Katz MD [10], Jacob L McCauley PhD [11], Rosa Rademakers PhD [12], Andrea Malaspina MD PhD [13], Lyle W Ostrow MD PhD [14], and Joanne Wuu ScM [1], on behalf of the CReATe Consortium

Author Information
[1] University of Miami, Miller School of Medicine, Miami, FL, United States
[2] Kansas University Medical Center, Kansas City, KS, United States
[3] University of Iowa, Iowa City, IA, United States
[4] University of California San Diego, San Diego, CA, United States
[5] University of Texas Health Science Center San Antonio, San Antonio, TX, United States
[6] University of Virginia, Charlottesville, VA, United States
[7] UT Southwestern Medical Center, Dallas, TX, United States
[8] Cleveland Clinic, Cleveland, OH, United States
[9] Wake Forest School of Medicine, Winston-Salem, NC, United States
[10] California Pacific Medical Center, San Francisco, CA, United States
[11] John P Hussman Institute for Human Genomics, Miami, FL, United States
[12] Mayo Clinic Jacksonville, Jacksonville, FL, United States
[13] Blizard Institute, Queen Mary University of London, London, United Kingdom
[14] Johns Hopkins University, Baltimore, MD, United States

Title character count: 95

Abstract word count: 233 **Word count**: 3,029 **Reference count**: 27

Tables: 3 **Figure**: 2

Supplemental Data: Table e1. Figure e1

Corresponding Authors:

Name: Michael Benatar, Joanne Wuu

Address: 1120 NW 14th Street, Suite 1300, University of Miami, Miami, FL, 33136

Email: mbenatar@med.miami.edu; jwuu@med.miami.edu

Statistical Analysis: Lily Wang PhD, Lanyu Zhang MS, Joanne Wuu ScM (all affiliated with the University of Miami)

Disclosure of financial relationships:

Michael Benatar reports grants from National Institutes of Health, the ALS Association, the Muscular Dystrophy Association, the Centers for Disease Control and Prevention, the Department of Defense, and Target ALS during the conduct of the study; personal fees from Mitsubishi Tanabe Pharma, AveXis, and Genentech, outside the submitted work. In addition, Dr. Benatar has a provisional patent entitled 'Determining Onset of Amyotrophic Lateral Sclerosis'. Dr. Benatar also serves as a site investigator on clinical trials funded by Biogen and Orphazyme.

Lanyu Zhang reports no disclosures.

Lily Wang reports grants from the National Institutes of Health during the conduct of the study.

Volkan Granit reports no disclosures.

Jeffrey Statland reports grants from National Institutes of Health, MDA, and the FSH Society during the conduct of the study; and personal fees from Fulcrum, Acceleron, Strongbridge, Sarepta, and AveXis, outside the submitted work.

Richard Barohn reports grants from the National Institutes of Health during the conduct of the study; grants from FDA/OOPD , the ALS Association, the Muscular Dystrophy Association, Sanofi/Genzyme, Biomarin Pharmaceuticals, IONIS Pharmaceuticals, Teva Pharmaceuticals, Cytokinetics Pharmaceuticals, Eli Lilly and Company , PTC Therapeutics, Orphazyme, Neuraltus Pharmaceuticals, Alexion Pharmaceuticals, Sarepta Therapeutics, and The Marigold Foundation outside the submitted work; personal fees from NuFactor, Momenta and Plan 365; and other support from Novartis Pharmaceuticals, Option Care, Platform Q Health Education, Ra Pharma. In addition, Dr. Barohn has patents on the MG-ADL and QMG with royalties paid.

Andrea Swenson reports grants from the University of Iowa during the conduct of the study and other support from Amylyx Pharmaceuticals, FlexPharm and Cytokinetics outside the submitted work.

John Ravits reports personal fees from MT Pharma and AveXis, outside the submitted work and serves as a site investigator on a clinical trial funded by Biogen.

Carlayne Jackson reports grants from the National Institutes of Health, Cytokinetics and Mitsubishi Tanabe Pharma America during the conduct of the study; and personal fees

from Mallinckrodt, Brainstorm, Mitsubishi Tanabe Pharma America, ITF Pharma, and Anelixis outside the submitted work.
Ted M Burns reports no disclosures.
Jaya Trivedi reports grants from National Institutes of Health, Sanofi-Genzyme and CSL Behring during the conduct of the study; as well as personal fees from CSL Behring outside the submitted work.
Erik P Pioro reports grants from National Institutes of Health and Center for Disease Control during the conduct of the study; grant support from the ALS Association and CDC/NIH outside the submitted work; as well as personal fees from Avanir Pharmaceuticals, Inc., Biohaven Pharmaceuticals, Inc., Cytokinetics, Inc., ITF Pharma, Inc., MT Pharma America, Inc., and Otsuka America, Inc.
James Caress reports grants from National Institutes of Health, Amylyx Pharmaceuticals Inc, Orion Corporation, Cytokinetics Inc, Flex-Pharma and Neurology Clinical Research Initiative TCZALS-001 outside the submitted work.
Jonathan Katz reports no disclosures.
Jacob L McCauley reports no disclosures.
Rosa Rademakers reports no disclosures.
Andrea Malaspina reports no disclosures.
Lyle W Ostrow reports grants from The Target ALS Foundation and the ALS Association during the conduct of the study.
Joanne Wuu reports grants and funding from National Institutes of Health, the ALS Association, the Centers for Disease Control and Prevention, the Department of Defense, and Target ALS during the conduct of the study.

Running Title: Validation of serum neurofilaments as ALS biomarkers

Search Terms: [177] Anterior nerve cell disease; [178] Amyotrophic lateral sclerosis; [21] Clinical trials Methodology/study design; Biomarkers

Abstract (233 words)

Objective. Identify preferred neurofilament assays, and clinically validate serum NfL and pNfH as prognostic and potential pharmacodynamic biomarkers relevant to ALS therapy development.

Methods. Prospective, multi-center, longitudinal observational study of patients with ALS (n=228), primary lateral sclerosis (PLS, n=20) and progressive muscular atrophy (PMA, n=11). Biological specimens were collected, processed and stored according to strict standard operating procedures (SOPs). Neurofilament assays were performed in a blinded manner by independent contract research organizations (CROs).

Results. For serum NfL and pNfH measured using the Simoa assay, missing data (i.e. both technical replicates below the lower limit of detection (LLD) was not encountered. For the Iron Horse and Euroimmun pNfH assays, such missingness was encountered in ~4% and ~10% of serum samples respectively. Mean coefficients of variation (CVs) for pNfH in serum and CSF were ~4-5% and ~2-3% respectively in all assays. Baseline NfL concentration, but not pNfH, predicted the future ALSFRS-R slope and survival. Incorporation of baseline serum NfL into mixed effects models of ALSFRS-R slopes yields an estimated sample size saving of ~8%. Depending on the method used to estimate effect size, use of serum NfL as a pharmacodynamic biomarker, instead of the ALSFRS-R slope, yields significantly larger sample size savings.

Conclusions. Serum NfL (Simoa assay) may be considered a clinically validated prognostic biomarker for ALS. Both serum NfL and pNfH (using the Simoa assay) have potential utility as pharmacodynamic biomarkers of treatment effect.

Introduction

Therapy development for amyotrophic lateral sclerosis (ALS) is challenging for many reasons, with the design and interpretation of phase II studies especially so [1]. These "mid-development" studies typically employ traditional clinical measures such as survival or the rate of decline of the Revised ALS Functional Rating Scale (ALSFRS-R) as the principal measures of therapeutic effect, but are typically underpowered, making it difficult to decide which experimental therapeutics to advance from phase II to phase III [2]. Biomarkers with prognostic and potential pharmacodynamic utility have great potential to help overcome these challenges [2, 3]. Controlling for prognostic biomarkers, for example, might reduce phenotypic heterogeneity, thereby improving statistical power for a fixed sample size. Similarly, pharmacodynamic biomarkers may help to verify target engagement or demonstrate the presence of the intended biological effect. Neurofilaments – both light (NfL) and phosphorylated heavy (pNfH), in blood and cerebrospinal fluid (CSF) – have been proposed as potential ALS biomarkers with diagnostic value[4-8], utility in predicting prognosis[5-7, 9-14], and as possible pharmacodynamic biomarkers[15, 16]. The potential diagnostic utility aside, which is not the focus of this paper, the precise role of neurofilaments as prognostic or pharmacodynamic biomarkers in ALS trials is not yet fully defined. Several questions remain unresolved: Which pNfH assay, among the many available, should be selected? Do NfL and pNfH convey the same information, or should both be measured? Moreover, understanding the prognostic value of neurofilaments has been hampered by the use of *estimated* rather than *measured* rates of disease progression through longitudinal follow-up; and inconsistent consideration of the value added by neurofilaments to clinical predictors of prognosis (age, gender, bulbar onset, etc.). Similarly, studies of the longitudinal trajectories of neurofilaments have reached inconsistent conclusions about the relationship with clinical measures of disease progression, driven in part by relatively small numbers of participants with fast progressing disease.

Here we present NfL and pNfH data (obtained from serum and CSF) from a large cohort of patients with ALS and related disorders who have undergone careful longitudinal

clinical phenotyping along with serial collection of biological samples. Serum and CSF neurofilaments have been quantified using a variety of assays, performed by CROs blinded to participant identity and status. Combined with carefully collected phenotypic data, we have endeavored to address unanswered questions related to the prognostic and potential pharmacodynamic utility of neurofilament levels.

Materials and Methods

Study Population

Patients with ALS and related disorders were enrolled at multiple centers, through the *Phenotype-Genotype-Biomarker* study of the Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium. Rigorously standardized clinical assessments and biological sample collections were performed every 3-6 months. Samples included in this experiment were from study visits that took place between April 2015 and November 2017. The University of Miami IRB (the central-IRB for CReATe) approved the study; all participants provided written informed consent.

Sample Collection, Processing and Storage

Blood was collected in serum-separating BD-vacutainers and allowed to clot upright at room temperature for 1-2 hours. Following centrifugation (1750g for 10 minutes at 4°C) serum was aliquoted into cryogenic sterile freestanding conical microtubes (Nalgene or Bio Plas Inc.) and stored at -80°C. CSF was collected in polypropylene tubes, centrifuged (1750g for 10 minutes at 4°C), aliquoted into polypropylene cryogenic sterile freestanding conical microtubes, frozen within ~30 minutes of collection, and stored at -80°C.

Neurofilament Quantification

Serum and CSF neurofilament concentrations were quantified by Quanterix using their Simoa NfL[17, 18] and pNfH assays[19]; each plate contained calibrators (0-500 pg/mL for NfL and 0-2000pg/ml for pNfH) and quality controls. Samples were diluted 1:4 to fall within the range of the standard curve. Serum and CSF pNfH were also quantified by Iron Horse Diagnostics, using their in-house assay (which uses monoclonal capture and

polyclonal detection antibodies)[10] and the Euroimmun CE marked ELISA[20] (which reverses the capture and detection antibodies). All samples were measured in duplicate at the same dilution, and all assays were performed blind to clinical parameters.

Statistical Analysis

Symptom onset was defined as the first occurrence of limb weakness, dysarthria, dysphagia, or dyspnea. ΔFRS was calculated by subtracting baseline ALSFRS-R from 48 (maximum ALSFRS-R score), divided by time (months) from symptom onset to baseline [21]. Survival duration was defined as time from symptom onset to PAV (≥22 hours/day non-invasive ventilation), tracheostomy or death.

Each sample was assayed twice to generate two replicates on each platform. We then computed (a) the number of samples for which neurofilament levels were below the LLD, (b) the mean difference and 95% confidence interval of the differences between replicates (i.e. limits of agreement[22]), and (c) the CVs between replicates. The association between serum and CSF concentrations was assessed using Spearman's rank correlation. These analytic characteristics were qualitatively compared between assay platforms and guided our selection of one of the pNfH assays for further investigation.

To evaluate the prognostic utility of serum NfL and pNfH, we fitted a linear model to estimate the change in ALSFRS-R for each subject over time, using only the subset of participants with at least three ALSFRS-R scores. The associations between ALSFRS-R slope estimates and pre-specified clinical and baseline neurofilament predictor variables were investigated using multivariable regression. To assess the joint contribution of NfL and pNfH, a likelihood ratio test was used to compare the full model (including clinical and neurofilament variables) to a reduced model (including only clinical variables). The prognostic value of serum neurofilaments was also evaluated by Kaplan-Meier (univariate) and Cox regression (multivariate) methods, using PAV- and tracheostomy-free survival as the outcome.

We then estimated trajectories of changes in neurofilament levels over time using a linear regression fitted separately to data from each subject. This model, which was limited to subjects with at least three available neurofilament values, included neurofilament as the outcome variable and time as predictor variable. Because neurofilament values were heavily skewed, natural logarithm-transformation was used to help achieve normality.

To assess the impact of considering baseline serum NfL on the sample size for a placebo-controlled clinical trial with ALSFRS-R slope as its primary outcome, we conducted a simulation similar to that described by Kuffner[23]. Briefly, we compared two mixed effects models, typically used for assessing treatment effects on ALSFRS-R slopes in clinical trials: Model-1 included ALSFRS-R scores as the outcome variable, time, time x group (treatment vs. controls), log(baseline NfL), time x log(baseline NfL) as fixed effects as well as random subject effects to account for correlations from repeated measures in the same subjects; Model-2 is the same as Model 1, but without variables involving NfL. 1000 simulations were performed. For each simulation, treatment group assignment was simulated at random, which assigned half of the patients with ≥3 ALSFRS-R scores (n=53) to the treatment group and the remaining to the control group. With s_1 and s_2 representing the estimated standard errors of time x group effects in Model-1 and Model-2, respectively, the percent reduction in sample size is proportional to the reduction in variance given by $100 * [1 - ({s_1}^2/{s_2}^2)]$.

We also explored the utility of using serum NfL as a biomarker to detect pharmacodynamic effect in a phase-2 clinical trial. Sample size estimations for such a trial required us to define a difference (or change) in serum neurofilament that might be regarded as clinically meaningful and which a future trial would need to be powered to detect. First, we considered the difference in serum NfL between fast and slow progressing patients (ALSFRS-R rate of decline >1 and <0.5 point/month, respectively), reasoning that this difference in neurofilament could be considered meaningful since it corresponds to a clinically important difference in rates of disease progression. We also considered an alternative approach, in which we defined a meaningful change based on

a slope that is more negative (i.e. steeper decline) than the lower bound of the 95% confidence interval of NfL trajectories in untreated patients. All analyses were performed using SAS (version 9.4) and R software (https://www.r-project.org/).

Results

Study Population

A total of 259 patients (228 ALS, 11 PMA, and 20 PLS), with a total of 614 person-visits, were included in this study (Table-1). For evaluation of the pharmacodynamic potential of neurofilament measurements, analysis was restricted to the subset of 106 ALS, 3 PMA, and 4 PLS patients with at least 3 longitudinal biofluid collections (Table e1).

Analytic Characteristics of Neurofilament Assays in Serum and CSF

In CSF, both replicates were very infrequently below assay LLDs (i.e. undetectable); this was true for both NfL and pNfH and in all assays. For serum, undetectable values were infrequent using the Simoa NfL and pNfH assays, but were encountered in ~4% and ~10% respectively of samples assayed using the Iron Horse and Euroimmun pNfH platforms (Table-2). Agreement between technical replicates was high for all assays except for CSF analysis using the Euroimmun assay (Table-2b). The mean CVs for NfL were ~3% in serum and CSF; and for pNfH were ~4-5% and ~2-3% for all assays in serum and CSF, respectively. We selected the Simoa platform for subsequent analyses, given that it had the fewest undetectable values, and good replicate agreement (low CV) in both serum and CSF.

Baseline Neurofilament Concentrations

Initial serum concentrations of NfL were higher in the ALS group compared to those with PMA and PLS ($p<0.001$; Table-1). In the ALS group, higher serum NfL was associated with older age, higher ΔFRS, female, the presence of a *C9ORF72* repeat expansion, and bulbar symptom onset (all $p<0.001$), but not with baseline ALSFRS-R. Higher CSF NfL levels were similarly associated with older age, higher ΔFRS, and female gender, although not reaching statistical significance (possibly due to smaller sample size). Higher baseline serum pNfH among ALS patients was associated with older age

(p=0.046), higher ∆FRS (p<0.001), female gender (p=0.002), and bulbar onset disease (p=0.035), but not *C9ORF72* repeat expansion status or baseline ALSFRS-R. Higher CSF pNfH levels were also associated with a higher ∆FRS, although not reaching statistical significance.

Correlation between Serum and CSF Neurofilament Concentrations

While CSF could more directly reflect CNS pathophysiology, serum is more easily accessible. Therefore, we examined for each assay how well the value of serum measurements could serve as a proxy for CSF measurements. For pNfH, the serum-CSF correlations were comparable among the 3 assays but weak (r=0.16 to 0.22); the serum-CSF correlation was, however, much stronger for NfL (r=0.62, Simoa platform only).

Potential Prognostic Utility of Baseline Neurofilament Concentrations

The average (±SD) rate of decline of the ALSFRS-R was 0.65 (±0.65) points per month. Multivariate regression analysis, with the ALSFRS-R slope as the outcome, considering potential clinical predictors of prognosis (age, gender, *C9ORF72* status, site of disease onset and ∆FRS) as well as baseline serum NfL and pNfH, identified only ∆FRS as a meaningful clinical predictor. For every 1-point increase in the ∆FRS (pre-slope), the ALSFRS-R rate of decline is worsened by an additional 0.46 points/month (p=0.004) (Table-3). The addition of baseline serum NfL to the model shows that this biomarker adds prognostic value that is not already explained by known clinical predictors. For every 1-point increase in log serum NfL level, the ALSFRS-R rate of decline is worsened by an additional 0.35 points/month (p<0.001). By contrast, baseline serum pNfH provided little additional prognostic value (p=0.11). When both NfL and pNfH were jointly considered, the results showed that, together, these biomarkers added significantly more prognostic values independent of clinical predictors (p=0.002, likelihood ratio test); this result, however, was primarily driven by the effect of NfL.

In univariate survival analyses, the presence of *C9ORF72* repeat expansion, higher ∆FRS at baseline, and higher baseline serum NfL levels, were each associated with

worse prognosis (shorter survival), but baseline serum pNfH levels were not (Figure-1). In a multivariate Cox proportional hazards model that considered age, gender, *C9ORF72* status, site of disease onset, baseline ALSFRS-R, ∆FRS, and baseline serum levels of NfL and pNfH, only age, baseline ALSFRS-R, ∆FRS, and serum NfL were significant predictors of survival (Table-3).

Potential Pharmacodynamic Utility of Longitudinal Neurofilament Concentrations

To ready the use of serum NfL and pNfH as potential pharmacodynamic biomarkers of treatment effect in future trials, we estimated their longitudinal trajectories (and variability of slopes). For serum NfL, the average slope was 0.011 log units/month (95% confidence intervals [CI]: -0.054, 0.076). For serum pNfH, the average slope was 0.006 log units/month (95% CI: -0.063, 0.084) (Figure-2). We also quantified the magnitude change in biomarker concentration on the raw scale compared to baseline (Figure-e1). For serum NfL, the 90th and 95th percentiles of the maximum absolute changes in concentration from baseline are 15 and 22pg/ml respectively. For serum pNfH, the 90th and 95th percentiles of the maximum absolute changes in concentration from baseline are 146 and 196pg/ml respectively.

Utility of Neurofilament Biomarkers in Reducing Sample Size for Future Clinical Trials

Since baseline NfL level is predictive of ALSFRS-R trajectories, we hypothesized that the number of patients needed for a trial would be reduced when baseline NfL levels, as a *prognostic* biomarker, are considered. This may be done by including serum NfL as a covariate in the linear mixed models typically used to examine ALSFRS-R slopes. Using a simulation study similar to that described by Kuffner[23], we found that the sample size for an ALS trial would be reduced by about 8.2% with the addition of baseline NfL measurements.

Furthermore, we estimated the sample size required for a future phase-2 trial with serum NfL as a *pharmacodynamic* biomarker. We considered 2 different outcome measures and what a clinically meaningful treatment difference may be for each: First, assuming that an experimental therapeutic would need to change the *slope* of serum

NfL from our observed average in the untreated state (0.011 log units/month) to less than the lower bound of the 95% confidence interval (-0.054 log units/month) and thereby an estimated treatment difference of 0.065 log units/month, with an estimated standard deviation of 0.048 log units/month, a total sample size of N=26 (13 treatment, 13 placebo) would provide 90% power to detect such a treatment difference, using a two-sample t-test with 5% significance level. Alternatively, one may wish to achieve a clinically meaningful difference in NfL *levels* (rather than trajectories). Based on an estimated treatment difference of 0.67 log(pg/ml)—which was our observed difference in baseline log(NfL) levels between fast progressors (ALSFRS-R decline of >1 point/month) and slow progressors (ALSFRS-R decline of <0.5 points/month), which were 3.10 and 2.43 log(pg/ml), respectively—and a standard deviation of ~0.81 log(pg/ml), a total sample size of N=64 (32 treatment, 32 placebo) would provide 90% power to detect such a treatment difference, using a two sample t-test with 5% significance level.

For comparison, we also similarly estimated the required sample size for a clinically meaningful 20-30% reduction in *ALSFRS-R slope*, a common outcome measure employed in ALS clinical trials [24]. Based on our observed average ALSFRS-R slope of -0.648 points/month and an estimated standard deviation of 0.65 points/month, a total of N=1054 or N=470 patients would be needed respectively to provide power for detecting a 20% or 30% reduction in ALSFRS-R slopes.

Discussion

We have undertaken a head-to-head comparison of three different pNfH assays (performed by two different CROs) in serum and CSF, as well as an evaluation of the prognostic and potential pharmacodynamic utility of serum NfL and pNfH. For pNfH, the Simoa assay had the best sensitivity (fewest values below the assay LLD) and reproducibility (lowest CV between technical replicates in both serum and CSF). Using the Simoa assay, we also found a much stronger correlation between serum and CSF for NfL than for pNfH. The correlations between serum and CSF in this study are substantially lower than prior published reports [11, 20, 25]. One possible explanation for this

discrepancy is that we restricted our analyses to baseline data. As discussed by Bland and Altman, when multiple measurements taken from the same subjects are used to compute correlation coefficients, the result can appear artificially high because variability between measurements on the same subject are not properly accounted for and the independence assumption of correlation is violated [22].

We also found that serum NfL, but not serum pNfH, is prognostic of future ALSFRS-R decline and survival duration, providing information that is not captured by readily available clinical predictors. While absolute values can vary between patients, serum NfL and pNfH levels remain largely stable in each subject over time. This may portend clinical utility as pharmacodynamic biomarkers if there are detectable changes in levels following exposure to an experimental therapeutic. This expectation is supported by recent studies of Nusinersen for spinal muscular atrophy (SMA) [26, 27], though definitive proof in adult motor neuron diseases would necessarily require an effective therapeutic to test. Moreover, we have illustrated the ways in which use of serum NfL in a phase-2 trial might permit a reduction in sample size. Incorporation of serum NfL as a covariate in a trial that uses the ALSFRS-R rate of decline as the primary outcome measure would yield a modest benefit (~8% reduction in sample size), but use of NfL as a pharmacodynamic biomarker might offer meaningful sample size savings compared to more traditional phase-2 studies in which changes in the ALSFRS-R are used as the primary outcome measure.

The main limitation of this study is that the population may not be adequately representative of ALS patients who would ordinarily be enrolled in a clinical trial. Patients included in the current study were skewed towards those with more slowly progressive disease (evidenced by an ALSFRS-R slope of -0.65 points/month). Achieving better representation of a trial-like population is the goal of an ongoing study. A second limitation is that we have insufficient data to reliably comment on the value of CSF NfL and pNfH as either prognostic or pharmacodynamic biomarkers. Notwithstanding these considerations, this was a large multi-center study with prospective, systematic follow-up and strict SOPs for sample collection, processing and

storage. All assays were performed by independent CROs blinded to clinical data, with comparative data generated across assay platforms, CROs and neurofilament types.

These results support two conclusions. First, serum NfL, but not serum pNfH, may be considered a clinically validated prognostic biomarker. Second, both serum NfL and pNfH may have value as potential pharmacodynamic biomarkers, and both should be incorporated into ongoing and future phase-2 trials; the actual pharmacodynamic utility of these biomarkers, however, will only become apparent once we have truly effective treatments for ALS.

Data Sharing Statement

De-identified data will be shared immediately following publication.

Appendix 1

Authors

Name	Location	Role	Contribution
Michael Benatar, MD, PhD	University of Miami	Author	Study conceptualization and design; data acquisition; study oversight; statistical analysis and results interpretation; drafting of the manuscript for intellectual content; critical review and revision of manuscript
Lanyu Zhang, MS	University of Miami	Author	Statistical analysis; review of manuscript for intellectual content
Lily Wang, PhD	University of Miami	Author	Statistical analysis; review of manuscript for intellectual content
Volkan Granit, MD	University of Miami	Author	Data acquisition; critical review of manuscript for intellectual content
Jeffrey Statland, MD	United States, Kansas University Medical Center	Author	Data acquisition; critical review of manuscript for intellectual content
Richard Barohn, MD	United States, Kansas University Medical Center	Author	Data acquisition; critical review of manuscript for intellectual content
Andrea Swenson, MD	University of Iowa	Author	Data acquisition; critical review of manuscript for intellectual content
John Ravits, MD	University of California San Diego	Author	Data acquisition; critical review of manuscript for intellectual content
Carlayne Jackson,	University of Texas	Author	Data acquisition; critical review of

MD	Health Sciences Center, San Antonio		manuscript for intellectual content
Ted M Burns, MD	University of Virginia	Author	Data acquisition; critical review of manuscript for intellectual content
Jaya Trivedi, MD	University of Texas Southwestern Medical Center	Author	Data acquisition; critical review of manuscript for intellectual content
Erik P Pioro, MD, PhD	Cleveland Clinic	Author	Data acquisition; critical review of manuscript for intellectual content
James Caress, MD	Wake Forest School of Medicine	Author	Data acquisition; critical review of manuscript for intellectual content
Jonathan Katz, MD	California Pacific Medical Center	Author	Data acquisition; critical review of manuscript for intellectual content
Jacob L McCauley, PhD	John P Hussman Institute for Human Genomics	Author	Data acquisition; critical review of manuscript for intellectual content
Rosa Rademakers, PhD	Mayo Clinic, Jacksonville	Author	Data acquisition; critical review of manuscript for intellectual content
Andrea Malaspina, MD, PhD	Blizard Institute, Queen Mary University of London	Author	Results interpretation; critical review of manuscript for intellectual content
Lyle W Ostrow, MD, PhD	Johns Hopkins University	Author	Study design; results interpretation; critical review of manuscript for intellectual content
Joanne Wuu, ScM	University of Miami,	Author	Study conceptualization and design; data acquisition; study oversight; statistical analysis and results interpretation; drafting of the manuscript for intellectual content; critical review and revision of manuscript

Appendix 2

Co-Investigators (CReATe Consortium)

Name	Location	Role	Contribution
Sumaira Hussain	University of Miami, USA	Project Manager	Regulatory lead; multi-center oversight and project management
Anne Cooley	University of Miami, USA	Associate Project & Data Manager	Multi-center oversight and project management; data query resolution
Yindi Li	University of Miami, USA	Statistical programmer	Data management
Marielle Wallace	University of Miami, USA	Data Manager	Data management
Julie Steele	University of Miami,	Outcomes	Training and certification of

	USA	Coordinator	site coordinators in outcome measures
Jessica P. Hernandez	University of Miami, USA	Research Coordinator	Visit coordination and data collection
Jessica Medina	University of Miami, USA	Research Coordinator	Visit coordination and data collection
Maria Elena Paredes	University of Miami, USA	Research Coordinator	Visit coordination and data collection
Ashley Manso	University of Miami, USA	Research Coordinator	Data collection and review
Natalia Ravelo	University of Miami, USA	Research Coordinator	Visit coordination and data collection
Wendy Levy	University of Miami, USA	Research Coordinator	Visit coordination and data collection
Patrice Whitehead	John P Hussman Institute for Human Genomics	Biorepository Technical Director	CReATe Biorepository laboratory oversight
Stephan Zuchner, MD, PhD	University of Miami, USA	Co-investigator	CReATe Consortium Genetic Working Group member
Mamatha Pasnoor, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Omar Jawdat, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Duaa Jabari, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Constantine Farmakidis, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Melanie Glenn, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Mazen M Dimachkie, MD	University of Kansas Medical Center, USA	Sub-investigator	Data collection
Laura Herbelin	University of Kansas Medical Center, USA	Site Lead Coordinator	Visit coordination and data collection
Hellen Tanui	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Sherri Anderson	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Michaela Walker	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Tina Liu	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Ayla McCally	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Andrew Heim	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Melissa Currence	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection
Yolanda Harness	University of Kansas Medical Center, USA	Research Coordinator	Visit coordination and data collection

Jeri Sieren	University of Iowa, USA	Research Coordinator	Visit coordination and data collection
Emilee Gibson	University of Iowa, USA	Research Coordinator	Visit coordination and data collection
Gil Gutierrez	University of California San Diego, USA	Research Coordinator	Visit coordination and data collection
Danielle Bussey	University of California San Diego, USA	Research Coordinator	Visit coordination and data collection
Rose Previte	University of California San Diego, USA	Research Coordinator	Visit coordination and data collection
Pamella Kittrell	University of Texas Health Science Center at San Antonio, USA	Research Coordinator	Visit coordination and data collection
Amruta Joshi	University of Virginia, USA	Research Coordinator	Visit coordination and data collection
Amy Conger	UT Southwestern Medical Center, USA	Research Coordinator	Visit coordination and data collection
Debbie Hastings	Cleveland Clinic, USA	Research Coordinator	Visit coordination and data collection
Irys Caristo	Cleveland Clinic, USA	Research Coordinator	Visit coordination and data collection
Mozhdeh Marandi	Wake Forest School of Medicine, USA	Research Coordinator	Visit coordination and data collection
Simon Carty	California Pacific Medical Center, USA	Research Coordinator	Visit coordination and data collection
J Paul Taylor, MD, PhD	St. Jude Children's Research Hospital, USA	Co-investigator	CReATe Consortium genomics lead
Gang Wu, PhD	St. Jude Children's Research Hospital, USA	Co-investigator	CReATe Consortium genomic data analysis and Genetic Working Group member
Evadnie Rampersaud, PhD	St. Jude Children's Research Hospital, USA	Co-investigator	CReATe Consortium genomic data analysis and Genetic Working Group member
Rebecca Schule, MD, PhD	University of Tubingen, Germany	Co-investigator	CReATe Consortium Genetic Working Group member
Marka van Blitterswijk, MD, PhD	Mayo Clinic Jacksonville, USA	Co-investigator	CReATe Consortium Genetic Working Group member

Acknowledgements

We would like to acknowledge Joaquin del Cueto (Project Manager for Research Support, University of Miami); Dr. Lucie Bruijn (ALS Association Chief Scientist); Drs. Amelie Gubitz and Robin Conwit (NINDS Program Directors); Dr. Tiina Urv (NCATS Program Officer); Dr. Manish Raisinghani (CEO, Target ALS Foundation); and the Rare Diseases Clinical Research Network (RDCRN) Data Management and Coordinating Center (DMCC) at University of South Florida (PI: Dr. Jeffrey Krischer; regulatory specialists: Julie Martin and Kaleena Dezsi; and project managers: Callyn Kirk, Devon Rizzo, and Kristina Bowles). We would also like to thank Drs. Zaven Kaprielian (Amgen), Danielle Graham and Toby Ferguson (Biogen), Joe Lewcock (Denali), Felix Yeh (Genentech), Sophie Batteur Parmentier (Merck), Bryan Hill (Mitsubishi Tanabe Pharma), Dave Frendewey (Regeneron), Aarti Sharma (Regeneron), James Dodge (Sanofi), Laure Rosen (Takeda), Ian Reynolds and Neta Zach (Teva), Kelly Dakin (FBRI), Nadine Tatton (AFTD), Amanda Haidet-Phillips (MDA), Robert Bowser (Barrow Neurological Institute) and Leonard Petrucelli (Mayo Clinic Jacksonville) for helpful discussions around selecting biomarker candidates and contract research organizations. We would also like to acknowledge the staff and support of the John P. Hussman Institute for Human Genomics Biorepository Facility at the University of Miami School of Medicine. Most importantly, we are indebted to patients who participated in the CReATe Phenotype-Genotype-Biomarker (PGB) study for their altruism as well as their commitment and contribution to advancing therapeutic development for ALS and related disorders.

Funding

The CReATe Consortium (U54 NS090291) is part of the Rare Diseases Clinical Research Network (RDCRN), an initiative of the Office of Rare Diseases Research (ORDR), National Center for Advancing Translational Sciences (NCATS). CReATe is funded through collaboration between NCATS and the National Institute of Neurological Disorders and Stroke (NINDS). CReATe is also supported by a Clinical Trial Readiness Grant from NIH (U01NS107027). Supplementary support for the CReATe Biorepository was provided by the ALS Association (Grant ID 16-TACL-242). Target ALS provided

support to Quanterix and Iron Horse to cover assay costs as well as through a grant to Joanne Wuu. This work was also supported by CTSA grants (UL1TR002366 and UL1TR000001) from NCATS and awarded to the University of Kansas.

References

1. Benatar M. ALS Therapy Development: Challenges and Opportunities. Journal of Neuromuscular Diseases 2016;3:S10.
2. van den Berg LH, Sorenson E, Gronseth G, et al. Revised Airlie House consensus guidelines for design and implementation of ALS clinical trials. Neurology 2019;92:e1610-e1623.
3. Benatar M, Boylan K, Jeromin A, et al. ALS biomarkers for therapy development: State of the field and future directions. Muscle & nerve 2016;53:169-182.
4. Verde F, Steinacker P, Weishaupt JH, et al. Neurofilament light chain in serum for the diagnosis of amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2019;90:157-164.
5. Rossi D, Volanti P, Brambilla L, Colletti T, Spataro R, La Bella V. CSF neurofilament proteins as diagnostic and prognostic biomarkers for amyotrophic lateral sclerosis. Journal of neurology 2018;265:510-521.
6. Gaiani A, Martinelli I, Bello L, et al. Diagnostic and Prognostic Biomarkers in Amyotrophic Lateral Sclerosis: Neurofilament Light Chain Levels in Definite Subtypes of Disease. JAMA neurology 2017;74:525-532.
7. Steinacker P, Huss A, Mayer B, et al. Diagnostic and prognostic significance of neurofilament light chain NF-L, but not progranulin and S100B, in the course of amyotrophic lateral sclerosis: Data from the German MND-net. Amyotroph Lateral Scler Frontotemporal Degener 2017;18:112-119.
8. Oeckl P, Jardel C, Salachas F, et al. Multicenter validation of CSF neurofilaments as diagnostic biomarkers for ALS. Amyotroph Lateral Scler Frontotemporal Degener 2016;17:404-413.
9. Boylan KB, Glass JD, Crook JE, et al. Phosphorylated neurofilament heavy subunit (pNF-H) in peripheral blood and CSF as a potential prognostic biomarker in amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2012.

10. Gendron TF, Group CONS, Daughrity LM, et al. Phosphorylated neurofilament heavy chain: A biomarker of survival for C9ORF72-associated amyotrophic lateral sclerosis. Annals of neurology 2017;82:139-146.
11. Lu CH, Macdonald-Wallis C, Gray E, et al. Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis. Neurology 2015;84:2247-2257.
12. Lu CH, Petzold A, Topping J, et al. Plasma neurofilament heavy chain levels and disease progression in amyotrophic lateral sclerosis: insights from a longitudinal study. J Neurol Neurosurg Psychiatry 2015;86:565-573.
13. McCombe PA, Pfluger C, Singh P, Lim CY, Airey C, Henderson RD. Serial measurements of phosphorylated neurofilament-heavy in the serum of subjects with amyotrophic lateral sclerosis. Journal of the neurological sciences 2015;353:122-129.
14. Poesen K, De Schaepdryver M, Stubendorff B, et al. Neurofilament markers for ALS correlate with extent of upper and lower motor neuron disease. Neurology 2017;88:2302-2309.
15. Boylan K, Yang C, Crook J, et al. Immunoreactivity of the phosphorylated axonal neurofilament H subunit (pNF-H) in blood of ALS model rodents and ALS patients: Evaluation of blood pNF-H as a potential ALS biomarker. J Neurochem 2009.
16. Lu CH, Petzold A, Kalmar B, Dick J, Malaspina A, Greensmith L. Plasma neurofilament heavy chain levels correlate to markers of late stage disease progression and treatment response in SOD1(G93A) mice that model ALS. PLoS One 2012;7:e40998.
17. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clinical chemistry and laboratory medicine 2016;54:1655-1661.
18. Quanterix. Simoa NF-light [online]. Available at: https://www.quanterix.com/sites/default/files/assays/Simoa_NF-light_Data_Sheet_HD-1.pdf.

19. Quanterix. Simoa pNF_heavy [online]. Available at: https://www.quanterix.com/sites/default/files/assays/Simoa%20pNF-heavy%20Discovery%20data%20sheet.pdf.
20. De Schaepdryver M, Jeromin A, Gille B, et al. Comparison of elevated phosphorylated neurofilament heavy chains in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2018;89:367-373.
21. Kimura F, Fujimura C, Ishida S, et al. Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. Neurology 2006;66:265-267.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-310.
23. Kuffner R, Zach N, Norel R, et al. Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. Nat Biotechnol 2015;33:51-57.
24. Castrillo-Viguera C, Grasso DL, Simpson E, Shefner J, Cudkowicz ME. Clinical significance in the change of decline in ALSFRS-R. Amyotroph Lateral Scler 2010;11:178-180.
25. Weydt P, Oeckl P, Huss A, et al. Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis. Annals of neurology 2016;79:152-158.
26. Winter B, Guenther R, Ludolph AC, Hermann A, Otto M, Wurster CD. Neurofilaments and tau in CSF in an infant with SMA type 1 treated with nusinersen. J Neurol Neurosurg Psychiatry 2019.
27. Darras BT, Crawford TO, Finkel RS, et al. Neurofilament as a potential biomarker for spinal muscular atrophy. Annals of clinical and translational neurology 2019;6:932-944.

Figure Legends

Figure 1. Kaplan-Meier survival curves, showing the prognostic value of (A) *C9ORF72* repeat expansion, (B) ∆FRS (dichotomized at the median, 0.62points/month), (C) baseline serum NfL (dichotomized at the median, 17pg/ml), and (D) baseline serum pNfH (dichotomized at the median, 67pg/ml). The presence of a *C9ORF72* repeat expansion, higher ∆FRS, higher baseline serum NfL, and higher baseline serum pNfH are shown in red.

Figure 2. Longitudinal trajectories of serum neurofilaments: (A) spaghetti plot of log-transformed neurofilament light level (NfL), (B) spaghetti plot of log-transformed phosphorylated neurofilament heavy level (pNfH), and (C) boxplot of the distribution of NfL and pNfH slopes. Slope estimates were obtained from a linear model with log-transformed neurofilament level as the outcome, and time as the independent variable.

Supplemental Figure 1. Change in serum neurofilaments over time, as compared to baseline level: (A) Change from baseline serum neurofilament light (pg/ml), and (B) change from baseline in serum phosphorylated neurofilament heavy (pg/ml).

Tables

Table 1. Study Participant Characteristics

		ALS (N=228)	**PMA** (N=11)	**PLS** (N=20)
Baseline age (years)	Mean ± SD	60.1 ± 11.8	56.8 ± 11.6	58.8 ± 11.1
Male	N (%)	133 (58%)	8 (73%)	9 (45%)
Genotype: *C9ORF72* Other Unknown	N (%)	 27 (12%) 10 (4%) 191 (84%)	 0 0 11 (100%)	 0 0 20 (100%)
Site of onset: Bulbar only Limb only Other region(s) only Multiple regions Unknown	N (%)	 41 (18%) 151 (66%) 4 (2%) 24 (11%) 8 (4%)	 0 11 (100%) 0 0 0	 6 (30%) 9 (45%) 0 4 (20%) 1 (5%)
Years from onset of weakness to baseline	Median (IQR)	1.8 (1.0 – 3.2)	2.4 (1.7 – 9.0)	7.9 (3.8 – 11.6)
Years from diagnosis to baseline	Median (IQR)	0.8 (0.3 – 1.7)	1.0 (0.6 – 4.3)	2.5 (1.5 – 5.5)
Baseline ALSFRS-R	Mean (± SD)	34.7 ± 7.5	33.5 ± 8.5	35.2 ± 8.1
Baseline ΔFRS	Mean (± SD)	0.62 ± 0.48	0.34 ± 0.36	0.17 ± 0.12
Baseline serum NfL (pg/ml) [a]	Median (range)	17 (2, 369) [b]	7 (3, 26)	10 (2, 54)
Baseline serum pNfH (pg/ml) [a]	Median (range)	67 (1, 976) [b]	27 (2, 101)	30 (1, 328)
Baseline CSF NfL (pg/ml) [a]	Median (range)	107 (12, 269) [c]	N/A	N/A
Baseline CSF pNfH (pg/ml) [a]	Median (range)	103 (16, 216) [c]	N/A	N/A
ALSFRS-R slope [d]	Mean (± SD)	-0.65 ± 0.65	-0.57 ± 0.14	-0.09 ± 0.46
Survival time (months)	Median (IQR)	25.9 (14.4 - 43.3)	24.6 (24.6 - 33.7)	Not estimable [e]

Baseline = first visit at which serum sample was available (with or without contemporaneous CSF collection).
IQR = inter-quartile range. N/A = not available.
[a] Neurofilament data generated through Simoa assays
[b] Baseline serum data not available for 3 ALS participants
[c] Baseline CSF data only available for 29 ALS participants
[d] Calculated from the 106 ALS, 3 PMA and 4 PLS participants in the longitudinal subset (i.e. participants with at least three neurofilament measurements)
[e] Median survival time could not be reliably estimated because only 2 PLS patients reached survival endpoint during study follow-up

Table 2a. Analytical Characteristics of Neurofilament Light: Simoa Assay

	Simoa	
	Serum	**CSF**
Total number of samples	614	78
Number of samples with value(s) below LLD:		
In both replicates	0	0
In only one replicate	4	1
Difference (in pg/ml) between replicates, mean (95% CI) [a]	-0.05 (-4.0, 3.9)	-0.03 (-11.8, 11.7)
Coefficient of variation (CV) [b], mean	3.2	2.9
Number (proportion) of samples with CV > 10	12 (2.0%)	1 (1.3%)

LLD = lower limit of detection
[a] Difference = replicate 2 – replicate 1. The 95% CI of this difference is also known as the limits of agreement.
[b] CV = 100 x (standard deviation of replicates 1 and 2) / (mean of replicates 1 and 2).

Table 2b. Analytical Characteristics of Phosphorylated Neurofilament Heavy (pNfH): Simoa Assay, Iron Horse Assay, and Euroimmun Assay

	Simoa		Iron Horse		Euroimmun	
	Serum	**CSF**	**Serum**	**CSF**	**Serum**	**CSF**
Total number of samples	614	78	614	78	614	78
Number of samples with value(s) below LLD:						
In both replicates	0	1	26	0	64	0
In only one replicate	15	2	36	0	73	0
Difference (in pg/ml) between replicates, mean (95% CI) [a]	-0.4 (-26, 25)	1.0 (-7, 9)	-0.8 (-16, 14)	-4.9 (-90, 80)	-0.02 (-38, 38)	67.9 (-163, 298)
Coefficient of variation (CV) [b], mean	4.7	2.1	4.0	2.7	5.3	3.0
Number (proportion) of samples with CV > 10	51 (8.3%)	0 (0.0%)	63 (10.3%)	0 (0.0%)	88 (14.3%)	0 (0.0%)

LLD = lower limit of detection
[a] Difference = replicate 2 – replicate 1. The 95% CI of this difference is also known as the limits of agreement.
[b] CV = 100 x (standard deviation of replicates 1 and 2) / (mean of replicates 1 and 2).

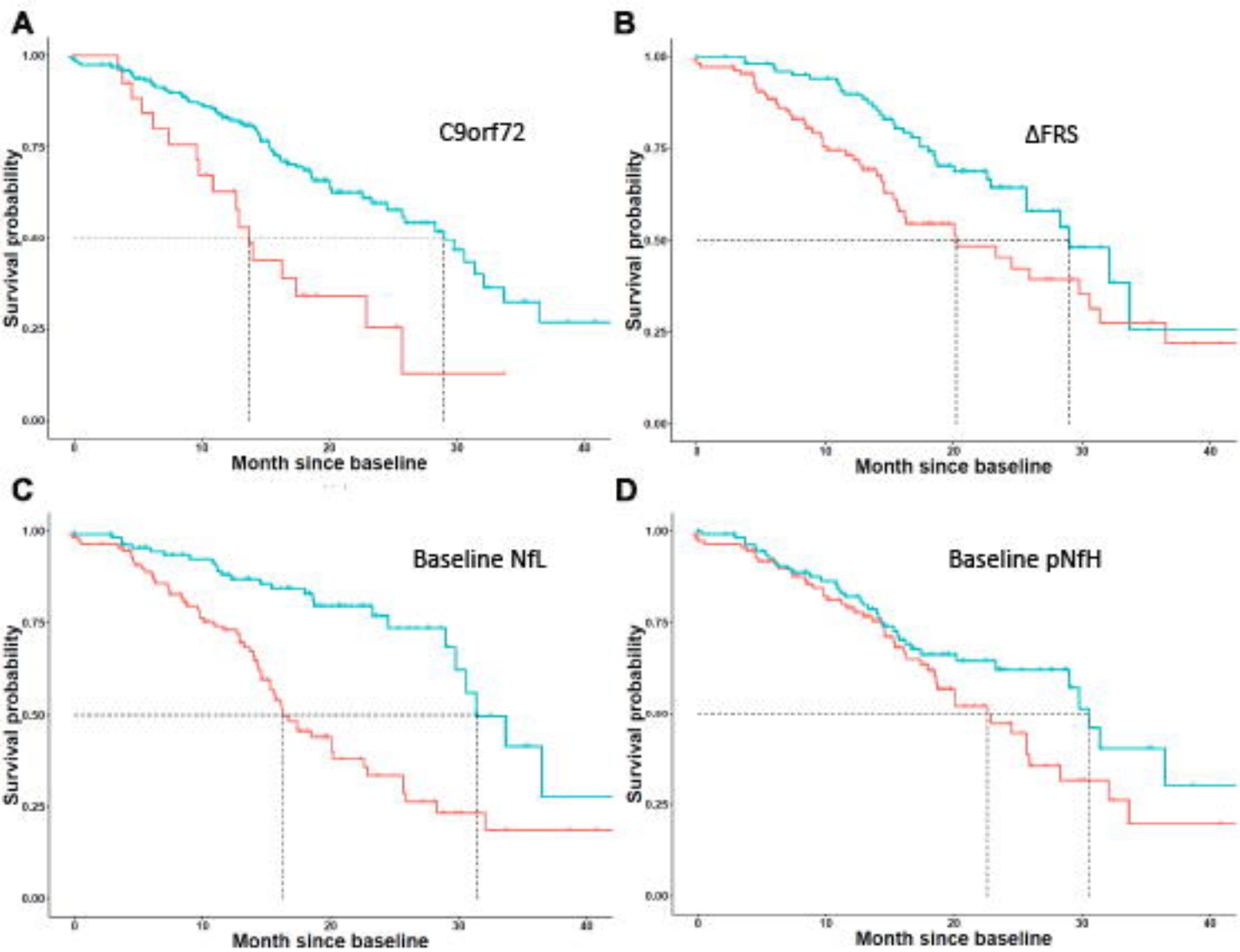
Table 3. Impact of Clinical and Neurofilament Characteristics on ALSFRS-R Rate of Decline and Survival

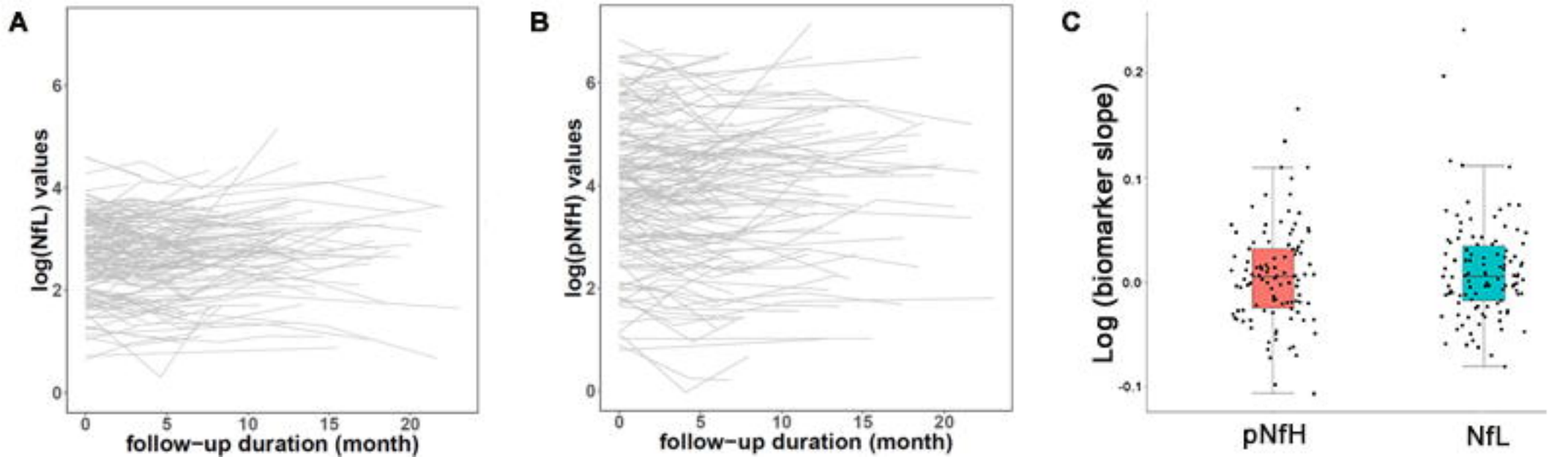
Covariates	Impact on ALSFRS-R Slope [a]		Impact on Survival [b]	
	Estimate (95% CI)	p-value	Hazard Ratio (95% CI)	p-value
Baseline age (years)	0.002 (-0.007, 0.012)	0.67	1.02 (1.00, 1.05)	**0.049**
Male	-0.09 (-0.33, 0.16)	0.49	1.37 (0.84, 2.23)	0.21
C9ORF72 HREM	-0.18 (-0.60, 0.24)	0.40	1.70 (0.91, 3.18)	0.10
Site of Onset: Bulbar	-0.01 (-0.43, 0.40)	0.94	1.10 (0.53, 2.29)	0.80
Site of Onset: Limb	0.03 (-0.44, 0.49)	0.92	1.04 (0.48, 2.27)	0.91
Baseline ALSFRS-R	N/A	N/A	0.96 (0.93, 0.99)	**0.02**
Baseline ΔFRS (points/month)	-0.46 (-0.76, -0.15)	**0.004**	1.67 (1.09, 2.57)	**0.02**
Baseline log(NfL) (log(pg/ml))	-0.35 (-0.55, -0.15)	**< 0.001**	2.12 (1.39, 3.23)	**< 0.001**
Baseline log(pNfH) (log(pg/ml))	0.08 (-0.02, 0.18)	0.11	1.01 (0.84, 1.22)	0.88

CI = confidence interval. N/A = not applicable.

[a] Based on multivariate regression analysis, with the ALSFRS-R slope (i.e. ALSFRS-R rate of decline) as outcome

[b] Based on Cox proportional hazards model, with tracheostomy- and permanent assisted ventilation (PAV)-free survival as outcome


A
C9orf72
Survival probability
Month since baseline
B
ΔFRS
Survival probability
Month since baseline
C
Baseline NfL
Survival probability
Month since baseline
D
Baseline pNfH
Survival probability
Month since baseline


A
log(NfL) values
follow-up duration (month)
B
log(pNfH) values
follow-up duration (month)
C
Log (biomarker slope)
pNfH
NfL