IsomiR Utility in ALS Prognostication

Cohen. Y.[1,2]*, Sinai. I.[1,2]*, Magen. I.[1,2*], Danino. M. Y.[1,2], Wuu. J.[3], Malaspina A.[4]., Benatar M.[3], Hornstein E.[1,2]

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel,
[2]Department of Molecular Neuroscience, Weizmann Institute of Science, Rehovot, Israel
[3]Department of Neurology and ALS Center, University of Miami Miller School of Medicine, Miami, FL, USA
[4]UCL Queen Square Motor Neuron Disease Center, UCL Queen Square Institute of Neurology, University College London, Queen Square, London, UK

* Authors contributed equally

Correspondence:
Andrea Malaspina (a.malaspina@ucl.ac.uk) and Michael Benatar (mbenatar@med.miami.edu) are corresponding authors for the ABC and PGB cohorts, respectively. Eran Hornstein (eran.hornstein@weizmann.ac.il) is the corresponding author for all other facets of the work.

Abstract

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease characterized by progressive motor neuron loss. IsomiRs are microRNA isoforms that arise from alternative processing or editing events during miRNA biogenesis. While isomiRs may carry distinct biological and clinical relevance, their potential as cell-free biomarkers in neurodegeneration remains largely unexplored. Intriguingly, loss of TAR DNA-binding protein 43 (TDP-43) nuclear function is a hallmark of disease and is known to impair isomiR expression. Here, we investigated the prognostic utility of plasma isomiRs in ALS, using next-generation sequencing. We profiled cell-free isomiRs in 154 ALS patients from a British cohort and identified higher levels of one isomiR, let-7g-5p.t, to be associated with longer survival. This finding was independently validated in an international ALS cohort of 200 patients and was in two orthogonal approaches. let-7g-5p.t prognostic utility was comparable to that of neurofilament light chain (NfL) or miR-181. These results establish isomiRs as a novel class of blood-based biomarkers in ALS with potential to refine prognostication in clinical trials for neurodegenerative diseases.

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of the human motor neuron system. Heterogeneity of patient progression rate[1] challenges the detection of therapeutic effects in clinical trials[2]. Biofluid-based biomarkers have garnered increasing attention as tools to monitor progression[3, 4]. These include urinary p75[5], troponin T2[6] and miRNAs[7]. Additionally, neurofilament light chain protein (NfL)[8] is a strong predictor of the future rate of disease progression, conveying meaningful prognostic information. Accordingly, we have recently demonstrated the utility of NfL in reducing the sample size required for clinical trials in a trial-like population[9].

microRNAs (miRNAs) are small endogenous RNAs that contribute to post-transcriptional silencing of mRNA[10]. miRNAs are critical for motor neuron viability [11] and their biogenesis is known to be dysregulated in ALS[12-15]. Previous studies have shown that cell-free miRNAs and in particular miR-181 might have prognostic value[7, 16]. miRNA biogenesis is a multistep process that starts with the transcription of a primary miRNA, which is cleaved by the nuclear microprocessor complex to yield an intermediate double-stranded hairpin precursor (pre-miRNA) [17]. After export of the pre-miRNA from nucleus to cytoplasm, it is cut by Dicer, creating a single strand 22nt guide that is loaded on to the RNA-induced silencing complex (RISC)[18]. IsomiRs display sequence variations that distinguish them from the reference sequence miRNAs (reference miRNA), including 3' or 5' sequence changes, or sequence polymorphism[19-21]. IsomiR production is often the result of aberrant miRNA biogenesis and isomiR levels in tissues or biofluids are often low and heterogeneous [22].

In ALS, cytoplasmic mislocalization of TAR DNA-binding protein 43 (TDP-43) and loss of its nuclear function, is reported in approximately 97% of patients [23] with cryptic exon splicing of dozens of mRNAs [24-27], some of which may have value as novel biomarkers relevant to ALS therapy development [28, 29]. Since loss of TDP-43 also leads to isomiR dysregulation [30], isomiRs may also be valuable as biomarkers. However, while the biomarker potential of cell-free isomiRs has been recently suggested for cancer and neurodevelopmental disorders [31-33], it was not yet tested in neurodegeneration.

Here, we test the potential value of isomiRs for ALS prognostication. We profiled cell-free isomiRs in a British cohort of 154 patients with ALS. Using Cox proportional hazard regression [34-36], and orthogonally machine learning (ML) survival prediction [36-40], we identified 18 isomiR predictors, which were successfully replicated in an independent international cohort of 200 patients. We further demonstrate that isomiR-let-7g-5p.t, particularly carries prognostic value with potential to complement established clinical features and previously described biomarkers, such as miR-181 and NfL. These findings establish isomiR as a novel family of biomarkers for neurodegeneration and highlight their potential as reliable predictors of ALS survival.

Methods

Clinical cohorts

Two ALS cohorts were included in this study. The ALS biomarker study (REC Reference: 09/H0703/27 IRAS ID: 47352) cohort which consists of 154 British patients (London, UK) matched in their inclusion criteria to a “trial-like” subset of the Phenotype Genotype Biomarker (PGB) study (registered at clinicaltrials.gov NCT02327845) consisting of 200 patients from 12 U.S. sites and one South African site, associated to the Clinical Research in ALS and Related Disorders for Therapeutic Development (CReATe) Consortium [9, 41]. Patients were diagnosed with ALS by experienced ALS neurologists. Inclusion criteria were as follows: time from symptom onset to first blood draw shorter than 36 months, a baseline slow vital capacity (SVC) equal to or above 50%, and at least two study visits with measurement of the revised amyotrophic lateral sclerosis functional rating scale (ALSFRS-R) [42]. Cognitive impairment (ALSci) or behavioral impairment (ALSbi) were determined by applying the revised Strong criteria [43] to Edinburgh Cognitive and Behavioral ALS Screen (ECAS [44]) data. Patients with either neuroinflammatory or neurodegenerative disorders, recent injuries, systemic or organ-specific autoimmune disorders, and recent treatment with steroids, immunosuppressants, or immunoglobulins were excluded from the British cohort. All participants provided written consent (or verbal permission for a carer to sign on their behalf) to be enrolled in the ALS biomarker study if they met inclusion criteria. Ethics approval was obtained from East London and the City Research Ethics Committee 1 or the University of Miami central IRB, which serves as the IRB of record for the CReATe Consortium (Internal Protocol ID #20160603). The University of Cape Town Health Sciences Human Research Ethics Committee approved the study in South Africa (REF number 165/2017). Patients were de-identified. Disease severity was assessed with the ALSFRS-R. The progression rate prior to enrollment was estimated using the DeltaFRS, calculated by subtracting the ALSFRS-R score at baseline assessment from a maximum score of 48, and dividing by the number of months passing from symptom onset to baseline assessment. Any use of riluzole was recorded. Index date (follow-up “time zero”) was defined as the date of enrollment and survival time was measured from the index date to events for either mortality, tracheostomy or permanent assisted ventilation (PAV). The censoring date for survival analysis was documented in a case where an event was not observed before the study ended. Experimenters were blinded during the steps of the molecular analysis. Whole blood was collected by venipuncture in EDTA tubes, and plasma was recovered by centrifugation at room temperature (RT) for 10 min at 1300g- 1750g. Plasma stored at < −70 °C prior to use.

Small RNA next-generation sequencing

The molecular experimenters were blinded to the identity of the samples. Total RNA was extracted from plasma using miRNeasy Micro Kit following the manufacturer's protocol except adding 2.5ml of Qiazol to each plasma sample (Qiagen, #217084). RNA was quantified with a Qubit fluorometer using the RNA Broad Range Assay Kit (Thermo Fisher Scientific, #Q10211). Small RNA next-generation sequencing libraries were prepared using the QIAseq miRNA Library Kit (Qiagen, #331502) from 7.5 ng of total RNA. If RNA concentration was below the limit of detection, 5µl of undiluted RNA were used. Precise linear quantification of miRNAs was achieved by 12-nucleotide long unique molecular identifiers (UMIs) [45]. Samples were randomly assigned to the preparation of libraries indexed with QIAseq miRNA NGS 48 Index IL (Qiagen, #331595) and sequenced in batches of 48 or fewer samples. Library concentration was determined with a Qubit fluorometer (dsDNA High Sensitivity Assay Kit, Thermo Fisher Scientific, #Q32854) and library size with TapeStation D1000 (Agilent). Libraries with different indices were multiplexed and sequenced on a NextSeq 500/550 v2 (Illumina, #20024906) or a NovaSeq 6000 flow cell (Illumina, #20028401), with a 75-bp single read or paired-end reads with 6-bp index. FASTQ files were de-multiplexed using the user-friendly transcriptome analysis pipeline [46]. Human miRNAs and isomiRs, as defined by miRBase - V22 [47], were mapped using CLC genomic workbench 22.0 (Qiagen) only to the positive strand. Reads were trimmed in two consecutive fashions; Phred score [48], based error ($p_{error} = 10^{-\frac{q}{10}}$) was calculated for every nucleotide, high values reflect poor read quality. A trimming score defined by the subtraction of the error from a threshold of 0.05 was set for each nucleotide. Then, a running sum of the nucleotide errors was calculated per read. The region between the first positive value of the running sum and the highest value of the running sum was kept with non-positive score reads being removed completely. Consecutively, Reads' ends were trimmed to keep the maximum length region containing 2 or fewer ambiguous nucleotides. This was followed by the library's 3'-adaptors (TGGAATTCTCGGGTGCCAAGG) trimming, with reads that did not contain any adaptors being excluded from downstream analyses. Reads were filtered based on lengths of 15-55 bps. IsomiRs were considered as 2bps changes (insertion, deletion, shift, and mismatch) up/downstream from MirBase reference sequence with a maximum of 2 mismatches allowed in alignment.

IsomiR definitions

Reference miRNA refers to the most abundantly expressed sequence variant among all isomiRs derived from a given miRNA gene. The collective set of all isomiRs—often comprising around 50 distinct variants—annotated for a specific miRNA gene is defined as the miRNA.

QC and pre-processing of sequencing data

Sequencing data were pre-processed separately for each cohort (British n=154 / PGB n=200) and yielded ~280K/~228K isomiRs respectively, but 99% of identified isomiRs were extremely rare and sparsely detected across samples. Therefore, isomiRs were included in the analysis only if their average expression level across cohort's samples was above 5 unique molecular identifiers (UMIs) yielding a total of 2046 and 2003 isomiRs in the British and PGB cohorts respectively. As geometric mean cannot be calculated with missing values, the remaining isomiR missing reads were imputed using multiple imputation with a denoising autoencoder algorithm (parameters: 40 forward passes, three layers with 512 nodes each, probability of corruption of 0.95, learning rate of $1e^{-6}$). Missing values were replaced by the mean isomiR levels across 50 imputed datasets. Clinical features used for estimation of similarity between patients: sex, site of symptom onset, and Riluzole treatment. Python-V3.9.2 and R-V4.0.4 package *rMIDAS*-V.0.3 [49] were used for the imputation. Using *DESeq2 [50]*, libraries were corrected for size by geometric mean, under the assumption that miRNA counts followed negative binomial distribution [50] Then, imputed values of original missing values were transformed back to zero to avoid bias in batch correction. Finally, data were corrected for the library preparation batch to reduce its potential bias by the R Combat function in SVA-V3.38 package [51]. Then, levels of 1480 isomiR identities shared across the two cohorts were standardized and isomiRs were considered for downstream analyses.

Machine learning-based isomiR selection for disease prognostication

Biomarker discovery analysis (training) was performed using the British cohort, that we have recently characterized [16] while the PGB cohort was reserved for external validation. The 1,480 isomiR candidates were randomly sampled with repetitions to generate 3,718 bootstrap subsets, each containing 20 isomiRs. The sampling process ensured that each isomiR was included in at least 50 distinct subsets. Importantly, any given subset contained one specific isomiR of interest alongside 19 other randomly selected isomiRs, resulting in at least 50 different subset compositions for each individual isomiR. For each bootstrap set, survival analysis was conducted using three survival models: Cox regression analysis [34], random survival forest (RSF) [36], and gradient survival

boosting (XGB) [35]. The analyses were performed using the scikit-survival library (version 0.23.0) in Python. Sequential forward selection (SFS) [52], adds iteratively isomiRs in different combinations to optimize the concordance index [53] score (C-index,mlxtend library in Python).

The process explored approximately 1,000,000 possible combinations of isomiRs for each subset $(\sum_{k=0}^{20} \binom{20}{k} = 2^{20} = 1,048,576)$, resulting in roughly 3.9 billion combinations across all 3,718 isomiRs subsets. Each isomiR combination was evaluated using 10-fold cross-validation (CV) on subsets of the UK (discovery) cohort within each of the three orthogonal survival models. A selection threshold of ≥90% in the SFS process yielded 84 isomiRs. Subsequently, additional iterations of feature selection were performed on the 84 isomiRs, generating 222 random subsets for further analysis. A C-index-based threshold, was determined using the elbow method [54], and applied across all random subsets. This process identified two final panels, together consisting of 18 unique isomiRs, which were selected for downstream analyses.

isomiRs panel prediction validation

After feature selection, an extreme gradient boosting (XGB) model was trained on the UK discovery cohort to predict patient survival time from enrollment. Hyperparameter tuning was conducted using 5-fold cross-validation over three rounds of Optuna optimization [55] (1000 trials per round, with a random sampler and seed = 42), followed by manual fine-tuning. The optimized model was trained on the selected isomiR levels from the UK cohort and validated using the corresponding isomiR levels from the PGB replication cohort. Shapley Additive Explanations (SHAP [56]) were used to identify feature importance within the model and were calculated on the individual predicted hazard function. Hazard ratios based on SHAP values were calculated as previously described [57]. Individual SHAP values were exponentiated across features. Values were averaged separately for all patients below feature mean level and those above it. The average value of above mean levels was divided by the average value of below mean level to yield the hazard ratio. The performance of the chosen isomiR-based model was compared against alternative models trained on: reference miRNA levels only, miRNAs levels (defined as the aggregate of all isomiRs within a miRNA annotation).

NfL assay

NfL levels were quantified using single molecule array (Simoa) technology on the Simoa HD-1 Analyzer (Quanterix) in plasma or serum from 152/200 UK discovery or PGB replication cohorts respectively. The assay utilized the Simoa Nf-L Advantage Kit (Quanterix, #102258) with standards,

primary and secondary antibodies, and detection ranges specified by the manufacturer, including defined lower and upper detection limits. Univariate or multivariate Cox regression with or without clinical features and associations between these and mortality were analyzed as described below.

Feature engineering

For univariate/multivariate Cox regressions isomiR levels and NfL levels were dichotomized into binary variables based on the median levels in the UK discovery cohort, whereas miR-181 levels were dichotomized using a previously published threshold [16]. Combinations of isomiR-let-7g-5p.t with NfL or miR-181 were evaluated by their ratios followed by a dichotomization based on the median levels in the UK discovery cohort for NfL combination or by an in-house algorithm that evaluated dichotomization cutpoint by optimizing univariate Cox mortality hazard ratio on the UK discovery cohort.

Statistical analysis

The Student's t-test was applied to assess significant differences in the age at symptom onset or study enrollment between cohorts. The Mann–Whitney U test was used to evaluate differences in ALSFRS-R and DeltaFRS scores between the cohorts. Differences in the proportions of censored events or sex prevalence were tested using chi-squared goodness-of-fit tests. Pearson correlation coefficients were calculated to examine the relationships between isomiR levels within/between the two cohorts. Kaplan–Meier survival estimators were analyzed using the log-rank Mantel–Cox test, with survival data censored at a fixed date. Univariate and multivariate Cox proportional hazard models were employed to estimate mortality hazard ratios from the time of enrollment, and model fit was assessed using the C-index. The performance of the XGB model was evaluated using the dynamic time-dependent area under the curve [58] (AUC(t)) across the range of survival periods, the average AUC(t), and the C-index. All statistical analyses were conducted, and visualizations were generated using Python (version 3.9.2), R (version 4.0.4) and BioRender.

Results

Study population and baseline characteristics

IsomiRs are sequence isoforms of miRNAs that arise from differences in processing, editing, or modifications of miRNA precursors (Figure 1A). In this study we use the term “reference miRNA” to refer to the most abundantly expressed (predominant) sequence variant among all isomiRs derived from a given miRNA gene. Noteworthy, the literature refers to a “miRNA” as the collective set of all dozens of annotated isomiRs [19].

Our study explores isomiRs as ALS prognostic biomarkers in two independent cohorts of patients with ALS. Clinical data and blood samples were obtained for 154 patients from the ALS biomarker cohort (UK) [59] and for 200 patients from the phenotype-genotype biomarker (PGB) study [41]. Clinical and demographic cohort characteristics are summarized in **Table 1.** Discovery and replication cohorts are clinically comparable. DeltaFRS (Mann-Whitney P-value = 0.18, **Figure S1A**), baseline ALSFRS-R score (Mann-Whitney P-value = 0.24, **Figure S1B**), site of symptom onset (χ^2 goodness of fit P-value = 0.08, **Figure S1C**) and Riluzole treatment (χ^2 goodness of fit P-value = 0.34, **Figure S1D**). Imbalances in patients’ sex distribution (χ^2 goodness of fit P-value = 0.009, **Figure S1E**) and age were noticed (t-test P-value ≤ 0.001, **Figure S1F**). The number of censored events were also statistically different between the cohorts (9% vs 54% in UK and PGB, respectively, χ^2 goodness of fit P-value ≤ 0.001, **Figure S1G**).

Sequencing and data preprocessing were conducted in parallel across 154 and 200 patient samples from the UK and PGB cohorts, respectively. Following noise filtration, 1,480 isomiRs from 287 miRNA genes were co-identified in both the discovery and replication cohorts. The discovery isomiRs study was performed on the British cohort. The PGB was held out as an unseen independent replication cohort.

	UK – discovery (N = 154)	PGB – Replication (N = 200)	Overall (N = 354)
Sex:			
Males	91 (59%)	90 (45%)	181 (51%)
Females	63 (41%)	110 (55%)	173 (49%)
Age at symptom onset (Years) (Mean ± SD)	62.3 (± 11)	57 (± 12)	59.3 (± 12)
Age at enrollment (Years) (Mean ± SD)	63.9 (± 11)	58.4 (± 12)	60.8 (± 12)
Disease Duration at enrollment (Months) (Mean ± SD)	19.3 (± 10)	16.5 (± 8)	17.7 (± 9)
Baseline ALSFRS-R (AU) (Median [Q1, Q3])	38 [33, 43]	39 [35, 41]	39 [34, 42]
DeltaFRS (-ALSFRS-R unit / month) (Median [Q1, Q3])	0.6 [0.3, 0.9]	0.6 [0.4, 0.9]	0.6 [0.3, 0.9]
Site of symptom onset			
Bulbar	52 (34%)	55 (28%)	107 (30%)
Non-bulbar	102 (66%)	145 (72%)	247 (70%)
Riluzole treatment			
Yes	109 (71%)	134 (67%)	243 (69%)
No	45 (29%)	66 (33%)	111 (31%)
Survival from symptom onset (Months) (Median [Q1, Q3])	33 [26.7, 46.7]	33.4 [22.5, 43.9]	33 [23.7, 44.5]
Survival from enrollment (Months) (Median [Q1, Q3])	13.9 [8.9, 23.9]	14.7 [8.5, 24.1]	14.3 [8.7, 24.1]
Censored event			
Yes	9 (6%)	108 (54%)	117 (33%)
No	145 (94%)	92 (46%)	237 (67%)

Table 1. Clinical and demographic description of study cohorts. Abbreviations: Baseline ALSFRS-R - ALSFRS-R score at baseline assessment. DeltaFRS – change of ALSFRS-R score from symptom onset to baseline assessment divided by months passed. Age at enrollment – age at baseline assessment. Survival from symptom onset / enrollment – time from symptom onset or baseline assessment, to death/event censored. Censored event – number and percentage of right-censored cases. SD standard deviation. Q1 - 25th percentile , Q3 - 75th percentile.

isomiR-let-7g-5p.t levels differentiate between short and long survival

To identify potential isomiRs associated with patient survival we dichotomized the 1480 isomiRs that were shared between cohorts, based on their median abundance in the UK discovery cohort. A univariate Cox regression—used to assess the association between isomiR expression and time to death—revealed 290 statistically significant isomiRs (P-value ≤ 0.05), of which 16 remained significant after correction for multiple comparisons (FDR ≤ 0.05, **Figure 1B**). Levels above median of 15 isomiRs were associated with decreased rate of mortality in patients (univariate Cox, HR < 1, 95% confidence interval (CI) = 0.34–0.77, FDR ≤ 0.05; **Figure 1C**) and levels above median of one isomiR were associated with increased rate of mortality (univariate Cox, HR = 1.8, 95% (CI) = 1.32–2.6, FDR ≤ 0.05; **Figure 1C**). Next, we tested these associations in the PGB replication cohort by dichotomizing isomiR levels by the threshold obtained in the UK discovery cohort. Only levels above median of isomiR-let-7g-5p.t were found to be associated with decreased rate of mortality in the PGB replication cohort (univariate Cox, HR = 0.6, 95% CI = 0.39–0.91, P-value = 0.015; **Figure 1D**). In addition, levels above median of isomiR-let-7g-5p.t were significantly associated with longer survival in the UK discovery cohort (log rank, χ^2 = 18.7, P-value ≤ 0.001, median survival difference = 8.2 months; **Figure 1E**) and in the PGB replication cohort (log rank, χ^2 = 5.99, P-value ≤ 0.01, median survival difference = 9.1 months; **Figure 1F**).

We sought to compare the mortality rate associated with isomiRs and the respective miRNAs, which reflect the aggregate of dozens of expressed isomiRs, assigned to a miRNA gene. While isomiR-let-7g-5p.t levels were associated with decreased mortality rate, no cognate miRNA was significantly associated with mortality rate in both cohorts including let-7g-5p (univariate Cox, HR = 0.87, 1.1; 95% CI = 0.63–1.2, 0.65-1.8; P-value > 0.05 in both, respectively **Figure S2**). Accordingly, plasma levels of cognate miRNA let-7g-5p did not differentiate between short and long survival in the UK discovery or the PGB replication cohorts (log rank χ^2 = 0.63, 0.09; P-value > 0.05 in both, median survival difference = 0.3 , 3.8 months respectively ; **Figure 1G, H**).

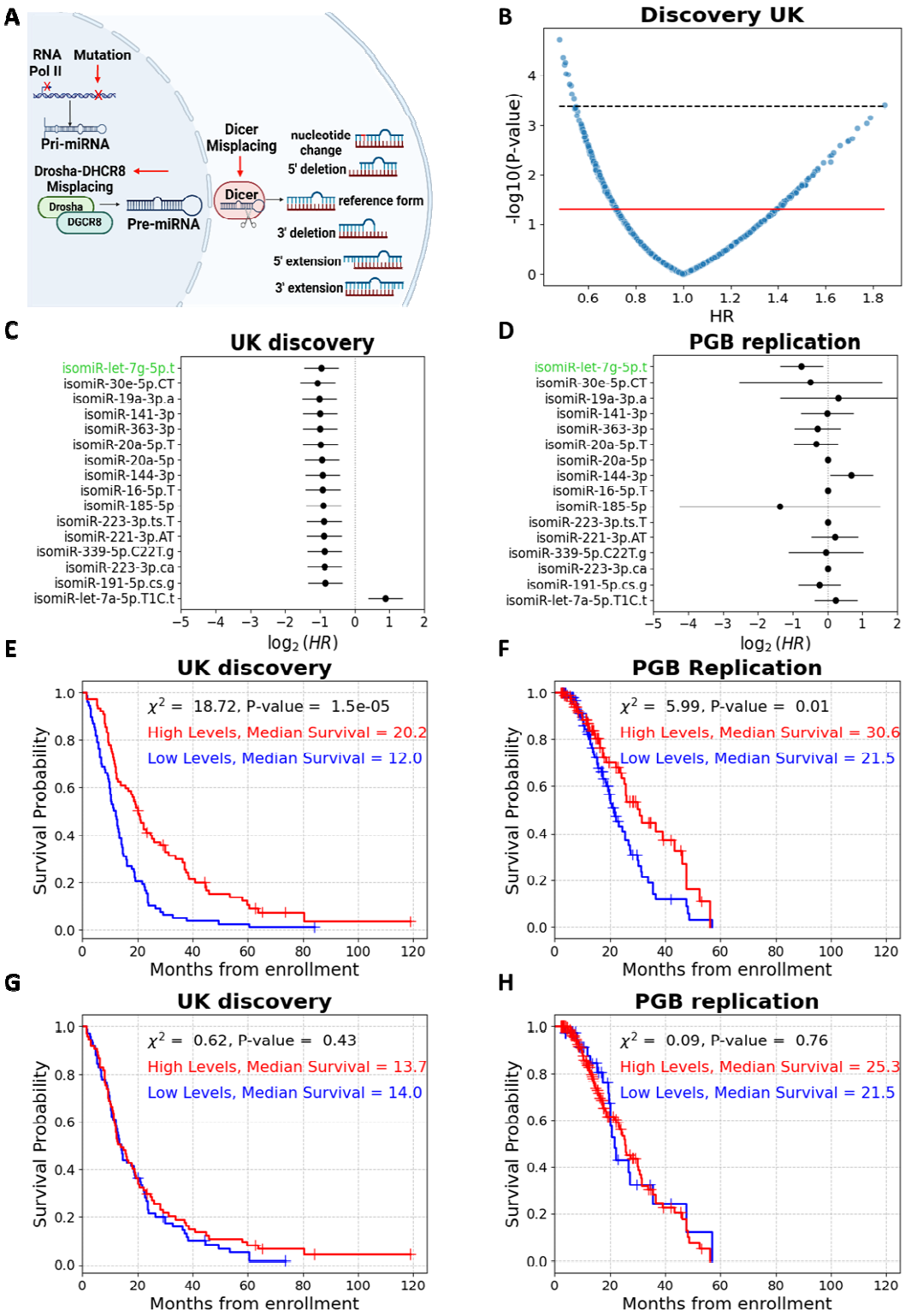

Figure 1 – High plasma isomiR-let-7g-5p.t levels predict longer survival in patients with ALS. (A) miRNA biogenesis yields several forms of miRNA isoforms (isomiRs). Following transcription, the primary miRNA is cleaved by the nuclear microprocessor, exported as a precursor to the cytoplasm, and further processed by Dicer into a ~22-nt guide RNA loaded onto the silencing complex (RISC) for target mRNA binding and silencing.

Variability in biogenesis, such as misplacement of processing factors, leads to the generation of isomiRs. **(B)** Volcano plot of 1480 plasma isomiRs univariate Cox Hazard ratios (HR, x axis) and their significance (y axis). Red line denotes significance threshold (P-value = 0.05). Only 16 isomiRs remained significant post multiple comparison correction (black line denotes FDR = 5.2e-4). Forest plots of univariate Cox proportional hazard ratios of these 16 significant isomiRs in UK discovery **(C)** or PGB replication **(D)** cohorts (HR ≤ 0.6 or HR = 1.8, P-value ≤ 0.015). Only isomiR-let-7g-5p.t remained significant in both cohorts. The Cumulative survival (Kaplan–Meier) curves for isomiR-let-7g-5p.t in the **(E)** UK discovery or **(F)** PGB replication cohorts. (UK: 77/77 patients in sub- vs. supra-threshold cohorts. PGB: 104/96 patients in sub- vs. supra-threshold cohorts). Kaplan–Meier curves for the mature miRNA counterpart, let-7g-5p, in the **(G)** UK discovery or **(H)** PGB replication cohorts. (UK: 77/77 patients in sub- vs. supra-threshold cohorts. PGB: 40/160 patients in sub- vs. supra-threshold cohorts patients subthreshold using the same threshold for dichotomization.)

IsomiR-let-7g-5p.t is the best prognostication isomiR in two independent ALS cohorts

Although isomiR-let-7g-5p.t emerged as a promising candidate for ALS prognostication, the high dimensionality of our dataset motivated us to explore the predictive performance of isomiR combinations and their potential non-linear interactions, using a machine learning, as employed in our previous studies [60]. We sampled at random subsets of 20 isomiRs, with repetition, using bootstrapping so each isomiR within the pool of 1480 would be included in at least 50 distinct subsets. To mitigate the risk of false positive discoveries inherent in large-scale analyses, the 154 patients of the UK discovery cohort were divided randomly into 10 bins and a sequential forward selection (SFS) with 10 fold cross-validation technique was performed for each of the isomiR subset [61, 62]. 84 isomiRs were selected within 90% of the subsets in which they were included after SFS (≥45 out of 50, with Cox regression [34], random survival forest (RSF) [36], or survival gradient boosting (XGB) [35]**, Figure S3A**). Of these three orthogonal survival models, XGB demonstrated the highest mean concordance index (C-index) across the 84 isomiRs (**Figure S3B**).

Next, we modeled survival of the 84 isomiRs by XGB. IsomiRs included if their subsets displayed a C-index higher than a threshold that was defined empirically by the elbow method[54] (> C-index of 0.74) Reassuringly, isomiR-let-7g-5p.t was detected, along with 17 other isomiRs: isomiR-let-7e-5p.G9C, isomiR-let-7f-5p.T16C.t, isomiR-let-7i-5p.G5A, isomiR-130b-5p.T, isomiR-142-3p.a, isomiR-451a.T.A, isomiR-144-3p, isomiR-146a-5p.T22A, isomiR-151a-3p.A9G, isomiR-25-3p.A22T, isomiR-30d-5p.A20G.CT, isomiR-425-5p.as, isomiR-432-5p, isomiR-486-5p.C, isomiR-486-5p.T9G.T, isomiR-92a-3p.T1C, isomiR-92a-3p.T22C (**Table S1** for nomenclature; **Figure S3C**).

After hyperparameter tuning with 5-fold cross-validation (Supplementary Methods), a model based on these 18 isomiRs, demonstrated strong prognostic capacity across patient survival duration

of 84 months (time dependent AUC (AUC(t)) values of 0.7 to 0.95; average 0.78; C-index = 0.7; survival duration 2 - 85 months; **Figure 2A**), but survival prediction was less effective in the PGB replication cohort, (AUC(t) 0.49- 0.9, average, 0.62; C-index = 0.55; survival duration 3–54 months; **Figure 2A**). Notably, prediction was better for patients with longer survival times.

We demonstrate higher performance with the model trained on the 18 selected isomiRs, relative to models trained on the collective set of all annotated isomiRs (the"miRNA", average AUC(t) = 0.64/0.53, C-index= 0.59/0.52 in the UK discovery and PGB replication cohorts, respectively) or the most abundant isomiR ("reference miRNAs" average AUC(t)= 0.68/0.46, and C-index= 0.63/0.47) **Figure 2B, C**). Additionally, the summation of all isomiRs annotated to a miRNA gene and ther most abundant isomiR (miRNA and reference miRNA), displayed poor performance over time (**Figure S4**).

To identify the isomiRs most valuable for survival prediction, we calculated Shapley Additive Explanations (SHAP) for the model average risk function. isomiR-let-7g-5p.t levels above median were associated with reduced risk of death and longer survival (**Figure 2D**). In contrast, levels above median of isomiR-92a-3p.T22C, isomiR-92a-3p.T1C, isomiR-let-7f-5p.T16C.t, and isomiR-let-7i-5p.G5A associate with higher risk and shorter survival. A similar SHAP analysis of the PGB replication cohort was highly correlated, highlighting the model's robustness (**Figure 2E**). We then derived the monthly hazard ratios (HRs) from SHAP values as previously described[57]. IsomiR-92a-3p.T22C and isomiR-let-7g-5p.t levels were associated the most with increased / decreased mortality, respectively (**Figure 2F**).

As isomiR expression can vary across cohorts, we assessed abundance and inter-cohort correlation to evaluate biological reproducibility. Among the 18 selected isomiRs, mean expression ranged from 4–9 UMIs (UK cohort) and 5–14 UMIs (PGB cohort) and the detection rates were comparable between cohorts (**Figure S5A, B**). IsomiR-142-3p.a had the highest expression (UK = 974 ± 833 UMIs, PGB = 1984 ± 945 UMIs), followed by isomiR-let-7g-5p.t (UK discovery cohort: 264 UMIs ± 129; PGB replication cohort: 259 UMIs ± 94). IsomiR abundance was well correlated between cohorts ($r = 0.74$, $P = 8.1e{-28}$ **Figure S5C, D**) but no correlation was noted between the abundance of isomiRs and their cognate "miRNA" (collective set of all annotated isomiRs) nor with the "reference miRNA" (predominant isomiR, **Table S2**).

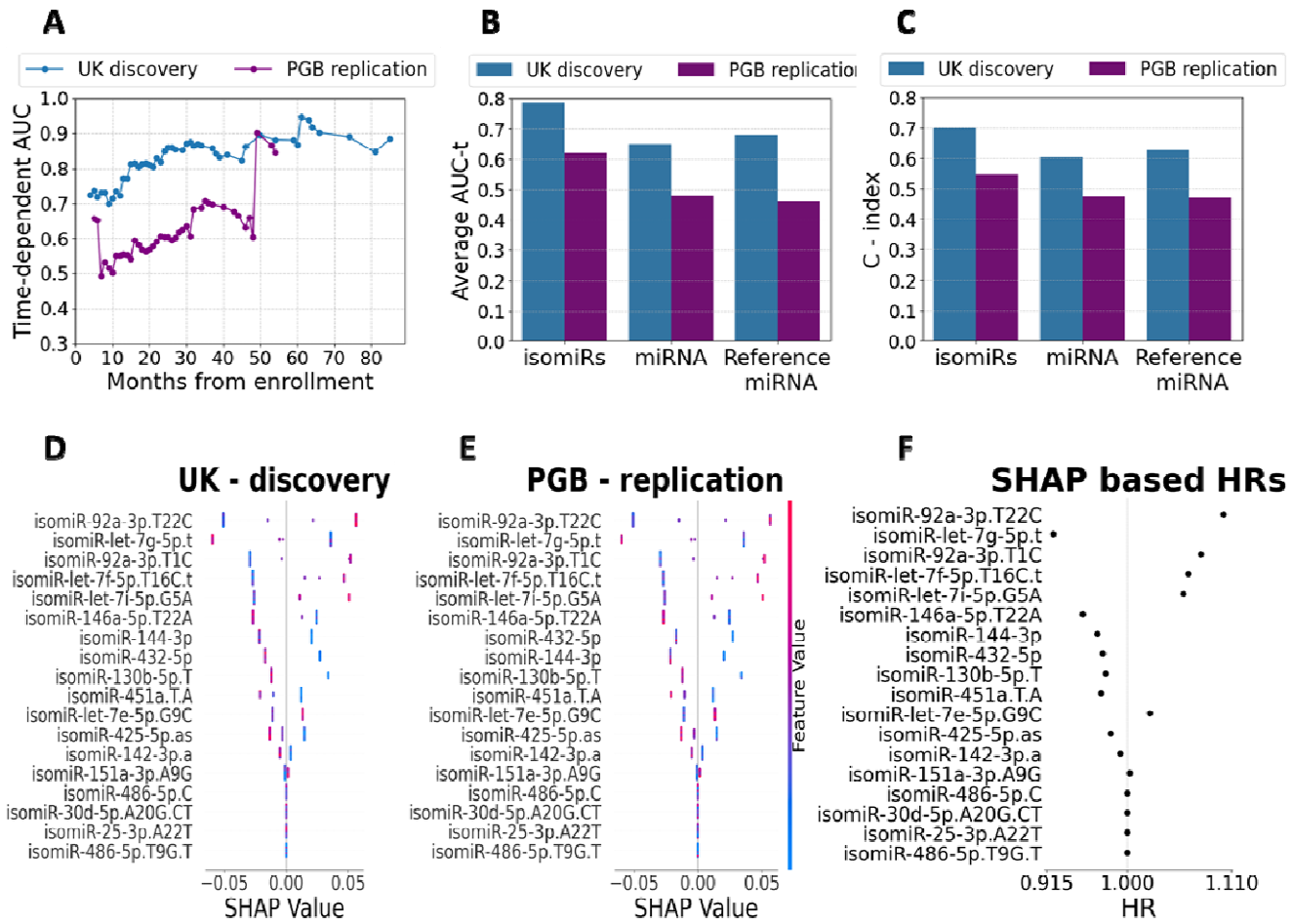

Figure 2 – IsomiR based survival prediction in two independent cohorts. **(A)** Plot of AUC (y axis) in the UK discovery (blue) or PGB replication (purple) cohorts for patient survival, as a function of duration from enrolment (x axis). Bar graphs of **(B)** average AUC across time **(C)** or C-index for a model based on 18 selected isomiRs. SHAP analysis the contribution of 18 isomiRs to risk prediction in the **(D)** UK discovery or **(E)** PGB replication cohorts. Range of isomiR levels (low–blue to high–red). Positive or negative SHAP values reveal higher or lower effect size of isomiR on the mortality risk (x axis), respectively. **(F)** Plot of HR, on the basis of XGB SHAP values calculated according to Sundrani & Lu[57].

Evaluating plasma isomiR-let-7g-5p.t alongside NfL in ALS prognostication

As we previously reported [16, 63], high (above median) blood NfL levels were significantly associated with shorter survival in the UK discovery cohort (log-rank, χ^2 = 9.4, P-value = 2e-3, median survival difference = 7.1 months; **Figure 3A**) and this was consistent in the PGB replication cohort, when NfL dichotomized according to the median plasma value in the UK cohort (log rank, χ^2 = 17.7, P-value = 2.7e-5, median survival difference = 13.7 months; **Figure 3B**). Above median NfL levels were associated with an increased mortality rate in both cohorts (UK discovery cohort: univariate Cox, HR = 1.7, 95% CI = 1.2-2.3, P-value = 2.4e-4, C-index = 0.56 ; PGB replication cohort Cox HR = 2.4, 95% CI 1.6-3.7, P-value = 4.5e-6, C-index = 0.63; **Figure 3C, D**). In contrast, higher levels of seven (out of eighteen selected by machine learning, **Figure S3C**) isomiRs were associated with a decreased mortality rate (univariate Cox, HR < 0.72, 95% CI = 0.37-0.99, P-value ≤ 0.05, C-index > 0.54 ;**Figure 3C**). However, only isomiR-let-7g-5p.t retained significant HR in univariate COX regression in the PGB replication cohort (UK discovery cohort: HR = 0.58 95% CI = 0.41-0.81, P-value = 1.4e-3, C-index = 0.57; PGB replication cohort: univariate Cox, HR = 0.6 95% CI = 0.4-0.9, P-value = 0.015, C-index = 0.55; **Figure 3C, D**). Accordingly, isomiR-let-7g-5p.t was the only isomiR that significantly distinguishes short and long survival times in both cohorts using the log-rank test (**Figures S6, S7**).

In multivariate Cox analysis NfL remained significantly associated with increased mortality rate in both cohorts (UK discovery HR = 1.7, 95% CI = 1.2-2, P-value = 6e-4; PGB replication HR = 3.3, 95% CI = 2.4-5.3, P-value = 2e-7), whereas isomiR-let-7g-5p.t was significant only in the PGB cohort (multivariate Cox, HR = 0.52, 95% CI = 0.32-0.86, P-value = 0.01; **Figure S8A, B**). Noteworthy, isomiRs modestly the prognostic capacity, compared to a model which only includes NfL, in both cohorts (C-index = 0.67/0.72 in UK discovery or PGB replication respectively, compared to 0.56/0.63 for NfL alone).

When considering only isomiR-let-7g-5p.t and NfL together, isomiR-let-7g-5p.t showed a stronger association with risk of mortality than NfL in the UK discovery cohort (HR of NfL in a multivariate Cox model = 1.64, 95% CI = 1.2-2.3, P-value = 3.9e-4; HR of isomiR-let-7g-5p.t in a multivariate Cox model = 0.49, 95% CI = 0.35-0.69, P-value = 3.8e-6; **Figure S8C**). Together NfL and isomiR-let-7g-5p.t yield better prognostication than either marker alone (multivariate Cox C-index = 0.63 vs. C-index = 0.56 NfL alone or C-index = 0.59 isomiR-let-7g-5p.t alone).

In the PGB cohort, NfL had a stronger association with mortality rate (NfL multivariate Cox, HR = 3, 95% CI = 1.9-4.6, P-value = 1e-7; isomiR-let-7g-5p.t multivariate Cox, HR = 0.46, 95% CI = 0.29-0.7, P-value = 4.8e-5; **Figure S8D**) but again, both markers together improved survival prediction compared to either one individually (multivariate Cox C-index = 0.67, NfL alone C-index = 0.63, isomiR-let-7g-5p.t alone C-index = 0.56). Moreover, when combined with relevant clinical features (baseline ALSFRS-R, site of symptom onset, age at onset), only isomiR-let-7g-5p.t (UK discovery multivariate Cox, HR = 0.53, 95% CI = 0.38-0.75, P-value = 3e-5; PGB replication multivariate Cox, HR = 0.52, 95% CI = 0.33-0.83, P-value =6.5e-3) and baseline ALSFRS-R (UK discovery multivariate Cox, HR = 0.97, 95% CI = 0.94-0.99, P-value = 2.6e-3; PGB replication multivariate Cox, HR = 0.96, 95% CI = 0.92-0.99, P-value = 0.013) were significantly associated with mortality risk in both cohorts (**Figure S8E,F**). In contrast, NfL levels association with mortality were only approaching significance in the UK discovery cohort (multivariate Cox, HR = 1.4, 95% CI = 0.97-1.94, P-value = 0.08 ;**Figure S8E**) while being significantly associated with mortality in the PGB cohort with an effect size larger than that of isomiR-let-7g-5p.t (multivariate Cox, HR = 2.5, 95% CI = 1.5-4, P-value ≤ 0.001 ;**Figure S8F**).

Since above-median levels of isomiR-let-7g-5p.t and NfL were associated with opposing effects on mortality risk, we next assessed their joint prognostic capacity by calculating the ratio between isomiR-let-7g-5p.t and NfL levels . High values of this ratio slightly improved distinction between short and long survival compared to NfL alone (UK discovery log-rank, $\square^2$ = 11.8, P-value = 5.9e-4, median survival difference = 8 months; PGB replication log-rank, $\square^2$ = 19.8, P-value = 8.8e-6, median survival difference = 14 months; **Figures 3E,F**). Notably, the ratio showed a stronger inverse association with mortality risk compared to either marker alone and improved overall prognosis (UK discovery univariate Cox, HR = 0.56, 95% CI = 0.4-0.78, P-value = 6.2e-4, C-index = 0.58; PGB replication univariate Cox, HR = 0.4, 95% CI = 0.26-0.6, P-value = 1.8e-5, C-index =0.64 ;**Figure S9**). Taken together, these results suggest that isomiR-let-7g-5p.t performs comparably to NfL and that combining isomiR-let-7g-5p.t with NfL improved predictive performance, at least as reflected by the C-index. These findings highlight the value of isomiR-let-7g-5p.t as a potential prognostic biomarker in ALS.

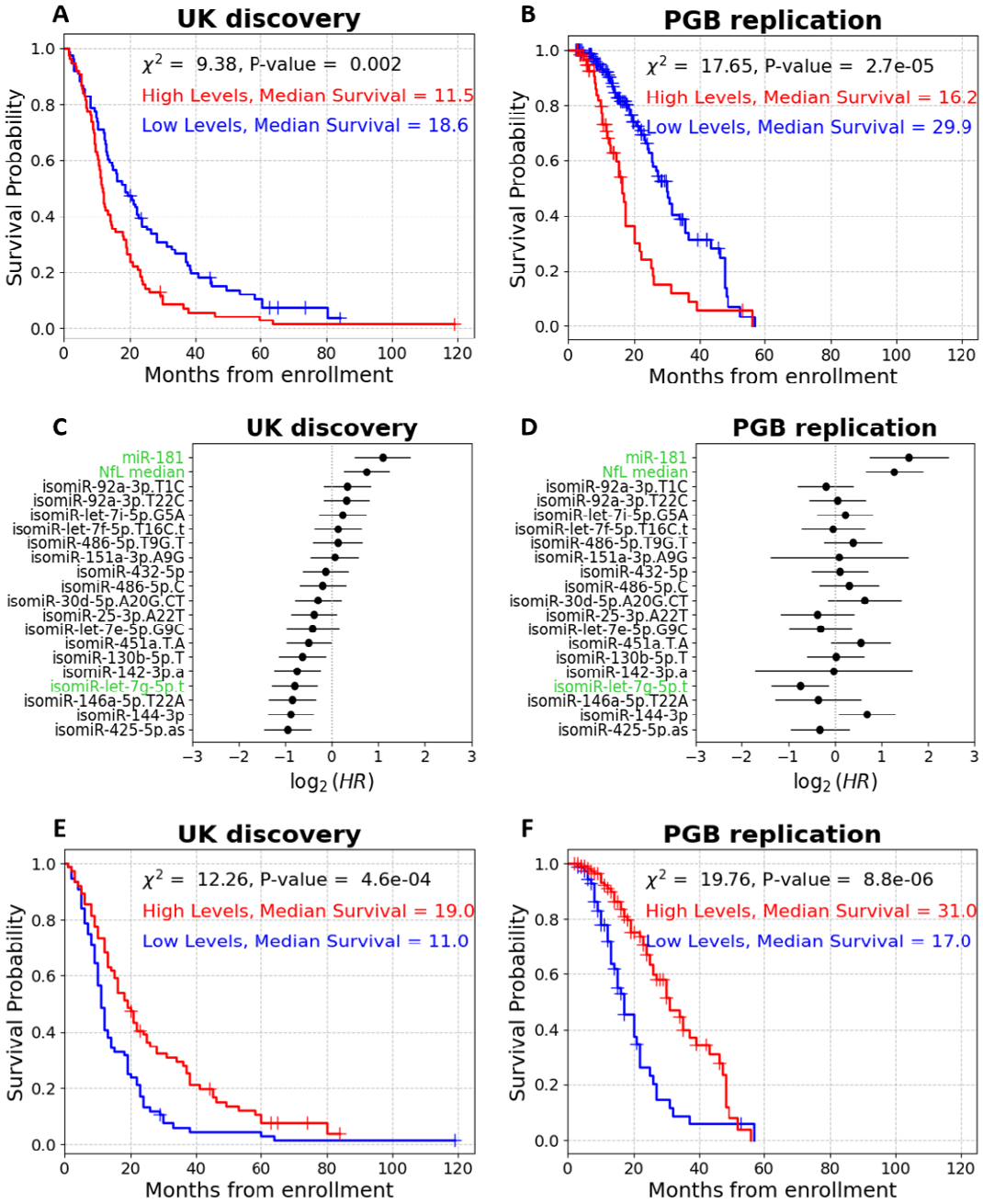

Figure 3 – Prognostication by NfL and by isomiR-let-7g-5p.t. **(A)** Cumulative survival (Kaplan–Meier) curves for NfL plasma levels (76/ 68 subthreshold (blue) /suprathreshold (red) patients, respectively, from enrollment in the UK discovery cohorts and similarly, NfL serum levels (60/140 subthreshold (blue) /suprathreshold (red) patients respectively) in **(B)** PGB replication cohort, respectively. Forest plots of univariate Cox proportional hazard ratios of 18 isomiRs, NfL and miR-181 in **(C)** the UK discovery cohort or, **(D)** PGB replication cohort. NfL, miR-181 and isomiR-let-7g-5p.t are highlighted in green. **(E)** Cumulative survival (Kaplan–Meier) for isomiR-let-7g-5p.t/NfL ratio (76/76 subthreshold (blue) / suprathreshold (red) patients respectively) from enrollment in the UK discovery cohorts and **(F)** PGB replication (right 147/53 subthreshold (blue) / suprathreshold (red) patients respectively)

Evaluating plasma isomiR-let-7g-5p.t alongside miR-181 in ALS prognostication

Our UK cohort was a subset of the cohort where we originally discovered miR-181, a multiplication of sister miRNAs miR-181a-5p and miR-181b-5p, as a prognostic marker [16]. A univariate Cox analysis revealed that high miR-181 levels were associated with a >2.1-fold increased risk of mortality, again demonstrating that isomiR-let-7g-5p.t has an equivalent prognosis capacity to previously published markers (UK discovery cohort HR = 2.1, 95% CI = 1.4–3.2, P-value ≤ 0.001, C-index = 0.58; PGB replication cohort HR = 3.0, 95% CI = 1.7–5.4, P-value ≤ 0.001, C-index = 0.55; **Figure 3C, D**).

miR-181 levels distinguished between short and long survivors in the UK cohort (log rank on dichotomized miR-181 levels: χ^2 = 13.7, P-value = 2.1e-4, median survival difference = 7.2 months; **Figure 4A**) and also in the PGB cohort (log rank: χ^2 = 15.1, P-value = 1e-4, median survival difference = 9.7 months; **Figure 4B**). Notably, isomiR-let-7g-5p.t was the only isomiR to remain statistically significant in both cohorts, in a combined analysis of miR-181 and the 18 isomiR chosen from the XGB model (UK discovery multivariate Cox, HR = 0.63, 95% CI = 0.4-0.98, P-value = 0.04; PGB replication multivariate Cox, HR = 0.57, 95% CI = 0.36-0.9, P-value = 0.017; **Figure 4C, D**). To assess joint prognostic capacity of isomiR-let-7g-5p along with miR-181 we tested isomiR-let-7g-5p / miR-181 ratio capacity to distinguish between short and long survival. We Found a significant association of the ratio levels to patient survival in both cohorts (UK discovery log rank: χ^2 = 20.6, P-value = 5.6e-6, median survival difference = 8 months; PGB replication log rank: χ^2 = 7, P-value = 8e-3 median survival difference = 6 months; **Figures 4E,F**). Moreover, the hazard ratio of isomiR-let-7g-5p / miR-181 ratio was 0.38 in the discovery and 0.44 in the replication cohort. (UK discovery univariate Cox, 95% CI = 0.19–0.77, C-index = 0.58, P-value = 8e-6; PGB replication univariate Cox, 95% CI = 0.18-0.8, C-index = 0.53, P-value = 0.01, **Figure S9)**. Taken together, these results highlight the potential of isomiR-let-7g-5p.t as a robust biomarker for ALS survival.

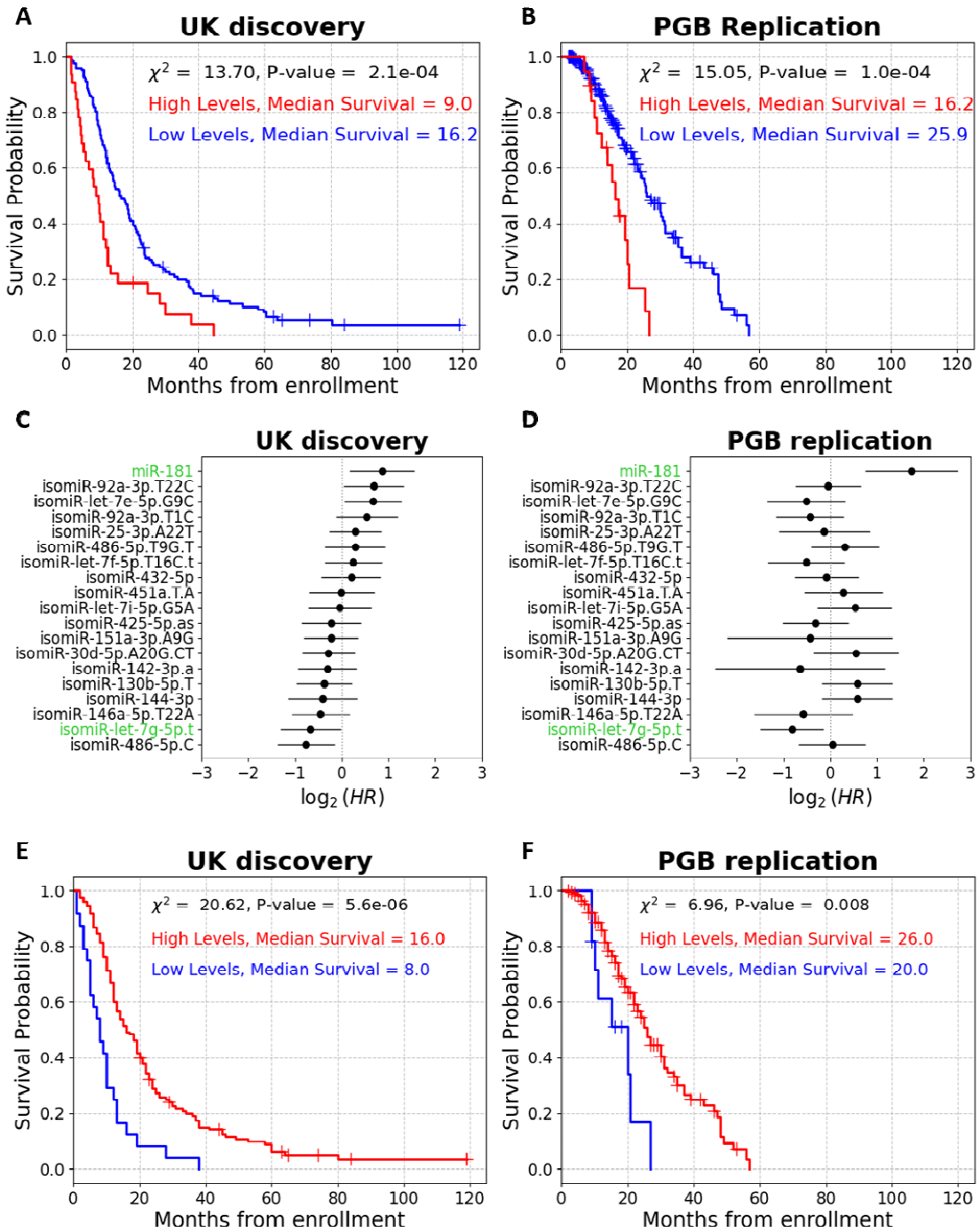

Figure 4 – isomiR-let-7g-5p.t based prognostication compared to miR-181. Cumulative survival (Kaplan–Meier) for miR-181. (122/32 subthreshold (blue) / suprathreshold (red) patients respectively) in the UK discovery **(A)** / PGB replication (181 /19 patients subthreshold (blue) / suprathreshold (red) patients respectively) **(B)** cohorts. Forest plots of multivariate Cox proportional hazard ratios of 18 isomiRs and miR-181 in the UK discovery **(C)** or PGB replication **(D)** cohorts. isomiR-let-7g-5p.t is highlighted in green. **(E)** Cumulative survival (Kaplan–Meier) for isomiR-let-7g-5p.t/miR-181 ratio (130/24 subthreshold (blue) / suprathreshold (red) patients respectively) from enrollment in the UK discovery cohorts and **(F)** PGB replication (right 89/11 subthreshold (blue) / suprathreshold (red) patients respectively)

Discussion

IsomiRs display small variations of either sequence insertion, deletion or single nucleotide polymorphism that distinguish them from the reference miRNA form. While miRNAs are extensively studied in ALS [12, 14-16, 64-69], we are first to examine the value of isomiRs as biomarkers for neurodegeneration. We tested the prognostic capacity of 1480 isomiRs in two cohorts and concluded that isomiR-let-7g-5p.t is a promising prognostic biomarker. High plasma levels of isomiR-let-7g-5p.t are significantly associated with lower mortality rate and longer survival. Higher plasma levels of isomiR-let-7g-5p.t are comparable but reciprocal to miR-181 and NfL, which we previously associated with ALS prognosis [16, 63]. Additionally, the incorporation of isomiR-let-7g-5p.t to either miR-181 or NfL enhances their prognostic value.

Dysregulation of miRNAs in ALS has previously been demonstrated in mice and humans [11, 12, 14, 15]. numerous studies have proposed miRNA utility as potential biomarkers for prognosis for ALS [7, 64, 69]. However, isomiRs have been overlooked, in part due to their sequence heterogeneity and their low abundance [22]. However, dysregulation of RNA-binding proteins such as TDP-43, DICER or DROSHA in ALS [25, 70-74], should prompt additional studies of isomiRs in ALS. In addition, as TDP-43 is also dysregulated in other neurodegenerative diseases [75], it is possible that isomiRs play a role not only in ALS but also in other neurodegenerative disorders, with implications ranging from basic research to clinical applications.

microRNA let-7g-5p has previously been reported to be lowered in abundance in the blood of patients with sporadic ALS (sALS) [65] and in patients with frontotemporal dementia (FTD) due to a *C9orf72* repeat expansion [76], compared to healthy controls. Interestingly, isomiR-let-7g-5p.t accounts for approximately 40% of the reads of let-7g-5p in the plasma. Therefore, isomiR-let-7g-5p.t is likely associated and perhaps predominates the let-7g-5p repertoire. Seventeen other isomiRs displayed potential relevance, though to a weaker extent. Some of the miRNAs associated with these isomiRs (miR-451, miR-144-3p and miR-486-5p) are linked to oxidative stress [77-86] and inflammation [77, 78, 81, 82, 84, 85] which potentially play a role in neurodegeneration. However, because we analyze the isomiRs in plasma, the tissue expression origin is unclear

Despite the promising prognostic potential of isomiRs in ALS, several limitations must be noted. First, the mechanisms underlying isomiR levels including isomiR-let-7g-5p.t and their

functional role in disease progression is yet to be fully elucidated. Second, while this study included two independent cohorts, additional validation in larger and more diverse populations would strengthen the generalizability of the findings. Larger cohorts could also help assess isomiRs robustness across different ALS subtypes and disease stages. Third, given the availability and strong prognostic performance of NfL, the practical utility of isomiRs as standalone biomarkers may be limited until an effective way to measure them is developed. These limitations highlight the need for continued research to establish the role of isomiRs in ALS and their potential integration into clinical practice.

Together, our results highlight the potential of isomiR-let-7g-5p.t as a prognostic biomarker in ALS, potentially offering comparable value to clinical features, miR-181, and NfL. Moreover, isomiR-let-7g-5p.t enhances effective prognostication in conjunction with these markers. These findings warrant further exploration of the utility of isomiR-let-7g-5p.t and other isomiRs in neurodegenerative diseases.

Supplementary tables and figures

Mutation	Nomenclature	Reference sequence	Mutated sequence	Comment
Reference	Let-7g-5p	TGAGGTAGTAGTTTGTACAGTT	TGAGGTAGTAGTTTGTACAGTT	No change
3' deletion	Let-7g-5p.tt	TGAGGTAGTAGTTTGTACAG**TT**	TGAGGTAGTAGTTTGTACAG	Lowercase letter (n)
3' insertion	Let-7g-5p.CC	TGAGGTAGTAGTTTGTACAG	TGAGGTAGTAGTTTGTACAG**CC**	Uppercase letter (N)
5' deletion	Let-7g-5p.stg	**TG**AGGTAGTAGTTTGTACAGTT	AGGTAGTAGTTTGTACAGTT	Addition of "s"
5' insertion	Let-7g-5p.sGG	TGAGGTAGTAGTTTGTACAGTT	**GG**TGAGGTAGTAGTTTGTACAGTT	Addition of "s"
Base exchange	Let-7g-5p.G2C.T16G	TGAGGTAGTAGTTTGTACAGTT	T**C**AGGTAGTAGTTTG**G**ACAGTT	Position relative to reference sequence

Table S1 - IsomiR Nomenclature

Mutations are separated by a dot. Examples of mutations are highlighted in bold and red.

Base exchanges are represented by the original nucleotide, followed by its position in the reference sequence, and then the new nucleotide (e.g., A12G).

Reference miRNAs do not have any suffixes. Deletions are indicated by lowercase letter suffixes. Insertions are indicated by uppercase letter suffixes.

5′ indels are prefixed with the letter "s"; otherwise, indels are assumed to be at the 3′ end. The position of base exchange in indels remains relative to the reference sequence.

	miRNA UK discovery	miRNA PGB replication	reference miRNA UK discovery	reference miRNA PGB replication
hsa-let-7e-5p	545.76 ± 350.51	561.41 ± 244.80	290.93 ± 225.13	303.17 ± 137.28
hsa-let-7f-5p	9434.27 ± 4849.83	11554.54 ± 4143.22	5908.90 ± 4210.59	7920.46 ± 3218.58
hsa-let-7g-5p	992.89 ± 372.96	1173.12 ± 307.08	555.03 ± 385.22	713.87 ± 264.32
hsa-let-7i-5p	12677.13 ± 5929.79	14073.11 ± 4257.76	9116.78 ± 5659.07	9864.22 ± 3331.09
hsa-miR-130b-5p	8.26 ± 5.39	9.40 ± 4.97	nan	nan
hsa-miR-142-3p	1180.20 ± 776.49	2430.42 ± 1049.35	61.75 ± 44.68	123.52 ± 63.52
hsa-miR-144-3p	24.58 ± 27.28	16.31 ± 18.94	8.63 ± 9.66	5.56 ± 6.45
hsa-miR-146a-5p	2270.01 ± 1118.95	6507.35 ± 1719.87	1200.03 ± 835.11	3592.34 ± 1241.20
hsa-miR-151a-3p	1480.06 ± 677.72	4782.99 ± 1337.09	802.72 ± 467.30	2963.98 ± 869.57
hsa-miR-25-3p	1865.72 ± 1072.00	2025.82 ± 669.98	1363.26 ± 744.40	1630.85 ± 599.57
hsa-miR-30d-5p	2920.10 ± 861.54	4553.77 ± 999.95	275.84 ± 146.96	399.31 ± 133.34
hsa-miR-425-5p	865.75 ± 427.35	914.48 ± 244.07	129.94 ± 90.52	213.61 ± 80.55
hsa-miR-432-5p	653.12 ± 706.14	640.87 ± 407.83	241.62 ± 214.92	266.80 ± 176.47
hsa-miR-451a	1416.05 ± 1993.90	550.35 ± 745.91	257.16 ± 461.63	133.97 ± 225.36
hsa-miR-486-5p	25375.59 ± 29731.65	7693.64 ± 7295.35	9280.17 ± 8542.15	2768.21 ± 2999.47
hsa-miR-92a-3p	7033.59 ± 6283.45	4454.01 ± 2302.21	4381.68 ± 3579.75	3375.72 ± 1824.18

Table S2 – plasma levels of reference miRNA associated to the 18 isomiRs studied. Levels of reference miRNAs and miRNA (integrated of all isomiRs associated to a defined miRNA gene), across patients in the UK discovery and the PGB replication cohorts. Units are UMIs, unique molecular identifiers. Mean plasma levels and standard deviation.

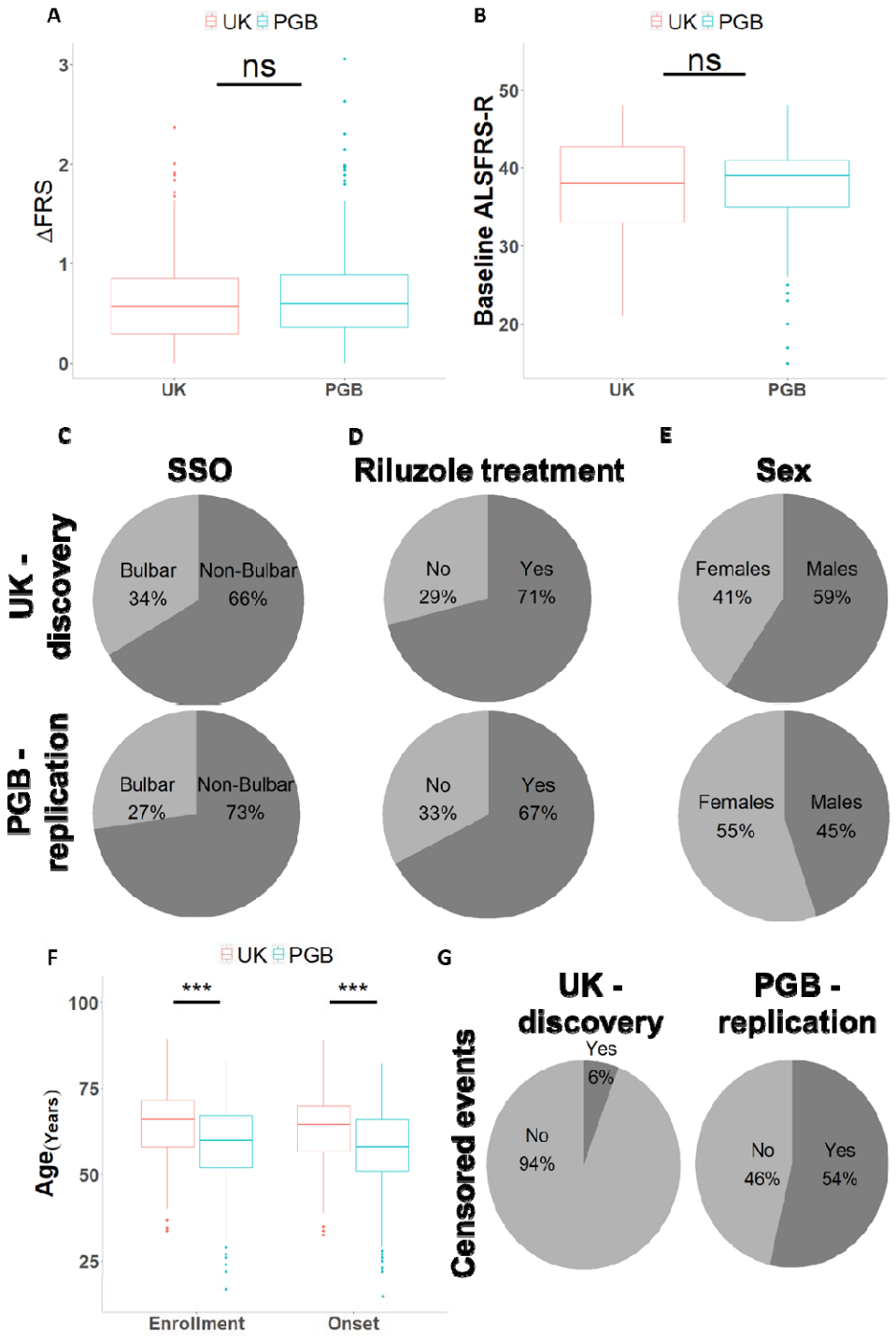

Figure S1 – Comparison of the UK discovery and PGB replication cohorts metadata. Boxplot of **(A)** DeltaFRS, measured as the reduction in the ALSFRS-R score per month or **(B)** baseline ALSFRS-R score at study enrollment. Pie charts depicting **(C)** sex **(D)**, SSO - site of symptom onset or **(E)** Riluzole treatment for the cohorts. **(F)** Boxplots of age at study enrollment and age at symptom onset. *** P-value ≤ 0.001, two-sided Student's t-test. Pie charts showing the number of censored and deceased events for the UK discovery (left) and the PGB replication (right) cohorts.

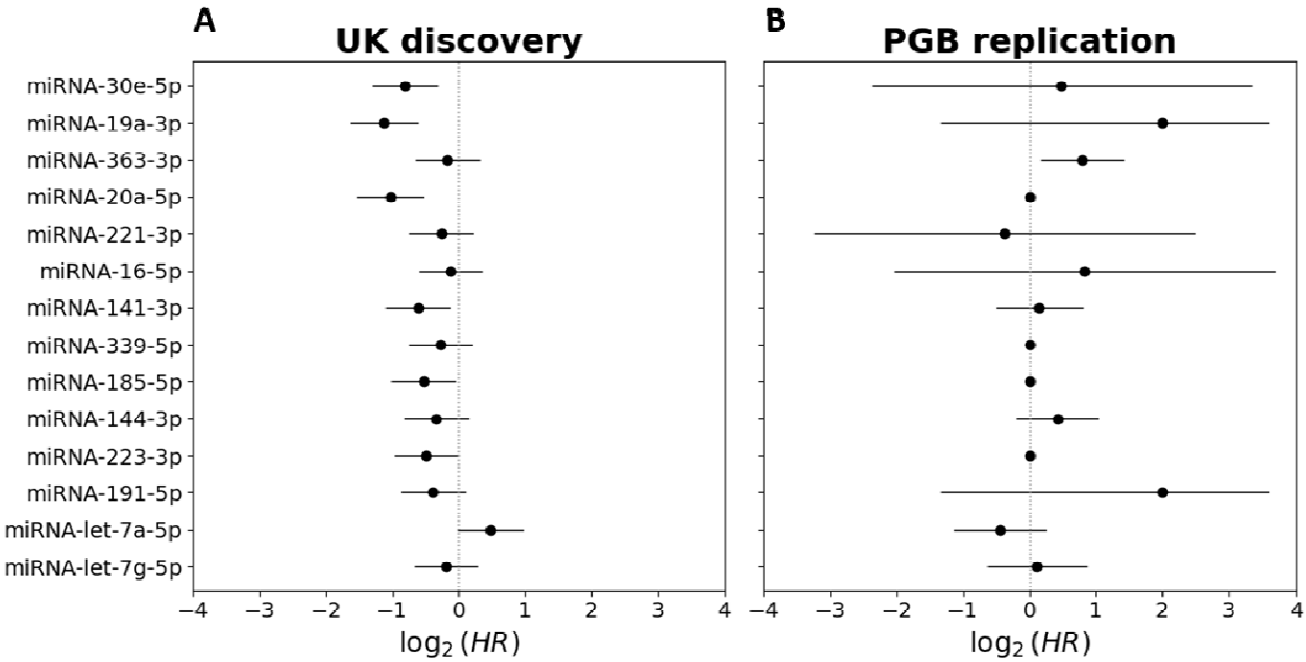

Figure S2 – miRNAs, a term reflecting on the integration of all isomiRs associated to a miRNA gene display poor prognostication capacity. Forest plots of univariate Cox proportional hazard ratios of 14 miRNA form of the 16 FDR significant isomiRs (FDR ≤ 0.05) in UK discovery **(A)** or PGB replication **(B)** cohorts.

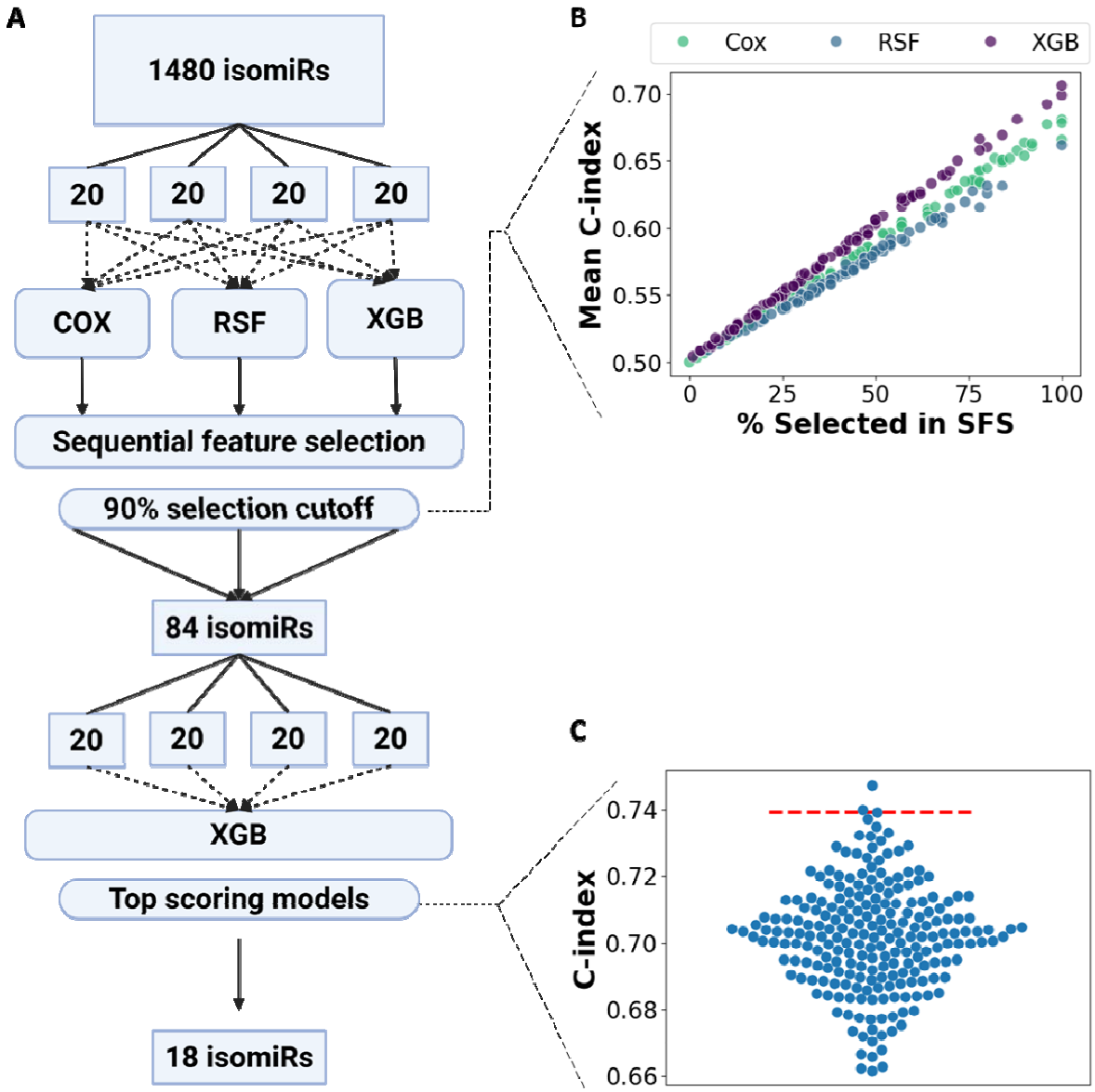

Figure S3 - Feature selection process yielded 18 candidate isomiRs as optimal predictors for survival. (A) The workflow outlines the multi-step feature selection strategy applied to 1,480 isomiRs profiled in a UK discovery cohort of 154 ALS patients. Initially, three different survival models—Cox proportional hazards (COX), Random Survival Forest (RSF), and survival-boosting (XGB) —were each applied across repeated data splits, composed of 20 isomiR. For each model, sequential feature selection (SFS) was performed to identify isomiRs contributing most to predictive performance. IsomiRs selected in at least 90% of the SFS runs were retained, yielding a set of 84 highly stable candidates. This refined set underwent a second round of XGB modeling. The top-scoring model from this phase, based on concordance index (C-indexI), identified a panel of 18 isomiRs that supports survival outcomes. **(B)** Scatter plot of the mean concordance index (C-index) across 3,718 isomiR subsets (y axis) and percentage of times isomiRs were selected in sequential forward selection (SFS) in three survival analyses: Cox (green), RSF (blue), or XGB (purple). **(C)** A plot of C-index (y axis) for sampling of 20 isomiRS from a total of 84, with repetition, using XGB. 18 isomiRs within two subsets, above the red dashed line, displayed best C-index.

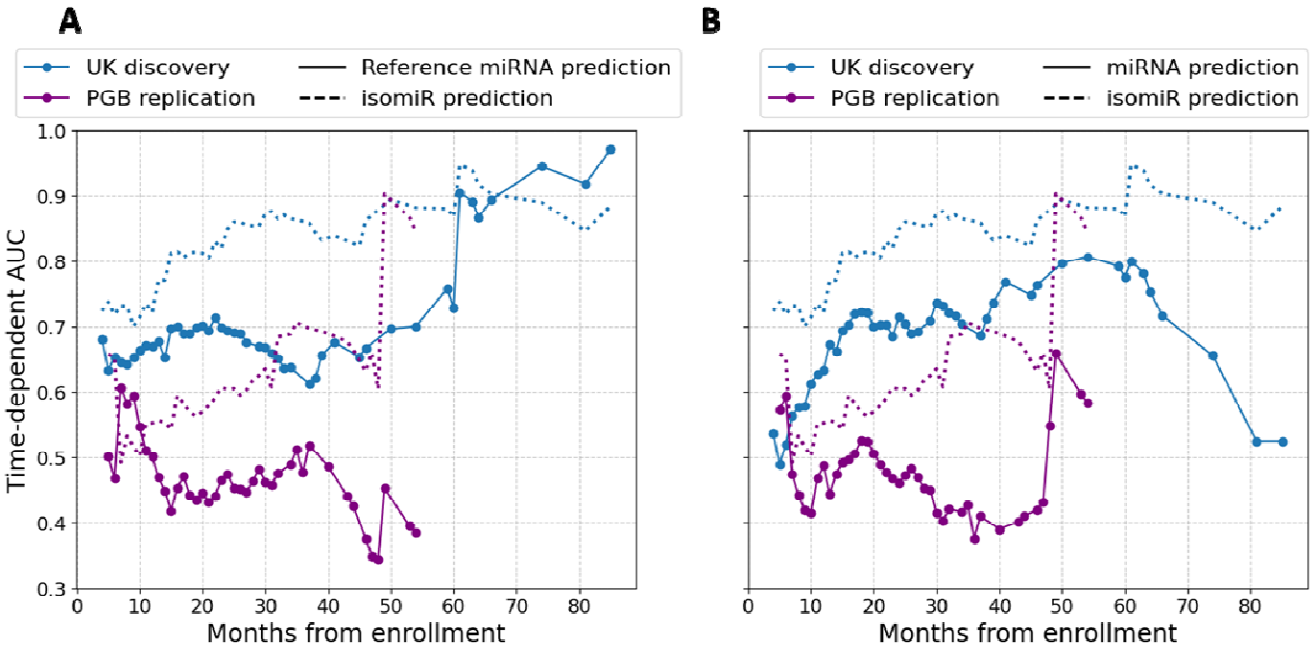

Figure S4 - **Prognostic capacity of selected isomiRs and their cognate miRNAs.** Time-dependent AUC (y axis) of **(A)** reference miRNA (predominant isomiR), **(B)** miRNA (integration of all isomiRs associated to a miRNA gene) or isomiRs (dashed line) in the UK discovery (blue) or PGB replication (purple) across cohort patient survival duration (x axis).

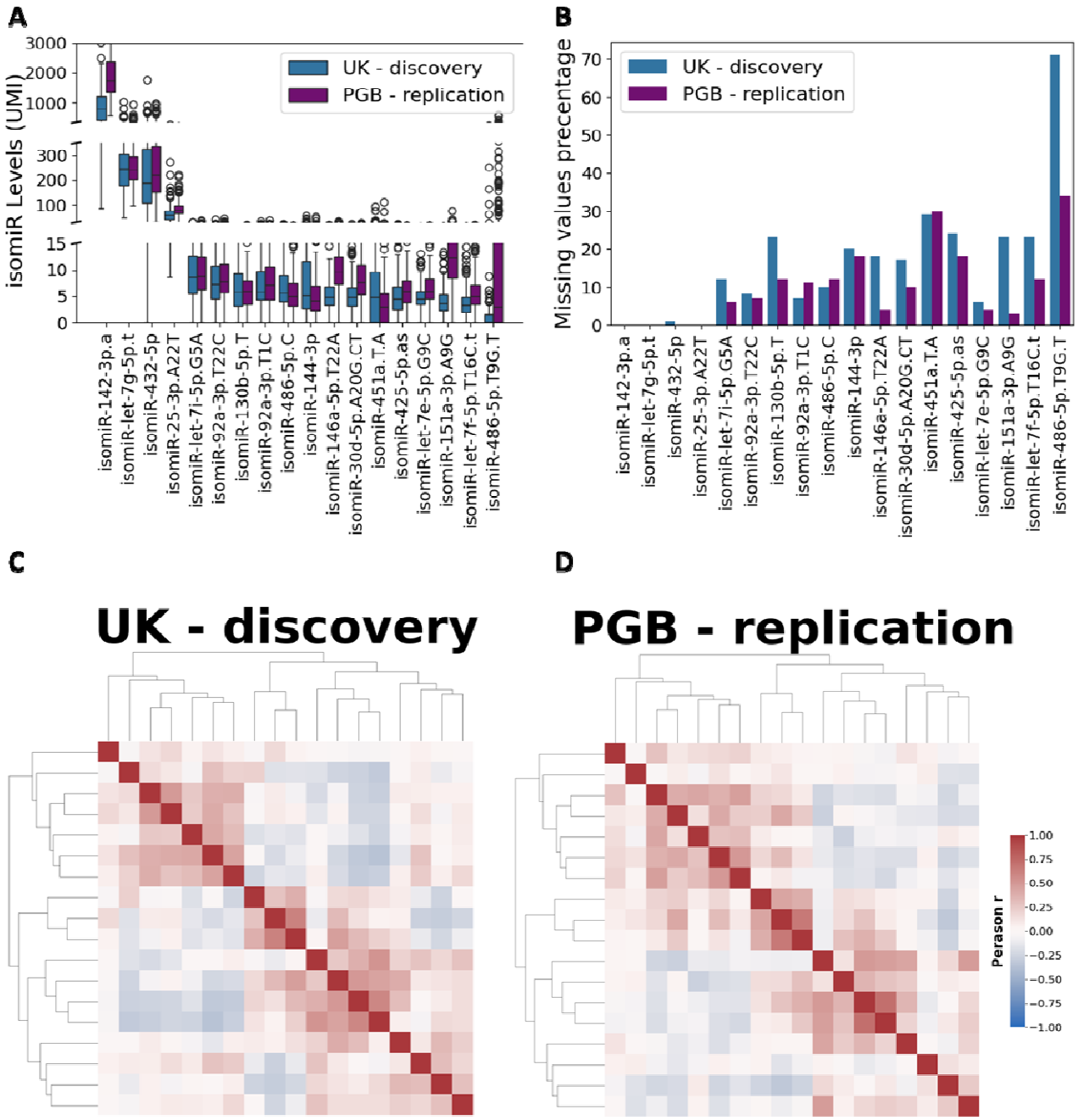

Figure S5 – isomiR characteristics across the discovery and replication cohorts. (A) Boxplots of corrected isomiR levels (y-axis) in the UK discovery cohort (blue) and the PGB replication cohort (purple). The 1st and 99th percentiles isomiR's levels are denoted by whiskers. Extreme values are marked by circles **(B)** Bar plot of missing isomiR values in patients from the UK discovery (blue) and PGB replication (purple) cohorts. Heatmaps of the correlation between plasma isomiR levels in **(C)** the UK discovery cohort and **(D)** the PGB replication cohort clustered by hierarchical clustering according to euclidean distance and average linkage by UK discovery data. Pearson's correlation r from -1.0 (complete negative correlation, clue) to 1.0 (complete correlation, red).

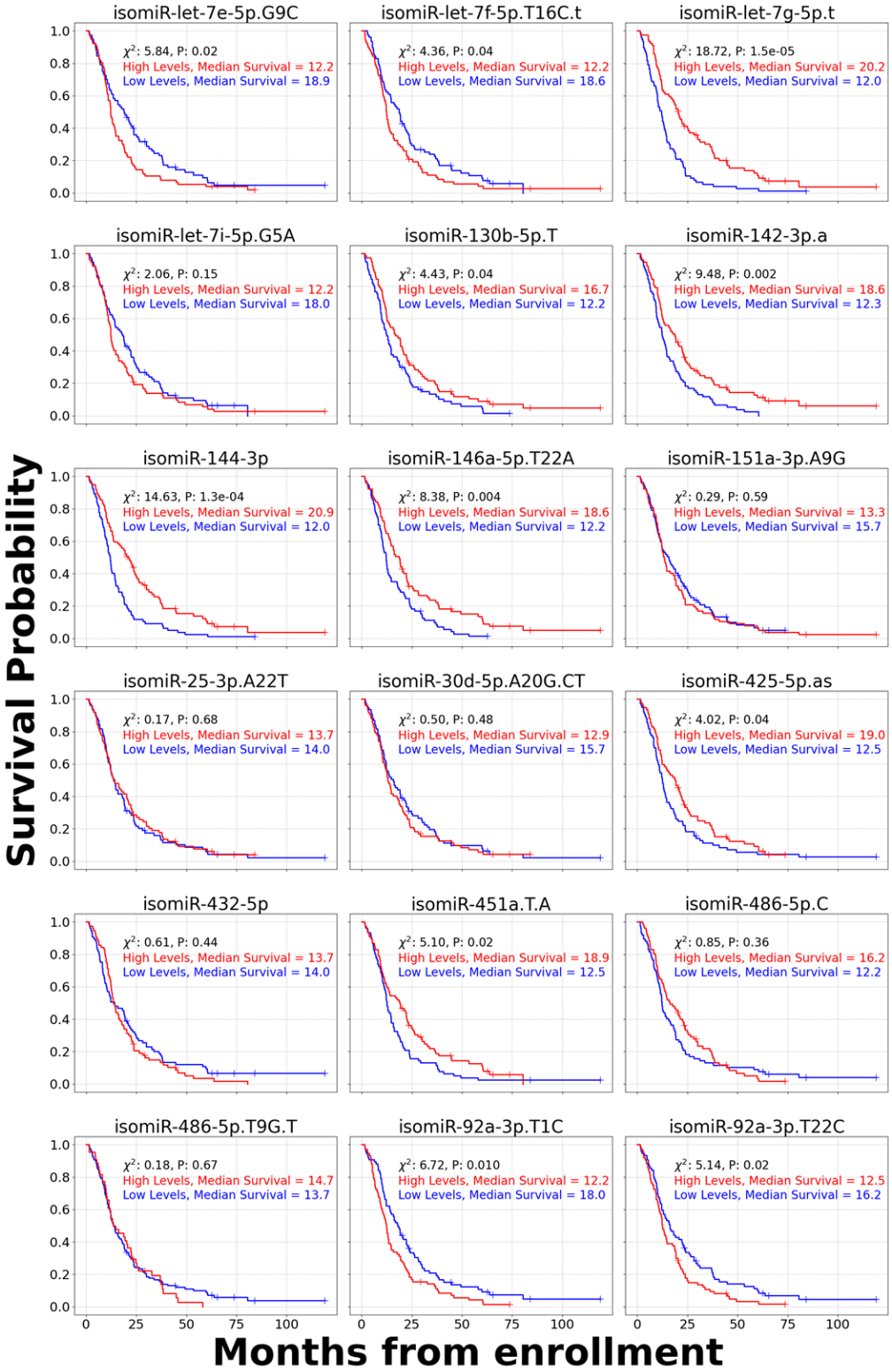

Figure S6 - Cumulative survival (Kaplan–Meier) curves for selected isomiRs in the UK discovery cohort. Data dichotomized by UK discovery isomiR median values. Subthreshold or suprathreshold groups are denoted in blue / red, respectively. Group median survival and Logrank statistics are presented.

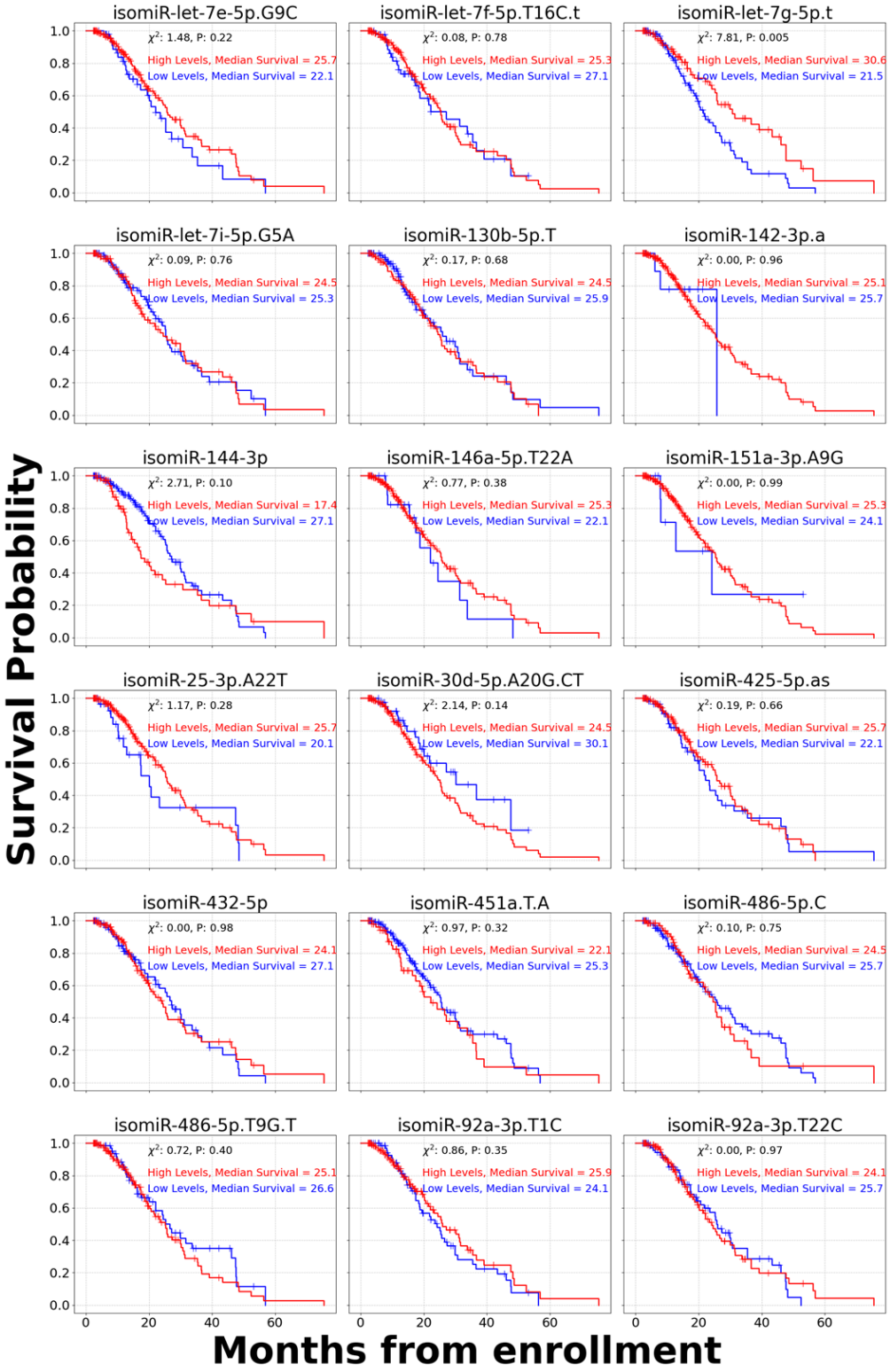

Figure S7 - Cumulative survival (Kaplan–Meier) curves for selected isomiRs in the PGB replication cohort. Data dichotomized by UK discovery isomiR median values. Subthreshold or suprathreshold groups are denoted in blue / red, respectively. Group median survival and Logrank statistics are presented.

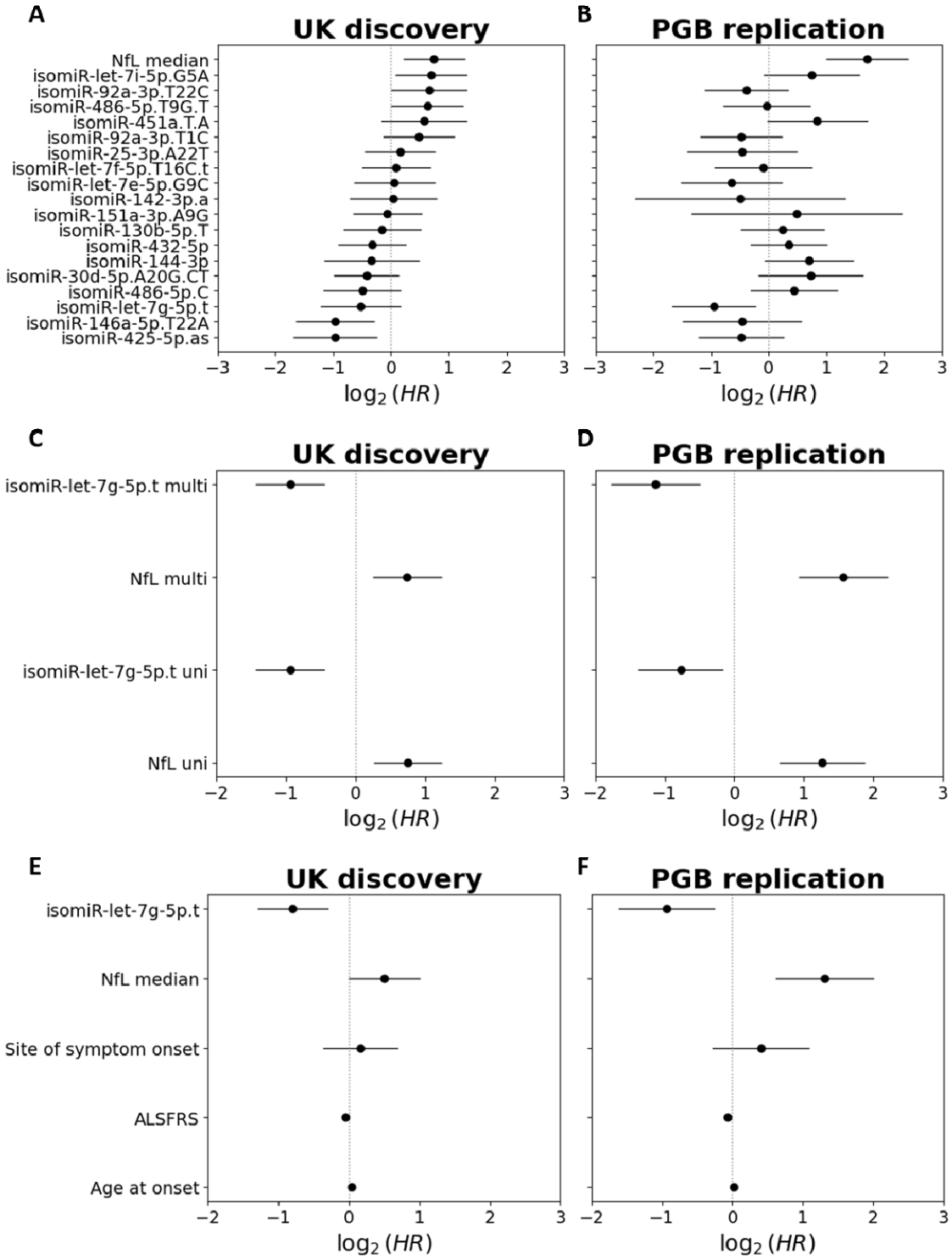

Figure S8 - Prognostication performance of isomiR-let-7g-5p.t. Forest plots of multivariate Cox proportional hazard ratios of 18 isomiRs and miR-181 in **(A)** the UK discovery or **(B)**PGB replication cohorts. Multivariate and univariate Cox hazard ratios of NfL and isomiR-let-7g-5p.t in the UK discovery **(C)** or PGB replication **(D)** cohorts Multivariate Cox hazard ratios of NfL, isomiR-let-7g-5p.t and clinical features in the UK discovery **(E)** or PGB replication **(F)**.

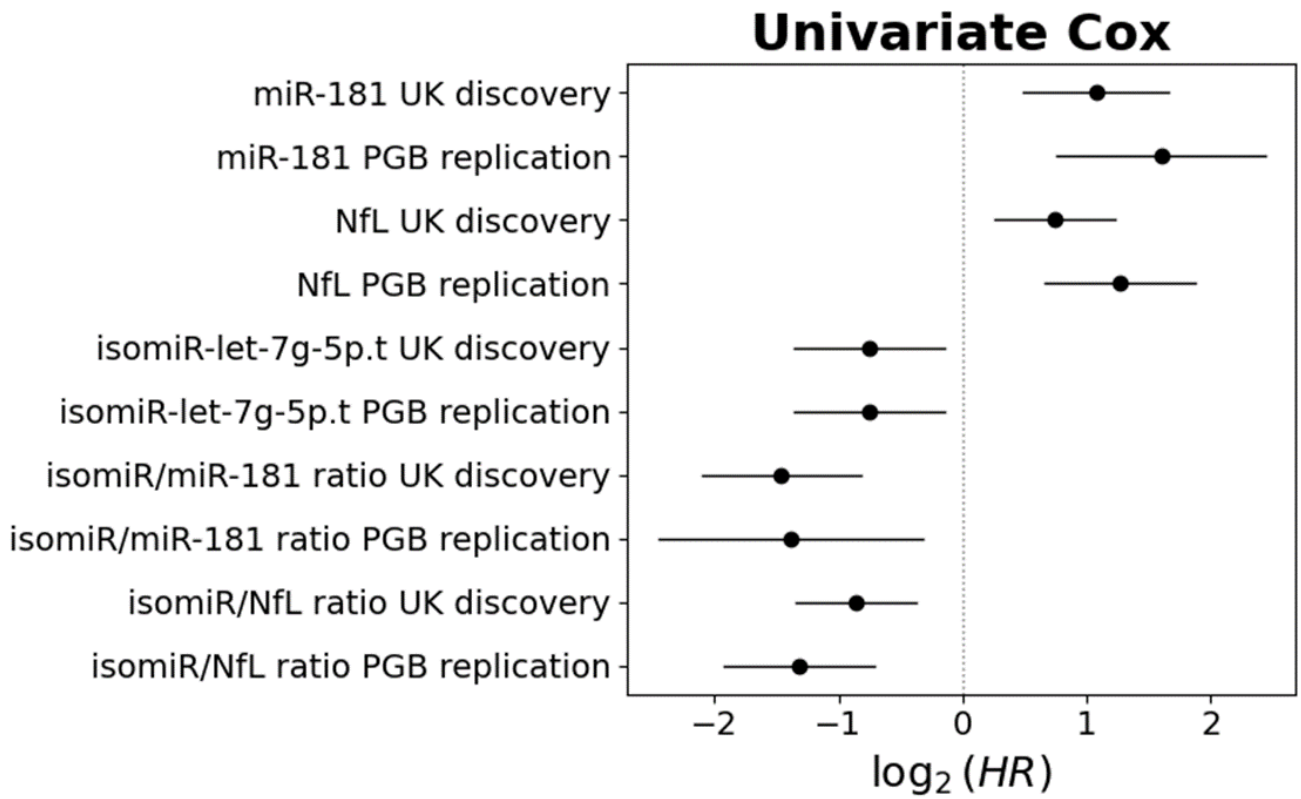

Figure S9 - Ratios of isomiR-let-7g-5p.t to NfL or miR-181 improve mortality association over single markers Forest plot of univariate Cox proportional hazard ratios of isomiR-let-7g-5p.t, NfL, miR-181 and the ratios of isomiR-let-7g-5p.t,/NfL or isomiR-let-7g-5p.t/miR-181, in both cohorts.

Supplementary methods

Models for feature selection were fitted with the following parameters:

- Cox: without regularization, ties = 'breslow' and a tolerance of $1e^{-9}$.
- Random Survival Forest: n_estimators = 2, max_depth = 4, max_samples = 0.25, n_estimators = 20, min_samples_split = 2, min_samples_leaf = 2, max_features = 0.5, max_samples = 0.5, min_weight_fraction_leaf = 0.0, max_leaf_nodes = None, bootstrap = True, random_state = 42).
- Survival Gradient Boosting: n_estimators = 20, learning_rate = 0.5, max_depth = 2, subsample = 0.3, loss = 'coxph', criterion = 'friedman_mse', min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, min_impurity_decrease = 0.0, max_features = None, max_leaf_nodes = None, validation_fraction = 0.1, n_iter_no_change = None, tol = 0.0001, dropout_rate = 0.3, ccp_alpha = 0.0,random_state = 0)

Survival Gradient Boosting hyperparameters post optuna and manuel tuning: n_estimators = 95, learning_rate = 0.1, max_depth = 35 subsample = 0.4545454545454545, loss = 'coxph', criterion = ‘friedman_mse’, min_samples_split = 60, min_samples_leaf = 20, min_weight_fraction_leaf = 0.060461799614030046, min_impurity_decrease = 0.6, max_features = 0.1111111111111111, max_leaf_nodes = 100, validation_fraction = 0.1, n_iter_no_change = None, tol = 0.0001, dropout_rate = 0.016, verbose = 1, ccp_alpha = 0.1, random_state = 42.

Acknowledgments

Eran Honrstein is the Mondry Family Professorial Chair and Head of the Nella and Leon Benoziyo Centre for Neurological Diseases and of the Andi and Larry Wolfe Centre for Neuroimmunology and Neuromodulation. We acknowledge patients for their contributions. We thank ALS biomarker and the Phenotype-Genotype biomarker studies co-workers and their contributions to the biobanking project, the Crown Genomics Team of Nancy and Stephen Grand Israel National Center for Personalized Medicine for small RNA sequencing. The authors thank participants in the CReATe Consortium's Phenotype-Genotype-Biomarker (PGB1) study; research staff at each clinical site; and the CReATe Consortium project management and data management teams, genomics sub-core, and biorepository

Funding

Research at the Hornstein lab is funded by Andi and Larry Wolfe Centre for Neuroimmunology and Neuromodulation; the Binational Science Foundation (BSF); Association Française Contre les Myopathies (AFM) grants 24882, 28680; Muscular Dystrophy Association (MDA) grant 1280000; Target ALS; Israel Science Foundation (ISF 3497/21, 424/22); ALS Canada; Minna-James-Heineman Stiftung through Minerva, Minerva Foundation, with funding from the Federal German Ministry for Education and Research; Robert Packard Center for ALS Research at Johns Hopkins; McGill University; EU - ERA-Net; Radala Foundation for ALS Research; Additional support generously provided by the Kekst Family Institute for Medical Genetics. Weizmann SABRA - Yeda-Sela - WRC Program, the Estate of Emile Mimran, and The Maurice and Vivienne Wohl Biology Endowment. Nella and Leon Benoziyo Center for Neurological Diseases, Goldhirsh-Yellin Foundation. Dr. Sydney Brenner and friends, Weizmann - Center for Research on Neurodegeneration,. Redhill Foundation – Sam and Jean Rothberg Charitable Trust Dr. Dvora and Haim Teitelbaum Endowment Fund and Weizmann – Institute for Artificial Intelligence - Acceleration Grant. The CReATe Consortium (U54NS092091) is part of the NIH Rare Diseases Clinical Research Network (RDCRN), an initiative of the Office of Rare Diseases Research (ORDR), NCATS. This Consortium is funded through a collaboration between NCATS and the NINDS. This work was also supported by a Clinical Trial Readiness grant (U01NS107027) from NINDS and by a grant from the ALS Association to support the CReATe Biorepository (grant ID 16-TACL-242).

References

1. Pfohl SR, Kim RB, Coan GS, Mitchell CS: **Unraveling the Complexity of Amyotrophic Lateral Sclerosis Survival Prediction.** *Front Neuroinform* 2018, **12:**36.
2. Beghi E, Chio A, Couratier P, Esteban J, Hardiman O, Logroscino G, Millul A, Mitchell D, Preux PM, Pupillo E, et al: **The epidemiology and treatment of ALS: focus on the heterogeneity of the disease and critical appraisal of therapeutic trials.** *Amyotroph Lateral Scler* 2011, **12:**1-10.
3. Benatar M, Boylan K, Jeromin A, Rutkove SB, Berry J, Atassi N, Bruijn L: **ALS biomarkers for therapy development: State of the field and future directions.** *Muscle Nerve* 2016, **53:**169-182.
4. Turner MR, Bowser R, Bruijn L, Dupuis L, Ludolph A, McGrath M, Manfredi G, Maragakis N, Miller RG, Pullman SL, et al: **Mechanisms, models and biomarkers in amyotrophic lateral sclerosis.** *Amyotroph Lateral Scler Frontotemporal Degener* 2013, **14 Suppl 1:**19-32.
5. Shepheard SR, Wuu J, Cardoso M, Wiklendt L, Dinning PG, Chataway T, Schultz D, Benatar M, Rogers M-L: **Urinary p75ECD: A prognostic, disease progression, and pharmacodynamic biomarker in ALS.** *Neurology* 2017, **88:**1137-1143.
6. Dergai O, Wuu J, Koziczak-Holbro M, Malaspina A, Granit V, Hernandez JP, Cooley A, Sachdev R, Yu L, Bidinosti M, et al: **Skeletal muscle biomarkers of amyotrophic lateral sclerosis: a large-scale, multi-cohort proteomic study.** *medRxiv* 2025.
7. Dobrowolny G, Martone J, Lepore E, Casola I, Petrucci A, Inghilleri M, Morlando M, Colantoni A, Scicchitano BM, Calvo A, et al: **A longitudinal study defined circulating microRNAs as reliable biomarkers for disease prognosis and progression in ALS human patients.** *Cell Death Discov* 2021, **7:**4.
8. Benatar M, Wuu J, Turner MR: **Neurofilament light chain in drug development for amyotrophic lateral sclerosis: a critical appraisal.** *Brain* 2023, **146:**2711-2716.
9. Benatar M, Macklin EA, Malaspina A, Rogers M-L, Hornstein E, Lombardi V, Renfrey D, Shepheard S, Magen I, Cohen Y, et al: **Prognostic clinical and biological markers for amyotrophic lateral sclerosis disease progression: validation and implications for clinical trial design and analysis.** *EBioMedicine* 2024, **108:**105323.
10. Lee RC, Feinbaum RL, Ambros V: **The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.** *Cell* 1993, **75:**843-854.
11. Haramati S, Chapnik E, Sztainberg Y, Eilam R, Zwang R, Gershoni N, McGlinn E, Heiser PW, Wills A-M, Wirguin I, et al: **miRNA malfunction causes spinal motor neuron disease.** *Proc Natl Acad Sci U S A* 2010, **107:**13111-13116.
12. Emde A, Eitan C, Liou L-L, Libby RT, Rivkin N, Magen I, Reichenstein I, Oppenheim H, Eilam R, Silvestroni A, et al: **Dysregulated miRNA biogenesis downstream of cellular stress and ALS-causing mutations: a new mechanism for ALS.** *EMBO J* 2015, **34:**2633-2651.
13. Katsu M, Hama Y, Utsumi J, Takashina K, Yasumatsu H, Mori F, Wakabayashi K, Shoji M, Sasaki H: **MicroRNA expression profiles of neuron-derived extracellular vesicles in plasma from patients with amyotrophic lateral sclerosis.** *Neurosci Lett* 2019, **708:**134176.
14. Toivonen JM, Manzano R, Oliván S, Zaragoza P, García-Redondo A, Osta R: **MicroRNA-206: a potential circulating biomarker candidate for amyotrophic lateral sclerosis.** *PLoS One* 2014, **9:**e89065.
15. Wakabayashi K, Mori F, Kakita A, Takahashi H, Utsumi J, Sasaki H: **Analysis of microRNA from archived formalin-fixed paraffin-embedded specimens of amyotrophic lateral sclerosis.** *Acta Neuropathol Commun* 2014, **2:**173.

16. Magen I, Yacovzada NS, Yanowski E, Coenen-Stass A, Grosskreutz J, Lu C-H, Greensmith L, Malaspina A, Fratta P, Hornstein E: **Circulating miR-181 is a prognostic biomarker for amyotrophic lateral sclerosis.** *Nat Neurosci* 2021, **24:**1534-1541.
17. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN: **The nuclear RNase III Drosha initiates microRNA processing.** *Nature* 2003, **425:**415-419.
18. Bernstein E, Caudy AA, Hammond SM, Hannon GJ: **Role for a bidentate ribonuclease in the initiation step of RNA interference.** *Nature* 2001, **409:**363-366.
19. Guo L, Chen F: **A challenge for miRNA: multiple isomiRs in miRNAomics.** *Gene* 2014, **544:**1-7.
20. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu A-L, Zhao Y, McDonald H, Zeng T, Hirst M, et al: **Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells.** *Genome Res* 2008, **18:**610-621.
21. Newman MA, Mani V, Hammond SM: **Deep sequencing of microRNA precursors reveals extensive 3' end modification.** *RNA* 2011, **17:**1795-1803.
22. Neilsen CT, Goodall GJ, Bracken CP: **IsomiRs--the overlooked repertoire in the dynamic microRNAome.** *Trends Genet* 2012, **28:**544-549.
23. Hardiman O, Al-Chalabi A, Chio A, Corr EM, Logroscino G, Robberecht W, Shaw PJ, Simmons Z, van den Berg LH: **Amyotrophic lateral sclerosis.** *Nature Reviews Disease Primers* 2017, **3:**1-19.
24. Brown A-L, Wilkins OG, Keuss MJ, Hill SE, Zanovello M, Lee WC, Bampton A, Lee FCY, Masino L, Qi YA, et al: **TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A.** *Nature* 2022, **603:**131-137.
25. Ling S-C, Albuquerque CP, Han JS, Lagier-Tourenne C, Tokunaga S, Zhou H, Cleveland DW: **ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS.** *Proc Natl Acad Sci U S A* 2010, **107:**13318-13323.
26. Ma XR, Prudencio M, Koike Y, Vatsavayai SC, Kim G, Harbinski F, Briner A, Rodriguez CM, Guo C, Akiyama T, et al: **TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A.** *Nature* 2022, **603:**124-130.
27. Melamed Ze, López-Erauskin J, Baughn MW, Zhang O, Drenner K, Sun Y, Freyermuth F, McMahon MA, Beccari MS, Artates JW, et al: **Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration.** *Nature Neuroscience* 2019, **22:**180-190.
28. Calliari A, Daughrity LM, Albagli EA, Castellanos Otero P, Yue M, Jansen-West K, Islam NN, Caulfield T, Rawlinson B, DeTure M, et al: **HDGFL2 cryptic proteins report presence of TDP-43 pathology in neurodegenerative diseases.** *Molecular Neurodegeneration* 2024, **19:**1-5.
29. Irwin KE, Jasin P, Braunstein KE, Sinha IR, Garret MA, Bowden KD, Chang K, Troncoso JC, Moghekar A, Oh ES, et al: **A fluid biomarker reveals loss of TDP-43 splicing repression in presymptomatic ALS–FTD.** *Nature Medicine* 2024, **30:**382-393.
30. Chen X, Fan Z, McGee W, Chen M, Kong R, Wen P, Xiao T, Chen X, Liu J, Zhu L, et al: **TDP-43 regulates cancer-associated microRNAs.** *Protein Cell* 2018, **9:**848-866.
31. Ibuki Y, Nishiyama Y, Tsutani Y, Emi M, Hamai Y, Okada M, Tahara H: **Circulating microRNA/isomiRs as novel biomarkers of esophageal squamous cell carcinoma.** *PLoS One* 2020, **15:**e0231116.
32. Koppers-Lalic D, Hackenberg M, de Menezes R, Misovic B, Wachalska M, Geldof A, Zini N, de Reijke T, Wurdinger T, Vis A, et al: **Non‐invasive prostate cancer detection by measuring miRNA variants (isomiRs) in urine extracellular vesicles.** *Oncotarget* 2016, **7:**22566-22578.

33. Shepheard SR, Wuu J, Cardoso M, Wiklendt L, Dinning PG, Chataway T, Schultz D, Benatar M, Rogers ML: **Urinary p75(ECD): A prognostic, disease progression, and pharmacodynamic biomarker in ALS.** *Neurology* 2017, **88:**1137-1143.
34. Cox DR: **Regression Models and Life-Tables.** *J R Stat Soc Series B Stat Methodol* 1972, **34:**187-202.
35. Hothorn T, Bühlmann P, Dudoit S, Molinaro A, Van Der Laan MJ: **Survival ensembles.** *Biostatistics* 2005, **7:**355-373.
36. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS: **Random survival forests.** *aoas* 2008, **2:**841-860.
37. Breiman L: **Bagging predictors.** *Mach Learn* 1996, **24:**123-140.
38. Džeroski S, Ženko B: **Is Combining Classifiers with Stacking Better than Selecting the Best One?** *Mach Learn* 2004, **54:**255-273.
39. Malviya R, Ghinea G, Dhanaraj RK, Balusamy B, Sundram S: *Deep Learning for Targeted Treatments: Transformation in Healthcare.* John Wiley & Sons; 2022.
40. Van Belle V, Pelckmans K, Van Huffel S, Suykens JAK: **Support vector methods for survival analysis: a comparison between ranking and regression approaches.** *Artif Intell Med* 2011, **53:**107-118.
41. **ClinicalTrials.gov/study/NCT02327845.**
42. Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, Nakanishi A: **The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III).** *J Neurol Sci* 1999, **169:**13-21.
43. Strong MJ, Abrahams S, Goldstein LH, Woolley S, McLaughlin P, Snowden J, Mioshi E, Roberts-South A, Benatar M, HortobáGyi T, et al: **Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria.** *Amyotroph Lateral Scler Frontotemporal Degener* 2017, **18:**153-174.
44. Abrahams S, Newton J, Niven E, Foley J, Bak TH: **Screening for cognition and behaviour changes in ALS.** *Amyotroph Lateral Scler Frontotemporal Degener* 2014, **15:**9-14.
45. Coenen-Stass AML, Magen I, Brooks T, Ben-Dov IZ, Greensmith L, Hornstein E, Fratta P: **Evaluation of methodologies for microRNA biomarker detection by next generation sequencing.** *RNA Biol* 2018, **15:**1133-1145.
46. Kohen R, Barlev J, Hornung G, Stelzer G, Feldmesser E, Kogan K, Safran M, Leshkowitz D: **UTAP: User-friendly Transcriptome Analysis Pipeline.** *BMC Bioinformatics* 2019, **20:**154.
47. Kozomara A, Griffiths-Jones S: **miRBase: annotating high confidence microRNAs using deep sequencing data.** *Nucleic Acids Res* 2014, **42:**D68-73.
48. Ewing B, Hillier L, Wendl MC, Green P: **Base-Calling of Automated Sequencer Traces UsingPhred. I. Accuracy Assessment.** *Genome Res* 1998, **8:**175-185.
49. Lall R, Robinson T: **Efficient Multiple Imputation for Diverse Data in Python and R: MIDASpy and rMIDAS.**
50. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biol* 2014, **15:**550.
51. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD: **The sva package for removing batch effects and other unwanted variation in high-throughput experiments.** *Bioinformatics* 2012, **28:**882-883.
52. Raschka S: **MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack.** *Journal of Open Source Software* 2018, **3:**638.
53. Harrell FE, Jr., Lee KL, Mark DB: **Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.** *Stat Med* 1996, **15:**361-387.
54. Thorndike RL: **Who belongs in the family?** *Psychometrika* 1953, **18:**267-276.

55. **Optuna** [https://dl.acm.org/doi/10.1145/3292500.3330701 http://dx.doi.org/10.1145/3292500.3330701]
56. Lundberg SM, Lee S-I: **A Unified Approach to Interpreting Model Predictions.** *Advances in Neural Information Processing Systems* 2017, **30**.
57. Sundrani S, Lu J: **Computing the hazard ratios associated with explanatory variables using machine learning models of survival data.** *JCO Clin Cancer Inform* 2021, **5:**364-378.
58. Lambert J, Chevret S: **Summary measure of discrimination in survival models based on cumulative/dynamic time-dependent ROC curves.** *Stat Methods Med Res* 2016, **25:**2088-2102.
59. **ISRCTN.**
60. Magen I, Yacovzada N-S, Warren JD, Heller C, Swift I, Bobeva Y, Malaspina A, Rohrer JD, Fratta P, Hornstein E: **microRNA-based predictor for diagnosis of frontotemporal dementia.** *Neuropathol Appl Neurobiol* 2023**:**e12916.
61. Charilaou P, Battat R: **Machine learning models and over-fitting considerations.** *World J Gastroenterol* 2022, **28:**605-607.
62. Taylor JMG, Ankerst DP, Andridge RR: **Validation of biomarker-based risk prediction models.** *Clin Cancer Res* 2008, **14:**5977-5983.
63. Lu C-H, Macdonald-Wallis C, Gray E, Pearce N, Petzold A, Norgren N, Giovannoni G, Fratta P, Sidle K, Fish M, et al: **Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis.** *Neurology* 2015, **84:**2247-2257.
64. Koval ED, Shaner C, Zhang P, du Maine X, Fischer K, Tay J, Chau BN, Wu GF, Miller TM: **Method for widespread microRNA-155 inhibition prolongs survival in ALS-model mice.** *Hum Mol Genet* 2013, **22:**4127-4135.
65. Liguori M, Nuzziello N, Introna A, Consiglio A, Licciulli F, D'Errico E, Scarafino A, Distaso E, Simone IL: **Dysregulation of MicroRNAs and Target Genes Networks in Peripheral Blood of Patients With Sporadic Amyotrophic Lateral Sclerosis.** *Front Mol Neurosci* 2018, **11:**288.
66. Maimon R, Ionescu A, Bonnie A, Sweetat S, Wald-Altman S, Inbar S, Gradus T, Trotti D, Weil M, Behar O, Perlson E: **miR126-5p Downregulation Facilitates Axon Degeneration and NMJ Disruption via a Non-Cell-Autonomous Mechanism in ALS.** *J Neurosci* 2018, **38:**5478-5494.
67. Raheja R, Regev K, Healy BC, Mazzola MA, Beynon V, Von Glehn F, Paul A, Diaz-Cruz C, Gholipour T, Glanz BI, et al: **Correlating serum micrornas and clinical parameters in amyotrophic lateral sclerosis.** *Muscle Nerve* 2018, **58:**261-269.
68. Reichenstein I, Eitan C, Diaz-Garcia S, Haim G, Magen I, Siany A, Hoye ML, Rivkin N, Olender T, Toth B, et al: **Human genetics and neuropathology suggest a link between miR-218 and amyotrophic lateral sclerosis pathophysiology.** *Sci Transl Med* 2019, **11**.
69. Saucier D, Wajnberg G, Roy J, Beauregard A-P, Chacko S, Crapoulet N, Fournier S, Ghosh A, Lewis SM, Marrero A, et al: **Identification of a circulating miRNA signature in extracellular vesicles collected from amyotrophic lateral sclerosis patients.** *Brain Res* 2019, **1708:**100-108.
70. Kapeli K, Martinez FJ, Yeo GW: **Genetic mutations in RNA-binding proteins and their roles in ALS.** *Hum Genet* 2017, **136:**1193-1214.
71. Kawahara Y, Mieda-Sato A: **TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes.** *Proc Natl Acad Sci U S A* 2012, **109:**3347-3352.
72. Kour S, Fortuna T, Anderson EN, Mawrie D, Bilstein J, Sivasubramanian R, Ward C, Roy R, Rajasundaram D, Sterneckert J, Pandey UB: **Drosha-dependent microRNAs modulate FUS-mediated neurodegeneration in vivo.** *Nucleic Acids Res* 2023, **51:**11258-11276.

73. Ma H, Wu Y, Choi J-G, Wu H: **Lower and upper stem-single-stranded RNA junctions together determine the Drosha cleavage site.** *Proc Natl Acad Sci U S A* 2013, **110:**20687-20692.
74. Ma H, Wu Y, Niu Q, Zhang J, Jia G, Manjunath N, Wu H: **A sliding-bulge structure at the Dicer processing site of pre-miRNAs regulates alternative Dicer processing to generate 5'-isomiRs.** *Heliyon* 2016, **2:**e00148.
75. Jo M, Lee S, Jeon Y-M, Kim S, Kwon Y, Kim H-J: **The role of TDP-43 propagation in neurodegenerative diseases: integrating insights from clinical and experimental studies.** *Exp Mol Med* 2020, **52:**1652-1662.
76. Kmetzsch V, Latouche M, Saracino D, Rinaldi D, Camuzat A, Gareau T, Le Ber I, Colliot O, Becker E: **MicroRNA signatures in genetic frontotemporal dementia and amyotrophic lateral sclerosis.** *Annals of Clinical and Translational Neurology* 2022, **9:**1778-1791.
77. Lin Z, Ge J, Wang Z, Ren J, Wang X, Xiong H, Gao J, Zhang Y, Zhang Q: **Let-7e modulates the inflammatory response in vascular endothelial cells through ceRNA crosstalk.** *Scientific Reports* 2017, **7:**1-12.
78. Tan W, Gu Z, Leng J, Zou X, Chen H, Min F, Zhou W, Zhang L, Li G: **Let-7f-5p ameliorates inflammation by targeting NLRP3 in bone marrow-derived mesenchymal stem cells in patients with systemic lupus erythematosus.** *Biomed Pharmacother* 2019, **118:**109313.
79. Ma N, Li S, Lin C, Cheng X, Meng Z: **Mesenchymal stem cell conditioned medium attenuates oxidative stress injury in hepatocytes partly by regulating the miR-486-5p/PIM1 axis and the TGF-β/Smad pathway.** *Bioengineered* 2021.
80. Liu Q, Hu Y, Zhang M, Yan Y, Yu H, Ge L: **microRNA-451 protects neurons against ischemia/reperfusion injury-induced cell death by targeting CELF2.** *Neuropsychiatric Disease and Treatment* 2018.
81. Luo Q, Zhu J, Zhang Q, Xie J, Yi C, Li T: **MicroRNA-486-5p Promotes Acute Lung Injury via Inducing Inflammation and Apoptosis by Targeting OTUD7B.** *Inflammation* 2020, **43:**975-984.
82. Mandolesi G, De Vito F, Musella A, Gentile A, Bullitta S, Fresegna D, Sepman H, Di Sanza C, Haji N, Mori F, et al: **miR-142-3p Is a Key Regulator of IL-1β-Dependent Synaptopathy in Neuroinflammation.** *The Journal of Neuroscience* 2017, **37:**546.
83. Li K, Zhang J, Ji C, Wang L: **MiR-144-3p and Its Target Gene beta-Amyloid Precursor Protein Regulate 1-Methyl-4-Phenyl-1,2-3,6-Tetrahydropyridine-Induced Mitochondrial Dysfunction.** *Mol Cells* 2016, **39:**543-549.
84. Thomas JJ, Harp KO, Bashi A, Hood JL, Botchway F, Wilson MD, Thompson WE, Stiles JK, Driss A: **MiR-451a and let-7i-5p loaded extracellular vesicles attenuate heme-induced inflammation in hiPSC-derived endothelial cells.** *Front Immunol* 2022, **13:**1082414.
85. Gan X, Zhao H, Wei Y, Jiang Q, Wen C, Ying Y: **Role of miR-92a-3p, oxidative stress, and p38MAPK/NF-κB pathway in rats with central venous catheter related thrombosis.** *BMC Cardiovascular Disorders* 2020, **20:**1-8.
86. Strycharz J, Wróblewski A, Zieleniak A, Świderska E, Matyjas T, Rucińska M, Pomorski L, Czarny P, Szemraj J, Drzewoski J, Śliwińska A: **Visceral Adipose Tissue of Prediabetic and Diabetic Females Shares a Set of Similarly Upregulated microRNAs Functionally Annotated to Inflammation, Oxidative Stress and Insulin Signaling.** *Antioxidants* 2021, **10:**101.