

Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model



Henk-Jan Westeneng, Thomas P A Debray, Anne E Visser, Ruben P A van Eijk, James P K Rooney, Andrea Calvo, Sarah Martin, Christopher J McDermott, Alexander G Thompson, Susana Pinto, Xenia Kobeleva, Angela Rosenbohm, Beatrice Stubendorff, Helma Sommer, Bas M Middelkoop, Annelot M Dekker, Joke J F A van Vugt, Wouter van Rheenen, Alice Vajda, Mark Heverin, Mbombe Kazoka, Hannah Hollinger, Marta Gromicho, Sonja Körner, Thomas M Ringer, Annkathrin Rödiger, Anne Gunkel, Christopher E Shaw, Annelien L Bredenoord, Michael A van Es, Philippe Corcia, Philippe Couratier, Markus Weber, Julian Grosskreutz, Albert C Ludolph, Susanne Petri, Mamede de Carvalho, Philip Van Damme, Kevin Talbot, Martin R Turner, Pamela J Shaw, Ammar Al-Chalabi, Adriano Chiò, Orla Hardiman, Karel G M Moons, Jan H Veldink, Leonard H van den Berg

Summary

Background Amyotrophic lateral sclerosis (ALS) is a relentlessly progressive, fatal motor neuron disease with a variable natural history. There are no accurate models that predict the disease course and outcomes, which complicates risk assessment and counselling for individual patients, stratification of patients for trials, and timing of interventions. We therefore aimed to develop and validate a model for predicting a composite survival endpoint for individual patients with ALS.

Methods We obtained data for patients from 14 specialised ALS centres (each one designated as a cohort) in Belgium, France, the Netherlands, Germany, Ireland, Italy, Portugal, Switzerland, and the UK. All patients were diagnosed in the centres after excluding other diagnoses and classified according to revised El Escorial criteria. We assessed 16 patient characteristics as potential predictors of a composite survival outcome (time between onset of symptoms and non-invasive ventilation for more than 23 h per day, tracheostomy, or death) and applied backward elimination with bootstrapping in the largest population-based dataset for predictor selection. Data were gathered on the day of diagnosis or as soon as possible thereafter. Predictors that were selected in more than 70% of the bootstrap resamples were used to develop a multivariable Royston-Parmar model for predicting the composite survival outcome in individual patients. We assessed the generalisability of the model by estimating heterogeneity of predictive accuracy across external populations (ie, populations not used to develop the model) using internal-external cross-validation, and quantified the discrimination using the concordance (*c*) statistic (area under the receiver operator characteristic curve) and calibration using a calibration slope.

Findings Data were collected between Jan 1, 1992, and Sept 22, 2016 (the largest data-set included data from 1936 patients). The median follow-up time was 97.5 months (IQR 52.9–168.5). Eight candidate predictors entered the prediction model: bulbar versus non-bulbar onset (univariable hazard ratio [HR] 1.71, 95% CI 1.63–1.79), age at onset (1.03, 1.03–1.03), definite versus probable or possible ALS (1.47, 1.39–1.55), diagnostic delay (0.52, 0.51–0.53), forced vital capacity (HR 0.99, 0.99–0.99), progression rate (6.33, 5.92–6.76), frontotemporal dementia (1.34, 1.20–1.50), and presence of a *C9orf72* repeat expansion (1.45, 1.31–1.61), all $p < 0.0001$. The *c* statistic for external predictive accuracy of the model was 0.78 (95% CI 0.77–0.80; 95% prediction interval [PI] 0.74–0.82) and the calibration slope was 1.01 (95% CI 0.95–1.07; 95% PI 0.83–1.18). The model was used to define five groups with distinct median predicted (SE) and observed (SE) times in months from symptom onset to the composite survival outcome: very short 17.7 (0.20), 16.5 (0.23); short 25.3 (0.06), 25.2 (0.35); intermediate 32.2 (0.09), 32.8 (0.46); long 43.7 (0.21), 44.6 (0.74); and very long 91.0 (1.84), 85.6 (1.96).

Interpretation We have developed an externally validated model to predict survival without tracheostomy and non-invasive ventilation for more than 23 h per day in European patients with ALS. This model could be applied to individualised patient management, counselling, and future trial design, but to maximise the benefit and prevent harm it is intended to be used by medical doctors only.

Funding Netherlands ALS Foundation.

Copyright © 2018 Elsevier Ltd. All rights reserved.

Introduction

Neurodegenerative diseases impose an enormous clinical and economic burden on patients and health systems.¹ Development of disease-modifying therapies

and strategies for effective palliative care have been limited by disease heterogeneity and the presence of overlapping phenotypes.² Therefore, models that can reliably predict outcomes at an individual patient level

Lancet Neurol 2018; 17: 423–33

Published Online

March 26, 2017

[http://dx.doi.org/10.1016/S1474-4422\(18\)30089-9](http://dx.doi.org/10.1016/S1474-4422(18)30089-9)

See [Comment](#) page 386

Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, Netherlands (H-J Westeneng MD, A E Visser MD, R P A van Eijk MD, B M Middelkoop, A M Dekker MD, J J F A van Vugt PhD, W van Rheenen MD, M A van Es MD, Prof J H Veldink MD, Prof L H van den Berg MD); Department of Epidemiology, Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, Netherlands (T P A Debray PhD, A L Bredenoord PhD, Prof K G M Moons PhD); Cochrane Netherlands, University Medical Centre Utrecht, Utrecht, Netherlands (T P A Debray, Prof K G M Moons); Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College, Dublin, Ireland (J P K Rooney MSc, A Vajda PhD, M Heverin MSc, Prof O Hardiman MD); 'Rita Levi Montalcini' Department of Neuroscience, ALS Centre, University of Torino, Torino, Italy (A Calvo MD, Prof A Chiò MD); Maurice Wohl Clinical Neuroscience Institute, and United Kingdom Dementia Research Institute at the Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK (S Martin BSc, Prof C E Shaw MD, Prof A Al-Chalabi PhD); Sheffield Institute for Translational Neuroscience, University of

Sheffield, Sheffield, UK (Prof C J McDermott PhD, M Kazoka, H Hollinger MA, Prof P J Shaw MD); **Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK** (A G Thompson BMBCh, Prof K Talbot DPhil, Prof M R Turner MD); **Institute of Physiology-Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, Lisbon, Portugal** (S Pinto PhD, M Gromicho PhD, Prof M de Carvalho MD); **Department of Neurology, Hannover Medical School, Hannover, Germany** (X Kobleva MD, S Körner MD, Prof S Petri MD); **Department of Neurology, University of Ulm, Ulm, Germany** (A Rosenbohm MD, Prof A C Ludolph MD); **Hans-Berger Department of Neurology, Jena University Hospital, Jena, Germany** (B Stübendorff PhD, T M Ringer MD, A Rödiger MD, A Gunkel MD, J Grosskreutz MD); **Neuromuscular Diseases Centre/ALS Clinic, Kantonsspital St Gallen, St Gallen, Switzerland** (H Sommer, Prof M Weber MD); **Centre de compétence SLA-fédération Tours-Limoges, CHU de Tours, Tours, France** (Prof P Corcia MD); **Centre de compétence SLA-fédération Tours-Limoges, CHU de Limoges, Limoges, France** (Prof P Couratier MD); **Department of Neurology, University Hospital Leuven, Leuven, Belgium, and Department of Neurosciences, Katholieke Universiteit, Leuven (University of Leuven) and Centre for Brain and Disease Research, VIB, Leuven, Belgium** (Prof P Van Damme MD); and **Department of Neurology, Beaumont Hospital, Beaumont, Ireland** (Prof O Hardiman)

Correspondence to: Prof Leonard H van den Berg, Department of Neurology, University Medical Centre Utrecht, Utrecht, 3508 GA, Netherlands. l.h.vandenbergh@umcutrecht.nl

Research in context

Evidence before this study

Prediction of survival in individual patients with amyotrophic lateral sclerosis (ALS) is important but remains elusive. We searched PubMed without language restrictions using the terms “([amyotrophic lateral sclerosis] OR [motor neuron disease]) AND [survival] OR [mortality] OR [prognosis] OR [prognostic factor] OR [prediction model]” for articles on validated models that predict survival at an individual level and articles on clinical, cognitive, and genetic predictors of survival in ALS, published between Jan 1, 1960, and Dec 31, 2017. We searched for well studied predictors of survival that can be collected during diagnostic work-up and excluded factors that require longitudinal assessments (eg, slope of forced vital capacity) or that are rare (eg, *SOD A4V* mutations). We identified 16 variables as predictors of survival in ALS in at least one peer-reviewed article, but replication was frequently absent, outcomes of studies were contradictory, and studies used different sets of covariates in multivariable analyses or adopted different variable definitions, which complicated a formal synthesis of previous findings. Two studies about prediction of disease progression in ALS (but not survival) were identified. These studies used data from 1822 and 3742 patients from the Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) database. This database includes patients from 23 clinical trials, but not from population-based studies or patient registries. We identified no sufficiently powered, externally validated, prediction models of survival in individual patients with ALS.

Added value of this study

In our study of 11 475 patients with ALS from 14 European ALS centres, we derived and validated, according to the TRIPOD guidelines, a multivariable model for the prediction of survival

without tracheostomy or non-invasive ventilation for more than 23 h per day, for individual patients, based on clinical, cognitive, and genetic predictors that were defined at diagnosis. Most of these predictors can be collected cheaply, simply, and non-invasively, rendering them useful for the assessment of individual prognosis in an outpatient clinic setting. We assessed the external validity of our model across several populations, and showed that it had a probability of more than 95% for good performance. We defined five subgroups of patients with distinctive survival times (very short, short, intermediate, long, and very long), and illustrated the clinical applicability of our model to individual patients. All parameters and equations of our model are provided in our appendix to allow improvement of predictions through continuing research on other prognostic factors or outcome measures.

Implications of all the available evidence

Our study proposes a set of independent predictors of survival without tracheostomy or non-invasive ventilation for more than 23 h per day in patients with ALS and provides a validated prediction model. To minimise the risk of potential harm to patients if the prediction they receive is shorter than expected, the online model will be accessible only to medical doctors, who have to register and sign in before access is provided. We encourage implementation of the model in evidence-based, tailored counselling of individual patients and their caregivers, and in multidisciplinary care practices, by facilitating the selection of appropriate care pathways (eg, frequency of consultations involving specific disciplines), and in innovative trial design, by refining inclusion and exclusion criteria, stratification, and subgroup analyses.

could become important factors in a precision-medicine approach by improving the potential for prognostic counselling, stratification of patients for trials, and timing of interventions.²

Amyotrophic lateral sclerosis (ALS) is one of the most devastating neurodegenerative diseases. It predominantly affects motor neurons in the brain and spinal cord, leading to weakness of voluntary muscles.^{3,4} Muscle weakness progresses gradually, spreading from a site of clinical onset; patients eventually become paralysed and die, usually as the result of respiratory failure.^{3,4} The clinical features of ALS are heterogeneous: the disease can occur at any adult age, up to 15% of patients develop frontotemporal dementia,⁵ and 10–15% of patients have a family history of ALS or frontotemporal dementia.⁴

Survival for patients with ALS varies greatly, ranging from several months to more than 10 years. Determinants of survival at group level have been studied extensively:⁶ patient characteristics such as older age at onset of symptoms, the presence of frontotemporal dementia, or a repeat expansion in *C9orf72* have been associated with shorter survival.^{6,7} Unfortunately, despite the well

documented natural history of ALS,^{3,6} prediction of survival in individual patients remains elusive. It is important that such risk predictions are sufficiently accurate across different settings and populations. The development and validation of such models is, however, challenging because few large datasets with individual participant data and adequate follow-up times are available.^{8–10}

We therefore analysed clinical, cognitive, and genetic data of patients with ALS from ALS centres in Europe with a view to predicting a composite survival outcome—time between onset of symptoms and non-invasive ventilation for more than 23 h per day, tracheostomy, or death—for individual patients on the day of diagnosis. We aimed to develop and externally validate a prediction model in multiple cohorts.

Methods

Study design and participants

We did a two-stage study using clinical, cognitive, and genetic data for patients with ALS from 14 specialised ALS centres (each one designated as a cohort) in Belgium,

France, the Netherlands, Germany, Ireland, Italy, Portugal, Switzerland, and the UK. In the first stage, predictor selection was done in the largest cohort, with data from consecutive incident patients, diagnosed with ALS according to the revised El Escorial criteria, who participated in an ongoing prospective, population-based study in the Netherlands between Jan 1, 2006, and March 31, 2015.¹¹ In the second stage, data for identified predictors and survival outcomes were requested from 13 other centres to develop and externally validate the model. All patients were diagnosed in the centres after excluding other diagnoses, and classified according to the revised El Escorial criteria.¹² Written informed consent was obtained from all participants and institutional review boards approved this study (appendix). The study is reported in accordance with the TRIPOD guidance for transparent reporting of prediction models.¹⁰

Procedures

Survival analyses in ALS research generally use composite endpoints (ie, events) comprising both death and respiratory events.^{13,14} In our study, survival was defined as time between onset of symptoms and our composite endpoint of non-invasive ventilation for more than 23 h per day, or tracheostomy, or death. We assessed ten clinical, four cognitive, and two genetic characteristics. All predictors were selected on the basis of a PubMed search.

Clinical predictors were sex, site of onset (spinal vs bulbar), age at onset of weakness or bulbar symptoms, revised El Escorial criteria (definite vs probable or possible ALS),¹² diagnostic delay (time from onset of weakness or bulbar symptoms to diagnosis), forced vital capacity (FVC; percentage of predicted based on normative values for age, sex, and body height), progression rate defined by the slope on the revised ALS Functional Rating Scale (ALSFERS-R),¹⁵ premorbid body-mass index (BMI),¹⁶ current smoking,¹⁷ and cigarette pack-years.⁶

Cognitive predictors were the presence of frontotemporal dementia according to the Neary or Rascovsky criteria,¹⁸ and scores on the verbal fluency index,¹⁹ the Frontal Assessment Battery,²⁰ and the ALS Frontotemporal Dementia Questionnaire (ALSFTD-Q).²¹ For the verbal fluency index participants were asked to name as many words as possible beginning with the letter D in 3 min, following which the participants were instructed to read aloud the generated words. The index was calculated as the number of words generated divided by the time needed to read them aloud. The frontal assessment battery is a bedside test comprising six tasks (maximum of 3 points per task) measuring conceptualisation, mental flexibility, motor programming, sensitivity to interference, inhibition control, and environmental autonomy; this test is sensitive to frontal lobe dysfunction. The ALSFERS-Q is a 25-item questionnaire (maximum of 4 points per item) applied to a proxy (ie, a caregiver who is able to assess the patient's

behaviour) that is developed to screen for behavioural disturbances in the patient with ALS.

Genetic predictors were presence of a *C9orf72* mutation, and presence of the minor allele homozygous genotype (C/C) of the *UNC13A* single nucleotide polymorphism, which were determined as previously described.^{7,22–24} These two genetic predictors are the most frequent mutations in patients with ALS and were previously shown to be associated with survival outcomes.^{7,22}

Data were gathered on the day of diagnosis or as soon as possible thereafter. The ALSFERS-R slope was determined as follows: (48 minus ALSFERS-R score)/(date of the ALSFERS-R score minus date of onset).¹⁵ ALSFTD-Q scores were trichotomised into no, mild, and severe behavioural impairment.²¹ A detailed overview of predictors is reported in the appendix.

All cohorts used the ALSFERS-R score, except the cohort in Oxford (UK), which used the unrevised version of the ALSFERS. The ten questions of the unrevised version are identical to the first ten questions of the ALSFERS-R, but the ALSFERS-R has two additional questions about respiration. The maximum score per question is 4, resulting in a maximum score of 48 for the ALSFERS-R and 40 for the ALSFERS. To be able to compare scores, we transformed the ALSFERS score to the ALSFERS-R score by multiplying by 1.2. Furthermore, all cohorts used FVC to measure respiratory function, except the cohort in Dublin (Ireland), which used sniff nasal inspiratory pressure (SNIP), which is known to be correlated with FVC.²⁵ Based on patients in the Dublin cohort for whom we had information about both FVC and SNIP, we transformed SNIP data of all patients to FVC to allow comparability between cohorts.

We explored patient preferences with regard to knowing their personalised prognosis by doing an online survey in 242 Dutch patients with ALS (appendix).

Statistical analysis

We used a backward elimination procedure with bootstrapping to select predictors of the composite survival outcome (appendix).²⁶ Predictors that were selected in more than 70% of the bootstrap resamples entered a multivariable prediction model.²⁶ Data from the 14 cohorts were combined using the internal–external cross-validation framework.^{27,28} This developed a model for predicting our composite survival endpoint in all but one cohort, after which its external validity was evaluated in the omitted cohort. The process was repeated for all 14 cohorts (every cohort being omitted once), yielding multiple estimates of external validity for a given modelling strategy. A meta-analysis was done to assess the overall performance of the model and to identify sources of between-study heterogeneity,²⁹ such as differences in survival time or differences in the definition of predictors between different cohorts, which might affect the generalisability of a model. The internal–external cross-validation framework thus

See Online for appendix

	Utrecht, Netherlands	Dublin, Ireland	Torino, Italy	Sheffield, UK	London, UK	Oxford, UK	Leuven, Belgium	Lisbon, Portugal	Hannover, Germany	Ulm, Germany	Jena, Germany	St Gallen, Switzerland	Tours, France	Limoges, France	Total	Hazard ratio (95% CI)*	
Data source	Population-based	Population-based	Population-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Total	Hazard ratio (95% CI)*	
N	1936	1818	1022	1187	1266	849	833	594	506	443	343	286	190	202	11475 (100.0)	..	
Duration of data acquisition (years)	9.2	20.5	9.2	15.7	23.1	15.2	23.8	17.4	12.9	6.8	8.0	13.0	5.2	5.2	12.9 (8.3-17.0)	..	
Age at onset (years)	62.7 (55.4-69.8)	65.3 (56.9-72.4)	68.0 (60.2-74.3)	62.3 (52.6-69.9)	63.1 (54.1-70.7)	63.4 (53.8-71.1)	61.2 (51.4-67.8)	62.9 (54.9-70.8)	60.2 (51.9-67.0)	61.1 (52.0-69.3)	61.7 (52.5-68.8)	61.0 (50.0-70.0)	60.9 (52.0-69.3)	65.5 (59.9-73.6)	64.0 (56.9-73.6)	63.3 (54.8-70.7)	1.03 (1.03-1.03)
Diagnostic delay (months)	10.1 (6.6-16.3)	10.3 (6.0-16.9)	9.0 (5.0-13.9)	12.1 (7.8-21.3)	12.3 (8.0-21.0)	12.6 (8.4-20.9)	8.9 (5.3-13.0)	12.0 (7.5-17.7)	8.0 (5.0-14.9)	10.4 (6.3-17.1)	9.0 (5.0-17.0)	11.0 (6.1-21.9)	9.0 (6.0-14.6)	8.0 (5.0-14.0)	10.5 (6.3-17.6)	0.52 (0.51-0.53)†	
Progression rate (ALSFRS points per month)	0.6 (0.3-1.0)	0.6 (0.3-1.0)	0.7 (0.3-1.3)	0.6 (0.3-1.2)	0.6 (0.3-1.0)	0.5 (0.2-0.9)	0.6 (0.3-1.1)	0.6 (0.4-1.1)	0.6 (0.3-1.0)	0.5 (0.3-1.1)	0.6 (0.4-1.0)	0.4 (0.2-0.8)	0.9 (0.4-1.5)	0.9 (0.5-1.6)	0.6 (0.3-1.1)	6.33 (5.92-6.76)‡	
Forced vital capacity (%)	91.0 (77.0-105.0)	75.5 (45.7-109.8)	87.0 (66.8-102.0)	83.5 (68.0-102.2)	71.0 (57.0-90.0)	86.0 (68.0-100.0)	94.0 (77.0-110.0)	88.0 (74.0-102.0)	81.0 (67.0-95.0)	74.0 (55.0-92.0)	..	86.0 (66.0-101.0)	89.3 (70.0-104.2)	92.0 (67.0-107.0)	88.0 (69.0-103.0)	0.99 (0.99-0.99)	
Bulbar onset	610/1855 (33%)	690/1744 (39%)	347 (34%)	329 (28%)	411/1244 (33%)	200/788 (25%)	224/831 (27%)	196/592 (33%)	121/437 (28%)	130 (29%)	90/307 (29%)	82/278 (29%)	64/178 (36%)	59 (29%)	3553/11138 (32%)	1.71 (1.63-1.79)	
Definite ALS	354/1844 (19%)	942/1642 (57%)	342/1019 (34%)	298/1099 (27%)	329/1116 (28%)	..	64/313 (20%)	146 (25%)	49 (10%)	227 (51%)	40/239 (17%)	61/195 (31%)	68/175 (39%)	33/188 (18%)	2953/9424 (31%)	1.47 (1.39-1.55)	
Fronto-temporal dementia	60/1576 (4%)	86/366 (23%)	127 (12%)	18 (2%)	12 (<1%)	25/813 (3%)	40 (5%)	31 (5%)	4 (<1%)	21 (5%)	..	2 (<1%)	17 (9%)	13 (6%)	456/9284 (5%)	1.34 (1.20-1.50)	
C9orf72 repeat expansion	155/1690 (9%)	71/698 (10%)	50/794 (6%)	51/401 (13%)	5/59 (8%)	..	99/604 (16%)	17/121 (14%)	4/106 (4%)	13/274 (5%)	27 (14%)	15/25 (60%)	507/4962 (10%)	1.45 (1.31-1.61)	
Time to our composite endpoint (95% CI)	35.8 (34.8-37.6)	28.6 (27.3-29.8)	30.2 (28.3-31.4)	37.0 (35.4-39.5)	32.5 (31.3-34.5)	35.3 (33.5-38.3)	35.0 (32.7-38.2)	39.2 (36.2-42.0)	36.0 (33.6-39.0)	46.2 (40.9-51.7)	45.7 (42.0-52.9)	55.5 (46.3-64.0)	32.9 (26.9-40.5)	29.0 (26.5-32.6)	34.7 (34.2-35.3)¶	..	
since onset (months)¶¶	1460 (75%)	1586 (87%)	834 (82%)	850 (72%)	1131 (89%)	699 (82%)	645 (77%)	374 (63%)	443 (88%)	191 (43%)	189 (55%)	139 (49%)	114 (60%)	164 (81%)	8819 (77%)	..	
Achieved one of the composite survival endpoints**	

Data are median (IQR) or number (%) (except where denoted otherwise) for 14 cohorts, one at each ALS centre. --unavailable data. ALS=amyotrophic lateral sclerosis. *Hazard ratios were calculated with a univariable random-effects Cox model with the 14 ALS research centres as a grouping factor for the random effects. †Because of non-normally distributed data, a natural logarithm transformation was applied. ‡Because of non-normally distributed data, a square-root transformation was applied. §Definite ALS according to the revised El Escorial criteria. ¶Time between onset and our composite survival endpoint. ||Calculated with a random-effects Cox model with the 14 ALS research centres as a grouping factor for the random effects. **Composite endpoint of survival defined as the presence of non-invasive ventilation >23 h per day, or tracheostomy, or death. Different cohorts might have used different definitions (appendix).

Table 1: Patient characteristics

provided a means to learn about the generalisability of the model across different settings and populations.

For model development, we used multivariable Royston-Parmar survival models³⁰ rather than Cox survival models, to facilitate the calculation of absolute risks in individual patients when implementing the model in clinical practice (appendix). We assumed a common baseline hazard for all cohorts, but also reported values of cohort-specific baseline hazard functions, which might help to tailor predictions to different populations. We used fractional polynomials to identify non-linear relationships with our composite survival endpoint.³¹ Missing values were multiply imputed using multilevel joint modelling techniques in accordance with previous publications (appendix).^{32,33}

For model validation, we assessed discrimination (ie, the ability to differentiate between patients who reached our composite endpoint and those who did not) and calibration (ie, the agreement between observed and predicted times to our composite endpoint). Discrimination was quantified using the concordance (*c*) statistic, which is the area under the receiver operator characteristic (ROC) curve generalised to all possible survival timepoints as defined by the composite endpoint. The *c* statistic measured the ability of the model to differentiate between patients who reached our composite endpoint and those who did not: values close to 1 indicated very good discriminative ability, while values close to 0.5 indicated poor discriminative ability. We also measured time-dependent ROC curves 3 years after onset, because this is the period reported as median survival.³ Calibration was assessed by calibration plots and quantified by the calibration-in-the-large and calibration slope statistic.^{28,29,34} A calibration slope of 1 in combination with a calibration-in-the-large of 0 indicated good overall calibration (appendix).

We did a complete case analysis and sensitivity analyses omitting frontotemporal dementia from the model, omitting the *C9orf72* repeat expansion from the model, and using time of diagnosis as starting point for prediction. We compared predicted and observed curves for times to reach our composite endpoint for combined and individual cohorts as well as the effect of recalibration of the intercept on predictive performance in different cohorts.³⁴ Five equal-sized prognostic groups were created based on the linear predictor of the model (20% of data per group): very short, short, intermediate, long, and very long times to our composite survival endpoint. A computer algorithm was applied to the full dataset to randomly select five patients from the five prognostic groups (in total 25 patients), illustrating application of the model to individual patients.

Because the internal–external cross-validation approach allows evaluation of external performance across multiple studies, we meta-analysed estimates of model performance and calculated 95% CI, using random-effects meta-analyses.²⁹ We also calculated 95% prediction intervals (PI) to quantify the range of model performance

across different populations, which helped to assess the potential generalisability of the model.²⁹

Prognostic models for predicting survival and related outcomes require adequate discrimination and calibration performance to be clinically useful. We therefore calculated the joint 95% PI and estimated the probability that the model would achieve a certain predefined *c* statistic and calibration slope in future patients (appendix).

Data sharing

For medical doctors, we are making our model freely available online. We report all parameters of the model in the appendix and further supporting information can be requested online or from the corresponding author.

For the model see <http://www.encalsurvivalmodel.org>

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Data were collected between Jan 1, 1992, and Sept 22, 2016, for 11475 patients. The different cohorts had different start and end dates within this period (appendix). Total follow-up was 40016 years and median follow-up time was 97.5 months (IQR 52.9–168.5). Patient characteristics are shown in table 1 and the appendix.

Based on backward elimination in the largest dataset (1936 patients), eight of the 16 candidate predictors were selected for the multivariable prediction model: age at onset ($n=10000$, 100% of bootstraps), FVC (7598, 76%), diagnostic delay (10000, 100%), ALSFRS-R slope (10000, 100%), bulbar onset (8094, 81%), definite

	Univariable HR*	Multivariable HR†
Age at onset (years)	1.28 (1.25–1.31)	1.02 (1.02–1.03)
Diagnostic delay (months)	0.48 (0.46–0.49)	0.63 (0.59–0.68)‡
Progression rate (points per month)	3.05 (2.84–3.27)	3.19 (2.71–3.75)§
Forced vital capacity (%)	0.76 (0.72–0.80)	0.99 (0.99–0.99)
Bulbar onset	..	1.25 (1.17–1.33)
Definite ALS¶	..	1.25 (1.16–1.34)
Frontotemporal dementia	..	1.18 (1.01–1.37)
<i>C9orf72</i> repeat expansion	..	1.37 (1.19–1.57)

Data are hazard ratio (HR; 95% CI). All HRs were estimated using random-effects Cox models with the 14 ALS research centres as grouping factor for the random effects. This table is intended to provide an indication of the relative effects of the transformed predictors, and the random-effects Cox model differs from the Royston-Parmar proportional odds model that was used to generate the predictions. ALS=amyotrophic lateral sclerosis. ..=not transformed. *Using the different transformed continuous predictors, we made a comparison between the first and third quartile and report the univariable HRs (see appendix for applied transformations). This illustrates the relative effects of the continuous predictors after transformation. Because the last four variables were dichotomous, it was not possible to make a comparison between the first and third quartiles. †This column shows the HRs when all eight predictors were combined (more details in appendix). The HRs resulted from coefficients and thus did not rely on the quartiles that were used in the univariate HR column. ‡Because of non-normally distributed data, a natural logarithm transformation was applied. §Because of non-normally distributed data, a square-root transformation was applied. ¶According to revised El Escorial criteria.

Table 2: Relative effects of transformed predictors

ALS (7120, 71%), presence of frontotemporal dementia (7416, 74%), and the *C9orf72* repeat expansion (8679, 87%). Results for variables that were not included in the model are reported in the appendix. Age at onset, ALSFRS-R slope, diagnostic delay, and FVC were transformed because of non-linear relationships with our composite survival endpoint (appendix); the relative effects of the

predictors after transformation are presented in table 2. Using all imputed datasets, we showed that a proportional odds Royston-Parmer model with two internal knots, without time-dependent covariates, was most appropriate according to the Akaike information criterion and that a proportional odds model consistently outperformed a proportional hazard model (appendix).

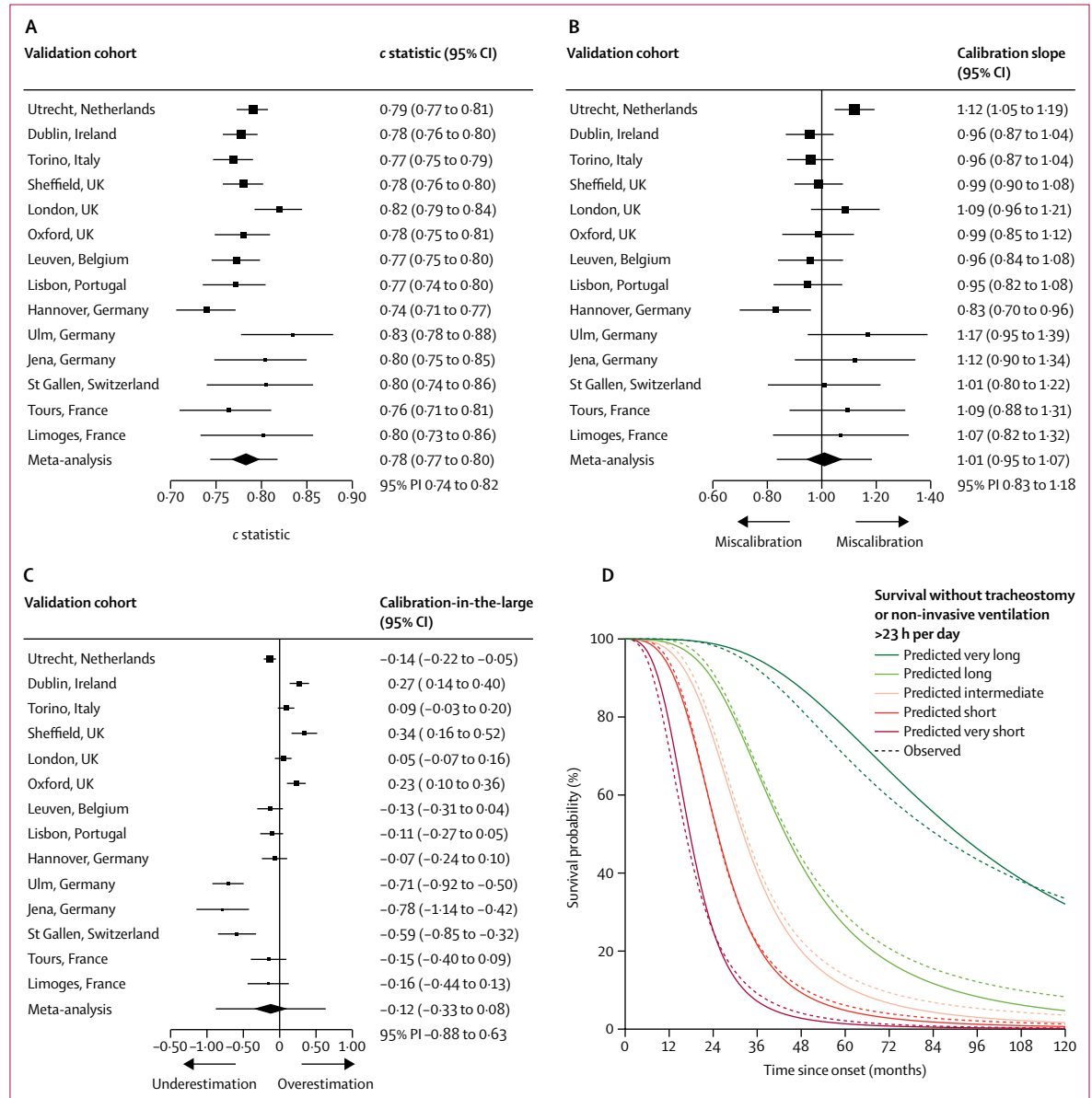


Figure 1: Univariate meta-analysis of predictive performance

Panel A shows the random-effects meta-analysis of discrimination as measured with the c statistic (time-dependent receiver operator characteristic [ROC] curves in the appendix). Panels B and C show the random-effects meta-analysis of calibration as measured with calibration slope (B) and calibration-in-the-large (C). A calibration slope of 1 in combination with a calibration-in-the-large of 0 indicated good overall calibration (calibration slope <1: predictions are too extreme; calibration slope >1: predictions are too similar; further details in the appendix). The black diamonds indicate the mean (95% CI) of the predictive accuracy. 95% prediction intervals (PI), which indicate predicted accuracy of the model in a single new dataset or patient, are presented as a numeric range (below the 95% CI). Panel D is the visual translation of A–C in prognostic curves showing the agreement between predicted and observed probability of reaching the composite endpoint as well as indicating good discriminative power of the model. The curves also illustrate the possibility to stratify patients in different groups based on their predicted prognosis on the day of diagnosis. Five equal-sized groups were created based on predicted time to our composite endpoint. Omitted cohort label indicates the cohort left out of the internal-external cross-validation, which was done once for all 14 different cohorts.

The internal-external cross-validation framework resulted in 14 cycles of model development and external validation (meta-analysis shown in figure 1 and appendix). The model was used to define five groups with distinct median predicted times (SE) and observed times (SE) in months to the composite survival outcome since symptom onset: very short 17.7 (0.20), 16.5 (0.23); short 25.3 (0.06), 25.2 (0.35); intermediate 32.2 (0.09), 32.8 (0.46); long 43.7 (0.21), 44.6 (0.74); and very long 91.0 (1.84), 85.6 (1.96; figure 1). Comparing the other groups with the group with very long times to our composite endpoint revealed hazard ratios (HR) of 15.29 (95% CI 13.92–16.79) for the very short group, 7.19 (6.53–7.91) for the short group, 4.30 (3.94–4.69) for the intermediate group, and 2.49 (2.30–2.70) for the long group). Visual inspection of calibration plots 3 years after onset showed good agreement between predicted and observed probability for most cohorts (n=11; appendix). Predicted and observed probabilities of reaching our composite endpoint for cohort-specific curves showed similar results as calibration plots (appendix). For the three cohorts in which agreement between observed and predicted times to our composite endpoint was suboptimal, recalibration of the baseline hazard improved calibration (appendix). Similar results were noted for complete case analysis (appendix). We assessed all possible two-way interactions (ie, eight covariates with one interaction comprising two of the eight covariates), which did not improve predictive accuracy (appendix). The probability of good external predictive performance of the model was estimated to be 100.0% for the *c* statistic and 97.1% for calibration slope. The joint (combined) probability of good performance was 98.3% (appendix). Compared to estimates of group level data, the model provides more accurate and precise predictions of times to our composite endpoint (figure 2).

Our prediction of time to our composite survival endpoint is based on the regression coefficients of the final prediction model (appendix) and can be tailored (ie, recalibration of the intercept) for each of the 14 participating cohorts. To illustrate its potential use in clinical practice, we have provided examples of patients selected from our dataset and their associated predictions, including uncertainty at the patient level (figure 3) and provided worked examples (appendix). A large proportion of patients deviated from the median time to reach the composite survival endpoint (at the group level): six patients (24%) had a 75% chance of dying, having a tracheostomy, or having non-invasive ventilation for more than 23 h per day before the median time, whereas eight other patients (32%) had a 75% chance of living without tracheostomy or non-invasive ventilation for more than 23 h per day for longer than the median time. The sensitivity analyses showed that the prediction model also provided accurate predictions when patient-level information about the presence of frontotemporal

dementia or a *C9orf72* repeat expansion was not available and similar results of predictions using date of onset versus date of diagnosis as starting point for prediction (appendix). A survey of patients' preferences with regard to knowing their personalised prognosis is reported in the appendix.

Discussion

On the basis of data collected from 11475 patients with ALS across Europe, we have developed and externally validated a model for prediction of survival without tracheostomy or non-invasive ventilation for more than 23 h per day in individual patients. Using recently developed methods, we derived ranges of predictive performance across different settings (ie, population-based vs referral-based cohorts) and patient populations, thereby providing evidence for the potential generalisability of our model. Medical doctors can use our model freely online to provide estimates of prognosis in individual patients with ALS, with the aim of facilitating its use in clinical practice and in innovative trial design.

Associations between predictors and survival outcomes in ALS have previously been reported at the group level.⁶ Attempts to translate predictive associations into estimates of survival outcomes in individual patients

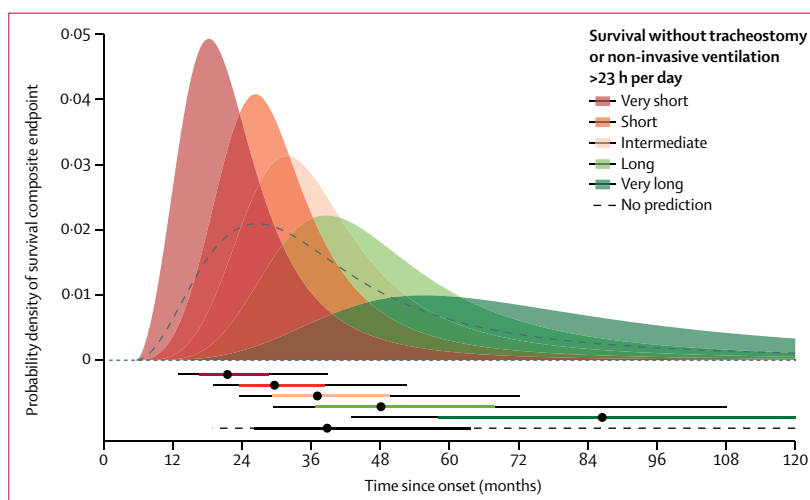


Figure 2: Group predictions

Probability densities for times to our composite endpoint (y-axis) are shown for the five groups from figure 1D. The probability density (or distribution) is not evenly distributed over time, meaning that the largest number of patients are predicted to die at the time when the curve has the highest density, but there is a small chance that patients will survive without tracheostomy or non-invasive ventilation for more than 23 h per day for much longer—this is reflected by the long, thin tails on the right side of this figure. These predictions curves provide guidance for discussing prognosis with patients. Furthermore, if the prediction model is not used (dashed curve), estimates of group-level data (ie, average or median survival in ALS, without using a personalised prediction) will generally be overly positive for groups with short or very short times to the composite endpoint, overly negative for groups with long and very long times to the composite endpoint, or overly uncertain (for groups with intermediate times to the composite endpoint). The horizontal bars at the bottom provide the same information for the different groups, with dots representing median times to our composite endpoint, thick lines representing probability IQRs, and thin lines representing 10–90% probability intervals to reach our composite survival endpoint. This figure quantifies the uncertainty of individual predictions because, for example, a participant with a predicted very short time to our composite endpoint still has a 2% chance of surviving up to 5 years after onset without tracheostomy or non-invasive ventilation for more than 23 h per day.

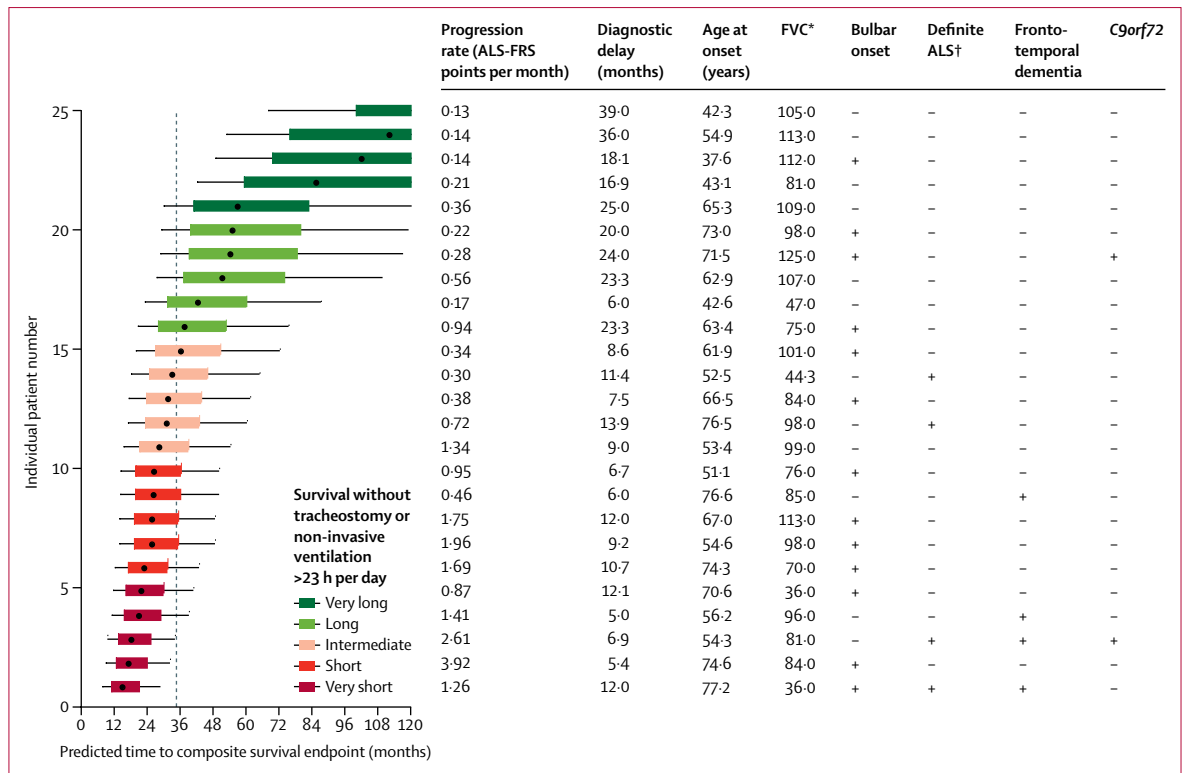


Figure 3: Clinical applicability to individual patients

Characteristics of and predictions for 25 randomly selected participants with amyotrophic lateral sclerosis (ALS) are presented to illustrate application of the model to predict time to the survival composite endpoint for individual patients. Predictions are shown on the left and patient characteristics on the right. Median predicted times to the composite endpoint for individual patients are represented by the dots. The thick coloured lines represent 25–75% probability IQRs and the thin black lines represent 10–90% probability intervals. The vertical dashed line represents the median time to the composite endpoint (34.7 months). Predicted times to our composite endpoint differed substantially between patients, reflecting both heterogeneity in ALS and the ability of the proposed model to stratify patients according to their characteristics. +=present. -=absent. *Forced vital capacity (FVC) expressed as percentage of predicted (based on normative values for age, sex, and height). †Definite ALS determined according to revised El Escorial criteria.

with ALS are, however, infrequent and external validation of such models remains elusive.³⁵ It is acknowledged that the development of reliable prediction models requires access to large datasets, with at least 200 events for validation.⁸ Hence, the current meta-analysis of individual participant data is the first study in ALS to enable the investigation of a large number of predictors, combining them in a prediction model. This is the first prediction model in ALS that rigorously assessed discrimination and calibration performance across different populations. We identified sources of heterogeneity across the cohorts, and addressed transportability issues of the model.^{27,28} This enabled us to determine the likely performance when applying the model in clinical practice, and thus to assess its clinical use in local circumstances. Dependent on requirements and clinicians' or patients' preferences, the model can be applied using probabilities of survival without tracheostomy or non-invasive ventilation for more than 23 h per day, prognostic groups, or point estimates. For clinical care, we prefer the first two options, because these provide the possibility of being realistic without being unduly negative (appendix). Because the probability

distributions for times to our composite endpoint were skewed, point estimates might not be preferred for predictions in individual patients. Ranges will become wider for patients with longer times to our composite endpoint, owing to less steep prognostic curves.

Large-scale, European Union funded, collaborative projects, such as Euro-MOTOR, SOPHIA, and NETCALC encourage standardised and harmonised data collection between the three population-based registries (n=4776, 42%) and 11 referral-based patient registries (n=6699, 58%) that contributed data to our study.^{36,37} Differences in standard of care and cultural or genetic background between registries were taken into account in our model by allowing cohort-specific adjustments for estimation of prognosis in individual patients.

Predictors and times to our composite endpoint are likely to differ between different populations, as shown in our study. This might be due to real differences between populations. Other possibilities might be referral bias in referral-based versus population-based registries, variable interpretation of diagnostic criteria, differences in standard care, cultural or genetic background, or a combination of these factors.

For more on Euro-MOTOR see www.euromotorproject.eu
 For more on SOPHIA see www.sophiaproject.eu
 For more on NETCALC see www.jpnd.eu

Importantly, internal–external cross-validation in our study did not reveal that differences between populations affected predictive performance of our model. Furthermore, differences in times to our composite endpoint between cohorts could not be interpreted as being caused by a specific cohort.

A weakness of our study is that we did not include treatment with riluzole, the only effective drug with proven effect on survival, as a predictor in our model. Reliable information about its usage is missing in most registries, but is estimated to be at least 75% in Europe.¹³ Although the effect of riluzole on survival (pooled HR 0·84, based on a Cochrane meta-analysis)¹⁴ is more than ten times smaller than the combined effect of predictors in our proposed model (HR up to 15·29), incorporating riluzole use as well as other treatments, such as ventilation or gastrostomy, into our model will allow further tailoring of predictions.

Missing data, which is inherent to observational data, can lead to biased results when not appropriately addressed. We therefore used imputation methods that are proven to prevent such biases, even if a predictor is completely missing from one or more cohorts.³² Using internal–external cross-validation, we rigorously assessed the model's predictive performance in external datasets. This approach not only allows assessment of the model's overall performance in new patients, but also to establish whether performance is consistent across the different cohorts. Internal–external cross-validation therefore helps to determine to what extent model generalisability might be affected by differences in definition of endpoint, in collection of FVC versus SNIP or ALSFRS versus ALSFRS-R data, in possible recall bias of disease onset, in recruitment of patients between cohorts, in heterogeneity of disease characteristics between cohorts, or in the presence of unmeasured confounders. Although we did not aim to study such differences between cohorts and we can only speculate about possible causes (eg, intrinsic differences of ALS between cohorts, or extrinsic differences due to possible differences in, eg, clinical care), these differences did not appear to reduce the predictive performance of our proposed prediction model.

Our online model is ethically sensitive and demands thoughtful implementation. It might support clinicians in helping patients to maintain a degree of autonomy and help them in planning their lives. Our survey showed that some patients preferred to be informed about their personalised life expectancy (appendix). This is in line with previous reports in cancer medicine.³⁸ The ethical aspects of tailored predictions in ALS or other neurodegenerative diseases have not been studied before, but from cancer medicine it is known that early discussions about goals of care are associated with better quality of life, reduced use of non-beneficial medical care near death, enhanced goal-consistent care, positive family outcomes, and reduced costs.³⁹ However, many cancers have several modes of therapy, whereas ALS is still an

incurable disease. Furthermore, the outcomes of the administered questionnaire might differ in patients from other cultural backgrounds; a selection bias might be present in the patients who responded to the survey; and patients might not foresee all consequences of the knowledge, and might regret asking to be informed if the predicted time to death, tracheostomy, or non-invasive ventilation for more than 23 h per day they receive is shorter than expected. To minimise the risk of such potential harm, the model will be accessible online only to medical doctors who have to register and sign in before access is provided. Professional health-care providers might use our model to determine the intensity of care pathways and to tailor counselling of individual patients and their caregivers. Prediction models might improve clinical trial design in two ways: use of the predicted outcome as an inclusion criterion instead of arbitrary eligibility criteria and use of predicted prognosis instead of a limited set of prognostic variables to stratify randomisation, thereby creating more homogeneous strata and greater statistical power to detect an effect. This was, for example, previously shown in patients with glioblastoma who have methylation of the *MGMT* (O⁶-methylguanine–DNA methyltransferase) gene.⁴⁰

In accordance with the TRIPOD statement,¹⁰ all parameters and equations of our model were provided (appendix) to allow improvement of predictions through continuing research into other prognostic factors (eg, wet or imaging biomarkers or newly discovered genetic mutations) or implementation of outcome measures other than our composite survival endpoint (eg, being wheelchair-bound or other specific loss of function). This should also facilitate validation of our model in registries from a non-European background and its application in clinical trial populations.

In conclusion, we have developed and externally validated a model for prediction of survival without tracheostomy or non-invasive ventilation more than 23 h per day in individual patients with ALS. We have shown the generalisable and robust predictive performance of this model and made it freely available online. The outcomes of this study have the potential to facilitate tailored care and trial design, which will hopefully lead to a more successful discovery of effective treatments for patients with this devastating disease.

Contributors

H-JW, TPAD, KGMM, JHV, and LHvdB contributed to the study design. H-JW did the literature search and created the figures. H-JW and TPAD analysed the data. All authors contributed to the data collection and interpretation of results, reviewed and critically revised the manuscript, and approved the final version for submission.

Declaration of interests

JPKR reports grants from the Irish Health Research Board (grant number HPF-2014-537). CJM was supported by the UK Motor Neurone Disease Association. MAvE serves on the Motor Neurone Disease Association biomedical research advisory panel, has consulted for Biogen and received travel grants from Baxalta, and funding sources include the Netherlands Organization for Health Research and Development (Veni scheme), The Thierry Latran Foundation, the ALS Foundation Netherlands, and the

European Union (EU) Joint Programme Neurodegenerative Disease Research (JPND). MW reports personal fees from Biogen Idec and Mitsubishi Tanabe Pharma. SP reports grants from the German Neuromuscular Society, the German Federal Ministry of Education and Research, clinical trial funding from Cytokinetics, GlaxoSmithKline, and Orion Pharma, and speaking fees from Teva Pharmaceutical Industries. PVD reports personal fees for advisory work from Cytokinetics, and personal fees for advisory work (paid to his institution) from Biogen, Pfizer, CSL Behring, and Treeway. KT is supported by the Motor Neurone Disease Association and the UK Medical Research Council (MRC). MRT is supported by the UK MRC and Motor Neurone Disease Association Lady Edith Wolfson Senior Clinical Fellowship (MR/K01014X/1). PJS reports grants from the UK National Institute of Health Research (NIHR; Senior Investigator Award), Sheffield NIHR Biomedical Research Centre for Translational Neuroscience, the Motor Neurone Disease Association, the EU JPND (STRENGTH project), the UK MRC, EU Horizon 2020 (Modifying Immune Response and Outcomes in ALS [MIROCALS] trial) and grants from ReNeuron; and clinical trial funding from Cytokinetics, Orion Pharma, and Biogen Idec. AAC is a consultant for Mitsubishi Tanabe Pharma, Chronos Therapeutics, Orion Pharma, Cytokinetics, and Treeway. AC reports grants from Italfarmaco, and personal fees from Biogen Idec and Mitsubishi. OH is a journal editor for *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*; and reports grants from the Irish Health Research Board (CSA 2012/11), Science Foundation Ireland (15/SPP/3244 and 16/ERC/3854), and the charity Research Motor Neurone. LHvdB reports grants from The Netherlands ALS Foundation, The Netherlands Organization for Health Research and Development (ZonMw; Vici scheme), the EU 7th framework programme (grant number 259867) for the Euro-MOTOR project, The Netherlands Organization for Health Research and Development (Sampling and biomarker Optimization and Harmonization In ALS and other motor neuron diseases [SOPHIA], funded through JPND), and Baxalta; and personal fees from Biogen, Cytokinetics, and Baxalta. All other authors declare no competing interests.

Acknowledgments

The Netherlands ALS Foundation funded this study. We thank all the patients with amyotrophic lateral sclerosis who, notwithstanding their devastating disease, participated in this study.

References

- GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016; **388**: 1459–544.
- Jameson JL, Longo DL. Precision medicine—personalized, problematic, and promising. *N Engl J Med* 2015; **372**: 2229–34.
- Kiernan MC, Vucic S, Cheah BC, et al. Amyotrophic lateral sclerosis. *Lancet* 2011; **377**: 942–55.
- van Es MA, Hardiman O, Chio A, et al. Amyotrophic lateral sclerosis. *Lancet* 2017; **390**: 2084–98.
- Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* 2005; **65**: 586–90.
- Chio A, Logrosino G, Hardiman O, et al. Prognostic factors in ALS: a critical review. *Amyotroph Lateral Scler* 2009; **10**: 310–23.
- van Rheenen W, van Blitterswijk M, Huisman MH, et al. Hexanucleotide repeat expansions in C9ORF72 in the spectrum of motor neuron diseases. *Neurology* 2012; **79**: 878–82.
- Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016; **35**: 214–26.
- Riley RD, Ensor J, Snell KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016; **353**: i3140.
- Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015; **350**: g7594.
- Huisman MH, de Jong SW, van Doormaal PT, et al. Population based epidemiology of amyotrophic lateral sclerosis using capture-recapture methodology. *J Neurol Neurosurg Psychiatry* 2011; **82**: 1165–70.
- Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron Disease. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000; **1**: 293–99.
- Cudkovic ME, van den Berg LH, Shefner JM, et al. Dexamipexole versus placebo for patients with amyotrophic lateral sclerosis (EMPOWER): a randomised, double-blind, phase 3 trial. *Lancet Neurol* 2013; **12**: 1059–67.
- Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev* 2012; **3**: CD001447.
- 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**: 265–67.
- Gallo V, Wark PA, Jenab M, et al. Prediagnostic body fat and risk of death from amyotrophic lateral sclerosis: the EPIC cohort. *Neurology* 2013; **80**: 829–38.
- Calvo A, Canosa A, Bertuzzo D, et al. Influence of cigarette smoking on ALS outcome: a population-based study. *J Neurol Neurosurg Psychiatry* 2016; **87**: 1229–33.
- Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011; **134**: 2456–77.
- Abrahams S, Leigh PN, Harvey A, Vythelingum GN, Grise D, Goldstein LH. Verbal fluency and executive dysfunction in amyotrophic lateral sclerosis (ALS). *Neuropsychologia* 2000; **38**: 734–47.
- Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology* 2000; **55**: 1621–26.
- Raaphorst J, Beeldman E, Schmand B, et al. The ALS-FTD-Q: a new screening tool for behavioral disturbances in ALS. *Neurology* 2012; **79**: 1377–83.
- Diekstra FP, van Vught PW, van Rheenen W, et al. UNC13A is a modifier of survival in amyotrophic lateral sclerosis. *Neurobiol Aging* 2012; **33**: 630 e3–8.
- Majounie E, Renton AE, Mok K, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* 2012; **11**: 323–30.
- van Blitterswijk M, DeJesus-Hernandez M, Rademakers R. How do C9ORF72 repeat expansions cause amyotrophic lateral sclerosis and frontotemporal dementia: can we learn from other noncoding repeat expansion disorders? *Curr Opin Neurol* 2012; **25**: 689–700.
- Fregonezi G, Araujo PR, Macedo TL, Dourado Junior ME, Resqueti VR, Andrade Ade F. Monitoring respiratory muscle strength assists in early diagnosis of respiratory dysfunction as opposed to the isolated use of pulmonary function evaluation in amyotrophic lateral sclerosis. *Arq Neuropsiquiatr* 2013; **71**: 146–52.
- Heymans MW, van Buuren S, Knol DL, van Mechelen W, de Vet HC. Variable selection under multiple imputation using the bootstrap in a prognostic study. *BMC Med Res Methodol* 2007; **7**: 33.
- Debray TP, Moons KG, Ahmed I, Koffijberg H, Riley RD. A framework for developing, implementing, and evaluating clinical prediction models in an individual participant data meta-analysis. *Stat Med* 2013; **32**: 3158–80.
- Debray TP, Riley RD, Rovers MM, Reitsma JB, Moons KG, Cochrane IPDM-aMg. Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use. *PLoS Med* 2015; **12**: e1001886.
- Snell KI, Hua H, Debray TP, et al. Multivariate meta-analysis of individual participant data helped externally validate the performance and implementation of a prediction model. *J Clin Epidemiol* 2016; **69**: 40–50.
- Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002; **21**: 2175–97.
- Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *J R Stat Soc Series A Stats Soc* 1999; **162**: 71–94.
- Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. *Stat Med* 2016; **35**: 2938–54.

- 33 Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; **338**: b2393.
- 34 Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013; **13**: 33.
- 35 Elamin M, Bede P, Montuschi A, Pender N, Chio A, Hardiman O. Predicting prognosis in amyotrophic lateral sclerosis: a simple algorithm. *J Neurol* 2015; **262**: 1447–54.
- 36 D'Ovidio F, Rooney JPK, Visser AE, et al. Critical issues in ALS case-control studies: the case of the Euro-MOTOR study. *Amyotroph Lateral Scler Frontotemporal Degener* 2017; **18**: 411–18.
- 37 Hardiman O, Al-Chalabi A, Brayne C, et al. The changing picture of amyotrophic lateral sclerosis: lessons from European registers. *J Neurol Neurosurg Psychiatry* 2017; **88**: 557–63.
- 38 Innes S, Payne S. Advanced cancer patients' prognostic information preferences: a review. *Palliat Med* 2009; **23**: 29–39.
- 39 Bernacki RE, Block SD, American College of Physicians High Value Care Task Force. Communication about serious illness care goals: a review and synthesis of best practices. *JAMA Intern Med* 2014; **174**: 1994–2003.
- 40 Hegi ME, Diserens A, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005; **352**: 997–1003.