
Article:
Westeneng, H.-J., Debray, T.P.A., Visser, A.E. et al. (2018) Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. Lancet Neurology, 17 (5). pp. 423-433. ISSN: 1474-4422

https://doi.org/10.1016/S1474-4422(18)30089-9

Title:
Prediction of personalised prognosis in patients with amyotrophic lateral sclerosis: development and validation of a prediction model

Authors:
Henk-Jan Westeneng[1] MD, Thomas P.A. Debray[2,3] PhD, Anne E. Visser[1] MD, Ruben P.A. van Eijk[1] MD, James P.K. Rooney[4] MSc, Andrea Calvo[5] MD, Sarah Martin[6] BSc, Prof. Christopher J. McDermott[7] PhD, Alexander G. Thompson[8] BMBCh, Susana Pinto[9] PhD, Xenia Kobeleva[10] MD, Angela Rosenbohm[11] MD, Beatrice Stubendorff[12] PhD, Helma Sommer[13], Bas M. Middelkoop[1], Annelot M. Dekker[1] MD, Joke J.F.A. van Vugt[1] PhD, Wouter van Rheenen[1] MD, Alice Vajda[4] PhD, Mark Heverin[4] MSc, Mbombe Kazoka[7], Hannah Hollinger[7] MA, Marta Gromicho[9] PhD, Sonja Körner[10] MD, Thomas M. Ringer[12] MD, Annekathrin Rödiger[12] MD, Anne Gunkel[12] MD, Prof. Christopher E. Shaw[6] MD, Prof. Annelien L. Bredenoord[2] PhD, Michael A. van Es[1] MD, Prof. Philippe Corcia[14] MD, Prof. Philippe Couratier[15] MD, Prof. Markus Weber[13] MD, Julian Grosskreutz[12] MD, Prof. Albert C. Ludolph[11] MD, Prof. Susanne Petri[10] MD, Prof. Mamede de Carvalho[9] MD, Prof. Philip Van Damme[16] MD, Prof. Kevin Talbot[8] DPhil, Prof. Martin R. Turner[8] MD, Prof. Pamela J. Shaw[7] MD, Prof. Ammar Al-Chalabi[6] PhD, Prof. Adriano Chiò[5] MD, Prof. Orla Hardiman[4,17] MD, Prof. Karel G.M. Moons[2,3] PhD, Prof. Jan H. Veldink[1] MD, Prof. Leonard H. van den Berg[1] MD

Affiliations:
1 Department of Neurology, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, The Netherlands.
2 Department of Epidemiology, Julius Centre for Health Sciences & Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands.
3 Cochrane Netherlands, University Medical Centre Utrecht, Utrecht, The Netherlands.
4 Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College, Dublin, Ireland.
5 'Rita Levi Montalcini' Department of Neuroscience, ALS Centre, University of Torino, Torino, Italy.
6 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.
7 Sheffield Institute for Translational Neuroscience, University of Sheffield, UK.
8 Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK.
9 Institute of Physiology-Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, Portugal.
10 Department of Neurology, Hannover Medical School, Hannover, Germany.
11 Department of Neurology, University of Ulm, Germany.
12 Hans-Berger Department of Neurology, Jena University Hospital, Jena, Germany.
13 Neuromuscular Diseases Centre/ALS Clinic, Kantonsspital St. Gallen, St.Gallen, Switzerland.
14 Centre de compétence SLA-fédération Tours-Limoges, CHU de Tours, France..
15 Centre de compétence SLA-fédération Tours-Limoges, CHU de Limoges, France.
16 Department of Neurology, University Hospital Leuven; Department of Neurosciences, KU Leuven (University of Leuven) and Centre for Brain & Disease Research, VIB, Leuven, Belgium.
17 Department of Neurology, Beaumont Hospital, Beaumont, Ireland.

Corresponding author:
Leonard H. van den Berg, MD PhD,
Department of Neurology, G03.228, University Medical Centre Utrecht,
P.O. Box 85500, 3508 GA Utrecht, The Netherlands.
Telephone: +31 88 7557939.
Fax number: +31 30 2542100.
Email: L.H.vandenBerg@umcutrecht.nl

Keywords: survival, prediction, amyotrophic lateral sclerosis

Abstract

Background
Amyotrophic lateral sclerosis (ALS) is a relentlessly progressive, fatal motor neurone disease with a variable natural history. The lack of accurate models that predict the disease course and outcomes complicates individual risk assessment and counselling, stratification of patients for trials, and timing of interventions. We, therefore, aimed to develop and validate a model for predicting a composite survival endpoint of individual ALS patients.

Methods
Using clinical, cognitive and genetic data from 14 cohorts across nine European countries, we performed an individual participant data meta-analysis of 11475 patients. All patients were diagnosed in specialized ALS centres after excluding other diagnoses, and classified according to the El Escorial criteria. Backward elimination with bootstrapping was applied in the largest population-based dataset (n=1936); eight out of 16 predictors were selected and used to develop a multivariable model for predicting a composite survival outcome (time between onset of symptoms and non-invasive ventilation >23 hours/day, tracheostomy, or death) in individual patients. We assessed the generalisability of the model by estimating heterogeneity of predictive accuracy across external populations (i.e. populations not used to develop the model) using internal-external cross-validation.

Findings
Data were collected between January 1, 1992, and September 22, 2016. The median follow-up time was 97·5 months (IQR, 52·9 to 168·5). Eight of 16 candidate predictors entered the prediction model: bulbar versus non-bulbar onset (univariable hazard ratio (HR (95% confidence interval)) 1·71 (1·63 to 1·79)), age at onset (HR 1·031 (1·029 to 1·033)), 'definite' versus 'probable' or 'possible' ALS according to the revised El Escorial criteria (HR 1·47 (1·39 to 1·55)), diagnostic delay (HR 0·52 (0·51 to 0·53)), forced vital capacity (HR 0·988 (0·987 to 0·989)), progression rate (HR 6·33 (5·92 to 6·76)), frontotemporal dementia (HR 1·34 (1·20 to 1·50)), *C9orf72* repeat expansion (HR 1·45 (1·31 to 1·61)), all $p<0·0001$. Our model achieved good and consistent external predictive accuracy: the concordance (c)-statistic was 0·78 (95% confidence interval: 0·77 to 0·80; 95% prediction interval: 0·74 to 0·82) and calibration slope was 1·01 (95% confidence interval: 0·95 to 1·07; 95% prediction interval: 0·74 to 0·82). The model was used to define five groups with distinct predicted (and observed (standard error)) median time to our composite survival outcome since symptom onset: 17·7 (16·5 (0·23)), 25·3 (25·2 (0·35)), 32·2 (32·8 (0·46)), 43·7 (44·6 (0·74)) and 91·0 (85·6 (1·96)) months.

Interpretation
The present large-scale individual participant data meta-analysis proposes an externally validated model to predict survival free of tracheostomy and non-invasive ventilation >23 hours/day of European ALS patients. This model is implemented in an online tool that can be applied to individualised patient management, counselling, and future trial design. This model is ethically sensitive, because it might cause harm if applied inappropriately, and is, therefore, intended to be used by medical doctors only, thereby maximizing benefit and preventing harm to patients.

Funding
Netherlands ALS Foundation

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] Models which reliably predict outcomes at an individual patient level may become an important factor in a precision medicine approach to find the most effective treatment for patients with neurodegenerative disease, by improving 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 neurones in brain and spinal cord, leading to weakness of voluntary muscles.[3,4] Muscle weakness progresses gradually, spreading from a site of clinical onset to other regions of the body; patients eventually become paralysed and die, usually as the result of respiratory failure.[3,4] The clinical features of ALS are heterogeneous; it can occur at any adult age, up to 15% of patients develop frontotemporal dementia (FTD),[5] and 10-15% of patients have a family history of ALS or FTD.[4]

Survival varies greatly, ranging from several months to more than ten 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 FTD, or a repeat expansion in *C9orf72*, have been shown to be associated with shorter survival.[6,7] Unfortunately, despite the well-documented natural history of ALS,[3,6] prediction of survival in individual ALS 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 due to the limited availability of large datasets with individual participant data (IPD) and adequate follow-up time.[8-10]

We therefore analysed clinical, cognitive and genetic data of 11475 ALS patients originating from 14 ALS centres in Europe with a view to predicting survival, defined as time between onset of symptoms and non-invasive ventilation >23 hours/day, tracheostomy, or death in individual patients on the day of diagnosis. We aim to develop and externally validate a prediction model in multiple cohorts; the model is available as a free online tool, initially only for health care providers.

Methods

Study design and participants

The study was carried out in two stages. In the first stage, predictor selection was performed in 1936 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 January 1st 2006 and March 31th 2015).[11] In the second stage, data on identified predictors and survival outcomes were requested from 13 European ALS centres to develop and externally validate the model. Patients were classified according to the revised El Escorial criteria.[12] Written informed consent was obtained from all participants and institutional review boards approved this study (Tables S1 to S13 of the appendix). The study is reported in accordance with the TRIPOD guidance for transparent reporting of prediction models.[10]

Procedures

We assessed ten clinical, four cognitive and two genetic characteristics, and used a backward elimination procedure with bootstrapping for predictor selection (appendix section III.1).[13] Predictors that were selected in >70% of the bootstrap resamples entered the multivariable prediction model.[13] All predictors were selected based on a PubMed search (see research in context).

Clinical predictors were: sex, site of onset (spinal vs. bulbar), age at onset of weakness or bulbar symptoms, 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, body height), progression rate defined by the slope on the revised ALS functional rating scale (ALSFRS-R; see below),[14] premorbid body mass index (BMI),[15] current smoking,[16] cigarette pack years.[6] Survival analyses in ALS research generally use composite endpoints (i.e. events) comprising both death and respiratory events.[17,18] In our study, survival was defined as time between onset of symptoms and a composite endpoint, which we defined as non-invasive ventilation >23 hours/day, or tracheostomy, or death.

Cognitive predictors were: the presence of FTD according to the Neary or Rascovsky criteria,[19] and scores on verbal fluency index (VFI),[20] frontal assessment battery (FAB)[21] and ALSFTD questionnaire (ALSFTD-Q).[22] For the VFI, participants were asked to name as many words as possible beginning with the letter 'D' in three minutes. After these three minutes the participants were instructed to read aloud the generated words. The VFI was calculated as the number of words generated divided by the time needed to read them aloud. The FAB is a bedside test comprising six tasks (maximum of three points per task) measuring conceptualisation, mental flexibility, motor programming, sensitivity to interference, inhibition control and environmental autonomy, and is sensitive to frontal lobe dysfunction. The ALSFRS-Q is a 25-item questionnaire (maximum of four points per item) applied to a proxy (i.e. a caregiver who is able to assess the patient's behaviour) that is developed to screen for behavioural disturbances in ALS.

Genetic predictors were presence of a *C9orf72* mutation, and the minor allele homozygous genotype (C/C) of the *UNC13A* single nucleotide polymorphism, which were determined as previously described.[7,23-25] These two genetic predictors are the most frequent mutations in ALS patients and were previously shown to be associated with survival outcomes in ALS.[7,23]

Data were gathered on the day of diagnosis or as soon as possible thereafter. The ALSFRS-R slope was determined as follows: ALSFRS-R slope = (48 – ALSFRS-R score)/(date of the ALSFRS-R score – date of onset).[14] ALSFTD-Q scores were trichotomised into 'no', 'mild' and 'severe' behavioural impairment.[22] A detailed overview of predictors is provided in Tables S1 to S13 of the appendix.

All cohorts used the ALSFRS-R score, except cohort 6, which used the ALSFRS (i.e. the unrevised version of the ALSFRS). The ten questions of the ALSFRS are identical to the first ten questions of the ALSFRS-R, but the latter has two additional questions about respiration. The maximum score per question is four, resulting in a maximum score of 48 for the ALSFRS-R and 40 for the ALSFRS. To be able to compare scores, we transformed the ALSFRS score to the ALSFRS-R score by multiplying by 1·2 (1·2 times 40 = 48). Furthermore, all cohorts used FVC to measure respiratory function, except cohort 2, which used sniff nasal inspiratory pressure (SNIP), known to be correlated with FVC.[26] Based on subjects in cohort 2 for whom we had information about both FVC and SNIP, we transformed SNIP of all subjects to FVC to allow comparability between cohorts.

Fractional polynomials were used to identify non-linear relationships with our composite survival endpoint.[27] Missing values were multiple imputed using multilevel joint modelling techniques in accordance with previous publications (appendix sections III.2 and III.3).[28,29]

We explored patient preferences with regard to knowing their personalised prognosis by conducting an online survey among 242 Dutch ALS patients (appendix section VII).

Statistical analysis

Data from the 14 cohorts were combined using the internal-external cross-validation (IECV) framework.[30,31] This develops a model for predicting our composite survival endpoint in all but one cohort, after which its external validity is evaluated in the omitted cohort. The process is 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 performed to assess overall performance of the model and to identify sources of between-study heterogeneity.[32] Between-study heterogeneity, such as differences in survival time or differences in (definition of) predictors between different cohorts might affect the generalisability of a model. The IECV framework thus provides a means to learn about the model's generalisability across different settings and populations.

For model development, we used multivariable Royston-Parmar survival models (RP-models)[33] rather than Cox survival models to facilitate the calculation of absolute risks in individual patients when implementing the model in clinical practice (appendix section III.4). 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.

For model validation, we assessed discrimination, i.e. the ability to differentiate between patients who reached our composite endpoint and those who did not, and calibration, i.e. the agreement between observed and predicted risk. Discrimination was quantified using the c-statistic, which is the area under the receiver operator characteristic (ROC) curve generalised to all survival times as defined by the composite endpoint. A c-statistic indicates very good discriminative ability for values close to 1 and poor discriminative ability for values close to 0·5. We also measured time-dependent ROC curves three 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.[31,32,34] A calibration slope of 1 in combination with a calibration-in-the-large of 0 indicates good overall calibration (appendix sections III.5 to III.7).

We performed a complete case analysis and sensitivity analyses for omitting FTD or 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 out of the five prognostic groups (in total 25 patients), illustrating application of the model to individual patients.

Because the IECV approach allows evaluation of external performance across multiple studies, we meta-analysed estimates of model performance and calculated 95% confidence intervals (CI), using random-effects meta-analyses.[32] More importantly, we calculated 95% prediction intervals (PI) to quantify the range of model performance across different populations which helps to assess the model's potential generalisability.[32]

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 will achieve a certain predefined c-statistic *and* calibration slope in new (future) patients (see appendix sections III.8 and V for details and definitions of 'good' performance).

Data sharing

For medical doctors, we are making our developed and validated model freely available via an online tool (www.encalssurvivalmodel.org). We report all parameters of the model in the appendix and further supporting information can be requested via the online tool or corresponding author.

Role of the funding source
The funders of the study had no role in study design; data collection, analysis, or interpretation; or writing of the Article or the decision to submit the paper for publication. 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 January 1, 1992, and September 22, 2016, the different cohorts having different start and end dates within this period (Supplementary Tables S1 to S13). 11475 ALS patients from 14 European ALS centres across 9 countries participated in this study. Total follow-up was 40016 years and median follow-up time was 97·5 months (IQR, 52·9 to 168·5). Patient characteristics are summarised in Table 1 and Figures S1 and S2 of the appendix.

Based on backward elimination, eight of the 16 candidate predictors were selected for the multivariable prediction model: age at onset (n=10000, 100% of bootstraps), FVC (n=7598, 76%), diagnostic delay (n=10000, 100%), ALSFRS-R slope (n=10000, 100%), bulbar onset (n=8094, 81%), 'definite' ALS (n=7120, 71%), presence of FTD (n=7416, 74%) or the *C9orf72* repeat expansion (n=8679, 87%); see appendix section III.1 for variables that were not included in the model. Age at onset, ALSFRS-R slope, diagnostic delay and FVC were transformed because of non-linear relationships with our composite survival endpoint (appendix section III.2); the relative effects of the predictors after transformation are shown in Table 2. Using all imputed datasets, we found that a proportional odds RP-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 outperforms a proportional hazard model (appendix section III.4).

Cohort	1	2	3	4	5	6	7	8
ALS centre	Utrecht, NLD	Dublin, IRL	Torino, ITA	Sheffield, UK	London, UK	Oxford, UK	Leuven, BEL	Lisbon, PRT
Data source	Population-based	Population-based	Population-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based
N	1936	1818	1022	1187	1266	849	833	594
Duration of acquisition (years)	9·2	20·5	9·2	15·7	23·1	15·2	23·8	17·4
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)
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)
Progression rate (points/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)
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)
Bulbar onset	610 (32·9)	690 (38·9)	347 (34·0)	329 (27·7)	411 (33·0)	200 (25·4)	224 (27·0)	196 (33·1)
'Definite' ALS[a]	354 (19·2)	942 (57·4)	342 (33·6)	298 (27·1)	329 (28·2)	--	64 (20·4)	146 (24·6)
Frontotemporal dementia	60 (3·8)	86 (23·5)	127 (12·4)	18 (1·5)	12 (0·9)	25 (3·1)	40 (4·8)	31 (5·2)
***C9orf72* repeat expansion**	155 (8·8)	71 (10·2)	50 (6·3)	51 (12·7)	5 (8·5)	--	99 (16·4)	17 (14·0)
Survival since onset (months)[b]	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)
Composite endpoint of survival[c]	1460 (75·4)	1586 (87·2)	834 (81·6)	850 (71·6)	1131 (89·3)	699 (82·3)	645 (77·4)	374 (63·0)

(Continuing below)

Cohort	9	10	11	12	13	14	Total	HR (95% CI)[d]
ALS centre	Hannover, GER	Ulm, GER	Jena, GER	St. Gallen, CHE	Tours, FRA	Limoges, FRA		
Data source	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based	Referral-based		
N	506	443	343	286	190	202	11475 (100·0)	
Duration of acquisition (years)	12·9	6·8	8·0	13·0	5·2	5·2	12·9 (8·3-17·0)	
Age at onset (years)	60·2 (51·9-67·0)	61·1 (52·0-69·3)	61·7 (52·5-68·8)	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)	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)[e]
Progression rate (points/month)	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)[f]
Forced vital capacity (%)	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	121 (27·7)	130 (29·3)	90 (29·3)	82 (29·5)	64 (36·0)	59 (29·2)	3553 (31·9)	1·71 (1·63-1·79)
'Definite' ALS[a]	49 (9·7)	227 (51·2)	40 (16·7)	61 (31·3)	68 (38·9)	33 (17·6)	2953 (31·3)	1·47 (1·39-1·55)
Frontotemporal dementia	4 (0·8)	21 (4·7)	--	2 (0·7)	17 (8·9)	13 (6·4)	456 (4·9)	1·34 (1·20-1·50)
***C9orf72* repeat expansion**	--	--	4 (3·8)	13 (4·7)	27 (14·2)	15 (60·0)	497 (10·0)	1·45 (1·31-1·61)
Survival since onset (months)[b]	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)[g]	
Composite endpoint of survival[c]	443 (87·5)	191 (43·1)	189 (55·1)	139 (48·6)	114 (60·0)	164 (81·2)	8819 (76·9)	

Table 1. Patient characteristics.

Data are median (interquartile range (IQR)) or number (%).Numbers 1-14 indicate the 14 different participating ALS research centres. "--" Indicates unavailable data. [a]'Definite' ALS according to the El Escorial criteria. [b]Data are median time between onset and composite survival endpoint (95% confidence interval). [c]Composite endpoint of survival was defined as the presence of non-invasive ventilation (NIV) >23 hours/day, or tracheostomy, or death. Different cohorts might use different definitions as shown in the appendix (Supplementary Tables S1-S13). [d]Data are hazard ratio (95% confidence interval); hazard ratios are calculated with a univariable random-effects Cox model with 14 ALS research centres as grouping factor for the random effects. [e]Because of non-normal distributed data, a natural logarithm transformation was applied. [f]Because of non-normal data, a square-root transformation was applied. [g]Calculated with a random-effects Cox model with 14 ALS research centres as grouping factor for the

random effects. Abbreviations: ALSFRS-R=revised version of the ALS functional rating scale. HR=hazard ratio. 95% CI=95% confidence interval. NA=not available. NLD=The Netherlands. IRL=Ireland. ITA=Italy. UK=United Kingdom. BEL=Belgium. PRT=Portugal. GER=Germany. CHE=Switzerland. FRA=France.

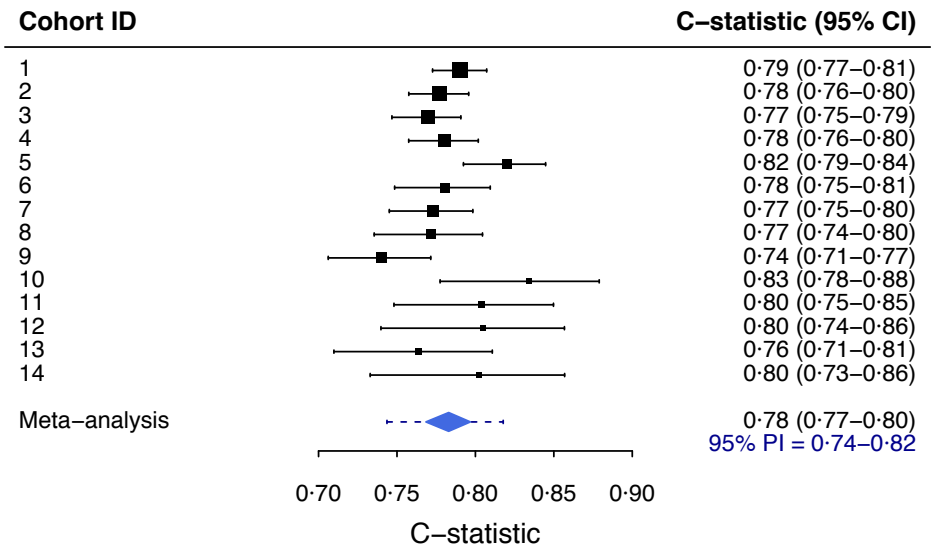
Predictors	Univariable HR[b]	Multivariable HR[c]
Age at onset (years)	1.28 (1.25-1.31)	1.02 (1.02-1.03)
Diagnostic delay (months)	0.48 (0.46-0.49)	0.63 (0.59-0.68)[d]
Progression rate (points/month)	3.05 (2.84-3.27)	3.19 (2.71-3.75)[e]
Forced vital capacity (%)	0.76 (0.72-0.80)	0.99 (0.99-0.99)
Bulbar onset	--	1.25 (1.17-1.33)
'Definite' ALS[a]	--	1.25 (1.16-1.34)
Frontotemporal dementia	--	1.18 (1.01-1.37)
***C9orf72* repeat expansion**	--	1.37 (1.19-1.57)

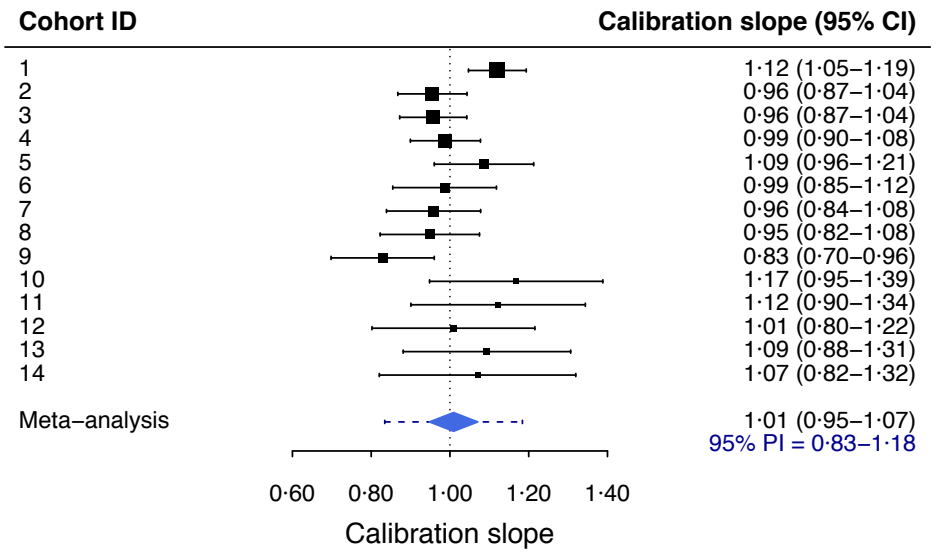
Table 2. Relative effects of transformed predictors.
Data are hazard ratio (95% confidence interval). All HRs were estimated using random-effects Cox models with the 14 ALS research centres as grouping factor for the random effects. Please note that this Table is intended to provide an indication of the relative effects of the transformed predictors and that the used random-effects Cox model differs from the Royston-Parmar proportional odds model that was used to generate the predictions. [a]'Definite' ALS according to the El Escorial criteria. [b]Using the different transformed continuous predictors, we made a comparison between the first and third quartile and report the univariable HRs (see appendix III.2 for applied transformations). This illustrates the relative effects of the continuous predictors after transformation. As the last four variables are dichotomous, it is not possible to make a comparison between a first and third quartile because this does not exist for dichotomous variables. [c]This column shows the HRs when all eight predictors were combined (See Supplementary Figure S2 for more details). HRs result from coefficients and thus do not rely on the quartiles that are used in the univariate HR column. [d]Because of non-normal distributed data, a natural logarithm transformation was applied. [e]Because of non-normal data, a square-root transformation was applied. HR = hazard ratio.

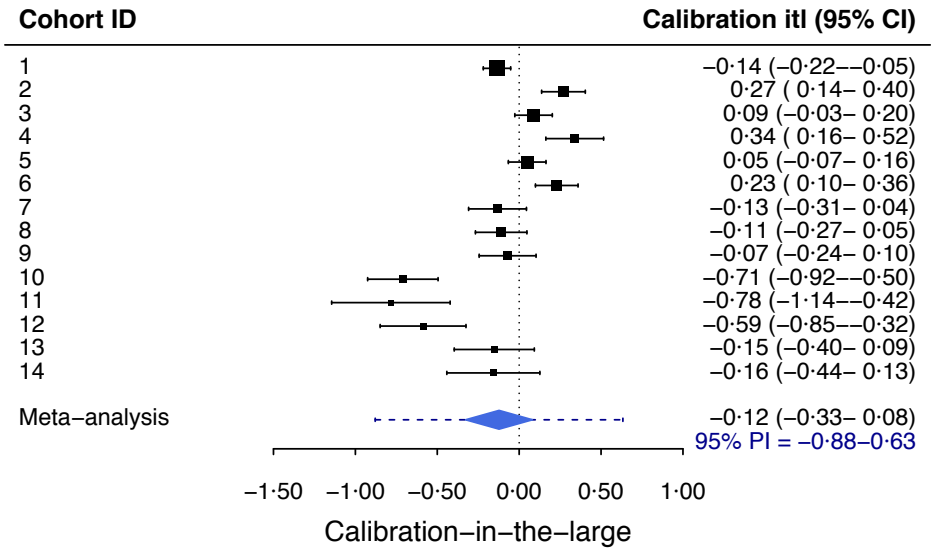
The IECV framework resulted in 14 cycles of model development and external validation. The meta-analysis of these 14 external validations showed a summary c-statistic of 0·78 (95% CI, 0·77 to 0·80; 95% PI, 0·74 to 0·82), an area under the time-dependent ROC curve three years after onset of 0·86 (95% CI, 0·85 to 0·88; 95% PI, 0·82 to 0·90), a calibration slope of 1·01 (95% CI, 0.95 to 1·07; 95% PI, 0·83 to 1·18) and calibration-in-the-large of -0·12 (95% CI, -0·33 to 0·08; 95% PI, -0·88 to 0·63) (Figure 1A-C and Figure S4 of the appendix). The five prognostic groups that were created, based on this prediction model, showed good agreement: observed median times to our composite endpoint (standard error) were 16·5 (0·23), 25·2 (0·35), 32·8 (0·46), 44·6 (0·74) and 85·6 (1·96) months and predicted median times to our composite endpoint (standard error) were 17·7 (0·20), 25·3 (0·06), 32·2 (0·09), 43·7 (0·21) and 91·0 (1·84) months (Figure 1D). Comparing 'very short', 'short', 'intermediate' and 'long' groups with the group with very long times to our composite endpoint revealed a hazard ratio (HR) of 15·29 (95% CI, 13·92 to 16·79), 7·19 (95% CI, 6·53 to 7·91), 4·30 (95% CI, 3·94 to 4·69) and 2·49 (95% CI, 2·30 to 2·70). Visual inspection of cohort-specific calibration plots three years after onset showed good agreement between predicted and observed probability for most cohorts (N=11; Figure S7 of the appendix). Predicted and observed probabilities of reaching our composite endpoint for cohort-specific curves showed similar results as calibration plots (Figure S5 of the appendix). For three cohorts where agreement between observed and predicted times to our composite endpoint was suboptimal, recalibration of the baseline hazard notably improved calibration (Figure S6 of the appendix). Similar results were found for complete case analysis (Figure S14 of the appendix). We assessed all possible two-way interactions (i.e. eight covariates with one interaction comprising two of the eight covariates), which did not improve predictive accuracy (Table S20 of the appendix). The probability of 'good' external predictive performance of the model was estimated to be 100·0% for c-statistic and 97·1% for calibration slope. The joint (combined) probability of 'good' performance was 98·3% (Supplementary Figure S8). Compared to estimates of group level data, the model provides more accurate and precise predictions of times to our composite endpoint (Figure 2).

An online tool for prediction of time to our composite survival endpoint is available at www.encalssurvivalmodel.org. This instrument is based on the regression coefficients of the final prediction model (appendix (section III.9)) and can be tailored per population. To illustrate its usage in clinical practice, we provide examples of patients selected from our dataset and their associated predictions, including the level of uncertainty at patient level (Figure 3) and provide worked examples (appendix section VIII). The sensitivity analyses demonstrated that the prediction model also provides accurate predictions when patient-level information about the presence of FTD or a *C9orf72* repeat

expansion is not available and similar results of predictions using date of onset versus date of diagnosis as starting point for prediction (appendix section IV). Patients' preferences with regard to knowing their personalised prognosis are presented in appendix section VII.


Cohort ID
C–statistic (95% CI)
1 0·79 (0·77–0·81)
2 0·78 (0·76–0·80)
3 0·77 (0·75–0·79)
4 0·78 (0·76–0·80)
5 0·82 (0·79–0·84)
6 0·78 (0·75–0·81)
7 0·77 (0·75–0·80)
8 0·77 (0·74–0·80)
9 0·74 (0·71–0·77)
10 0·83 (0·78–0·88)
11 0·80 (0·75–0·85)
12 0·80 (0·74–0·86)
13 0·76 (0·71–0·81)
14 0·80 (0·73–0·86)
Meta–analysis
0·78 (0·77–0·80)
95% PI = 0·74–0·82
0·70 0·75 0·80 0·85 0·90
C–statistic


Cohort ID
Calibration slope (95% CI)
1 1·12 (1·05–1·19)
2 0·96 (0·87–1·04)
3 0·96 (0·87–1·04)
4 0·99 (0·90–1·08)
5 1·09 (0·96–1·21)
6 0·99 (0·85–1·12)
7 0·96 (0·84–1·08)
8 0·95 (0·82–1·08)
9 0·83 (0·70–0·96)
10 1·17 (0·95–1·39)
11 1·12 (0·90–1·34)
12 1·01 (0·80–1·22)
13 1·09 (0·88–1·31)
14 1·07 (0·82–1·32)
Meta–analysis
1·01 (0·95–1·07)
95% PI = 0·83–1·18
0·60 0·80 1·00 1·20 1·40
Calibration slope


Cohort ID
Calibration itl (95% CI)
1 –0·14 (–0·22––0·05)
2 0·27 (0·14– 0·40)
3 0·09 (–0·03– 0·20)
4 0·34 (0·16– 0·52)
5 0·05 (–0·07– 0·16)
6 0·23 (0·10– 0·36)
7 –0·13 (–0·31– 0·04)
8 –0·11 (–0·27– 0·05)
9 –0·07 (–0·24– 0·10)
10 –0·71 (–0·92––0·50)
11 –0·78 (–1·14––0·42)
12 –0·59 (–0·85––0·32)
13 –0·15 (–0·40– 0·09)
14 –0·16 (–0·44– 0·13)
Meta–analysis
–0·12 (–0·33– 0·08)
95% PI = –0·88–0·63
–1·50 –1·00 –0·50 0·00 0·50 1·00
Calibration–in–the–large

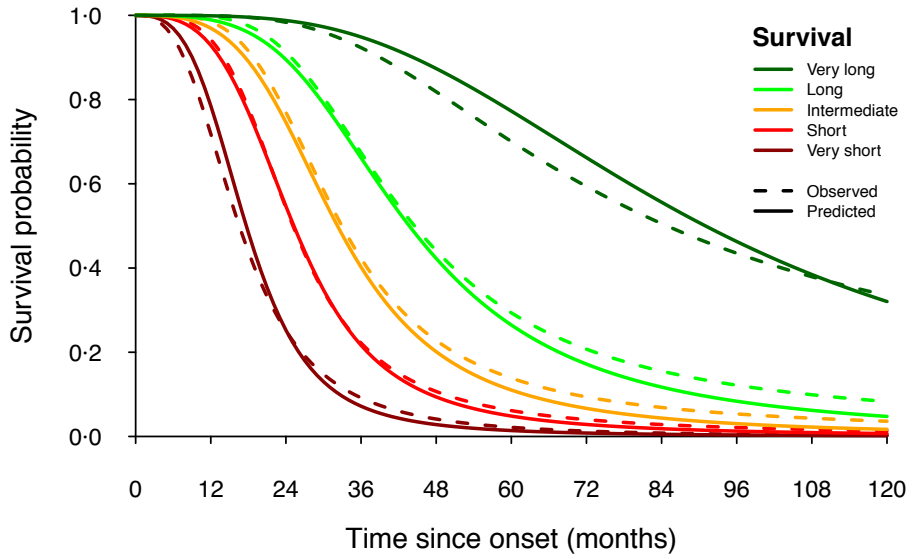

Survival probability
1·0
0·8
0·6
0·4
0·2
0·0
0 12 24 36 48 60 72 84 96 108 120
Time since onset (months)
Survival
Very long
Long
Intermediate
Short
Very short
Observed
Predicted

Figure 1. Univariate meta-analysis of predictive performance.
Panel A shows the random effects meta-analysis of discrimination as measured with the c-statistic. C-statistic can be viewed as a generalised area under the receiver operator curve (ROC) for all possible time-points. It measures the ability of the model to differentiate between patients who reached our composite endpoint and those who did not. See Figure S4 of the appendix for time-dependent ROC curves. Panels B and C show the random effects meta-analysis of calibration as measured with calibration slope (panel B) and calibration-in-the-large (panel C). Calibration measures the agreement between observed and predicted time to our composite endpoint. A calibration slope of 1 in combination with a calibration-in-the-large of 0 indicates good overall calibration (see appendix section III.7 for further details). The blue 'diamond' indicates the mean and 95%-confidence interval (95% CI) of the predictive accuracy. 95%-prediction intervals (95% PI), which indicate predicted accuracy of the model in a single new dataset or patient, are depicted as dashed blue lines and as a numeric range (in blue, below 95% CI). Panel D is the visual translation of panels A-C in prognostic curves showing the agreement between predicted and observed probability of reaching our 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. Cohort ID indicates the cohort left out in the internal-external cross-validation (IECV), which was performed once for all 14 different datasets. Abbreviations: calibration itl = calibration-in-the-large, 95% PI = 95% prediction interval (reflecting 95% range of accuracy of new predictions).

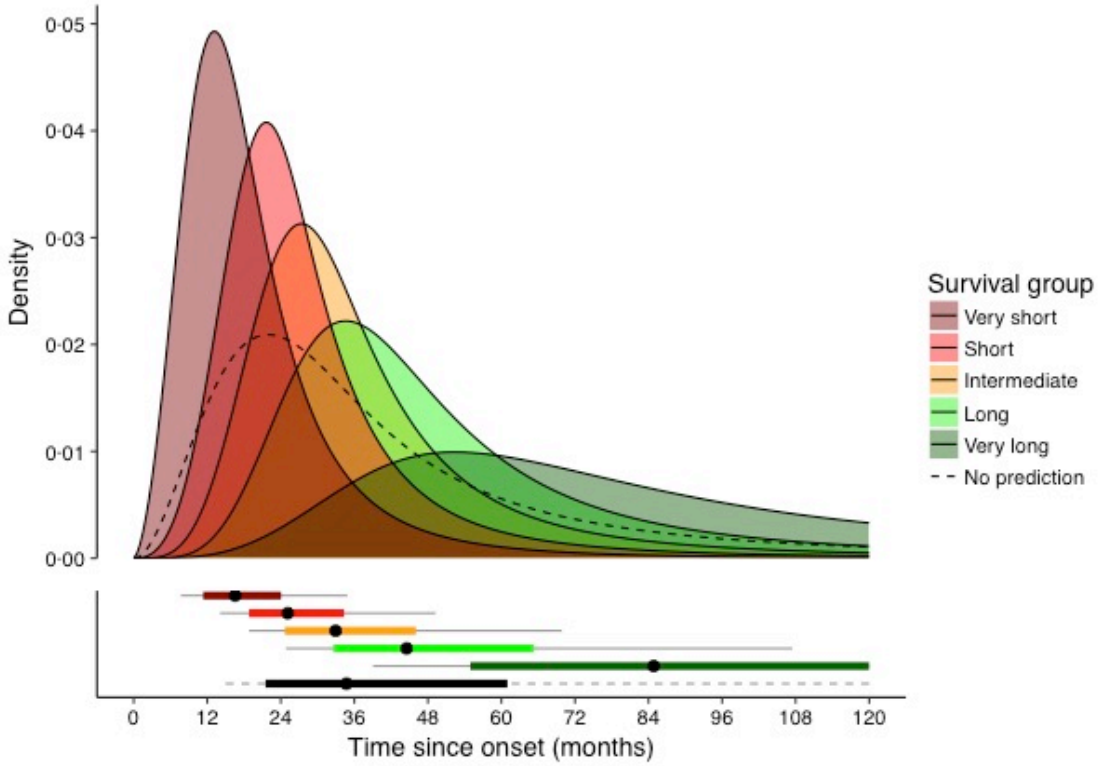
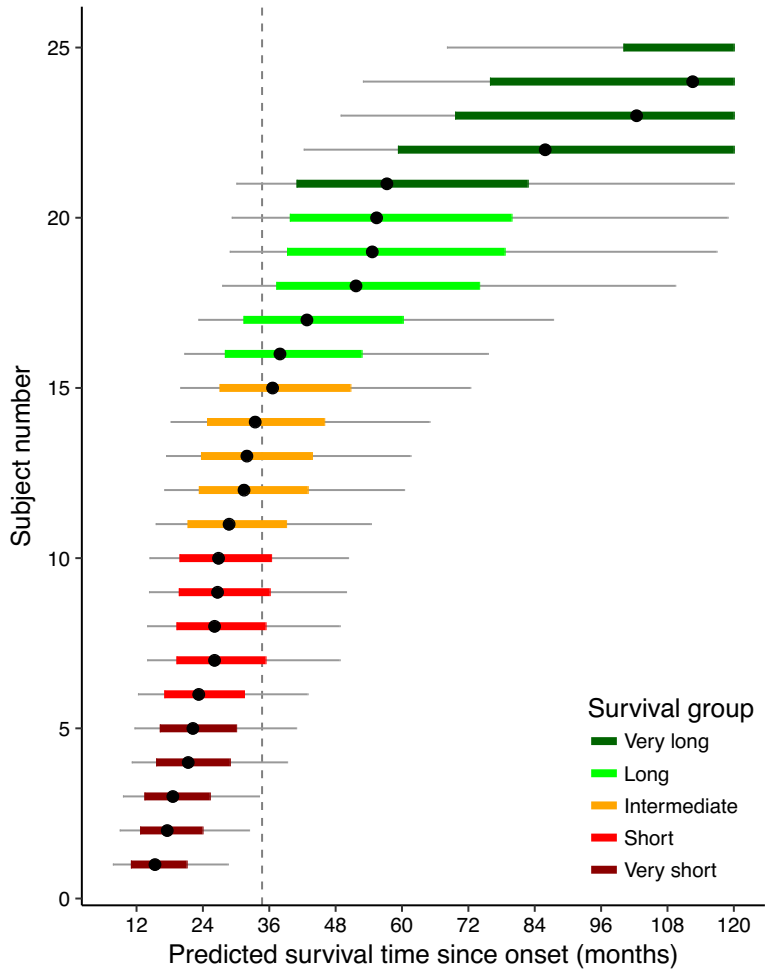


Figure 2. Risk group predictions.
This Figure shows probability densities for times to our composite endpoint (y-axis) of the five risk groups in Figure 1D on a time scale (x-axis). Density reflects the probability density (or distribution), which is not evenly distributed over time, meaning that the largest number of patients will die where the curve has the highest density, but that there is a small chance that patients survive without tracheostomy or non-invasive ventilation for >23 hours/day much longer; this is reflected by the long, thin tails on the right side of this Figure. This provides guidance for discussing prognosis with patients. Furthermore, if the prediction model is not used (dashed curve), predictions will generally be overly positive (for groups with short and very short times to our composite endpoint), overly negative (for groups with long and very long times to our composite endpoint), or overly uncertain (for group with intermediate times to our 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, boundaries of thick lines representing 25 to 75% probability intervals and boundaries of thin lines representing 10 to 90% probability intervals to reach our composite survival endpoint. This Figure quantifies the uncertainty of individual predictions because, for example, a subject with a predicted very short time to our composite endpoint still has a 2% chance of surviving up to 5 years after onset.



Progression rate	Diagnostic delay	Age at onset	FVC	Bulbar onset	'Definite' ALS	FTD	C9
0·13	39·0	42·3	105·0	–	–	–	–
0·14	36·0	54·9	113·0	–	–	–	–
0·14	18·1	37·6	112·0	+	–	–	–
0·21	16·9	43·1	81·0	–	–	–	–
0·36	25·0	65·3	109·0	–	–	–	–
0·22	20·0	73·0	98·0	+	–	–	–
0·28	24·0	71·5	125·0	+	–	–	+
0·56	23·3	62·9	107·0	–	–	–	–
0·17	6·0	42·6	47·0	–	–	–	–
0·94	23·3	63·4	75·0	+	–	–	–
0·34	8·6	61·9	101·0	+	–	–	–
0·30	11·4	52·5	44·3	–	+	–	–
0·38	7·5	66·5	84·0	+	–	–	–
0·72	13·9	76·5	98·0	–	+	–	–
1·34	9·0	53·4	99·0	–	–	–	–
0·95	6·7	51·1	76·0	+	–	–	–
0·46	6·0	76·6	85·0	–	–	+	–
1·75	12·0	67·0	113·0	+	–	–	–
1·96	9·2	54·6	98·0	+	–	–	–
1·69	10·7	74·3	70·0	+	–	–	–
0·87	12·1	70·6	36·0	+	–	–	–
1·41	5·0	56·2	96·0	–	–	+	–
2·61	6·9	54·3	81·0	–	+	+	+
3·92	5·4	74·6	84·0	+	–	–	–
1·26	12·0	77·2	36·0	+	+	+	–

Figure 3. Clinical applicability to individual patients.
Characteristics and predictions of 25 randomly selected ALS patients (using a computer algorithm developed to select five patients per risk group from all available patients) are presented to illustrate application of the tool to predict time to our composite endpoint to individual patients. Predictions are provided on the left and patient characteristics on the right. Median predicted time to our composite endpoint of individual ALS patients is represented by the dots. The boundaries of the thick lines represent 25 to 75% probability intervals and boundaries of the thin lines represent 10 to 90% probability intervals. The vertical dashed line represents the median time to our composite endpoint (34·7 months). A large proportion of patients deviates from the median time to reach our composite survival endpoint (at group level): six patients (24%) have a 75% chance of dying, having tracheostomy, or having NIV >23 hours/day before the median time to our composite endpoint while eight other patients (32%) have a 75% chance of living without thracheostomy or NIV >23 hours/day longer than median time to our composite endpoint. Predicted times to our composite endpoint differ significantly between patients, reflecting both heterogeneity in ALS as well as the ability of the proposed model to stratify patients according to their characteristics.
Abbreviations: + = present; - = absent; ALSFRS-R slope is in points decrease of the ALSFRS-R score per month; diagnostic delay is in months; age at onset is in years; FVC is forced vital capacity and is expressed as percentage of predicted (corrected for age, sex and body height); 'definite' ALS is according to the El Escorial criteria; FTD is presence of frontotemporal dementia; C9 is presence of a *C9orf72* repeat expansion.

Discussion

On the basis of data from 11,475 ALS patients, collected across Europe, we have developed and externally validated a model for prediction of survival without tracheostomy or non-invasive ventilation more than 23 hours/day in individual ALS patients. Using recently developed state-of-the-art methods, we derived ranges of predictive performance across different settings and patient populations, thereby providing evidence for the potential generalisability of our model. We implemented our model in a free online tool for medical doctors to provide estimates of prognosis in individual ALS patients, with the aim of facilitating its use in clinical practice and in innovative trial design.

Associations between predictors and survival ourcomes in ALS have previously been reported at group level.[6] Attempts to translate predictive associations into estimates of survival outcomes in individual ALS patients 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 IPD meta-analysis, including more than 11000 subjects, is the first study in ALS to enable the investigation of a relatively 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 implemented recent methodology to identify sources of heterogeneity across the European cohorts, and addressed transportability issues of the model.[30,31] This enabled us to determine the likely performance when applying the model in clinical practice, and thus to assess its clinical utility 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 more than 23 h per day, prognostic groups or point estimates. For care, we prefer the first two options, because these provide the possibility of being realistic without driving a patient to despair (see appendix section VIII for worked examples). Because of the probability distributions for times to our composite endpoint (Figure 2), which are also skewed, point estimates may not be preferred for predictions in individual patients. Ranges will become wider for subjects with longer times to our composite endpoint (Figure 2) due to less steep prognostic curves.

Large-scale European Union funded collaborative projects, such as Euro-MOTOR (www.euromotorproject.eu), SOPHIA (www.sophiaproject.eu) and NETCALS (www.jpnd.eu), encourage standardised and harmonised data collection between three population-based (n=4776, 42%) and 11 referral-based patient registries in this study (Table 1).[36,37] Differences in standard of care, or 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 (Table 1). This might be due to real differences between populations. Other possibilities may 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. In our study we focussed on developing and validating a generalisable prediction model for ALS. Importantly, state-of-the-art IECV 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.[17] Although the effect of riluzole on survival (pooled HR 0·84, based on a Cochrane meta-analysis)[18] 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 recently developed imputation methods, which are proven to prevent such biases, even if a predictor is completely missing from one or more cohorts.[28] Using IECV, we rigorously assessed the model's predictive performance in external datasets. This approach not only allows one to evaluate the model's overall performance in new patients, but also to establish whether

performance is consistent across the different cohorts. It therefore helps to determine to what extent model generalisability may 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 (e.g. 'intrinsic' differences of ALS between cohorts, or 'extrinsic' due to possible differences in, for example, clinical care), these differences did not appear to reduce the predictive performance of our proposed prediction model, as demonstrated by IECV.

Our newly developed online tool is ethically sensitive and demands thoughtful implementation. The tool may support patients in maintaining a degree of autonomy and help them in planning their lives. Our survey demonstrated that there is a group of patients who prefer to be informed about their personalised life expectancy. This is in line with previous publications 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] On the other hand, many cancers have multiple 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; patients may not oversee all consequences of the acquired knowledge and may regret asking to be informed if the predicted time to death, tracheostomy, or NIV >23 hours/day they receive is shorter than expected. To minimise the risk of such potential harm, the online tool will only be accessible to medical doctors, who have to register and sign before access is provided (i.e. controlled access). Professional health care providers may 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 (1) by using the predicted outcome as inclusion criterion instead of arbitrary eligibility criteria (such as pulmonary function), and (2) by improving stratified randomisation using predicted prognosis instead of a limited set of prognostic variables, thereby creating more homogeneous strata and greater statistical power to detect an effect. This was, for example, previously shown in glioblastoma patients with methylation of the MGMT (O^6-methylguanine–DNA methyltransferase) gene, illustrating that prediction might be a promising tool for future clinical trials.[40]

In accordance with the TRIPOD statement,[10] all parameters and equations of our model are provided (appendix section III.4 and III.9) to allow improvement of predictions through continuing research into other prognostic factors, such as wet or imaging biomarkers, or newly discovered genetic mutations, or to implement outcome measures other than our composite survival endpoint, e.g. being wheelchair-bound or other specific loss of function. This also facilitates validation of our model in registries from a non-European background and its application in trial populations.

In conclusion, we have developed a model for prediction of survival free of tracheostomy or non-invasive ventilation more than 23 hours/day in individual ALS patients and externally validated this model multiple times. We have demonstrated the generalisable and robust predictive performance of this model and made it freely available as an easy-to-use online tool (http://www.encalssurvivalmodel.org). The outcomes of this study will facilitate tailored care and trial design, which will hopefully lead to a more successful discovery of effective treatments for patients suffering from this devastating disease.

Declaration of interests

Dr. Rooney reports grants from Health Research Board (grant number HPF-2014-537), during the conduct of the study.

Dr. McDermott was supported by the MND association.

Dr. van Es serves on the Motor Neurone Disease Association biomedical research advisory panel, has consulted for Biogen and has 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 Joint Programme Neurodegenerative Disease Research (JPND).

Dr. Weber reports personal fees from Biogen Idec and Mitsubishi Tanabe Pharma, outside the submitted work.

Dr. Petri reports grants from the German Neuromuscular Society, the German Federal Ministry of Education and Research (BMBF), clinical trial funding from Cytokinetics, GlaxoSmithKline, and Orion Pharma, and speaking fees from Teva Pharmaceutical Industries, outside the submitted work.

Dr. Van Damme reports personal fees for advisory work from Cytokinetics, outside the submitted work; personal fees for advisory work (paid to institution) from Biogen, Pfizer, CSL Behring and Treeway, outside the submitted work.

Dr. Talbot is supported by the Motor Neurone Disease Association and the UK Medical Research Council.

Dr. Turner is supported by the Medical Research Council & Motor Neurone Disease Association Lady Edith Wolfson Senior Clinical Fellowship (MR/K01014X/1).

Dr. Pamela J. Shaw reports grants from the NIHR (Senior Investigator Award), Sheffield NIHR Biomedical Research Centre for Translational Neuroscience, the Motor Neurone Disease Association, the European Union Joint Programme – Neurodegenerative Disease Research, (JPND; STRENGTH project), UK Medical Research Council, European Union Horizon 2020 (Modifying Immune Response and Outcomes in ALS [MIROCALS] trial), during the conduct of the study; reports grants from ReNeuron, clinical trial funding from Cytokinetics, Orion Pharma, and Biogen Idec, outside the submitted work.

Dr. Al-Chalabi is a consultant from Mitsubishi Tanabe Pharma, Chronos Therapeutics, Orion Pharma, Cytokinetics, and Treeway, outside the submitted work.

Dr. Chio reports grants from Italfarmaco, and personal fees from Biogen Idec and Mitsubishi, outside the submitted work.

Dr. Hardiman is a journal editor for Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration; reports grants from Health Research Board (CSA 2012/11), Science Foundation Ireland (15/SPP/3244 and 16/ERCD/3854), and the charity Research Motor Neurone, outside the submitted work.

Dr. van den Berg reports grants from The Netherlands ALS Foundation, The Netherlands Organization for Health Research and Development (ZonMw; Vici scheme), the European Union 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 the European Union Joint Programme – Neurodegenerative Disease Research (JPND), during the conduct of the study; grants from Baxalta, personal fees from Biogen, Cytokinetics and Baxalta, outside the submitted work.

All other authors declare no competing interests.

Acknowledgements
The Netherlands ALS Foundation funded this study. We thank all ALS patients who, notwithstanding their devastating disease, participated in this study.

Author contributions

Henk-Jan Westeneng, Thomas P.A. Debray, Carl K.G.M. Moons, Jan H. Veldink and Leonard H. van den Berg contributed to the study design. Henk-Jan Westeneng did the literature search. Henk-Jan Westeneng and Thomas P.A. Debray analysed the data. Henk-Jan Westeneng created the figures. All authors contributed to the data collection, interpretation of results, reviewed and critically revised the Article, and approved the final version for submission.

References

1. Mortality GBD, Causes of Death C. 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**(10053): 1459-544.
2. Jameson JL, Longo DL. Precision medicine--personalized, problematic, and promising. *N Engl J Med* 2015; **372**(23): 2229-34.
3. Kiernan MC, Vucic S, Cheah BC, et al. Amyotrophic lateral sclerosis. *Lancet* 2011; **377**(9769): 942-55.
4. van Es MA, Hardiman O, Chio A, et al. Amyotrophic lateral sclerosis. *Lancet* 2017.
5. Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* 2005; **65**(4): 586-90.
6. Chio A, Logroscino G, Hardiman O, et al. Prognostic factors in ALS: A critical review. *Amyotroph Lateral Scler* 2009; **10**(5-6): 310-23.
7. 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**(9): 878-82.
8. 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**(2): 214-26.
9. 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.
10. 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.
11. 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**(10): 1165-70.
12. Brooks BR, Miller RG, Swash M, Munsat TL, World Federation of Neurology Research Group on Motor Neuron D. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000; **1**(5): 293-9.
13. 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.
14. 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**(2): 265-7.
15. 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**(9): 829-38.
16. 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**(11): 1229-33.

17. Cudkowicz ME, van den Berg LH, Shefner JM, et al. Dexpramipexole versus placebo for patients with amyotrophic lateral sclerosis (EMPOWER): a randomised, double-blind, phase 3 trial. *Lancet Neurol* 2013; **12**(11): 1059-67.
18. Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev* 2012; (3): CD001447.
19. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011; **134**(Pt 9): 2456-77.
20. 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**(6): 734-47.
21. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. *Neurology* 2000; **55**(11): 1621-6.
22. Raaphorst J, Beeldman E, Schmand B, et al. The ALS-FTD-Q: a new screening tool for behavioral disturbances in ALS. *Neurology* 2012; **79**(13): 1377-83.
23. Diekstra FP, van Vught PW, van Rheenen W, et al. UNC13A is a modifier of survival in amyotrophic lateral sclerosis. *Neurobiol Aging* 2012; **33**(3): 630 e3-8.
24. 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**(4): 323-30.
25. 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**(6): 689-700.
26. 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**(3): 146-52.
27. Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *Journal of the Royal Statistical Society Series a-Statistics in Society* 1999; **162**: 71-94.
28. Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. *Stat Med* 2016; **35**(17): 2938-54.
29. 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.
30. 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**(18): 3158-80.
31. 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**(10): e1001886.

32. 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.
33. 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**(15): 2175-97.
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**(6): 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. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 2017: 1-8.
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.
38. Innes S, Payne S. Advanced cancer patients' prognostic information preferences: a review. *Palliat Med* 2009; **23**(1): 29-39.
39. Bernacki RE, Block SD, American College of Physicians High Value Care Task F. Communication about serious illness care goals: a review and synthesis of best practices. *JAMA Intern Med* 2014; **174**(12): 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**(10): 997-1003.