

# Refining eligibility criteria for amyotrophic lateral sclerosis clinical trials

Ruben P.A. van Eijk, MD, MSc, Henk-Jan Westeneng, MD, Stavros Nikolakopoulos, PhD, Iris E. Verhagen, MD, Michael A. van Es, MD, PhD, Marinus J.C. Eijkemans, PhD, and Leonard H. van den Berg, MD, PhD

*Neurology*® 2019;92:e451-e460. doi:10.1212/WNL.0000000000006855

## Correspondence

Dr. van den Berg  
lberg@umcutrecht.nl

## Abstract

### Objective

To assess the effect of eligibility criteria on exclusion rates, generalizability, and outcome heterogeneity in amyotrophic lateral sclerosis (ALS) clinical trials and to assess the value of a risk-based inclusion criterion.

### Methods

A literature search was performed to summarize the eligibility criteria of clinical trials. The extracted criteria were applied to an incidence cohort of 2,904 consecutive patients with ALS to quantify their effects on generalizability and outcome heterogeneity. We evaluated the effect of a risk-based selection approach on trial design using a personalized survival prediction model.

### Results

We identified 38 trials. A large variability exists between trials in all patient characteristics for enrolled patients ( $p < 0.001$ ), except for the proportion of men ( $p = 0.21$ ). Exclusion rates varied widely (from 14% to 95%; mean 59.8%; 95% confidence interval 52.6%–66.7%). Stratification of the eligible populations into prognostic subgroups showed that eligibility criteria lead to exclusion of patients in all prognostic groups. Eligibility criteria neither reduce heterogeneity in survival time (from 22.0 to 20.5 months,  $p = 0.09$ ) nor affect between-patient variability in functional decline (from 0.62 to 0.65,  $p = 0.25$ ). In none of the 38 trials were the eligibility criteria found to be more efficient than the prediction model in optimizing sample size and eligibility rate.

### Conclusions

The majority of patients with ALS are excluded from trial participation, which questions the generalizability of trial results. Eligibility criteria only minimally improve homogeneity in trial endpoints. An individualized risk-based criterion could be used to balance the gains in trial design and loss in generalizability.

## RELATED ARTICLE

### Editorial

Heterogeneity, urgency, generalizability, and enrollment: The HUGE balance in ALS trials

Page 215

From the Department of Neurology (R.P.A.v.E., H.-J.W., I.E.V., M.A.v.E., L.H.v.d.B.), Brain Center Rudolf Magnus, and Biostatistics & Research Support (R.P.A.v.E., S.N., M.J.C.E.), Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, the Netherlands.

Go to [Neurology.org/N](https://www.neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

The Article Processing Charge was funded by The Netherlands ALS Foundation.

## Glossary

**ALS** = amyotrophic lateral sclerosis; **ALSFRS** = Amyotrophic Lateral Sclerosis Functional Rating Scale; **ALSFRS-R** = Amyotrophic Lateral Sclerosis Functional Rating Scale–revised; **CI** = confidence interval; **ENCALS** = European Network for the Cure of ALS; **FVC** = forced vital capacity.

The clinical heterogeneity of amyotrophic lateral sclerosis (ALS) makes conducting clinical trials complex.<sup>1</sup> To increase the potential to demonstrate therapeutic efficacy, investigators apply eligibility criteria to enroll a more homogeneous population, to improve protocol adherence, or to exclude patients who are unlikely to benefit.<sup>2–4</sup> For ALS, this often means excluding those patients with long disease durations or those who are unlikely to survive the follow-up period. Thus, many patients are excluded,<sup>2</sup> which could have important consequences for the generalizability of a trial (i.e., “to whom do the results of this trial apply?”).<sup>4</sup> A low generalizability potentially provides limited information about safety or efficacy for the general population.<sup>2,4,5</sup>

Nevertheless, a pragmatic approach (i.e., no selection) may not be feasible in ALS because this increases endpoint variability and may inflate sample size.<sup>6</sup> It is therefore important to balance the generalizability of a trial (i.e., selecting a population that represents the general population) and its endpoint heterogeneity (i.e., selecting a sensitive population to show efficacy). Prediction models such as the recently validated ALS survival model<sup>7</sup> might improve the selection of patients by using the predicted outcome as an inclusion criterion instead of a set of arbitrary criteria.<sup>8</sup>

Virtually all ALS clinical trials have imposed various sets of eligibility criteria, but little is known about the consequences.<sup>2,9</sup> In this study, we review the current practices of participant selection and assess the effects on trial populations, efficacy endpoints, and generalizability. We then gauge the value of an individualized risk-based selection criterion in balancing outcome heterogeneity and patient exclusion rates.

## Methods

The effect of eligibility criteria on population characteristics and trial endpoints was estimated in a 2-step approach. First, we systematically reviewed the literature to compose a list of commonly used criteria and meta-analyzed the baseline characteristics of the included trials. Subsequently, we applied the extracted criteria to an incidence cohort to estimate their effects on survival and functional decline.

### Search strategy and trial selection

Two authors (R.P.A.v.E. and I.E.V.) individually searched the PubMed and Embase database for publications dating from January 1, 2000, up to and including November 2017 using the following search terms: amyotrophic lateral sclerosis or motor neuron disease and clinical trial. To harmonize the

comparison between clinical trials, we included only randomized, placebo-controlled, clinical trials evaluating the efficacy of a single pharmacologic agent. We excluded clinical trials investigating multiple agents, exclusively aiming to determine safety (phase I), having a nonclinical primary endpoint, or starting enrollment before the approval of riluzole (1996).

### Data extraction and harmonization

For each trial, we extracted the eligibility criteria and baseline characteristics. If a trial included multiple dosing groups, groups receiving the experimental agent were collapsed by a fixed-effects within-study meta-analysis. Predicted vital capacity was recorded as either the forced vital capacity (FVC) or slow vital capacity because there is no true difference between the 2 values.<sup>10</sup> When the mean Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) score was reported instead of the revised ALSFRS (ALSFRS-R), we added 7.76 to the mean and 0.23 to the reported SD because this procedure provided the least systematic error during simulations. The ALSFRS-R slope was calculated as the average monthly change from randomization. Studies reporting the ALSFRS slope were transformed to the ALSFRS-R slope by multiplying the mean ALSFRS rate of decline by 1.1 and its standard error by 1.14. If data were not reported as mean with SD (e.g., median and range), we transformed the reported estimates according to the method of Hozo et al.<sup>11</sup> Finally, if only the mean was reported without SD (6% of the studies), we imputed the natural logarithm of the SD from the natural logarithm of the mean using linear regression.<sup>12</sup>

### Exclusion rate and incidence cohort of patients with ALS

Our primary goal was to estimate, per set of eligibility criteria, the number of patients from the general population who would be considered ineligible to participate (i.e., exclusion rate). The exclusion rate per trial was defined as the number of ineligible patients divided by the total number of patients (population size). An exclusion rate of 60%, for example, indicates that only 40% of the general population was potentially evaluated in a given trial; the effectiveness and safety are unknown for 60% of the population. Because eligibility criteria are applied in a step-wise manner, the exclusion rate is a cumulative buildup of several percentages. An example of how this rate is calculated is presented in table 1.

The denominator is crucial for the accuracy of the exclusion rate. We therefore approximated the population size by including all consecutive patients with ALS in the Netherlands diagnosed between January 2006 and December 2016 in the

**Table 1** Example of the calculation of the exclusion rate for the Pentoxifylline 2006 trial

Criteria	Limits	Excluded, n (%)	Cumulative, n (%)
Age	18–80 y	205 (7.1)	205 (7.1)
Symptom duration	6–47 mo	773 (26.6)	934 (32.2)
Vital capacity	≤100% predicted	818 (28.2)	1,500 (51.7)
El Escorial criteria	Definite or probable	1,146 (39.5)	2,003 (69.0)

Using an incident cohort of patients with amyotrophic lateral sclerosis (ALS; n = 2,904), we calculated the individual and combined effects of 4 criteria on the exclusion rate. The cumulative, multivariate exclusion rate in this example is 69.0%. The cumulative exclusion rate is calculated by the sequential application of all 4 criteria (the default method in current ALS trials).

analysis. A detailed description of the incidence cohort of patients with ALS in the Netherlands, which has a coverage rate of 81%, has been given elsewhere.<sup>13</sup> In short, patients with ALS in the Netherlands diagnosed with possible, probable (laboratory supported), or definite ALS according to the revised El Escorial criteria are registered centrally at The Netherlands ALS Center.<sup>14</sup> Patients are located by annual screening of large medical center registries and by individually contacting Dutch neurologists.<sup>13</sup> For each patient, we determined whether he or she participated in a clinical trial in the same time period. In addition, clinical data were collected to determine an individual's eligibility for trial participation at the day of diagnosis (figure e-1 available from Dryad, doi.org/10.5061/dryad.86f1m6g). Time-varying variables such as ALSFRS-R total score and FVC were, ideally, collected at diagnosis or within a 3-month time interval; otherwise, the score was recorded as missing. Complete survival data (date of death or last follow-up) were obtained by checking the online municipal population register at 3-month intervals. In total, 39 surviving patients (1.3%) were followed up for <6 months and 139 (4.7%) for <12 months.

### Standard protocol approvals, registrations, and patient consents

The medical ethics committee and institutional review board of the University Medical Center Utrecht approved this study.

### Statistical analysis

Variables necessary to determine eligibility contained, on average, 36% missing values; kidney function was most often missing (in 82% of the cases). Eligibility was based on all variables provided (figure e-1 available from Dryad, doi.org/10.5061/dryad.86f1m6g). Missing data were accounted for by creating multiple imputed datasets (n = 100) using predictive mean matching and bootstrapping, discarding the first 100 iterations (burn-in). The imputation model contained all covariates; survival time was modeled with Nelson-Aalen estimates.<sup>15</sup> Subsequently, the extracted eligibility criteria were applied to each imputed dataset. Using the incidence

cohort as reference, we estimated the proportion of patients eligible for each trial individually. Results across imputations were pooled with Rubin rules.<sup>16</sup> Baseline characteristics and eligibility percentages were summarized by meta-analysis. Continuous variables were meta-analyzed using the raw reported means; count variables were recoded to logit-transformed proportions. Subsequently, between-population heterogeneity was evaluated with the Cochrane Q test and the  $I^2$  statistic.<sup>17</sup> All meta-analytical models were based on a random-effects model, fitted with restricted maximum-likelihood estimation. Standard errors of coefficients were adjusted according to the methodology given by Knapp and Hartung.<sup>18</sup>

For each patient, we determined his or her risk profile by taking the exponent of the centered linear predictor of the European Network for the Cure of ALS (ENCALS) survival model.<sup>7</sup> An estimated risk score of 0.5 indicates that the risk of dying during follow-up is half the risk of dying for the average patient with ALS, whereas a score of 2 means that this risk is twice the average. Using the quintiles of the individual risk scores, we defined 5 prognostic groups: very long, long, intermediate, short, and very short. To estimate the underlying survival time distribution, we fitted a parametric Royston-Parmar proportional hazard models with 3 internal knots (based on Akaike information criterion) per prognostic subgroup.<sup>19,20</sup> The interquartile range of the survival times was used to define survival time variability. Bootstrapping (n = 10,000) was used to estimate 95% confidence intervals (CIs) around the interquartile range. Variation in ALSFRS-R rates of decline was estimated by linear mixed-effects models, as described previously.<sup>6</sup> Longitudinal sample calculations by Eland (2009) determined the number of patients necessary to detect a 25% reduction in slope after 12 months (monthly follow-up) with 80% power and a 2-sided  $\alpha$  of 5%.<sup>6</sup> The sample size was used as an estimate of the sensitivity of the populations to detect a given treatment effect. All meta-analyses were performed with the R metafor package (version 2.0-0, Viechtbauer<sup>21</sup>). Linear mixed-effects models were fitted with the lmer function (lme4, version 1.1-12).<sup>22</sup> Imputation of missing data was performed using the aregImpute function (Hmisc, version 4.0-3).

### Data availability statement

All protocols, analyses, and anonymized data will be shared by request from any qualified investigator.

## Results

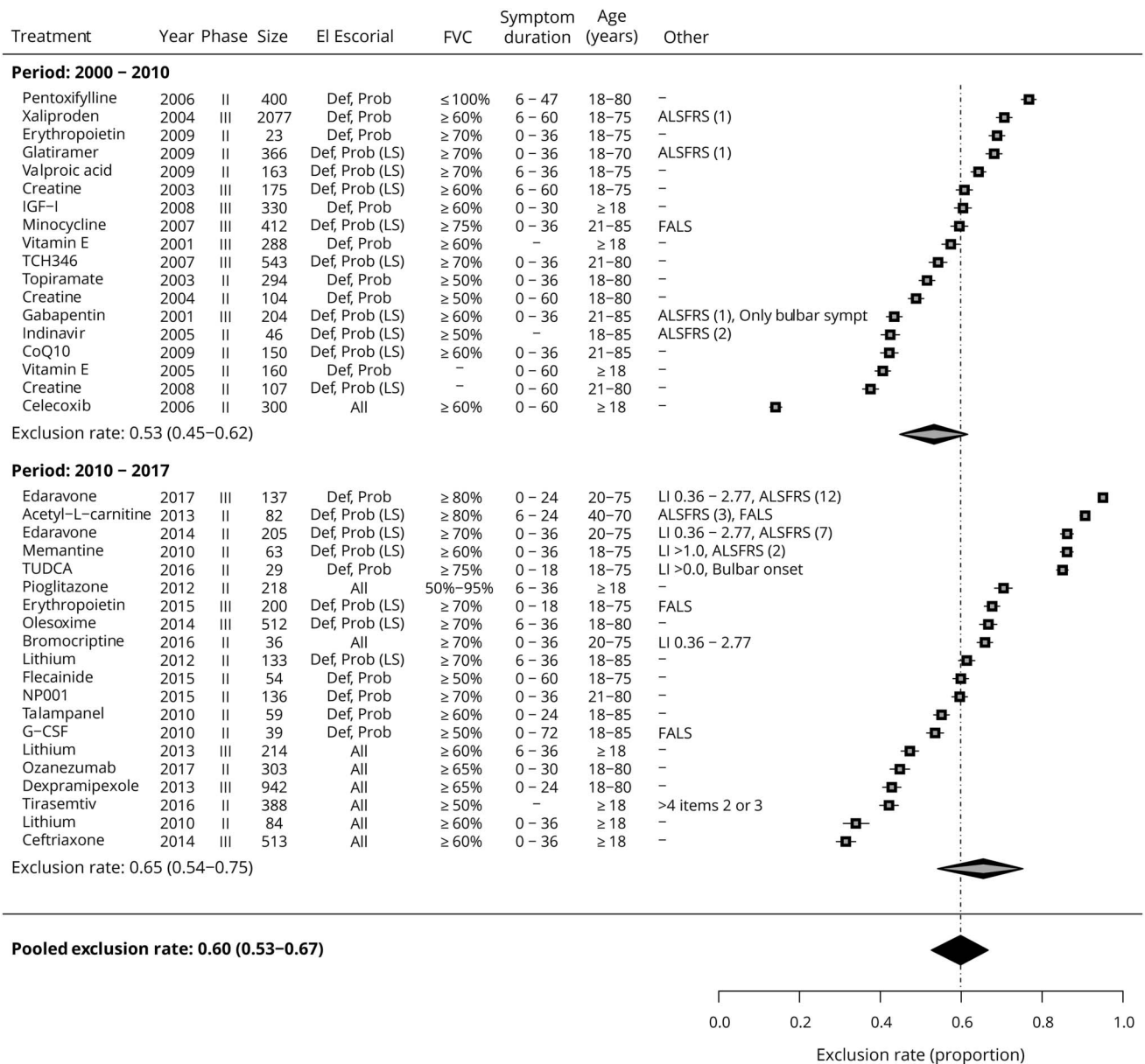
A total of 38 randomized placebo-controlled clinical trials were included in this study (figure e-2, available from Dryad doi.org/10.5061/dryad.86f1m6g); their eligibility criteria were applied to an incidence cohort of 2,904 patients with ALS. Figure 1 summarizes the primary eligibility criteria and exclusion rates per trial. On average, 59.8% (95% CI 52.6%–66.7%) of the patients would be excluded from trial participation on the day of diagnosis in the incidence cohort.

Exclusion rates varied widely between trials, ranging from 14% to 95%. Between 2000 and 2010, 53.3% of the patients were excluded, which increased to 65.5% between 2010 and 2017 ( $p = 0.08$ ). Not meeting a specific El Escorial category is the most important reason for exclusion (23%, 95% CI 18%–28%), followed by FVC (17%, 95% CI 14%–20%) and disease duration (12%, 95% CI 8%–15%) (figure e-1 available from Dryad, doi.org/10.5061/dryad.86f1m6g).

The enrolled populations of the 38 trials are presented in figure 2 and reveal a large variability between trials of all

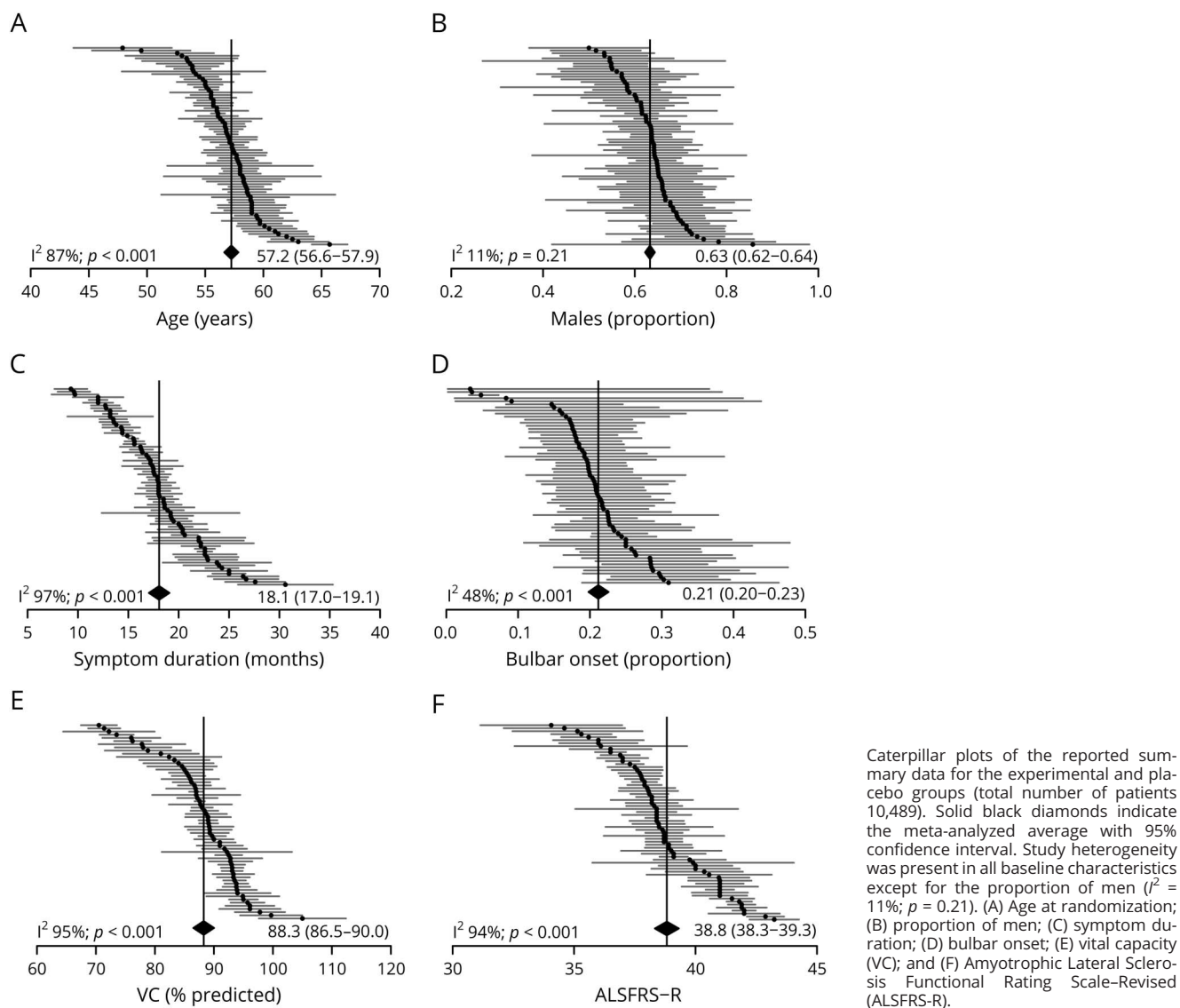
patient characteristics ( $p < 0.001$ ), except for the proportion of men ( $p = 0.21$ ). If we select those patients who are eligible for >50% of the trials from our general population (eligible population in table 2) and compare them with actual trial participants at our center (trial participants in table 2), large differences can be seen. Actually enrolled patients ( $n = 260$ , 9.0%, table 2, column 4) differ from their eligible population in sex (more men,  $p = 0.019$ ), age (younger,  $p < 0.001$ ), progression rate (slower,  $p < 0.001$ ), ENCALs risk profile (better,  $p < 0.001$ ), and survival (longer,  $p < 0.001$ ). This suggests that, despite the already applied eligibility criteria,

**Figure 1** Overview of eligibility criteria and exclusion rates of ALS trials conducted between 2000 and 2017



Caterpillar plot of the exclusion rates at diagnosis per trial ( $n = 38$ ). Exclusion rates ranged from 14% to 95%; numerical results per trial are given in table e-1 (available from Dryad, doi.org/10.5061/dryad.86f1m6g). ALS = amyotrophic lateral sclerosis; ALSFRS ( $n$ ) = number of selection items on Amyotrophic Lateral Sclerosis Functional Rating Scale; Def = definite; FALS = familial ALS; FVC = forced vital capacity, percent predicted; G-CSF = granulocyte colony-stimulating factor; IGF-1 = insulin-like growth factor-1; LI = lead-in ALSFRS-R slope; LS = probable, laboratory supported; Prob = probable.

**Figure 2** Between-trial variability in population characteristics of enrolled participants



only a selective subset of the eligible patients will participate in trials. This finding is indicative of an additional latent selection process.

Eligible patients are, compared to the general ALS population, 11% less likely to die during the first 24 months after diagnosis (pooled hazard ratio 0.89, 95% CI 0.84–0.93,  $p < 0.001$ ). When we stratify the eligible population into prognostic subgroups by applying a personalized prediction model, we can see that defining eligibility leads primarily to a temporary survival difference in the poorest prognostic group (6-month survival increases from 65.5% to 81.6%, figure 3). Large proportions (45%–66%) of other prognostic subgroups are also excluded, while the effect on survival is minimal (an absolute 2.3%–4.6% increase at 12 months and 1.5%–3.6% increase at 18 months). Within the eligible population, 12.7% of patients still have a very poor prognosis and 16.3% are very

long survivors. Results were similar for the individual trials (figure e-3 available from Dryad, doi.org/10.5061/dryad.86f1m6g), revealing a rather random exclusion process within all prognostic subgroups. Overall, eligibility criteria reduced heterogeneity in survival time by only 6.9% (95% CI –3.5% to 16.3%,  $p = 0.09$ ), reducing survival time variability from 22.0 months (95% CI 20.8–23.4) to 20.5 months (95% CI 18.8–22.4). This further underscores that both very short- and long-surviving patients remain in the trial population.

Longitudinal ALSFRS-R scores up to 24 months after diagnosis were available for 696 patients. Patients with a poor prognosis are relatively underrepresented in the ALSFRS-R data (bar charts, figure 4A). Between-patient ALSFRS-R variability does not differ between unselected and trial-eligible patients (slope variance 0.62 vs 0.65,  $p = 0.25$ ). The average ALSFRS-R rate of decline in eligible patients, however, is

**Table 2** Population characteristics of various populations

	General population	Eligible patients	Trial participants	
	(n = 2,904)	(n = 1,194)	Trial participants all trials (n = 10,489)	Trial participants UMCU (n = 260)
Age at onset, y	63.8 (63.4–64.2)	61.4 (60.9–62.0)	55.9 (55.2–56.7)	55.8 (54.4–57.3)
Male, % (n)	57.8% (1,678)	54.4% (649)	63.3% (6,640)	62.3% (162)
Bulbar site of onset, % (n)	37.1% (1,078)	35.1% (420)	21.2% (2,220)	31.4% (82)
Diagnostic delay, mo	10.0 (9.7–10.3)	9.4 (9.1–9.7)	—	8.7 (8.1–9.3)
ALSFRS-R score at diagnosis	39.0 (38.8–39.3)	39.7 (39.4–40.0)	—	41.5 (41.0–42.0)
ΔFRS, mo	0.82 (0.79–0.85)	0.81 (0.78–0.85)	—	0.68 (0.62–0.74)
FVC at diagnosis, % predicted	86.5 (85.5–87.4)	92.3 (91.1–93.5)	—	99.4 (97.4–101.4)
ENCALS risk profile	1.0 (0.94–1.07)	0.98 (0.91–1.05)	—	0.62 (0.51–0.74)
Survival after diagnosis (median), mo	17.7 (17.0–18.4)	19.4 (18.6–20.4)	—	27.8 (25.6–31.4)
ALSFRS-R slope after enrollment	—	—	0.97 (0.92–1.02)	1.11 (1.01–1.21)

Abbreviations: ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised; ΔFRS = (48 – ALSFRS-R score diagnosis)/symptom duration; ENCALs = European Network for the Cure of ALS; FVC = forced vital capacity; UMCU = University Medical Center Utrecht. ENCALs risk profiles are relative risk estimated by the ENCALs model and can be interpreted as hazard ratio. A relative risk of 0.62 means that, on average, the risk of dying during follow-up is 38% lower compared to the general population. General population includes all consecutive patients diagnosed with amyotrophic lateral sclerosis (ALS) in the Netherlands between 2006 and 2016. Eligible patients are patients in the Netherlands who are eligible for >50% of the trials. Trial participants all trials are meta-analyzed population characteristics of the 38 trials. Trial participants UMCU are patients from the general ALS population who participated in a clinical trial during the same time period at the UMCU. Data are mean (95% confidence interval) or percent (n).

higher (0.91 vs 0.98,  $p = 0.14$ ), which could reduce the sample size by 5.3% (from 125 to 118 patients per arm). Figure 5 shows the sample size reductions for the 38 trials relative to their eligibility rate, revealing a strong relationship between both variables (Pearson  $r$  0.58, 95% CI 0.32–0.76,  $p < 0.001$ ). The sample size estimates for trial populations defined by the ENCALs risk score are shown in black and gray. A sample size reduction of 34%, similar to Edaravone,<sup>23</sup> can be achieved by selecting only those patients with risk scores between 0.55 and 3.3. This single selection criterion would mean that ≈48% of the patients remain eligible, which would lead to an almost 5-fold increase in eligibility rate compared to Edaravone. None of the trials was more efficient than the prediction model in optimizing both the sensitivity and eligibility rate of the populations.

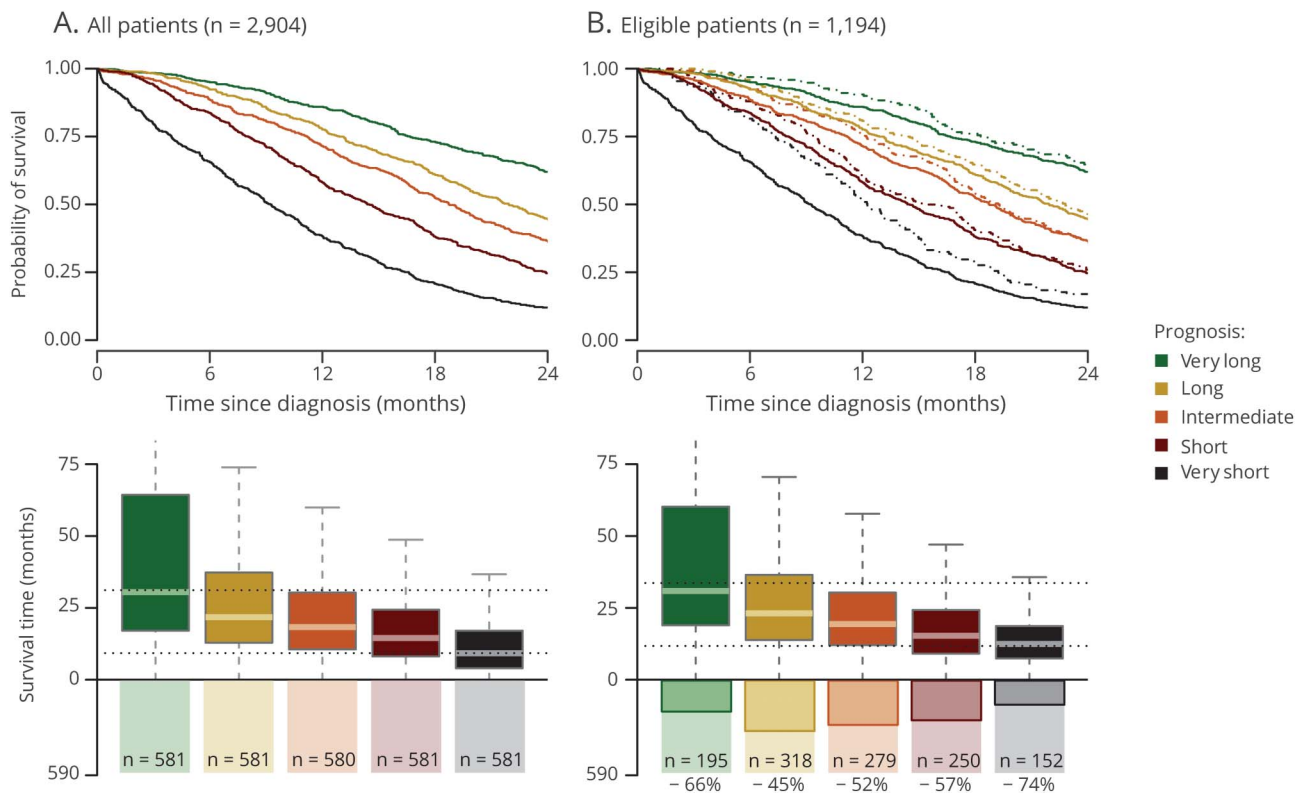
## Discussion

In this study, we show that on average 59.8% of the patients with ALS are found to be ineligible to participate in clinical trials. Although eligibility criteria reduce the number of patients with a poor prognosis, there is an adverse exclusion process that leads to a substantial untargeted exclusion of patients from all prognostic subgroups. Moreover, currently applied eligibility criteria select populations that may still contain a relatively high number of patients who show slow or fast progression rates and do not reduce between-patient variability. These findings raise questions regarding not only the value of currently applied eligibility criteria but also the generalizability of clinical trial results in ALS. Using prediction

models could individualize participant selection and optimize the balance between endpoint heterogeneity and the generalizability of trial results.

The concept of generalizability plays a central role in the translation of trial results to medical decision making.<sup>4</sup> Clinical trials with highly selected subgroups are difficult to interpret in real-world settings, and the safety or effectiveness of a drug may be unknown for the majority of the patients. We show that 59.8% of patients are excluded from participation at diagnosis. This percentage is, however, an underestimation because most patients will be enrolled a few months after diagnosis (6–9 months). At that time, 15% to 24% of our patients, who could theoretically be prescribed the drug at diagnosis, are deceased and are thus never evaluated in clinical trials. Moreover, a larger proportion of the remaining patients will fail the criteria due to disease progression. Together, these may lead to exclusion rates in real-world settings that approximate those reported in other fields (80%–96%).<sup>4,5,24,25</sup> These high exclusion rates could result in the indirect removal of patients with specific drug-responsive pathways. For example, patients with ALS–frontotemporal dementia and familial ALS are often excluded; however, because these subtypes are related to the *C9orf72* repeat expansion,<sup>26</sup> they are indirectly related to *C9orf72* disease pathways.<sup>27</sup> As was shown recently,<sup>28</sup> it is possible that the treatment effect is modified by pharmacogenetic interactions, which could be missed by the indirect exclusion of specific subgroups and may disguise important treatment clues. The same may hold true for the larger exclusion rates among bulbar-onset patients or women.

**Figure 3** Effect of selection criteria on overall mortality since diagnosis



Mortality since diagnosis for 5 prognostic subgroups defined by the European Network for the Cure of ALS personalized prediction model. Colors represent the 5 prognostic subgroups: very long (green), long (yellow), intermediate (orange), short (red), and very short (black) survival.<sup>7</sup> Solid lines in the Kaplan-Meier plots represent the survival patterns for all patients (n = 2,904; A), whereas the dotted lines are the survival curves for the eligible population (n = 1,194; B), defined by all patients who are eligible for > 50% of the trials. Boxplots provide the variability in survival time per prognostic subgroup (left, all patients; right, eligible patients). Dotted lines in boxplot are the unstratified interquartile ranges (25th and 75th percentile); overall survival time variability was reduced by 6.9% (95% confidence interval -3.5% to 16.3%,  $p = 0.09$ ). Bar charts provide the number of patients in each subgroup.

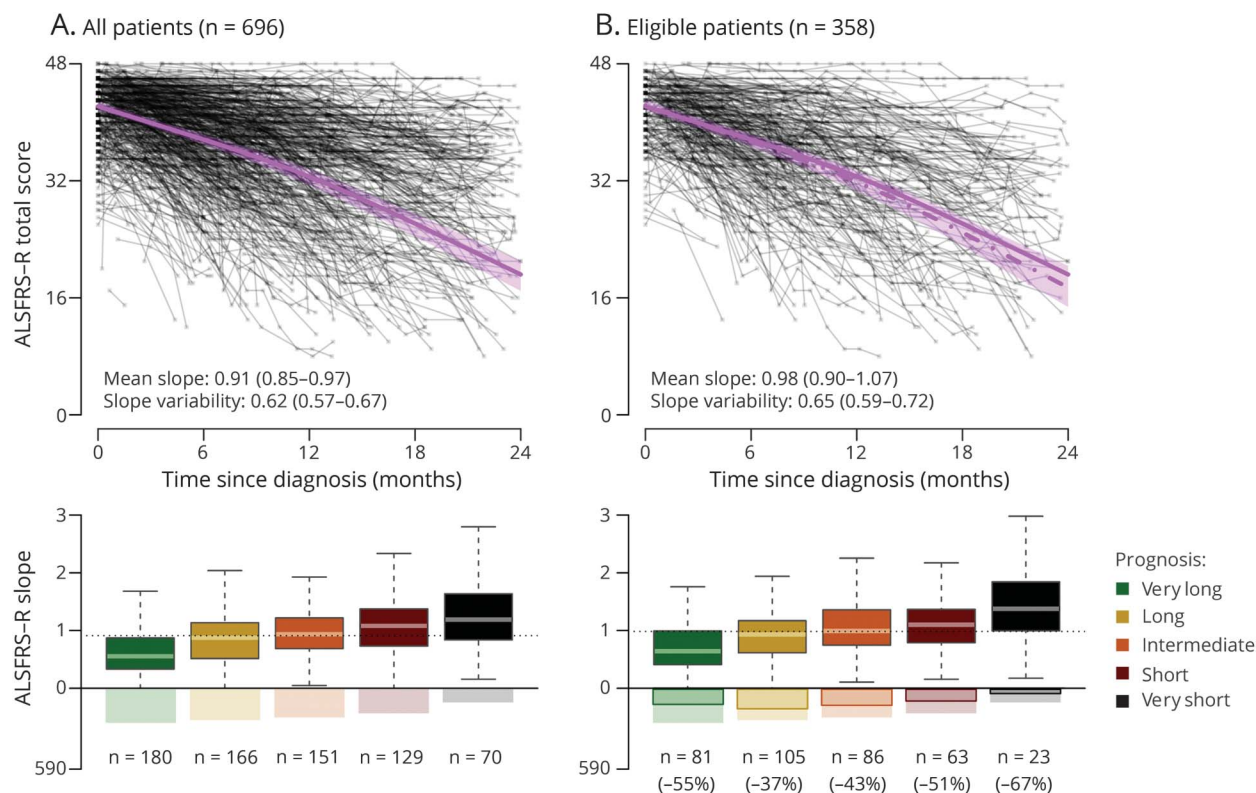
Because of the heterogeneous nature of ALS,<sup>29</sup> eligibility criteria aim primarily to reduce the amount of between-patient variation and to improve protocol adherence.<sup>2,3</sup> We show, however, that currently defined populations still contain fast- and slow-progressing patients and between-patient variation is virtually unaffected. A likely explanation is that selection criteria are applied in a step-wise, univariate manner (table 1), while progression rate is dictated by prognosis,<sup>30</sup> which is defined by a multivariate combination of predictors.<sup>7</sup> When a trial aims to exclude fast-progressing patients, patients who are on the lower limits of each criterion could still be enrolled. The sum of the lower limits, however, means that these patients have a poor prognosis and fast progression rates. To exemplify, a trial with only 2 criteria (age <75 years and FVC >60%) would enroll a 74-year-old patient with an FVC of 61% but exclude a 76-year-old with an FVC of 104%. The first patient is likely to exhibit a faster rate of decline and is more likely to die during follow-up. More important, the real-world effect is minimal. The dexrampipexole study,<sup>31</sup> for example, excluded patients with a disease duration >24 months and an FVC <65%. However, if none of the criteria had been applied, the statistical power would have been reduced from 90% to 88% (assuming an inflation of the

reported SD of 3%). It is doubtful whether a 2% gain in power justifies the exclusion of 25% to 40% of the patients.

This questions the value of currently used eligibility criteria, and a revision would seem to be indicated. There is a need to balance endpoint heterogeneity (or sample size) and the generalizability of trials. This balance could be achieved by using individual risk scores rather than group-level criteria. The risk score can be conceptualized as a summary of all available prognostic information per individual. Replacing sets of several eligibility criteria by a single risk estimate would allow investigators to select only those patients who are the most likely to exhibit the investigator-preferred disease pattern. This could reduce between-patient heterogeneity more effectively while balancing both generalizability and eligibility. In our ALSFRS-R example for Edaravone,<sup>23</sup> instead of excluding 90% of the patients, the same homogeneity effect could be reached using a risk-based selection with 48% of the patients remaining eligible (a nearly 5-fold increase).

Our study has several limitations that should be considered. First, because populations may differ between countries, our

**Figure 4** Longitudinal ALSFRS-R patterns in a subset of 696 patients



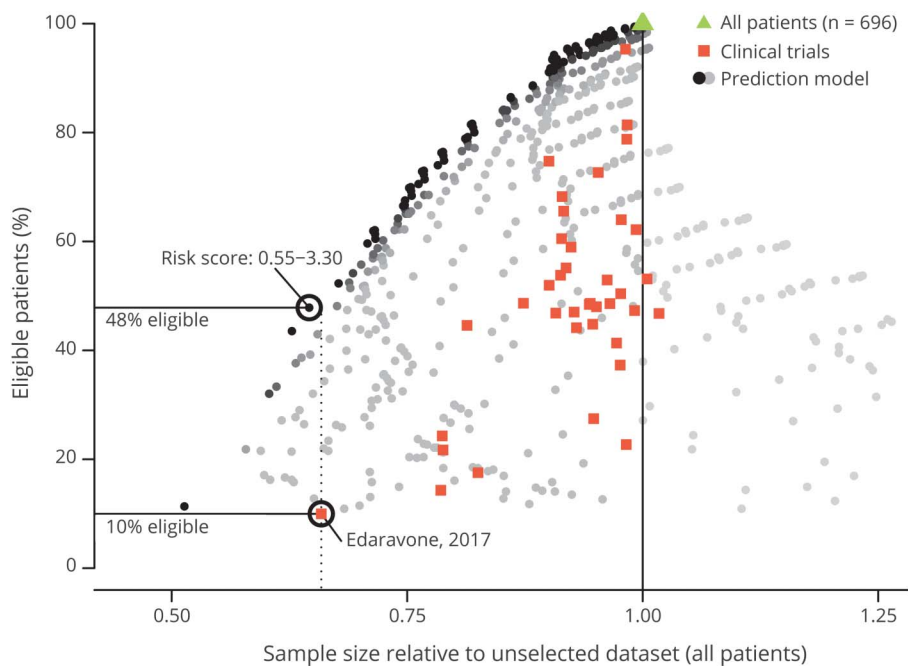
Colors of the boxplots represent the 5 prognostic subgroups: very long (green), long (yellow), intermediate (orange), short (red), and very short (black) survival.<sup>7</sup> Purple dotted line is the average pattern of decline in the eligible population (n = 356). Solid purple line is the average pattern in the full dataset (n = 696). Boxplots provide the variability in rates of decline over time (points per month), estimated by the best unbiased linear predictors (BLUPs) from a linear mixed-effects model. Dotted lines in boxplots represent the average rate of decline. (A) All patients and (B) eligible patients. ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

results are limited to a specific geographic area. In Italy and Ireland, for example, the age at onset is slightly higher, which would underestimate our current exclusion rate.<sup>32</sup> Second, this work focused primarily on optimizing the balance between endpoint heterogeneity and generalizability of trial results. However, ALS remains a complex disorder, and investigators may apply eligibility criteria with imperatives other than simply selecting the most responsive subgroup (e.g., safety, pharmacodynamics, or hypothesized mechanism or action). It may therefore be insightful to distinguish between biological eligibility criteria (those that are set for the hypothesized drug mechanism) and design criteria (those that are set to optimize the trial design). This is especially important when considering the current developments toward a personalized, genotype-oriented approach in ALS.<sup>28,29</sup> Genotype- or biomarker-oriented trials inherently have high exclusion rates and potentially investigate (ultra) rare subgroups, which compromises their feasibility. The antisense trial in *SOD1*-related ALS,<sup>33</sup> for example, had an enrollment rate of 1 patient per 4 months per site, which underscores the importance of optimizing the eligibility rate. Therefore, future clinical trials could combine both group-level biological (e.g., genetic marker or disease pathway) and individualized risk-based criteria to optimize both drug responsiveness and trial design.

As final note, our results, supported by previous studies,<sup>2,34</sup> indicate a latent selection of patients in which young male patients with a relatively mild disease are overrepresented in trial populations. This latent selection process is most probably the result of a multifactorial process, one that cannot be estimated in the meta-data of published clinical trials. The reason for eligible patients with ALS declining study participation is in 94% of cases the physical burden.<sup>9</sup> Given the relative underrepresentation of patients in advanced stages, it is plausible that the physical burden is an important latent factor, one that becomes more apparent as the disease progresses. This is supported by the observation that dropout and noncompliance are related to lower ALSFRS-R and FVC scores.<sup>3</sup> Similarly, the physician may deem the patient unfit to undergo the trial and not offer the option of participating.<sup>9</sup> Sex differences may also play a role; female patients are more often reluctant about medical testing or have an inability to cope with the protocol.<sup>35</sup> These latent factors may be overexpressed when 2 trials run simultaneously at the same site. The patient's choice for a certain trial could be related to a certain characteristic of the patient (e.g., sex or age). Dedicated studies are needed to determine the relevance of each factor and to potentially develop strategies to reduce this latent selection process (e.g., by reducing the physical burden



**Figure 5** Effect of trial eligibility criteria and the ENCALs risk scores on sample size



Sample size calculations were performed for various populations, either selected by the trial eligibility criteria ( $n = 38$ , orange) or defined by different cutoffs for the European Network for the Cure of ALS (ENCALS) risk scores ( $n = 1,050$  combinations). Results are expressed as sample size inflation factor (IF; x-axis). To exemplify, an IF of 0.79 means that the sample size is 21% smaller compared to the sample size necessary when all patients are included (green triangle, 100% eligibility). Populations defined by the ENCALs risk scores for all possible combinations of cutoff values are shown in black and gray. Black dots are the selected populations that resulted in the largest reduction in sample size and highest eligibility rate. To exemplify, the Edaravone 2017 trial resulted in the largest reduction in sample size ( $-34\%$ , IF 0.66), with a 10% eligibility rate. The ENCALs model could select a similarly sensitive population with 48% of the patients remaining eligible. Numerical effects of the different trials are provided in table e-1 (available from Dryad, doi.org/10.5061/dryad.86f1m6g).

using home-based outcome measures or patient-reported data).<sup>36</sup>

Our results reveal that the majority of patients with ALS are excluded from trial participation at diagnosis, which raises questions regarding the generalizability of current trials. Exclusion of ineligible patients only minimally improves homogeneity in trial endpoints. A risk-based selection criterion could individualize trial participant selection and may improve the balance between endpoint heterogeneity and exclusion rates.

### Author contributions

Design or conceptualization of the study: R.P.A.v.E., H.-J.W., L.H.v.d.B. Analysis or interpretation of the data: R.P.A.v.E., H.-J.W., S.N., I.E.V., M.J.C.E. Drafting or revising the manuscript for intellectual content: all authors.

### Study funding

Supported by the Netherlands ALS Foundation (Project TryMe).

### Disclosure

R. van Eijk, H. Westeneng, S. Nikolakopoulos, I. Verhagen, and M. van Es report no disclosures relevant to the manuscript. M. Eijkemans received grants from The Netherlands Organization for Health Research and Development (Veni scheme), the Thierry Latran Foundation, and The Netherlands ALS Foundation (Stichting ALS Nederland). He received travel grants from Baxalta and serves on the biomedical research advisory panel of the UK Motor Neurone Disease Association. L. van den Berg reports grants from The Netherlands ALS Foundation,

The Netherlands Organization for Health Research and Development (Vici scheme), and The Netherlands Organization for Health Research and Development (Sampling and Biomarker Optimization and Harmonization in ALS and Other Motor Neuron Diseases [SOPHIA], Survival, Trigger and Risk, Epigenetic, Environmental, and Genetic Targets for Motor Neuron Health [STRENGTH], A Programme for ALS Care in Europe [ALS-CarE] project), funded through the EU Joint Programme—Neurodegenerative Disease Research), and serving on the Scientific Advisory Board for Biogen, Cytokinetics, Prinses Beatrix SpierFonds, and the Thierry Latran Foundation. Go to Neurology.org/N for full disclosures.

### Publication history

Received by *Neurology* June 22, 2018. Accepted in final form September 19, 2018.

### References

1. Mitsumoto H, Brooks BR, Silani V. Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol* 2014;13:1127–1138.
2. Chio A, Canosa A, Gallo S, et al. ALS clinical trials: do enrolled patients accurately represent the ALS population?. *Neurology* 2011;77:1432–1437.
3. Atassi N, Yerramilli-Rao P, Szymonifka J, et al. Analysis of start-up, retention, and adherence in ALS clinical trials. *Neurology* 2013;81:1350–1355.
4. Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?” *Lancet* 2005;365:82–93.
5. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *Am J Psychiatry* 2002;159:469–473.
6. van Eijk RPA, Eijkemans MJC, Ferguson TA, Nikolakopoulos S, Veldink JH, van den Berg LH. Monitoring disease progression with plasma creatinine in amyotrophic lateral sclerosis clinical trials. *J Neurol Neurosurg Psychiatry* 2018;89:156–161.
7. Westeneng HJ, Debray TPA, Visser AE, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol* 2018;17:423–433.
8. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.

9. Bedlack RS, Pastula DM, Welsh E, Pulley D, Cudkowicz ME. Scrutinizing enrollment in ALS clinical trials: room for improvement? *Amyotroph Lateral Scler* 2008;9:257–265.
10. Pinto S, de Carvalho M. Comparison of slow and forced vital capacities on ability to predict survival in ALS. *Amyotroph Lateral Scler Frontotemporal Degener* 2017;18:528–533.
11. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol* 2005;5:13.
12. Fu R, Vandermeer BW, Shamliyan TA, et al. *Handling Continuous Outcomes in Quantitative Synthesis: Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville; Agency for Healthcare Research and Quality; 2008.
13. 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;82:1165–1170.
14. 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 Mot Neuron Disord* 2000;1:293–299.
15. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med* 2009;28:1982–1998.
16. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med* 1991;10:585–598.
17. Higgins JPT, Green S; Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley-Blackwell; 2008.
18. Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Stat Med* 2003;22:2693–2710.
19. 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;21:2175–2197.
20. Rooney J, Fogh I, Westeneng HJ, et al. C9orf72 expansion differentially affects males with spinal onset amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry* 2017; 88:281–287.
21. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Software* 2010;36:48.
22. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Software* 2015;67:1–48.
23. Edaravone Writing Group. Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol* 2017;16:505–512.
24. Zetin M, Hoepner CT. Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *J Clin Psychopharmacol* 2007;27:295–301.
25. Hoertel N, Le Strat Y, De Maricourt P, Limosin F, Dubertret C. Are subjects in treatment trials of panic disorder representative of patients in routine clinical practice? Results from a national sample. *J Affect Disord* 2013;146:383–389.
26. Umoh ME, Fournier C, Li Y, et al. Comparative analysis of C9orf72 and sporadic disease in an ALS clinic population. *Neurology* 2016;87:1024–1030.
27. Wen X, Westergard T, Pasinelli P, Trotti D. Pathogenic determinants and mechanisms of ALS/FTD linked to hexanucleotide repeat expansions in the C9orf72 gene. *Neurosci Lett* 2017;636:16–26.
28. van Eijk RPA, Jones AR, Sproviero W, et al. Meta-analysis of pharmacogenetic interactions in amyotrophic lateral sclerosis clinical trials. *Neurology* 2017;89: 1915–1922.
29. van Es MA, Hardiman O, Chio A, et al. Amyotrophic lateral sclerosis. *Lancet* 2017; 390:2084–2098.
30. van Eijk RPA, Eijkemans MJC, Rizopoulos D, van den Berg LH, Nikolakopoulos S. Comparing methods to combine functional loss and mortality in clinical trials for amyotrophic lateral sclerosis. *Clin Epidemiol* 2018;10:333–341.
31. Cudkowicz 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;12:1059–1067.
32. 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;18:411–418.
33. Miller TM, Pestronk A, David W, et al. An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *Lancet Neurol* 2013;12: 435–442.
34. 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;88:557–563.
35. Markanday S, Brennan SL, Gould H, Pasco JA. Sex-differences in reasons for non-participation at recruitment: Geelong Osteoporosis Study. *BMC Res Notes* 2013;6: 104.
36. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nat Biotechnol* 2011;29:411–414.